The Evolution of Labor Earnings Risk in the U.S. Economy*

Flavio Cunha[†] and James Heckman[‡]

December 29, 2005

Abstract

1 Introduction

A large body of literature has documented an increase in wage inequality over the 1970's and 1980's (see, for example, Levy and Murnane, 1992). This increase in wage inequality has occurred both within and between education-experience groups. While the former unfolds during the 1970's, the latter is experienced over the 1980's. Recent studies show that these trends are weakened in the 1990's.

The analysis of Juhn, Murphy and Pierce (1993) links the increase in wage inequality to technological progress. As technology advances the demand for ability grows at a faster pace than its supply causing skill prices to increase and the wage gap between skilled and unskilled workers to widen. To arrive to this conclusion they map quantiles of the distribution of the residuals in log wage equations to quantiles of the distribution of ability. This is possible provided the following four conditions hold: 1) there is one, and

*This research was supported by NIH R01-HD043411. Cunha is grateful to sponsorship from CAPES 1430/99-8 and the Claudio Haddad Dissertation Fund at the University of Chicago.

[†]University of Chicago, Department of Economics.

[‡]University of Chicago, Department of Economics and the American Bar Foundation.

only one, type of unobserved ability; 2) the distribution of this ability across individuals is invariant over time; 3) there are no unantecipated shocks in earnings so that the economy is described by a deterministic environment; 4) wages are measured without error.

Regarding the first condition, Heckman and Rubinstein (1997), Carneiro and Heckman (2003), and Heckman, Stixrud, and Urzua (2005) have shown that abilities are multiple in nature. These abilities may be combined in different quantities to produce the same outcome. For example, an individual with low stocks of cognitive ability but a considerable amount of persistence can be more successful than a smart person with no motivation. Hence, the one-to-one mapping from quantiles of the distribution of wage residuals to quantiles of the distribution of abilities is not meaningful.

The evidence presented by Gottschalk and Moffitt (1994) challenges the view that the distribution of ability has remained fixed over time. They show substantial increase in the variance of the distribution of unobserved heterogeneity when they compare the period 1970-1978 to the period 1979-1987.

Finally, Gottschalk and Moffitt (1994) documented an increase in "earnings instability": the variance of temporary shocks rose considerably from the period 1970-1978 to the period 1979-1987. Thus, their findings is consistent with uncertainty playing an important role in the economic envrironment that generated the increase in wage inequality.

All in all, any credible explanation for the rise in wage inequality has to consider the fact that individuals possess an array of abilities, that the joint distribution of these abilities could be changing over time, and that uncertainty is a major reality in the economic environment. In this paper, we build on the work of Carneiro, Hansen and Heckman (2003) and Cunha, Heckman and Navarro (2005) to separate heterogeneity from uncertainty in labor earnings and show how they have evolved over time. The essential idea is to model schooling and earnings equations jointly to identify the information set of the agent at the time of the schooling choices were made. Modelling schooling choices is not merely an econometric procedure to correct for selection in observed earnings. It is the source of information that allow us to separate what is known and acted on by individuals at the time the schooling choice is made (which we call heterogeneity) from what is not known (which we call uncertainty). Our approach can be viewed as an extension of the Granger-Sims causality test (in which future outcomes cause present decisions) to an economic setting where estimation of counterfactual outcomes must be made.

The intuition of our approach can be made clear in a few paragraphs. We seek to decompose the "returns" coefficient or the gross gains from schooling in an earnings-schooling model into components that are known at the time schooling choices are made and components that are not known. For simplicity assume that, for person i, returns are the same at all levels of schooling. Write discounted lifetime earnings of person i, E_i, as

$$E_i = \alpha + \rho_i S_i + U_i, \tag{1}$$

where ρ_i is the person-specific *ex post* return, S_i is years of schooling, and U_i is a mean zero unobservable. We need to separate ρ_i into two components $\rho_i = \eta_i + \nu_i$, where η_i is a component known to the agent when he/she makes schooling decisions and ν_i is revealed after the choice is made. Schooling choices are assumed to depend on what is known to the agent at the time decisions are made, $S_i = \lambda\left(\eta_i, Z_i, \tau_i\right)$, where the Z_i are other observed determinants of schooling and τ_i represents additional factors unobserved by the analyst but known to the agent. Both of these variables are in the agent's information set at the time schooling choices are made. We seek to determine what components of *ex post* lifetime earnings Y_i enter the schooling choice equation.

If η_i is known to the agent and acted on, it enters the schooling choice equation. Otherwise it does not. Component ν_i and any measurement errors in E_i should not be determinants of schooling choices. Neither should future skill prices that are unknown at the time agents make their decisions. If agents do not use η_i in making their schooling choices, even if they know it, η_i would not enter the schooling choice equation. Determining the correlation between realized Y_i and schooling choices based on *ex ante* forecasts enables economists to identify components known to agents and acted on in making their schooling decisions. Even if we cannot identify ρ_i, η_i, or ν_i for each person, under conditions specified in this chapter we can identify their distributions. The question is how to pick the information set.

Suppose that the model for schooling can be written in linear in parameters form:

$$S_i = \lambda_0 + \lambda_1 \eta_i + \lambda_2 \nu_i + \lambda_3 Z_i + \tau_i, \tag{2}$$

where τ_i has mean zero and is assumed to be independent of Z_i. The Z_i and the τ_i proxy costs and may also be correlated with U_i and η_i and ν_i in (1). In this framework, the goal of the analysis is to determine the η_i and ν_i components. By definition, $\lambda_2 = 0$ if ν_i is not known when agents make their schooling choices.

The method developed by [**?**, **?**, **?**] exploits the covariance between S and the realized Y_t to determine which components of Y_t are known at the time schooling decisions are made. It explicitly models selection bias and allows for measurement error in earnings. It does not rely on linearity of the schooling model. Their method recognizes the discrete nature of the schooling decision. It builds on the modern literature on constructing counterfactual schooling models.

Use this paragraph to describe our main findings.

Use this paragraph to explain how the paper is organized.

2 The Economic Model

2.1 The Problem of the Agent

We describe an uncertain economy populated with I individuals. Each individual lives for T periods. Before any uncertainty is realized, agents choose their schooling level and how to allocate consumption across states of nature and over time. Individuals supply labor inelastically, but their labor productivity is subject to idiosyncratic uncertainty. The variance of labor productivity can vary across schooling levels and over time, but agents can protect themselves against this uncertainty because they have access to a complete set of Arrow-Debreu securities. In order to obtain the solution of the model we first compute the solution to the consumption allocation problem for each possible schooling level of the agent. Next, we solve the schooling choice problem taking as given the consumption allocation in each schooling level.

2.1.1 The Consumption Allocation Problem

In each period t there is a realization of a stochastic event $\omega_t \in \Omega$. Let the histories of events up to and until time t be denoted $\omega^t = \{\omega_1, \omega_2, ..., \omega_t\}$. We use $\pi_t(\omega^t)$ to denote the unconditional probability of observing a particular sequence of events ω^t. This history ω^t is publicly observable.

Agent i's labor productivity is stochastic. We use $Y_{i,s,t}(\omega^t)$ to denote the labor productivity of an individual i who has schooling level s given history ω^t. We denote by $S = \{0, 1\}$ the set of possible schooling levels that the agent can choose. In our empirical study, $S = 0$ indicates an individual who is a high-school graduate while $S = 1$ indicates an individual who is a college graduate.

Let $c_{i,s,t}(\omega^t)$ denote the consumption of an agent i with schooling level s at period t given history ω^t. Consumption goods can be produced according to a constant returns to scale technology that depends only on aggregate labor. Consequently, the price of the consumption good is the same as the price of labor at all periods t for all possible histories ω^t. In each period, at each state of nature, the feasibility condition is such that:

$$\sum_{s \in S} \sum_{i \in s} c_{i,s,t}\left(\omega^t\right) = \sum_{s \in S} \sum_{i \in s} Y_{i,s,t}\left(\omega^t\right)$$

We assume that there is neither aggregate uncertainty nor deterministic trends in labor productivity so that:

$$\sum_{s\in S}\sum_{i\in s} Y_{i,s,t}\left(\omega^{t}\right) = \bar{Y} \text{ for all } t \text{ and } \omega^{t}$$

Let $q_t\left(\omega^t\right)$ denote the price of an Arrow-Debreu security that delivers one unit of period-t consumption good if the history ω^t is realized and zero otherwise. The consumption allocation problem of the agent is:

$$V\left(s\right) = Max\sum_{t=1}^{T}\sum_{\omega^{t}}\left(\frac{1}{1+\rho}\right)^{t}\pi_{t}\left(\omega^{t}\right)u\left[c_{i,s,t}\left(\omega^{t}\right)\right] \tag{3}$$

subject to:

$$\sum_{t=1}^{T}\sum_{\omega^{t}} q_{t}\left(\omega^{t}\right)c_{i,s,t}\left(\omega^{t}\right) = \sum_{t=1}^{T}\sum_{\omega^{t}} q_{t}\left(\omega^{t}\right)Y_{i,s,t}\left(\omega^{t}\right) \tag{4}$$

Let $\lambda_{i,s}$ denote the Lagrange multiplier associated to the budget equation of an agent i with schooling level s. The first-order condition is:

$$\lambda_{i,s}q_{t}\left(\omega^{t}\right) = \left(\frac{1}{1+\rho}\right)^{t}\pi_{t}\left(\omega^{t}\right)u'\left[c_{i,s,t}\left(\omega^{t}\right)\right]$$

Because there is no aggregate uncertainty, the equilibrium consumption allocation must be such that $c_{i,s,t}\left(\omega^t\right) = c_{i,s}$ for all period t and possible history ω^t. Consequently:

$$q_{t}\left(\omega^{t}\right) = \left(\frac{1}{1+\rho}\right)^{t}\pi_{t}\left(\omega^{t}\right)\frac{u'\left[c_{i,s}\right]}{\lambda_{is}} \tag{5}$$

Replacing (5) into (4) one obtains:

$$c_{i,s} = A\left(\rho, T\right)Y_{i,s} \tag{6}$$

where $Y_{i,s} = E\left[\sum_t \left(\frac{1}{1+\rho}\right)^t Y_{i,s,t}\middle|\mathcal{I}\right]$ and $A\left(\rho,T\right) = \frac{\left(1-\left(\frac{1}{1+\rho}\right)\right)}{\left(\frac{1}{1+\rho}\right)\left(1-\left(\frac{1}{1+\rho}\right)^{T+1}\right)}$.

2.1.2 The Schooling Decision Problem

Given the solution to the consumption allocation problem, we can compute the lifetime utility of an agent that chooses s years of schooling by replacing (6) into (3):

$$V_i(s) = \frac{1}{A(\rho, T)} u\left[A(\rho, T) Y_{i,s}\right]$$

Let $\tilde{C}_i$ denote the (unobservable) psychic costs associated with schooling choices. An agent chooses to go to college, i.e., $s = 1$, if, and only if:

$$E\left\{V_i(1) - V_i(0) - \tilde{C}_i \middle| \mathcal{I}\right\} \geq 0$$

or

$$\frac{1}{A(\rho, T)} E\left\{u\left[A(\rho, T) Y_{i,1}\right] - u\left[A(\rho, T) Y_{i,0}\right] - C_i | \mathcal{I}\right\} \geq 0,$$

where $C_i = A(\rho, T) \tilde{C}_i$. At this point we should stress that given data on schooling choices and earnings, we can't separate the role of the utility function u from the role of unobservable heterogeneity in schooling preferences (captured by the psychic costs) C_i. To see this, suppose that we observed C_i. Then, holding C_i fixed, we would be able to infer preferences for consumption by looking at their schooling choices and earnings. However, because we do not observe C_i, we cannot compare people by holding C_i constant. Consequently, we cannot hope to identify both utility function u and psychic costs C_i. Because of this fact, we normalize the utility function $u(x) = x$. Under this normalization, the schooling choice becomes:

$$E\left\{\sum_t \left(\frac{1}{1+\rho}\right)^t Y_{i,1,t} - \sum_t \left(\frac{1}{1+\rho}\right)^t Y_{i,0,t} - C_i \middle| \mathcal{I}\right\} \geq 0$$

3 The Econometric Model

3.1 Model Specification

Our work consists on estimating the components of the information set of the agents. In particular, we are interested in pinning down unobservable components which are known and acted on by the agents. We can recover these components by looking both at the choices and the outcomes associated with choices made by the individuals.

3.1.1 Earnings Equations

To motivate our econometric procedures we start by describing the earnings equations. For $t = 1, 2, ..., T$ we assume that $(Y_{i,0,t}, Y_{i,1,t})$ have finite means and can be expressed in terms of conditioning variables $\boldsymbol{X}$ in the following manner:

$$Y_{i,0,t} = \boldsymbol{X}_i \beta_{0,t} + U_{i,0,t}, \tag{7}$$

$$Y_{i,1,t} = \boldsymbol{X}_i \beta_{1,t} + U_{i,1,t}, \tag{8}$$

The error terms $U_{i,s,t}$ satisfy $E\left(U_{i,s,t} \mid \boldsymbol{X}\right) = 0$.

3.1.2 Choice Equations

Write the index I as a net utility,

$$I_i = E\left[\sum_{t=1}^{T} \left(\frac{1}{1+\rho}\right)^t (Y_{i,1,t} - Y_{i,0,t}) - C_i \middle| \mathcal{I}\right], \tag{9}$$

where C_i is the utility costs of attending college. We denote by $\mathbf{Z}_i$ and $U_{C,i}$ the observable and unobservable determinants of psychic costs, respectively. We assume that psychic costs can be written as:

$$C_i = \mathbf{Z}_i\gamma + U_{C,i} \tag{10}$$

If we define $\mu_I(\boldsymbol{X}_i, \boldsymbol{Z}_i) = \sum_{t=1}^{T} \left(\frac{1}{1+\rho}\right)^t \boldsymbol{X}_i \left(\beta_{1,t} - \beta_{0,t}\right) - \boldsymbol{Z}_i\gamma$ and $U_{I,i} = \sum_{t=1}^{T} \left(\frac{1}{1+\rho}\right)^t (U_{i,1,t} - U_{i,0,t}) - U_{C,i}$, and replace (7), (8), and (10) into (9) we obtain:

$$I_i = E\left[\mu_I(\boldsymbol{X}_i, \boldsymbol{Z}_i) + U_{I,i} | \mathcal{I}\right]. \tag{11}$$

More generally, we define $U_{I,i}$ as the error in the choice equation and it may or may not include all future $U_{i,1,t}$, $U_{i,0,t}$, or $U_{C,i}$. Similarly, $\mu_I(\boldsymbol{X}_i, \boldsymbol{Z}_i)$ may only be based on expectations of future $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$ at the time schooling decisions are made. The schooling decision of the agents is such that:

$$S_i = 1 \text{ if } I_i \geq 0; \; S_i = 0 \text{ otherwise.} \tag{12}$$

3.1.3 Test Score Equations

Aside from earnings and choice equations we also estimate a set of cognitive test score equations. Let $M_{i,k}$, $k = 1, 2, ..., K$, denote the agent i's score on the k^{th} test. Assume that $M_{i,k}$ have finite means and can be expressed in terms of conditioning variables $\boldsymbol{X}^M$. Write:

$$M_{i,k} = \boldsymbol{X}^M \beta_{i,k}^M + U_{i,k}^M \tag{13}$$

The test equations are introduced here because we expect both the decision to graduate from college and realized earnings to depend on the stock of cognitive skills the agents has at the time of the schooling choice.

3.1.4 The goal of the paper

Consider the random variable college earnings $Y_{1,t}$. It is only observed for the agents who choose to graduate from college. Consequently, from observational data, we can compute the the cross-sectional mean college earnings conditional on explanatory variables $\mathbf{X}$ and $S = 1$:

$$E\left[Y_{1,t}\right| \mathbf{X}, S = 1] = \mathbf{X}\beta_{1,t} + E\left(U_{1,t}\right| \mathbf{X}, S = 1)$$

Assume that $\mathbf{X}, \mathbf{Z}, U_C \in \mathcal{I}$. The event $S = 1$ corresponds to the event:

$$E\left[U_I\right| \mathcal{I}] \geq -\mu_I(\mathbf{X}, \boldsymbol{Z})$$

Consequently, because $E\left[Y_{1,t}\right| \mathbf{X}, S = 1] = E\left[Y_{1,t}\right| \mathbf{X}, E\left[U_I\right| \mathcal{I}] \geq -\mu_I(\mathbf{X}, \boldsymbol{Z})]$, we can establish that:

$$E\left[Y_{1,t}\right| \mathbf{X}, E\left[U_I\right| \mathcal{I}] \geq -\mu_I(\mathbf{X}, \boldsymbol{Z})] = \mathbf{X}\beta_{1,t} + E\left(U_{1,t}\right| \mathbf{X}, E\left[U_I\right| \mathcal{I}] \geq -\mu_I(\mathbf{X}, \boldsymbol{Z}))$$

It is interesting to separate two components. The first component, $E\left[U_I\right| \mathcal{I}]$, is used by the agent to make schooling choices. This expectation is determined by the elements in the information set of the agent which, in the end, influences their schooling decision.

The second component, $U_I - E\left[U_I\right| \mathcal{I}]$, does not affect selection into schooling. To see why, add and subtract $E\left[U_I\right| \mathcal{I}]$ on the schooling choice equation:

$$I = E\left[\mu_I(\mathbf{X}, \boldsymbol{Z}) + U_I - E\left[U_I\right| \mathcal{I}] + E\left[U_I\right| \mathcal{I}]\right| \mathcal{I}] =$$

$$= \mu_I(\boldsymbol{X}, \boldsymbol{Z}) + E\left[U_I | \mathcal{I}\right] + E\left[\left(U_I - E\left[U_I | \mathcal{I}\right]\right) | \mathcal{I}\right] = \mu_I(\boldsymbol{X}, \boldsymbol{Z}) + E\left[U_I | \mathcal{I}\right]$$

because $E\left[\left(U_I - E\left[U_I | \mathcal{I}\right]\right) | \mathcal{I}\right] = 0$.

Under the assumption that $U_C \in \mathcal{I}$ we can write:

$$U_I - E\left[U_I | \mathcal{I}\right] = \sum_{t=1}^{T} \left(\frac{1}{1+\rho}\right)^t \left(U_{1,t} - E\left[U_{1,t} | \mathcal{I}\right]\right) + \sum_{t=1}^{T} \left(\frac{1}{1+\rho}\right)^t \left(U_{0,t} - E\left[U_{0,t} | \mathcal{I}\right]\right)$$

and it is easy to see that $\left(U_{s,t} - E\left[U_{s,t} | \mathcal{I}\right]\right)$ affects realized earnings. This can be seen by adding and subtracting $E\left[U_{s,t} | \mathcal{I}\right]$ to the earnings equation at school level s and period t :

$$Y_{s,t} = \boldsymbol{X}\beta_{s,t} + E\left[U_{s,t} | \mathcal{I}\right] + \left(U_{s,t} - E\left[U_{s,t} | \mathcal{I}\right]\right)$$

Consequently, we conclude that $E\left[U_{s,t} | \mathcal{I}\right]$ affects both earnings and school choice equations while the term $\left(U_{s,t} - E\left[U_{s,t} | \mathcal{I}\right]\right)$ affects only earnings in period t and school s equation. To determine the unobservable components that are in the information set of the agent we need to determine which specification of the information set $\mathcal{I}$ fits better the covariance between schooling choices and earnings. We can determine the components that are not in the information set of the agent by varying the specification of $\left(U_{i,s,t} - E\left[U_{s,t} | \mathcal{I}\right]\right)$ while keeping fixed $\mathcal{I}$, so we can get the best possible fit of the cross-section distribution of $Y_{s,t}$. In the next section we describe how we use factor models to represent both $E\left[U_{s,t} | \mathcal{I}\right]$ and $\left(U_{s,t} - E\left[U_{s,t} | \mathcal{I}\right]\right)$ in a convenient framework.

4 Factor Models

4.1 Test Score Equations

To demonstrate our approach to determining the elements in the information set of the agent, we start by considering the test score equations. We break the error term U_k^M in test score equations in two components. The first component is a factor, θ_1, that is common across all test score equations. The second component is uniquely attached to test score equation k, ε_k^M. Consequently, we rewrite equation (13) as

$$M_k = \boldsymbol{X}^M \beta_{i,k}^M + \alpha_1^M \theta_1 + \varepsilon_1^M \tag{14}$$

Following the psychometric literature, the factor θ_1 is a latent cognitive ability which potentially affects all test scores. We assume that θ_1 is independent from $\boldsymbol{X}^M$ and ε_k^M. The main advantage of modelling test scores in this fashion consists in the fact that we are allowing test scores to be a noisy measure of the cognitive skill.

4.2 Earnings and Choice Equations

We decompose the error terms in the earnings equations in three terms. The first term is the cognitive factor θ_1. The second term is a "productivity" factor θ_2 which affects earnings and schooling choices, but not test scores. The third term is the idiosyncratic error term which affects only the period-t schooling-s earnings equation, $\varepsilon_{s,t}$. We rewrite equations (7) and (8) as:

$$Y_{i,0,t} = \boldsymbol{X}_i \beta_{0,t} + \alpha_{0,t}\theta_{i,1} + \delta_{0,t}\theta_{i,2} + \varepsilon_{i,0,t} \tag{15}$$

and

$$Y_{i,1,t} = \boldsymbol{X}_i \beta_{1,t} + \alpha_{1,t}\theta_{i,1} + \delta_{1,t}\theta_{i,2} + \varepsilon_{i,1,t} \tag{16}$$

We assume that the factor θ_j is independent from $\boldsymbol{X}, \varepsilon_{s,t}$, and θ_l for $l \neq j$, for all s, t.

The cost equation is decomposed as the earnings equations, so that (10) can be rewritten as:

$$C_i = \boldsymbol{Z}_i \gamma + \alpha_C \theta_{i,1} + \delta_C \theta_{i,2} + \varepsilon_{C,i} \tag{17}$$

Given the specifications with the factors in (15), (16), and (17) we can rewrite the school choice equation as:

$$I_i = E \left[\begin{array}{c} \sum_{t=1}^{T} \left(\frac{1}{1+\rho}\right)^t \mathbf{X}_i \left(\beta_{1,t} - \beta_{0,t}\right) - \mathbf{Z}_i \gamma + \theta_{i,1} \left[\sum_{t=1}^{T} \left(\frac{1}{1+\rho}\right)^t (\alpha_{1,t} - \alpha_{0,t}) - \alpha_C\right] \\ +\theta_{i,2} \left[\sum_{t=1}^{T} \left(\frac{1}{1+\rho}\right)^t (\delta_{1,t} - \delta_{0,t}) - \delta_C\right] + \sum_{t=1}^{T} \left(\frac{1}{1+\rho}\right)^t (\varepsilon_{i,1,t} - \varepsilon_{i,0,t}) - \varepsilon_{C,i} \end{array} \middle| \mathcal{I} \right] \tag{18}$$

5 The Estimation of the Components in the Information Set

We show how we can determine the unobservable components of the information set $\mathcal{I}$ of the agent at the time of the schooling choice by exploring the convenient structure provided by the factor models. Assume that $\mathbf{X}, \mathbf{Z}$, and ε_C are in the information set $\mathcal{I}$. To save on notation, define:

$$\alpha_I = \sum_{t=1}^{T} \left(\frac{1}{1+\rho}\right)^t (\alpha_{1,t} - \alpha_{0,t}) - \alpha_C \tag{19}$$

and

$$\delta_I = \sum_{t=1}^{T} \left(\frac{1}{1+\rho}\right)^t (\delta_{1,t} - \delta_{0,t}) - \delta_C \tag{20}$$

Suppose we propose that $\{\theta_1, \theta_2\} \sqsubset \mathcal{I}$, but $\varepsilon_{i,s,t} \notin \mathcal{I}$. Given the definitions of α_I, δ_I and $\mu_I(\boldsymbol{X}_i, \boldsymbol{Z}_i)$, if the null hypothesis is true, the school index I is:

$$I = \mu_I(\boldsymbol{X}_i, \boldsymbol{Z}_i) + \alpha_I \theta_{i,1} + \delta_I \theta_{i,2} + \varepsilon_{C,i} \tag{21}$$

Assume, for a moment, that we know both $\mu_I(\boldsymbol{X}_i, \boldsymbol{Z}_i)$ and $\beta_{s,t}$ for all s and t. Given observations on $\mathbf{X}$

and $\mathbf{Z}$ we can obtain from the data the covariance between the terms $I-\ \mu_I(\boldsymbol{X}_i, \boldsymbol{Z}_i)$ and $Y_{1,1} - \boldsymbol{X}\beta_{1,1}$. Under the null, this covariance is equal to

$$Cov\left(I - \mu_I(\boldsymbol{X}, \boldsymbol{Z}), Y_{1,1} - \boldsymbol{X}\beta_{1,1})\right) = \alpha_I \alpha_{1,1} \sigma^2_{\theta_1} + \delta_I \delta_{1,1} \sigma^2_{\theta_2}. \tag{22}$$

We can test the null $\{\theta_1, \theta_2\} \sqsubset \mathcal{I}$ against many different alternative ones. To fix ideas, consider the alternative assumption that proposes $\theta_1 \in \mathcal{I}$, but $\theta_2 \notin \mathcal{I}$ and that $E\left[\theta_2 | \mathcal{I}\right] = 0$. If the alternative is valid, the school index (18) becomes:

$$I = \ \mu_I(\boldsymbol{X}_i, \boldsymbol{Z}_i) + \alpha_I \theta_{i,1} + \varepsilon_{C,i} \tag{23}$$

In this case, the covariance between the terms $I-\ \mu_I(\boldsymbol{X}_i, \boldsymbol{Z}_i)$ and $Y_{1,1} - \boldsymbol{X}\beta_{1,1}$ satisfy:

$$Cov\left(I - \mu_I(\boldsymbol{X}, \boldsymbol{Z}), Y_{1,1} - \boldsymbol{X}\beta_{1,1})\right) = \alpha_I \alpha_{1,1} \sigma^2_{\theta_1}, \tag{24}$$

and the difference between the school index generated by the null and the alternative hypothesis is the term $\delta_I \delta_{1,1} \sigma^2_{\theta_2}$ that appears in (22) but not in (24). This insight allow us to redefine the test by generating parameters Δ_{θ_1} and Δ_{θ_2} be such that:

$$Cov\left(I - \mu_I(\boldsymbol{X}, \boldsymbol{Z}), Y_{1,1} - \mu_1(\boldsymbol{X})\right) - \Delta_{\theta_1} \alpha_I \alpha_{1,1} \sigma^2_{\theta_1} - \Delta_{\theta_2} \alpha_I \delta_{1,1} \sigma^2_{\theta_1} = 0$$

It is easy to see how we can rewrite the test in terms of Δ_{θ_1} and Δ_{θ_2}. We conclude that agents know and act on the information contained in factors 1 and 2, so that $\{\theta_1, \theta_2\} \sqsubset \mathcal{I}$, if we reject both $\Delta_{\theta_1} = 0$ and $\Delta_{\theta_2} = 0$.

It remains to be shown that we can actually identify all of the parameters of the model, in particular, the function $\mu_I(\boldsymbol{X}_i, \boldsymbol{Z}_i)$, the parameters β and α in the test and earnings equations, the parameters δ

in the earnings equations, the distribution of the factors, F_θ, as well as the distribution of idiosyncratic components F_ε in test, earnings and cost equations. We show how to recover these objects from the data in the next section.

6 Identification of the Model

We focus our discussion of identification on the normal case. See Carneiro, Hansen and Heckman (2003) for proofs of nonparametric identification of the distributions of the factors θ and uniquenesses ε.

6.1 Test Scores

To motivate our identification analysis we start by considering the test score equations. It is convenient to do so because the test scores are available for all agents and are taken by the agent before he makes the schooling decision. Therefore, we do not have to worry about selection issues when discussing identification from test score equations. Three assumptions are crucial in securing identification through factor models. First, the explanatory variables $\boldsymbol{X}^M$ are independent from both θ_1 and ε_k^M, for $k = 1, ..., K$. Second, the factor θ_1 is independent from ε_k^M, for $k = 1, ..., K$. Third, the uniqueness ε_k^M is independent from ε_l^M for any $k \neq l$, for $k, l = 1, ..., K$. The first assumption allows us to conclude that β_k^M can be consistently estimated from a simple OLS regression of M_k against $\mathbf{X}^M$. Given knowledge of these parameters we can construct differences $M_k - \boldsymbol{X}^M \beta_k^M$ and compute the covariances:

$$Cov\left(M_1 - \boldsymbol{X}^M \beta_1^M, M_2 - \boldsymbol{X}^M \beta_2^M\right) = \alpha_1^M \alpha_2^M \sigma_{\theta_1}^2 \tag{25}$$

$$Cov\left(M_1 - \boldsymbol{X}^M \beta_1^M, M_3 - \boldsymbol{X}^M \beta_3^M\right) = \alpha_1^M \alpha_3^M \sigma_{\theta_1}^2 \tag{26}$$

$$Cov\left(M_2 - \boldsymbol{X}^M \beta_2^M, M_3 - \boldsymbol{X}^M \beta_3^M\right) = \alpha_2^M \alpha_3^M \sigma_{\theta_1}^2 \tag{27}$$

The left-hand side of $(25), (26)$, and (27) can be computed straight from the data. The right-hand side of $(25), (26)$, and (27) is implied by the factor model. As is common in the factor literature, we need to normalize one of the factor loadings. Let $\alpha_1^M = 1$. If we take the ratio of (27) to (25) we identify α_3^M. Analogously, the ratio of (27) to (26) allows us to recover α_2^M. Given the normalization of $\alpha_1^M = 1$ and identification of α_2^M, we rescue $\sigma_{\theta_1}^2$ from (25). Finally, we can identify the variance of ε_k^M from the variance of $M_k - \boldsymbol{X}\beta_k^M$. Because the factor θ_1 and uniquenesses ε_k are independently normally distributed random variables, we have identified their distribution.

6.2 Earnings and Choice Equations

To establish the identification of the objects of interest in earnings equations requires a little more work because of the selection problem. It is at this stage of the problem that fixing the discussion on the normally distributed factors and uniquenesses becomes convenient, as we can use the closed-form solutions to reduce the identification problem to the identification of a few parameters.

We rely on four important assumptions to secure identification. First, all of the observable explanatory variables $\boldsymbol{X}$ and $\boldsymbol{Z}$ are independent of the unobservable factors, θ_1 and θ_2, as well as uniquenesses $\varepsilon_{s,t}$ for all s, t. Second, θ_1 is independent of θ_2. Third, both θ_1 and θ_2 are independent of ε_C and $\varepsilon_{s,t}$ for all s, t. Fourth, $\varepsilon_{s,t}$ is independent from ε_C and $\varepsilon_{s',t'}$ for any pairs s, s' and t, t' such that $s \neq s'$ or $t \neq t'$. According to the last three assumptions, all of the the dependence among $U_{0,t}, U_{1,t}$, and U_C is captured through the factors θ_1 and θ_2, which, for simplicity, we assume that

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma_{\theta_1}^2 & 0 \\ 0 & \sigma_{\theta_2}^2 \end{bmatrix}\right),$$

Because of the loadings $\alpha_{s,t}, \delta_{s,t}, \alpha_C$, and δ_C the factors θ can affect $U_{0,t}, U_{1,t}$, and U_C differently. Therefore, by aadopting the factor structure we are not imposing, for example, perfect ranking in the sense that the best in the distribution of earnings in sector s at period t is the best (or the worst) in the distribution of earnings in sector s' at period t'. When the schooling choice problem is analyzed under the factor model, the joint distribution of the labor earnings $Y_{0,t}, Y_{1,t}$ conditional on $\boldsymbol{X}$ is:

$$\begin{bmatrix} Y_{0,t} \\ Y_{1,t} \end{bmatrix} \mid \boldsymbol{X} \sim N\left(\begin{bmatrix} \boldsymbol{X}\beta_{0,t} \\ \boldsymbol{X}\beta_{1,t} \end{bmatrix}, \begin{bmatrix} \alpha_{0,t}^2\sigma_{\theta_1}^2 + \delta_{0,t}^2\sigma_{\theta_2}^2 + \sigma_{\varepsilon_{0,t}}^2 & \alpha_{0,t}\alpha_{1,t}\sigma_{\theta_1}^2 + \delta_{0,t}\delta_{1,t}\sigma_{\theta_2}^2 \\ \alpha_{0,t}\alpha_{1,t}\sigma_{\theta_1}^2 + \delta_{0,t}\delta_{1,t}\sigma_{\theta_2}^2 & \alpha_{1,t}^2\sigma_{\theta_1}^2 + \delta_{1,t}^2\sigma_{\theta_2}^2 + \sigma_{\varepsilon_{1,t}}^2 \end{bmatrix} \right). \tag{28}$$

As a result, identification of the joint distribution $F(Y_{0,t}, Y_{1,t} \mid \boldsymbol{X})$ reduces to the identification of the parameters $\beta_{s,t}, \alpha_{s,t}, \delta_{s,t}, \sigma_{\varepsilon_{s,t}}$, and $\sigma_{\theta_j}^2$ for $s = 0, 1$; $t = 1, ..., T$ and $j = 1, 2$. From the observed data and the factor structure it follows that:

$$E(Y_{1,t}| \boldsymbol{X}, S = 1) = \boldsymbol{X}\beta_{1,t} + \alpha_{1,t}E[\theta_1| \boldsymbol{X}, S = 1] + \delta_{1,t}E[\theta_2| \boldsymbol{X}, S = 1] + E[\varepsilon_{1,t}| \boldsymbol{X}, S = 1] \tag{29}$$

The event $S = 1$ corresponds to the event $I = E\left(\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^t (Y_{1,t} - Y_{0,t}) - C \middle| \mathcal{I}\right) \geq 0$. At this point it is convenient to distinguish the role played by the factors θ from the one played by the uniquenesses $\varepsilon_{s,t}$. In tune with our intuitive discussion, we need to have terms that will affect the covariance between schooling and earnings equations by changing the components of the information set $\mathcal{I}$, which is captured by the term $E(U_{s,t}|\mathcal{I})$. We also need to have components that will affect earnings while holding constant the information set $\mathcal{I}$ and the covariance between earnings and schooling, which is captured by the term $U_{s,t} - E(U_{s,t}|\mathcal{I})$. The former role will be played by the factors in the information set of the agent. The latter will be played by the factors not in the information set of the agents as well as the uniquenesses $\varepsilon_{s,t}$. Consequently, we will construct $\varepsilon_{s,t}$ so that they satisfy the requirement $\varepsilon_{s,t} \notin \mathcal{I}$. As a result, we conclude that:

$$E\left(\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^t (Y_{1,t} - Y_{0,t}) - C \middle| \mathcal{I}\right) = \mu_I(\boldsymbol{X}, \boldsymbol{Z}) + \alpha_I\theta_1 + \delta_I\theta_2 - \varepsilon_C$$

Let η be the linear combination of three independent normal random variables: $\eta = \alpha_I\theta_1 + \delta_I\theta_2 - \varepsilon_C$. Then, $\eta \sim N\left(0, \sigma_\eta^2\right)$, with $\sigma_\eta^2 = \alpha_I^2\sigma_{\theta_1}^2 + \delta_I^2\sigma_{\theta_2}^2 + \sigma_{\varepsilon_c}^2$ and

$$S = 1 \Leftrightarrow \eta > -\mu_I(\boldsymbol{X}, \boldsymbol{Z}) \tag{30}$$

If we replace (30) in (29) and using the fact that $\varepsilon_{s,t}$ is independent from $\boldsymbol{X}, \boldsymbol{Z},$ and θ, we can show that:

$$E\left(Y_{1,t} \middle| \boldsymbol{X}, S = 1\right) = \boldsymbol{X}\beta_1 + \alpha_{1,t}E\left[\theta_1 \middle| \boldsymbol{X}, \eta > -\mu_I(\boldsymbol{X}, \boldsymbol{Z})\right] + \delta_{1,t}E\left[\theta_2 \middle| \boldsymbol{X}, \eta > -\mu_I(\boldsymbol{X}, \boldsymbol{Z})\right] \tag{31}$$

Second, because θ_1, θ_2 and η are normal random variables we can use the projection property:

$$\theta_j = \frac{Cov\left(\theta_j, \eta\right)}{Var\left(\eta\right)}\eta + \nu_j \text{ for } j = 1, 2 \tag{32}$$

where ν_j is a mean zero, normal random variable independent from η. Because $Cov\left(\theta_1, \eta\right) = \sigma_{\theta_1}^2\alpha_I$ and $Cov\left(\theta_2, \eta\right) = \sigma_{\theta_2}^2\delta_I$ it follows that:

$$E\left[\theta_1 \middle| \boldsymbol{X}, \eta > -\mu_I(\boldsymbol{X}, \boldsymbol{Z})\right] = \frac{\sigma_{\theta_1}^2\alpha_I}{\sigma_\eta^2}E\left[\eta \middle| \eta > -\mu_I(\boldsymbol{X}, \boldsymbol{Z})\right]$$

$$E\left[\theta_2 \middle| \boldsymbol{X}, \eta > -\mu_I(\boldsymbol{X}, \boldsymbol{Z})\right] = \frac{\sigma_{\theta_2}^2\delta_I}{\sigma_\eta^2}E\left[\eta \middle| \eta > -\mu_I(\boldsymbol{X}, \boldsymbol{Z})\right]$$

For any standard normal random variable μ, $E\left(\mu \middle| \mu \geq -c\right) = \frac{\phi(c)}{\Phi(c)}$ where $\phi\left(.\right)$ and $\Phi\left(.\right)$ are the density and distribution function of a standard normal random variable. Define, for $j = 0, 1$, $\pi_{j,t} = \left(\frac{\alpha_{j,t}\alpha_I\sigma_{\theta_1}^2 + \delta_{j,t}\delta_I\sigma_{\theta_2}^2}{\sigma_\eta}\right)$.These facts together allow us to rewrite (29) as:

$$E\left(Y_{1,t} \middle| \eta \leq -\mu_I(\boldsymbol{X}, \boldsymbol{Z})\right) = \boldsymbol{X}\beta_{1,t} + \pi_{1,t}\frac{\phi\left(\frac{\mu_I(\boldsymbol{X},\boldsymbol{Z})}{\sigma_\eta}\right)}{\Phi\left(\frac{\mu_I(\boldsymbol{X},\boldsymbol{Z})}{\sigma_\eta}\right)} \tag{33}$$

It is easy to follow the same steps and derive a similar expression for mean observed earnings in sector "0":

$$E\left(Y_{0,t}\right|\eta > -\mu_I(\boldsymbol{X},\boldsymbol{Z})) = \boldsymbol{X}\beta_{0,t} - \pi_{0,t}\frac{\phi\left(\frac{\mu_I(\boldsymbol{X},\boldsymbol{Z})}{\sigma_\eta}\right)}{\Phi\left(\frac{\mu_I(\boldsymbol{X},\boldsymbol{Z})}{\sigma_\eta}\right)} \tag{34}$$

We can apply the two-step procedure proposed in Heckman (1976) to identify $\beta_{0,t}, \beta_{1,t}, \pi_{0,t}$ and $\pi_{1,t}$. Given identification of $\beta_{s,t}$ for all s and t, we can construct the differences $Y_{s,t} - \boldsymbol{X}\beta_{s,t}$ and compute the covariances:

$$Cov\left(M_1 - \boldsymbol{X}^M\beta_1^M, Y_{0,t} - \boldsymbol{X}\beta_{0,t}\right) = \alpha_{0,t}\sigma^2_{\theta_1} \tag{35}$$

$$Cov\left(M_1 - \boldsymbol{X}^M\beta_1^M, Y_{1,t} - \boldsymbol{X}\beta_{1,t}\right) = \alpha_{1,t}\sigma^2_{\theta_1} \tag{36}$$

The left-hand side of (**??**) is available from the data. The right-hand side is implied by the factor model and its assumptions. We determined $\sigma^2_{\theta_1}$ from the analysis of the test scores. So from equations (**??**) and (**??**) we can recover $\alpha_{0,t}$ and $\alpha_{1,t}$ for all t. Note that we can also identify the $\frac{\alpha_C}{\sigma_\eta}$ by computing the covariance:

$$Cov\left(M_1 - \boldsymbol{X}\beta_1^M, \frac{I - \mu_I(\boldsymbol{X},\boldsymbol{Z})}{\sigma_\eta}\right) = \frac{\sum_{t=1}^T\left(\frac{1}{1+\rho}\right)^t(\alpha_{1,t} - \alpha_{0,t}) - \alpha_C}{\sigma_\eta}\sigma^2_{\theta_1} \tag{37}$$

The argument why $\frac{\alpha_C}{\sigma_\eta}$ can be recovered is simple: Using (**??**) and (**??**) we can identify $\alpha_{1,t}$ and $\alpha_{0,t}$ for all t.The only remaining term to be indentified is the ratio $\frac{\alpha_C}{\sigma_\eta}$, which we can from the covariance equation (**??**).

Note that if $T \geq 2$ then we can also identify the parameters related to factor θ_2, such as $\delta_{s,t}$ and $\sigma^2_{\theta_2}$. To see this, first normalize $\delta_{0,1} = 1$ and compute the covariances:

$$Cov\left(Y_{0,1} - \boldsymbol{X}\beta_{0,1}, Y_{0,2} - \boldsymbol{X}\beta_{0,2}\right) - \alpha_{0,1}\alpha_{0,2}\sigma^2_{\theta_1} = \delta_{0,2}\sigma^2_{\theta_2} \tag{38}$$

$$Cov\left(Y_{0,1} - \boldsymbol{X}\beta_{0,1}, \frac{I - \mu_I(\boldsymbol{X},\boldsymbol{Z})}{\sigma_\eta}\right) - \frac{\alpha_{0,1}\sigma^2_{\theta_1}\sum_{t=1}^T(\alpha_{1,t} - \alpha_{0,t} - \alpha_C)}{\sigma_\eta} = \frac{\sigma^2_{\theta_2}\sum_{t=1}^T(\delta_{1,t} - \delta_{0,t} - \delta_C)}{\sigma_\eta} \tag{39}$$

$$Cov\left(Y_{0,2}-\boldsymbol{X}\beta_{0,2},\frac{I-\mu_I(\boldsymbol{X},\boldsymbol{Z})}{\sigma_\eta}\right)-\frac{\alpha_{0,2}\sigma^2_{\theta_1}\sum_{t=1}^{T}\left(\alpha_{1,t}-\alpha_{0,t}-\alpha_C\right)}{\sigma_\eta}=\frac{\delta_{0,2}\sigma^2_{\theta_2}\sum_{t=1}^{T}\left(\delta_{1,t}-\delta_{0,t}-\delta_C\right)}{\sigma_\eta} \tag{40}$$

On the left-hand side of (38), (39), and (40) are terms that we can compute from the data or have already identified. If we compute the ratio of (40) to (39) we can recover $\delta_{0,2}$. From (38) we can recover $\sigma^2_{\theta_2}$. We now add the covariances from the college earnings:

$$Cov\left(Y_{1,1}-\boldsymbol{X}\beta_{1,1},Y_{1,2}-\boldsymbol{X}\beta_{1,2}\right)-\alpha_{1,1}\alpha_{1,2}\sigma^2_{\theta_1}=\delta_{1,1}\delta_{1,2}\sigma^2_{\theta_2} \tag{41}$$

$$Cov\left(Y_{1,1}-\boldsymbol{X}\beta_{1,1},\frac{I-\mu_I(\boldsymbol{X},\boldsymbol{Z})}{\sigma_\eta}\right)-\frac{\alpha_{1,1}\sigma^2_{\theta_1}\sum_{t=1}^{T}\left(\alpha_{1,t}-\alpha_{0,t}-\alpha_C\right)}{\sigma_\eta}=\frac{\delta_{1,1}\sigma^2_{\theta_2}\sum_{t=1}^{T}\left(\delta_{1,t}-\delta_{0,t}-\delta_C\right)}{\sigma_\eta} \tag{42}$$

$$Cov\left(Y_{1,2}-\boldsymbol{X}\beta_{1,2},\frac{I-\mu_I(\boldsymbol{X},\boldsymbol{Z})}{\sigma_\eta}\right)-\frac{\alpha_{1,2}\sigma^2_{\theta_1}\sum_{t=1}^{T}\left(\alpha_{1,t}-\alpha_{0,t}-\alpha_C\right)}{\sigma_\eta}=\frac{\delta_{1,2}\sigma^2_{\theta_2}\sum_{t=1}^{T}\left(\delta_{1,t}-\delta_{0,t}-\delta_C\right)}{\sigma_\eta} \tag{43}$$

Now, by computing the ratios of (43) to (41) and (42) to (41) we obtain $\delta_{1,2}$ and $\delta_{1,1}$ respectively. Finally, we use the information in $Var\left(Y_{0,t}|\,\boldsymbol{X},S=0\right)$ and $Var\left(Y_{1,t}|\,\boldsymbol{X},S=1\right)$ to compute $\sigma^2_{\varepsilon_{0,t}}$ and $\sigma^2_{\varepsilon_{1,t}}$, respectively. Note that we have identified all of the elements that characterize the joint distribution as specified in (28).

7 Empirical Results

7.1 The data, equations, and estimation

The first problem we have to overcome is that few data sets contain the full life cycle of earnings along with the test scores and schooling choices needed to directly estimate our model and extract components of uncertainty. We need to combine data sets. Otherwise, we can only obtain partial identification of the model. In our empirical analysis, we use a sample of white males from the NLSY data pooled with PSID data, as described in Appendix 3 (placed on our website), to produce life cycle data on earnings and schooling.

Following the preceding theoretical analysis, we consider only two schooling choices: high school and college graduation. As above, we use $s = 0$ to denote high school and $s = 1$ to denote college.

Table 2.1 presents descriptive statistics of the data used to estimate the model. College graduates have higher test scores, come from better family backgrounds, and are more likely to live in a location where college tuition is lower.

To simplify the empirical analysis, we divide the lifetimes of individuals into 24 periods. The first period covers ages 18 to 19, the second period covers from ages 20 to 21, and so forth until the twenty-fourth period which covers ages 64 to 65. Aggregating ages in these periods serves two purposes. The first is that they potentially reduce the problem of measurement error in earnings (assuming it is classical measurement error). The second is it reduces the number of parameters to estimate. For each schooling level s, $s \in \{0, 1\}$, and for each period t, we calculate the present value of earnings as of age 18, $Y_{s,t}$. We assume individuals have a constant intertemporal preference rate $\rho = 0.03$. We conduct robustness tests regarding both the aggregation of ages in periods and by changing ρ. To simplify notation drop the 'i' subscript. If $Y_{s,t}$ is generated by a three factor model, we would write:

$$Y_{s,t} = \boldsymbol{X}\boldsymbol{\beta}_{s,t} + \theta_1 \alpha_{s,t} + \theta_2 \delta_{s,t,2} + \varepsilon_{s,t} \text{ for } t = 1, 2, ..., 16, \;\; s \in \{0, 1\}\,. \tag{44}$$

It turns out that a two-factor model is fits the data. Since the scales of the factors are unknown, it is necessary to normalize some loadings (the $\boldsymbol{\alpha}$). In this paper, we set $\delta_{1,1} = 1$. The normalization for ability (associated with the measurements $\boldsymbol{M}$ based on test scores) is presented in the next paragraph.

For the measurement system for cognitive ability ($\boldsymbol{M}$ in the notation of section **??**) we use five components of the ASVAB test battery: arithmetic reasoning, word knowledge, paragraph comprehension, math knowledge and coding speed. We dedicate the first factor (θ_1) to this test system, and exclude the other from it. This justifies our interpretation of θ_1 as ability. We include family background variables among the covariates $\boldsymbol{X}_M$ in the ASVAB test equations. In Table 2.2 we list the elements of $\boldsymbol{X}_M$. Formally, let M_j denote the test score j,

$$M_j = \boldsymbol{X}^M \boldsymbol{\beta}_j^M + \theta_1 \alpha_j^M + \varepsilon_j^M. \tag{45}$$

To set the scale of θ_1, we normalize $\alpha_1^M = 1$.

The cost function C is given by

$$C = \boldsymbol{Z}\boldsymbol{\gamma} + \theta_1 \alpha_C + \theta_2 \delta_C + \varepsilon_C, \tag{46}$$

where the $\boldsymbol{Z}$ are variables that affect the costs of going to college and include variables that do not affect outcomes $Y_{s,t}$, such as local tuition. Table 2.2 shows the full set of covariates used, and the exclusions (the variables in $\boldsymbol{Z}$ not in $\boldsymbol{X}$.) We include tuition among the elements of $\boldsymbol{Z}$ but allow for a more general notion of costs in our empirical work, including psychic costs.

The valuation or net utility function for schooling choice is

$$I = E\left(\sum_{t=0}^{4} \frac{Y_{c,t} - Y_{h,t}}{(1+\rho)^t} \middle| \mathcal{I} \right) - E\left(C | \mathcal{I}\right), \tag{47}$$

where ρ is the intertemporal preference rate. Individuals go to college if $I > 0$. The individual decision maker is assumed to be the child although parental resources can affect C. Cost variable C also includes the effect of ability on reducing tuition costs. We test and do not reject the hypothesis that individuals, at the time they make college going decisions, know their cost functions, the $\boldsymbol{Z}$ and the $\boldsymbol{X}$, factors θ_1, θ_2, and unobservables in cost ε_C. However, they do not know $\varepsilon_{s,t}$, $s \in \{c, h\}$, $t \in \{1, 2, , ..., 16\}$, at the time they make their educational choices.

We assume that each factor θ_k, is generated by a mixture of J_k normal distributions,

$$\theta_k \backsim \sum_{j=1}^{J_k} p_{k,j}\phi\left(\theta_k \mid \mu_{k,j}, \tau_{k,j}\right),$$

where $\phi\left(\eta \mid \mu_j, \tau_j\right)$ is a normal density for η with mean μ_j and variance τ_j and $\sum_{j=1}^{J_k} p_{k,j} = 1$, and $p_{k,j} > 0$. As shown in [?], mixtures of normals with a large number of components approximate any distribution of θ_k arbitrarily well in the ℓ^1 norm. The $\varepsilon_{s,t}$ are also assumed to be generated by mixtures of normals. We estimate the model using Markov Chain Monte Carlo methods as described in [?]. In Tables $2.3-2.5$ we present estimated coefficients and factor loadings. For all factors, a three-component model ($J_k = 3$, $k = 1, 2$) is adequate.For all $\varepsilon_{s,t}$ we use a four-component model.[1]

7.1.1 How the model fits the data

7.1.2 Returns to College

7.1.3 How well can agents predict future earnings?

7.1.4 Robustness Checks

8 Summary and Conclusion

9 References

Levy and Murnane (1992)

Juhn, Murphy and Pierce (1993)

Heckman and Rubinstein (1997)

[1]Additional components do not improve the goodness of fit of the model to the data.

Carneiro and Heckman (2003)

Heckman, Stixrud and Urzua (2005)

Gottschalk and Moffitt (1994)

Carneiro, Hansen and Heckman (2003)

Cunha, Heckman and Navarro (2005)

Table

Descriptive Statistics for Explanatory Variables and Test Scores

	NLS/1966 and PSID[1]					
	High School Sample			College Sample		
Variables	Obs	Mean	Std Error	Obs	Mean	Std Error
Mother's Education	1454	3.2001	1.2195	1458	4.3601	1.5910
Father's Education	1454	2.8975	1.3006	1458	4.3491	1.9392
Number of Siblings	1454	3.2552	2.2056	1458	2.3813	1.8320
Urban Residence at 14	1454	0.6623	0.4731	1458	0.8169	0.3869
Local Tuition[3]	1454	0.1447	0.0325	1458	0.1453	0.0295
Born between 1915 and 1919	1454	0.0268	0.1616	1458	0.0062	0.0784
Born between 1920 and 1924	1454	0.0406	0.1974	1458	0.0213	0.1443
Born between 1925 and 1929	1454	0.0495	0.2170	1458	0.0391	0.1939
Born between 1930 and 1934	1454	0.0475	0.2127	1458	0.0309	0.1730
Born between 1935 and 1939	1454	0.0399	0.1958	1458	0.0357	0.1855
Born between 1940 and 1944	1454	0.2063	0.4048	1458	0.2188	0.4136
Born between 1945 and 1949	1454	0.3287	0.4699	1458	0.4444	0.4971
Born between 1950 and 1954	1454	0.2607	0.4391	1458	0.2037	0.4029
IQ Test Score[2]	506	-0.5560	0.8758	581	0.5036	0.8281

[1]The sample constitutes of white males with a high school or college degree born in 1952 or before.

[2]Only available for the respondents of the NLS/1966.

[3]In ten thousand dollars, CPI deflated, base year 2000.

Table - Continued

Descriptive Statistics for Explanatory Variables and Test Scores

NLS/1979 and PSID[1]

	High School Sample			College Sample		
Variables	Obs	Mean	Std Error	Obs	Mean	Std Error
Mother's Education	2054	3.6509	1.2431	1701	4.8236	1.6108
Father's Education	2054	3.4971	1.4738	1701	5.1329	2.0161
Number of Siblings	2054	3.2240	2.0913	1701	2.4774	1.6992
Urban Residence at 14	2054	0.7230	0.4476	1701	0.8430	0.3639
Local Tuition[3]	2054	0.1853	0.0716	1701	0.1786	0.0669
Born between 1915 and 1919	2054	0.0190	0.1365	1701	0.0053	0.0726
Born between 1920 and 1924	2054	0.0287	0.1671	1701	0.0182	0.1338
Born between 1925 and 1929	2054	0.0351	0.1840	1701	0.0335	0.1800
Born between 1930 and 1934	2054	0.0336	0.1802	1701	0.0265	0.1605
Born between 1935 and 1939	2054	0.0282	0.1657	1701	0.0306	0.1722
Born between 1940 and 1944	2054	0.0467	0.2111	1701	0.0682	0.2522
Born between 1945 and 1949	2054	0.0667	0.2496	1701	0.1405	0.3476
Born between 1950 and 1954	2054	0.1207	0.3259	1701	0.1146	0.3187
Born between 1955 and 1959	2054	0.2400	0.4272	1701	0.2187	0.4135
Born between 1960 and 1964	2054	0.3174	0.4656	1701	0.3133	0.4640
Born between 1965 and 1969	2054	0.0638	0.2444	1701	0.0306	0.1722
ASVAB - Arithmetic Reasoning[2]	640	-0.4696	0.9408	561	0.5731	0.7285
ASVAB - Word Knowledge[2]	640	-0.4334	1.0578	561	0.5277	0.5578
ASVAB - Paragraph Composition[2]	640	-0.4289	1.0770	561	0.5147	0.5363
ASVAB - Coding Speed[2]	640	-0.3195	0.9434	561	0.3948	0.8849
ASVAB - Math Knowledge[2]	640	-0.5993	0.7949	561	0.7190	0.6863

[1]This sample is composed of white males with a high school or college degree including individuals born in 1969 or before.

[2]Only available for the respondents of the NLSY/1979.

[3]In ten thousand dollars, CPI deflated, base year 2000.

Table

Descriptive Statistics for Period[1] Earnings[2]

	NLS/1966 and PSID[3]						NLSY/1979 and PSID[4]					
	High School Sample			College Sample			High School Sample			College Sample		
Period[1]	Obs	Mean	Std Error	Obs	Mean	Std Error	Obs	Mean	Std Error	Obs	Mean	Std Error
1	337	2.1646	1.3041	286	1.1625	0.6503	239	1.6916	1.5715	197	1.1343	1.0315
2	264	4.0672	1.7317	323	1.5348	1.1102	593	3.3588	1.9616	352	1.4789	1.2424
3	296	4.4909	1.5981	287	2.4612	1.6511	942	4.1271	2.0627	501	1.9592	1.5957
4	439	4.5930	1.7946	416	4.2581	1.9477	1144	4.4228	2.2118	777	4.1343	2.3594
5	442	4.5072	1.6270	465	4.9543	2.0551	1233	4.6200	2.2706	977	5.3490	2.6021
6	362	4.2284	1.4760	452	5.3063	2.2915	1246	4.7645	2.3471	1084	6.1144	3.5069
7	326	3.8237	1.4249	439	5.3222	2.4097	1132	4.8102	3.5125	1096	6.7664	4.1518
8	339	3.6576	1.4890	474	5.2057	2.3136	947	4.7765	3.3080	950	7.0273	4.1659
9	299	3.4736	1.3392	466	5.0740	2.7304	767	4.7781	3.0637	811	7.1873	4.2407
10	283	3.2443	1.3001	444	4.9818	3.1373	573	4.5688	2.2295	688	7.2931	4.4655
11	302	2.8339	1.0813	446	4.8957	3.3196	461	4.5444	2.5835	576	7.4099	4.9780
12	281	2.6737	1.0140	417	4.5845	2.8893	401	4.3017	2.1373	516	7.1851	4.5491
13	307	2.4513	0.9056	431	4.2969	2.8265	360	3.9853	1.5783	484	6.9696	5.0493
14	319	2.2130	0.8999	437	3.8572	2.2055	319	3.6831	1.4977	437	6.4211	3.6746
15	284	1.9804	0.7774	392	3.6110	2.0889	284	3.4255	1.3443	392	6.2447	3.6087
16	271	1.8666	0.7883	346	3.1939	1.6448	271	3.3554	1.4180	346	5.7412	2.9573
17	253	1.6870	0.6947	272	2.9175	1.6778	253	3.1509	1.2981	272	5.4509	3.1405
18	237	1.5343	0.6150	222	2.7632	2.1200	237	2.9780	1.1942	222	5.3637	4.1206
19	226	1.3797	0.5884	196	2.4868	1.9836	226	2.7833	1.1885	196	5.0154	3.9992
20	184	1.2289	0.5586	154	2.1344	1.5364	184	2.5767	1.1719	154	4.4739	3.2211
21	163	1.1069	0.5363	132	1.9147	1.5534	163	2.4112	1.1681	132	4.1704	3.3826
22	135	0.9919	0.4935	104	1.7729	1.3125	135	2.2459	1.1191	104	4.0129	2.9691
23	96	0.8470	0.4610	77	1.6022	1.3454	96	1.9922	1.0843	77	3.7677	3.1644
24	59	0.6200	0.4084	52	1.3325	1.1089	59	1.5156	0.9988	52	3.2560	2.7078

[1]The first period goes from ages 18 and 19, the second period goes from ages 20 and 21, so on and so forth until the thwienty-fourth period which correspondes to ages 64 and 65.

[2]All earnings figures have been inflation adjusted using 2000 as base year.

[3]The sample constitutes of white males with a high school or college degree born in 1952 or before.

[4]This sample is composed of white males with a high school or college degree including individuals born in 1969 or before.

Table
χ^2 Goodness of Fit Test*
NLSY/1966 - White Males

	High School		College		Overall	
Period	**χ^2 statistic**	**Critical Value**	**χ^2 statistic**	**Critical Value**	**χ^2 statistic**	**Critical Value**
1	35.1666	36.4150	30.9157	31.4104	86.8989	62.8296
2	16.2276	30.1435	35.9966	36.4150	62.2311	62.8296
3	23.2362	33.9244	33.7769	33.9244	55.1028	60.4809
4	28.4962	46.1943	41.7248	43.7730	82.2013	83.6753
5	38.5206	47.3999	25.7795	48.6024	79.0193	88.2502
6	27.2044	38.8851	37.7031	47.3999	68.8676	80.2321
7	30.0069	36.4150	32.3167	46.1943	72.6573	75.6237
8	30.8421	37.6525	44.8474	47.3999	75.5611	80.2321
9	25.4233	33.9244	37.9660	49.8018	69.4125	75.6237
10	22.4095	32.6706	34.9972	47.3999	70.0617	72.1532
11	32.4199	35.1725	37.1016	47.3999	94.5121	76.7778
12	36.3761	32.6706	38.7352	44.9853	102.0853	69.8322
13	30.5514	33.9244	43.0433	44.9853	93.0956	72.1532
14	36.3597	36.4150	38.7660	46.1943	124.8956	75.6237
15	27.3590	30.1435	46.8502	43.7730	117.4130	68.6693
16	25.3761	31.4104	31.2686	38.8851	51.9383	60.4809
17	25.7091	30.1435	18.1242	31.4104	44.8770	55.7585
18	14.6712	27.5871	24.0683	26.2962	27.5918	48.6024
19	12.0451	26.2962	28.2599	23.6848	56.2757	43.7730
20	14.3254	22.3620	14.1747	19.6751	35.3521	37.6525
21	18.4607	19.6751	15.6153	16.9190	48.4015	33.9244
22	15.9271	16.9190	6.3777	14.0671	24.2935	27.5871
23	16.9783	12.5916	10.4284	11.0705	33.5786	21.0261
24	7.7536	7.8147	2.6828	7.8147	12.1896	12.5916

* 95% Confidence, equiprobable bins with aprox. 13 people per bin. A χ^2 statistic lower than the critical value indicates a "good" fit.

Table

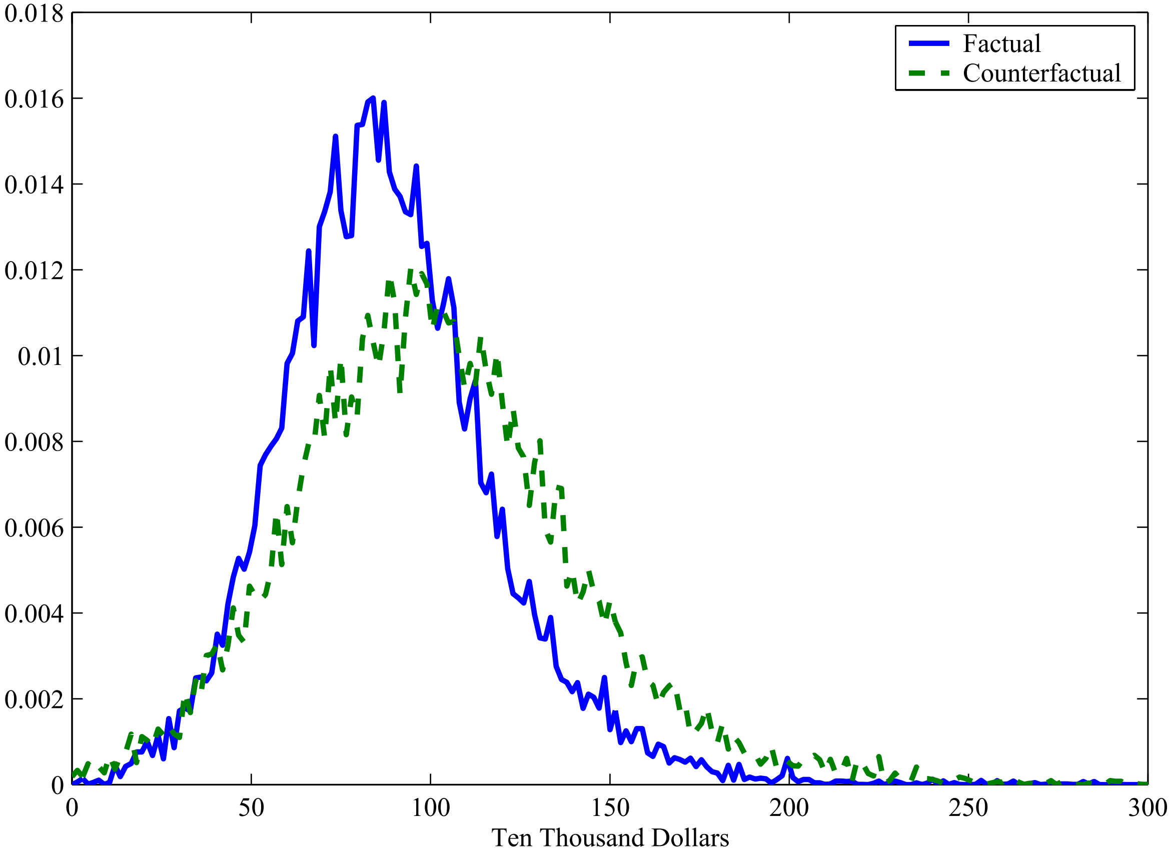
χ^2 Goodness of Fit Test*

NLSY/1979 - White Males

	High School		College		Overall	
Period	**χ^2 statistic**	**Critical Value**	**χ^2 statistic**	**Critical Value**	**χ^2 statistic**	**Critical Value**
1	21.8279	26.2962	33.9550	23.6848	97.3125	44.9853
2	149.2872	59.3035	34.0150	37.6525	140.0304	91.6702
3	84.5345	88.2502	51.6067	52.1923	201.1130	133.2569
4	147.2893	104.1387	110.0342	80.2321	283.6263	181.7702
5	122.4765	119.8709	146.5578	98.4844	276.9388	201.4234
6	112.2404	118.7516	95.2815	101.8795	247.3886	212.3039
7	104.7030	108.6479	97.7348	106.3948	206.2928	206.8668
8	76.3445	92.8083	91.5059	88.2502	173.4034	179.5806
9	62.2386	76.7778	74.1022	80.2321	142.0921	150.9894
10	55.7005	59.3035	73.3813	73.3115	123.6666	124.3421
11	36.4969	46.1943	54.4203	60.4809	111.0111	101.8795
12	37.6774	42.5570	53.5602	54.5722	112.9666	89.3912
13	30.8201	38.8851	50.5857	52.1923	59.8648	84.8206
14	27.8652	36.4150	50.7021	48.6024	75.1764	75.6237
15	27.5540	30.1435	63.8219	41.3371	86.4812	70.9935
16	41.0335	31.4104	42.9749	40.1133	46.7377	62.8296
17	26.4536	28.8693	30.7332	31.4104	53.0589	54.5722
18	18.8084	27.5871	31.1921	24.9958	41.7653	48.6024
19	15.1549	26.2962	38.3292	23.6848	46.3320	43.7730
20	9.0556	22.3620	17.0141	19.6751	28.4916	37.6525
21	18.5147	21.0261	20.4213	18.3070	46.6824	32.6706
22	16.6020	16.9190	11.5794	14.0671	20.8732	27.5871
23	18.0791	12.5916	2.9422	11.0705	21.9580	21.0261
24	8.4491	7.8147	0.1916	7.8147	13.4519	14.0671

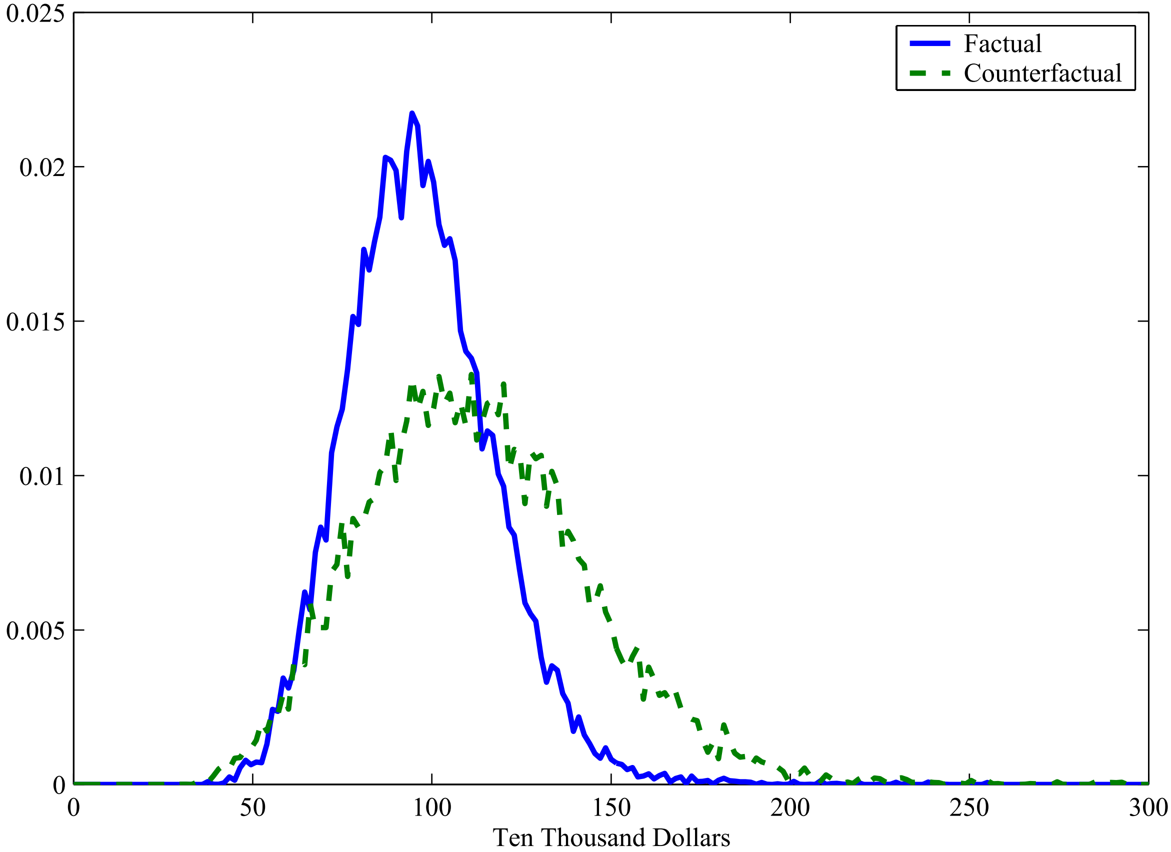
* 95% Confidence, equiprobable bins with aprox. 13 people per bin. A χ^2 statistic lower than the critical value indicates a "good" fit.

Figure
Densities of present value of lifetime earnings for High School Graduates
Factual and Counterfactual NLS/1966 Sample



Present Value of Lifetime Earnings from age 18 to 65 for high school graduates using a discount rate of 3%. Let Y_0 denote present value of earnings in high school sector. Let Y_1 denote present value of earnings in college sector. In this graph we plot the factual density function $f(y_0 \mid S=0)$ (the solid line), against the counterfactual density function $f(y_1 \mid S=0)$. We use kernel density estimation to smooth these functions.

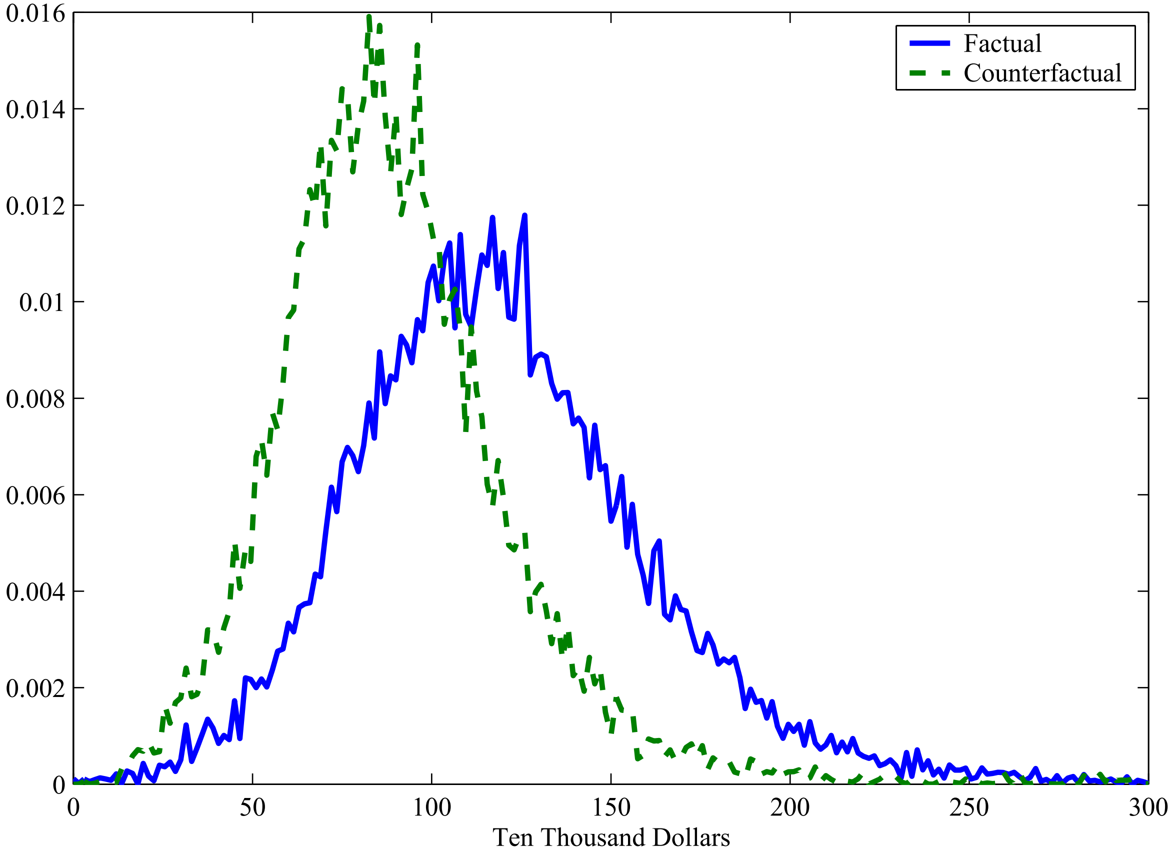
Figure
Densities of present value of lifetime earnings for High School Graduates
Factual and Counterfactual NLSY/1979 Sample



Present Value of Lifetime Earnings from age 18 to 65 for high school graduates using a discount rate of 3%. Let Y_0 denote present value of earnings in high school sector. Let Y_1 denote present value of earnings in college sector. In this graph we plot the factual density function $f(y_0 \mid S=0)$ (the solid line), against the counterfactual density function $f(y_1 \mid S=0)$. We use kernel density estimation to smooth these functions.

Figure

Densities of present value of earnings for College Graduates
Factual and Counterfactual NLS/1966 Sample

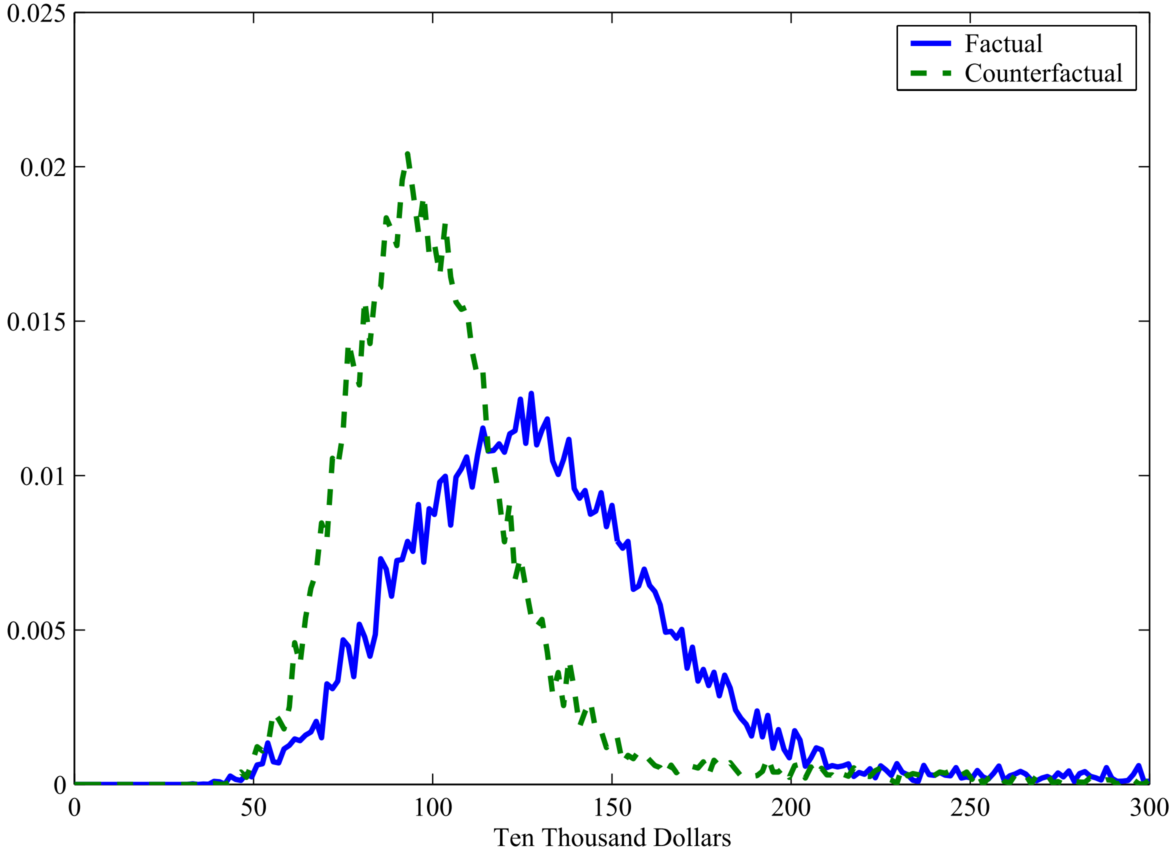


Present Value of Lifetime Earnings from age 18 to 65 for high school graduates using a discount rate of 3%. Let Y_0 denote present value of earnings in high school sector. Let Y_1 denote present value of earnings in college sector. In this graph we plot the counterfactual density function $f(y_1 \mid S=0)$ (the dashed line), against the factual density function $f(y_1 \mid S=0)$. We use kernel density estimation to smooth these functions.

Figure

Densities of present value of earnings for College Graduates

Factual and Counterfactual NLSY/1979 Sample



Present Value of Lifetime Earnings from age 18 to 65 for high school graduates using a discount rate of 3%. Let Y_0 denote present value of earnings in high school sector. Let Y_1 denote present value of earnings in college sector. In this graph we plot the counterfactual density function $f(y_1 \mid S=0)$ (the dashed line), against the factual density function $f(y_1 \mid S=0)$. We use kernel density estimation to smooth these functions.

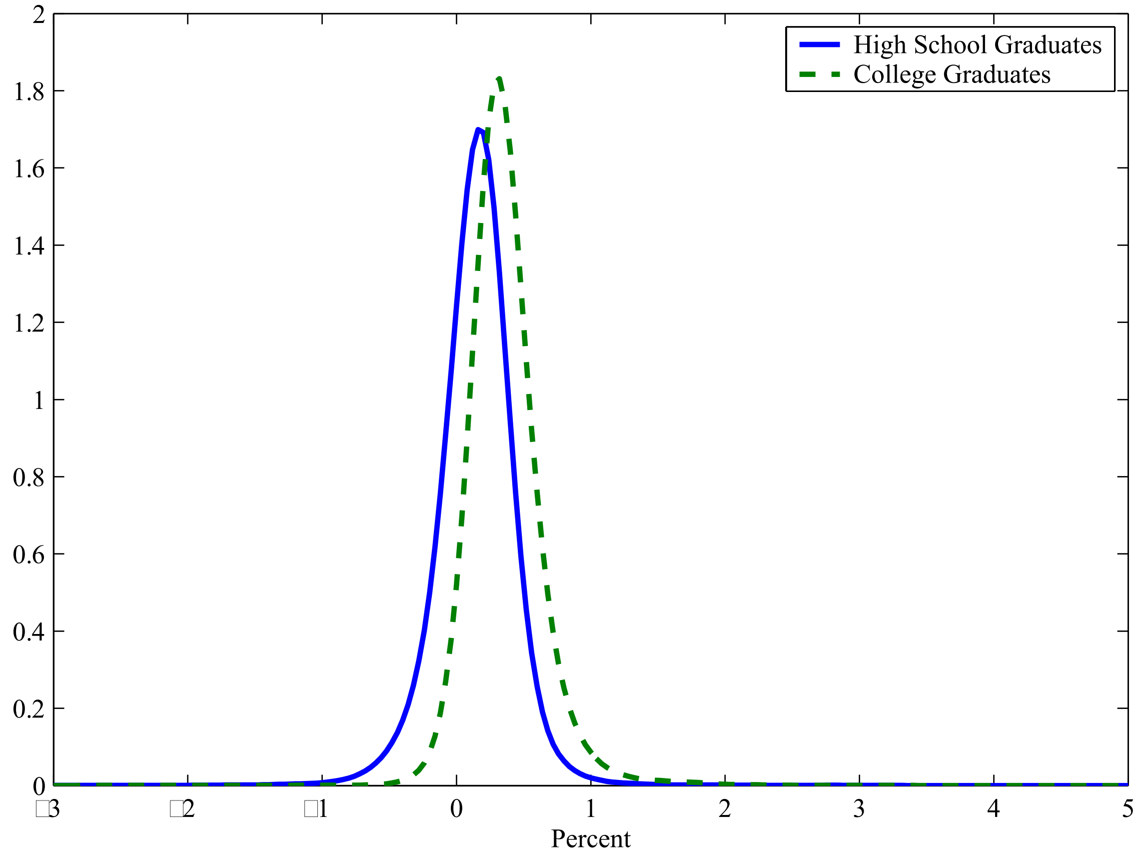
Table: Conditional Distributions for the NLS/1966 (College Earnings Conditional on High School Earnings)
$Pr(d_i<Yc<d_i+1 \mid d_j<Yh<d_j+1)$ where d_i is the ith decile of the College Lifetime Ex-Ante Earnings Distribution and d_j is the jth decile of the High School Ex-Ante Lifetime Earnings Distribution
Individual fixes known θ at their means, so Information Set=$\{\theta_1=0, \theta_2=0\}$
Corrrelation(Y_C,Y_H) = 0.9176

	College									
High School	1	2	3	4	5	6	7	8	9	10
1	0.5909	0.2479	0.0977	0.0353	0.0168	0.0072	0.0027	0.0008	0.0006	0.0000
2	0.2178	0.3016	0.2248	0.1376	0.0666	0.0318	0.0144	0.0030	0.0022	0.0002
3	0.0850	0.2020	0.2420	0.2072	0.1360	0.0752	0.0356	0.0124	0.0036	0.0008
4	0.0480	0.1242	0.1880	0.2108	0.1872	0.1256	0.0766	0.0268	0.0106	0.0022
5	0.0216	0.0674	0.1138	0.1700	0.2040	0.1910	0.1342	0.0696	0.0248	0.0034
6	0.0136	0.0346	0.0700	0.1186	0.1814	0.2088	0.1882	0.1348	0.0442	0.0058
7	0.0068	0.0166	0.0350	0.0708	0.1204	0.1824	0.2262	0.2222	0.1078	0.0118
8	0.0036	0.0094	0.0200	0.0338	0.0616	0.1186	0.2028	0.2752	0.2322	0.0428
9	0.0018	0.0022	0.0092	0.0144	0.0232	0.0526	0.1060	0.2120	0.3914	0.1872
10	0.0002	0.0002	0.0020	0.0024	0.0032	0.0070	0.0134	0.0432	0.1826	0.7458

Table: Ex-Ante Conditional Distributions for the NLSY/1979 (College Earnings Conditional on High School Earnings)
$Pr(d_i<Yc<d_i+1 \mid d_j<Yh<d_j+1)$ where d_i is the ith decile of the College Lifetime Ex-Ante Earnings Distribution and d_j is the jth decile of the High School Ex-Ante Lifetime Earnings Distribution
Individual fixes known θ at their means, so Information Set=$\{\theta_1=0,\theta_2=0\}$
Corrrelation(Y_C,Y_H) = 0.4083

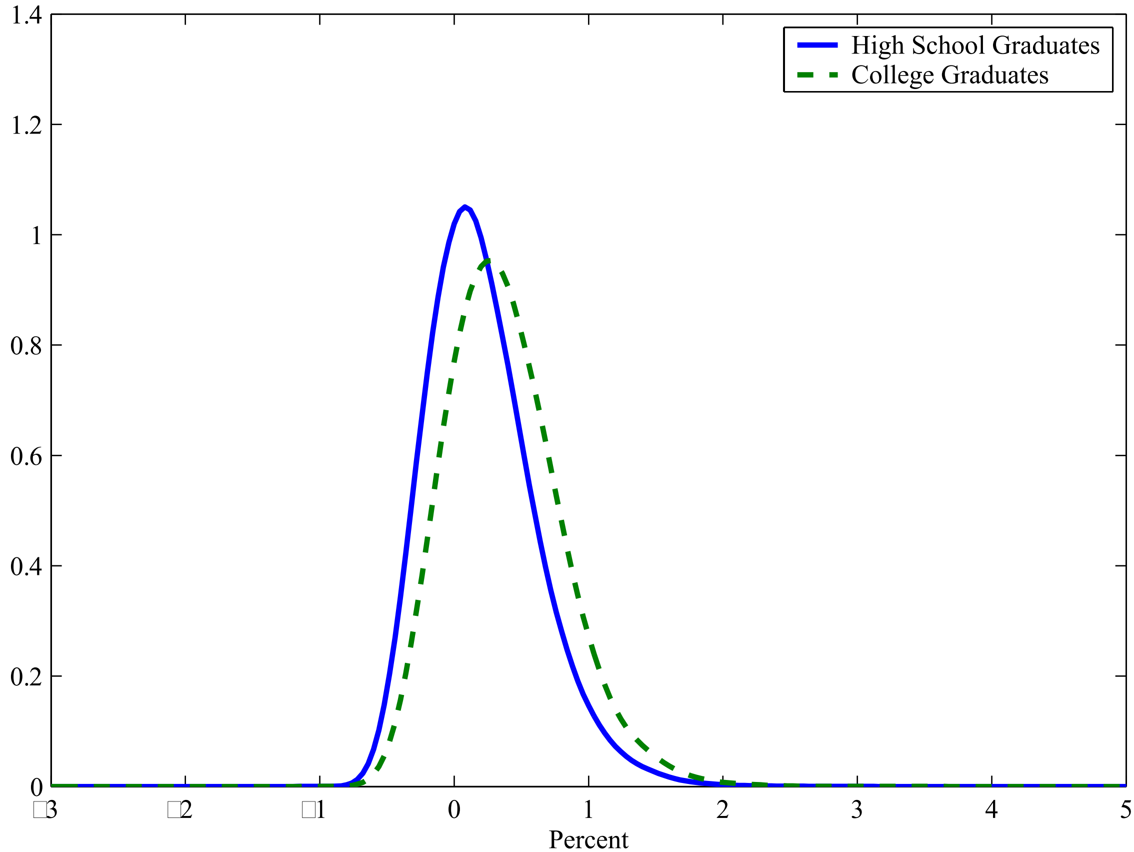
	College									
High School	1	2	3	4	5	6	7	8	9	10
1	0.1833	0.1631	0.1330	0.1066	0.0928	0.0758	0.0675	0.0630	0.0615	0.0535
2	0.1217	0.1525	0.1262	0.1139	0.1044	0.0979	0.0857	0.0796	0.0683	0.0498
3	0.1102	0.1263	0.1224	0.1198	0.1124	0.0970	0.0931	0.0907	0.0775	0.0506
4	0.0796	0.1083	0.1142	0.1168	0.1045	0.1034	0.1121	0.1006	0.0953	0.0652
5	0.0701	0.0993	0.1003	0.1027	0.1104	0.1165	0.1086	0.1112	0.1043	0.0768
6	0.0573	0.0932	0.1079	0.1023	0.1110	0.1166	0.1130	0.1102	0.1059	0.0825
7	0.0495	0.0810	0.0950	0.1021	0.1101	0.1162	0.1202	0.1174	0.1134	0.0950
8	0.0511	0.0754	0.0770	0.1006	0.1006	0.1053	0.1244	0.1212	0.1297	0.1147
9	0.0411	0.0651	0.0841	0.0914	0.1039	0.1117	0.1162	0.1216	0.1442	0.1206
10	0.0590	0.0599	0.0622	0.0645	0.0697	0.0782	0.0770	0.1028	0.1181	0.3087

Figure
Densities of Returns to College
NLS/1966 Sample



Let Y_0 denote present value of earnings in high school sector. Let Y_1 denote present value of earnings in college sector. Let $R = (Y_1 - Y_0)/Y_0$ denote the gross rate of return to college. In this graph we plot the density function of the returns to college conditional on being a high school graduate, $f(r \mid S=0)$ (the solid line), against the density function of returns to college conditional on being a college graduate, $f(r \mid S=1)$. We use kernel density estimation to smooth these functions.

Figure
Densities of Returns to College
NLSY/1979 Sample



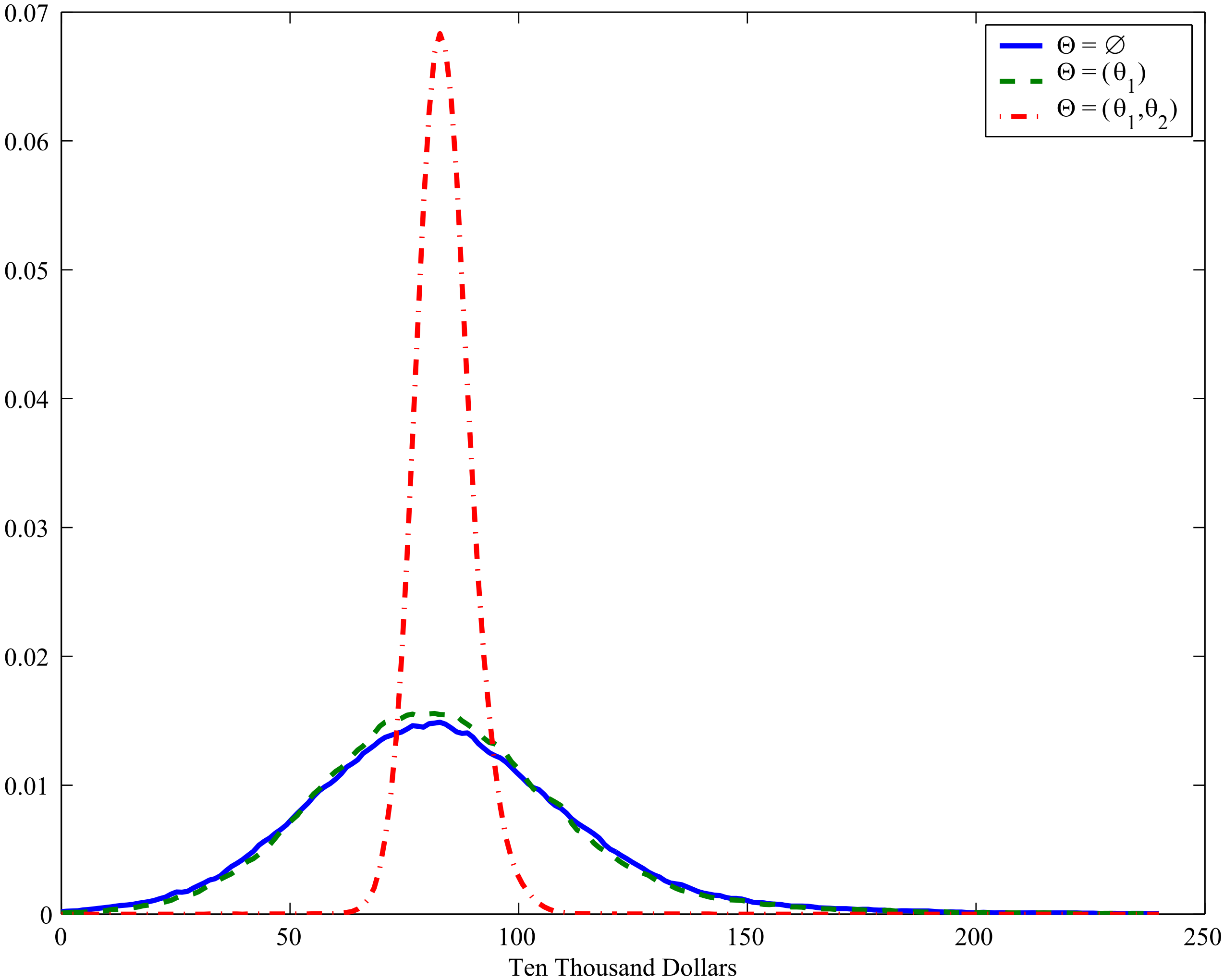
Let Y_0 denote present value of earnings in high school sector. Let Y_1 denote present value of earnings in college sector. Let $R = (Y_1 - Y_0)/Y_0$ denote the gross rate of return to college. In this graph we plot the density function of the returns to college conditional on being a high school graduate, $f(r \mid S=0)$ (the solid line), against the density function of returns to college conditional on being a college graduate, $f(r \mid S=1)$. We use kernel density estimation to smooth these functions.

Table
Lifetime Returns to College Conditional on Schooling Choices

Choice	NLSY/1966		NLSY/1979	
	Mean	Standard Error	Mean	Standard Error
High School Graduates	0.2284	0.0081	0.2055	0.0113
College Graduates	0.3421	0.0098	0.3740	0.0280
Individuals at the Margin	0.2800	0.0182	0.2828	0.0457

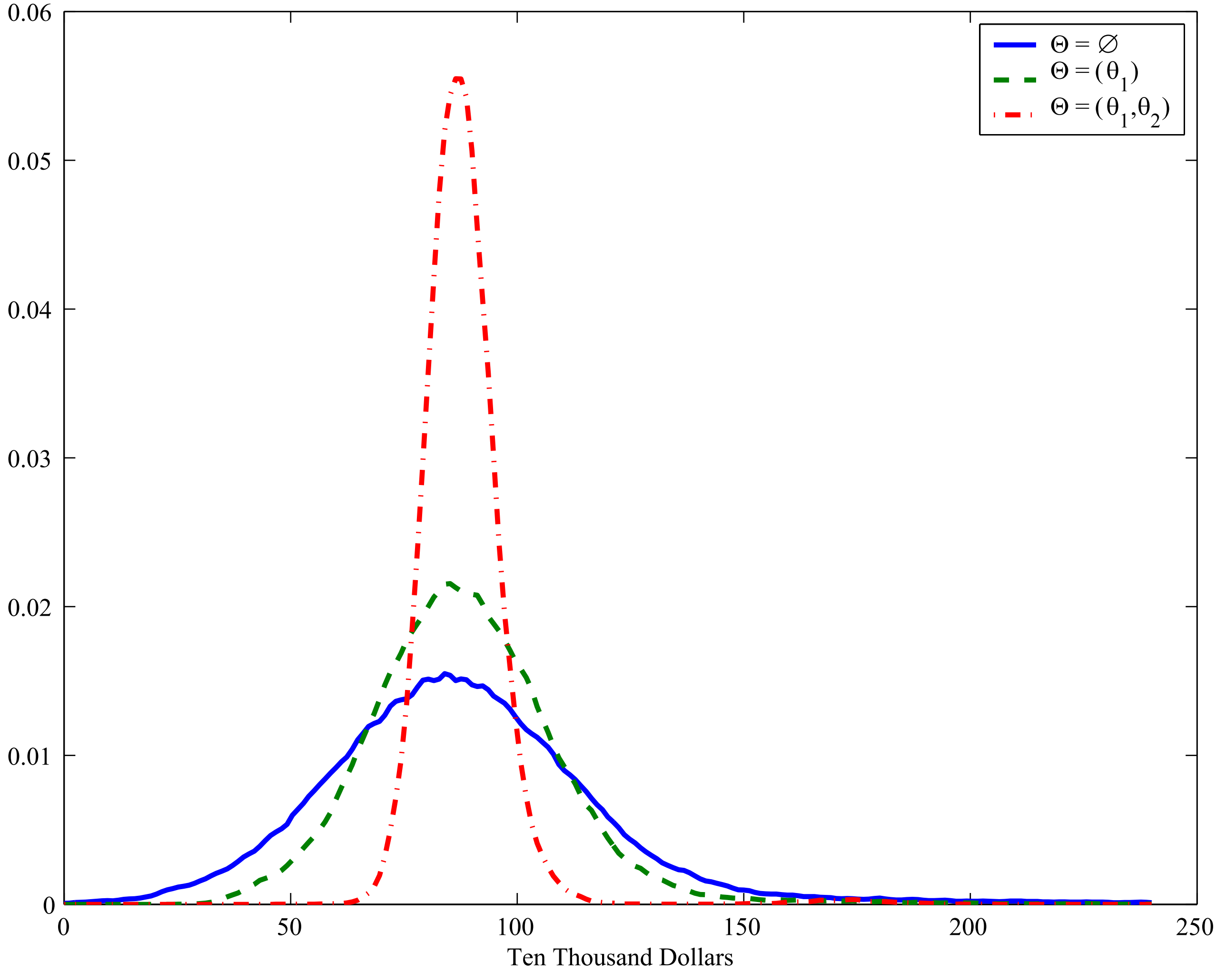
Let Y_0 denote lifetime present value of earnigs in the high school sector. Let Y_1 denote lifetime present value of earnigs in the college sector. Lifetime returns to college are $R = (Y_1 - Y_0)/Y_0$. In this table we show the mean and standard error of returns R conditional on schooling choices for both the NLSY/1966 and NLSY/1979 samples as estimated by the 2 factor model.

Figure
densities of present value of high school earnings - NLS/1966
under different information sets for the agent calculated
for the entire population irregardless of schooling choice



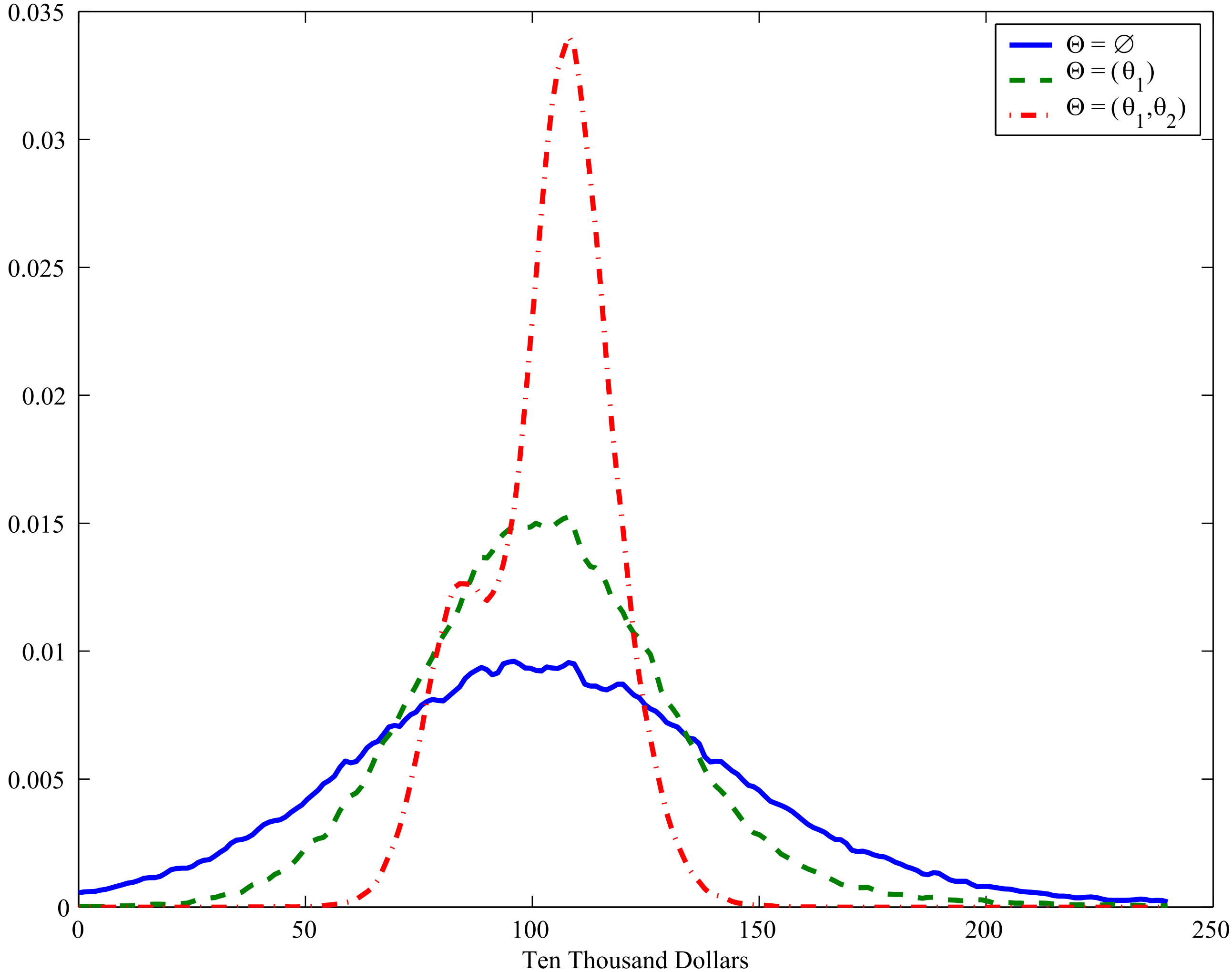
Let Θ denote the information set of the agent. Let Y_0 denote the present value of earnings in the high school sector (discounted at a 3% interest rate). Let $f(y_0|\Theta)$ denote the density of Y_0 conditional on information set Θ. The solid line plots the density of Y_0 when $\Theta=\phi$. The dashed line plots the density of Y_0 when $\Theta=\{\theta_1\}$. The dotted and dashed line plots the density of of Y_0 when $\Theta=\{\theta_1,\theta_2\}$. The X variables are in the information set of the agent. The factors θ, when known, are evaluated at their mean, which is zero.

Figure
densities of present value of high school earnings - NLSY/1979
under different information sets for the agent calculated
for the entire population irregardless of schooling choice



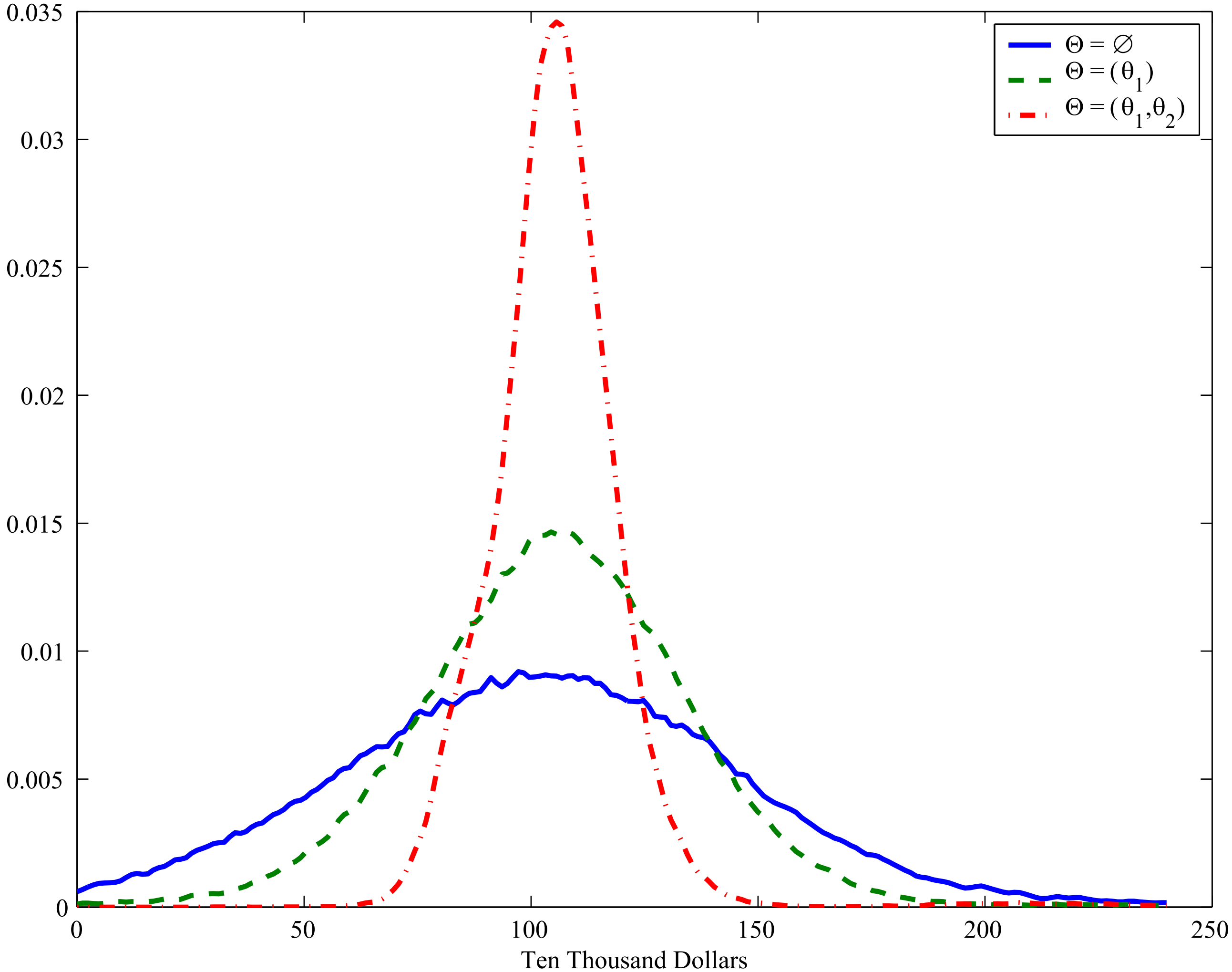
Let Θ denote the information set of the agent. Let Y_0 denote the present value of earnings in the high school sector (discounted at a 3% interest rate). Let $f(y_0|\Theta)$ denote the density of Y_0 conditional on information set Θ. The solid line plots the density of Y_0 when $\Theta=\phi$. The dashed line plots the density of Y_0 when $\Theta=\{\theta_1\}$. The dotted and dashed line plots the density of of Y_0 when $\Theta=\{\theta_1,\theta_2\}$. The X variables are in the information set of the agent. The factors θ, when known, are evaluated at their mean, which is zero.

Figure
densities of present value of college earnings - NLS/1966
under different information sets for the agent calculated
for the entire population irregardless of schooling choice



Let Θ denote the information set of the agent. Let Y_0 denote the present value of earnings in the college sector (discounted at a 3% interest rate). Let $f(y_0|\Theta)$ denote the density of Y_0 conditional on information set Θ. The solid line plots the density of Y_0 when $\Theta=\phi$. The dashed line plots the density of Y_0 when $\Theta=\{\theta_1\}$. The dotted and dashed line plots the density of of Y_0 when $\Theta=\{\theta_1,\theta_2\}$. The X variables are in the information set of the agent. The factors θ, when known, are evaluated at their mean, which is zero.

Figure
densities of present value of college earnings - NLSY/1979
under different information sets for the agent calculated
for the entire population irregardless of schooling choice



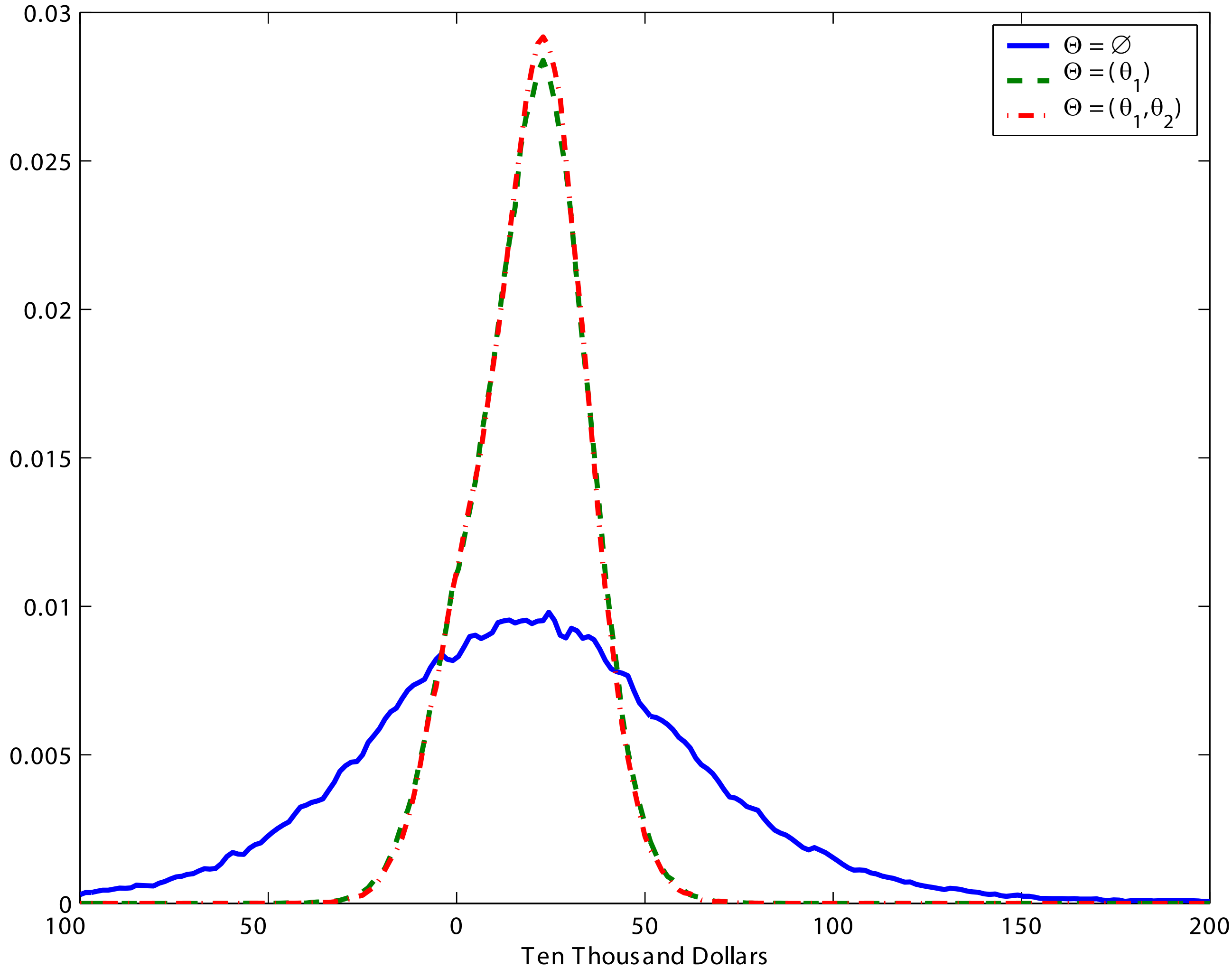
Let Θ denote the information set of the agent. Let Y_0 denote the present value of earnings in the college sector (discounted at a 3% interest rate). Let $f(y_0|\Theta)$ denote the density of Y_0 conditional on information set Θ. The solid line plots the density of Y_0 when $\Theta=\phi$. The dashed line plots the density of Y_0 when $\Theta=\{\theta_1\}$. The dotted and dashed line plots the density of of Y_0 when $\Theta=\{\theta_1,\theta_2\}$. The X variables are in the information set of the agent. The factors θ, when known, are evaluated at their mean, which is zero.

Figure

densities of present value of returns - NLS/1966

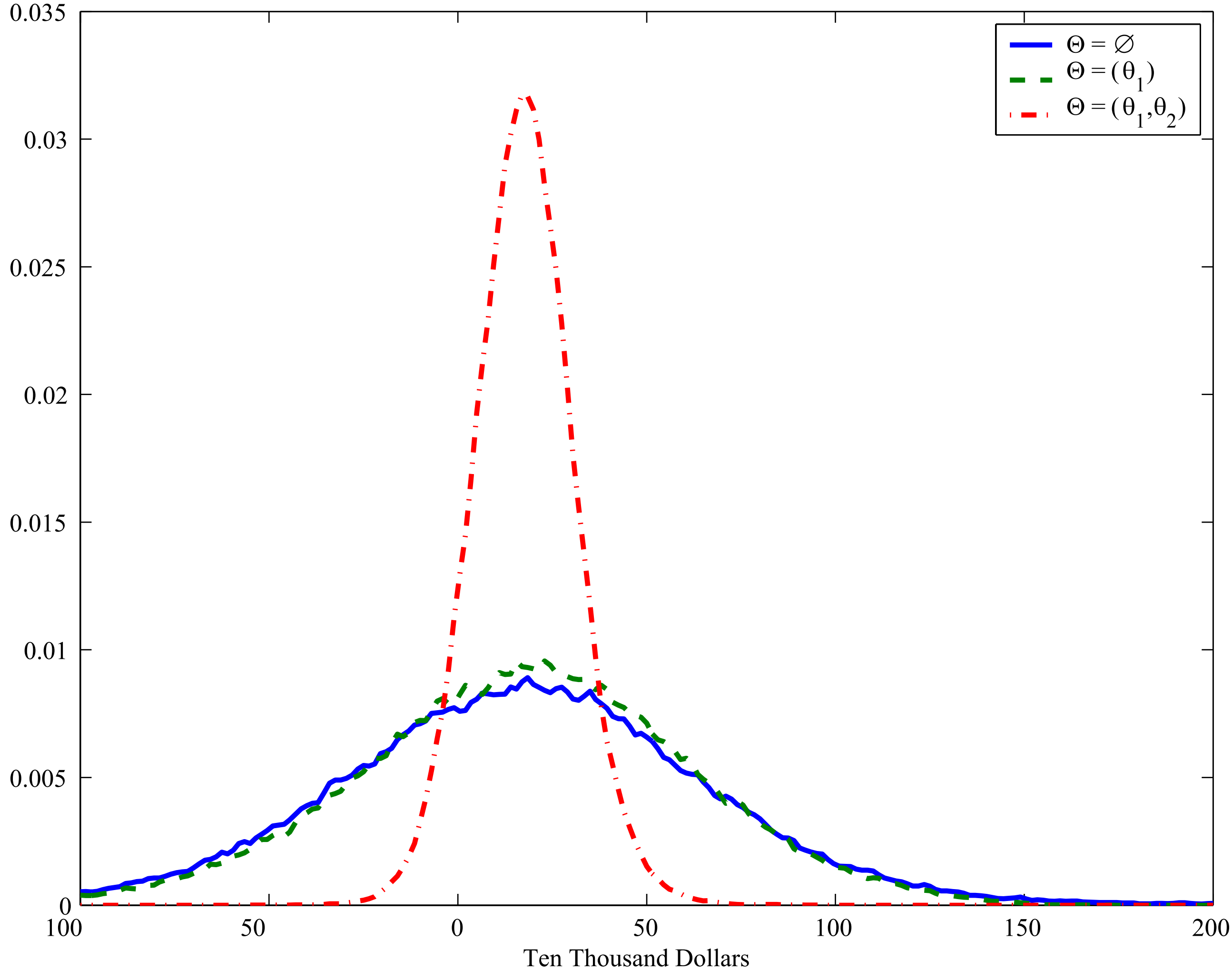
under different information sets for the agent calculated

for the entire population irregardless of schooling choice



Let Θ denote the information set of the agent. Let Y_0 denote the present value of returns (discounted at a 3% interest rate). Let $f(y_0|\Theta)$ denote the density of Y_0 conditional on information set Θ. The solid line plots the density of Y_0 when $\Theta=\phi$. The dashed line plots the density of Y_0 when $\Theta=\{\theta_1\}$.The dotted and dashed line plots the density of of Y_0 when $\Theta=\{\theta_1,\theta_2\}$. The X variables are in the information set of the agent. The factors θ, when known, are evaluated at their mean, which is zero.

Figure
densities of present value of returns - NLSY/1979
under different information sets for the agent calculated
for the entire population irregardless of schooling choice



Let Θ denote the information set of the agent. Let Y_0 denote the present value of returns (discounted at a 3% interest rate). Let $f(y_0|\Theta)$ denote the density of Y_0 conditional on information set Θ. The solid line plots the density of Y_0 when $\Theta=\phi$. The dashed line plots the density of Y_0 when $\Theta=\{\theta_1\}$.The dotted and dashed line plots the density of of Y_0 when $\Theta=\{\theta_1, \theta_2\}$. The X variables are in the information set of the agent. The factors θ, when known, are evaluated at their mean, which is zero.

Table
Agent's Forecast Variance of Present Value of Earnings*
Under Different Information Sets - NLS/1966
(fraction of the variance explained by Θ)**

	Var(Y_c)	Var (Y_h)	Var(Y_c-Y_h)
For lifetime:+			
Variance when $\Theta = \varnothing$	259.46	96.84	221.54
$\Theta = \{\theta_1\}$	16.83%	10.61%	34.90%
$\Theta = \{\theta_1, \theta_2\}$	66.99%	78.47%	51.93%

*We use an interest rate of 3% to calculate the present value of earnings.

+Variance of the unpredictable component of earnings between age 18 and 65 as predicted at age 18.

**So we would say that the variance of the unpredictable component of lifetime college earnings when the information set is $\times = \{\theta_1, \theta_2\}$ is (1-0.6699)*259.46

Table
Agent's Forecast Variance of Present Value of Earnings*
Under Different Information Sets - NLSY/1979
(fraction of the variance explained by Θ)**

	Var(Y_c)	Var (Y_h)	Var(Y_c-Y_h)
For lifetime:+			
Variance when $\Theta = \varnothing$	290.84	103.13	334.02
$\Theta = \{\theta_1\}$	46.06%	30.81%	10.68%
$\Theta = \{\theta_1, \theta_2\}$	65.13%	5594%	56.03%

*We use an interest rate of 3% to calculate the present value of earnings.

+Variance of the unpredictable component of earnings between age 18 and 65 as predicted at age 18.

**So we would say that the variance of the unpredictable component of lifetime college earnings when the information set is $\times = \{\theta_1, \theta_2\}$ is (1-0.6513)*290.84

Table
Agent's Forecast Variance of Present Value of Earnings*
Under Different Information Sets - NLS/1966

	Var(Y_c)	Var (Y_h)	Var(Y_c-Y_h)
For lifetime:			
Total Residual Variance	259.46	96.84	221.54
Share of Total Variance due to Forecastable Components	66.99%	78.46%	51.92%
Share of Total Variance due to Unforecastable Components	34.01%	21.54%	48.08%

*We use a rate of intertemporal preference ρ of 3% to calculate the present value of earnings.

Table
Agent's Forecast Variance of Present Value of Earnings*
Under Different Information Sets - NLSY/1979

	Var(Y_c)	Var (Y_h)	Var(Y_c-Y_h)
For lifetime:			
Total Residual Variance	290.84	103.13	334.02
Share of Total Variance due to Forecastable Components	65.13%	55.94%	56.04%
Share of Total Variance due to Unforecastable Components	34.87%	44.06%	43.94%

*We use a discount rate ρ of 3% to calculate the present value of earnings.

Table

Total Variance, Share of Total Variance due to Forecastable Component and Share of Total Variance due to Unforecastable Component for Different Values of the Rate of Intertemporal Preference

	White Males - NLS/1966				White Males - NLSY/1979			
	Total Variance				**Total Variance**			
Schooling Group	2%	3%	4%	5%	2%	3%	4%	5%
College	434.4680	259.4621	163.2094	107.0043	480.4024	290.8390	185.6628	120.8971
High School	149.1453	96.8433	67.6655	50.1924	158.1050	103.1273	71.7999	53.0408
Returns	342.4546	221.5384	151.7765	110.2629	481.3866	334.0230	239.6592	171.7202
	Share of Total Variance due to Forecastable Component				**Share of Total Variance due to Forecastable Component**			
Schooling Group	2%	3%	4%	5%	2%	3%	4%	5%
College	0.6860	0.6513	0.6520	0.6310	0.6793	0.6513	0.6224	0.5842
High School	0.8069	0.5594	0.7622	0.7342	0.6169	0.5594	0.5079	0.4605
Returns	0.5176	0.5604	0.5198	0.5209	0.5541	0.5604	0.5600	0.5406
	Share of Total Variance due to Unforecastable Component				**Share of Total Variance due to Unforecastable Component**			
Schooling Group	2%	3%	4%	5%	2%	3%	4%	5%
College	0.3140	0.3487	0.3480	0.3690	0.3207	0.3487	0.3776	0.4158
High School	0.1931	0.4406	0.2378	0.2658	0.3831	0.4406	0.4921	0.5395
Returns	0.4824	0.4396	0.4802	0.4791	0.4459	0.4396	0.4400	0.4594

Let $Y_{s,t}$ denote the real earnings in schooling group s at period t. Let C denote the psychic costs of attending college. Let $\mathcal{I}$ denote the information set of the agent at the time of the schooling choice. An agent decides to go to college if, and only if:

$$I = E\left(\sum_{t=1}^{24} \frac{Y_{1,t} - Y_{0,t}}{(1+\rho)^t} - C \middle| \mathcal{I}\right) \geq 0$$

where ρ is the intertemporal rate of preference. In this table we show total variance of $Y_{s,t}$, and the share of the total variance due to the variance of forecastable (but unobservable) components of $Y_{s,t}$, and the share of the total variance of unforecastable components of $Y_{s,t}$ for different values of ρ.

Table

Total Variance, Share of Total Variance due to Forecastable Component and Share of Total Variance due to Unforecastable Component for Distinct Lower Bound on Observable Earnings[1]

	White Males - NLS/1966			White Males - NLSY/1979		
	Total Variance			**Total Variance**		
Schooling Group	U$ 500.00	U$ 2,500.00	U$ 5,000.00	U$ 500.00	U$ 2,500.00	U$ 5,000.00
College	231.1051	259.4621	230.4558	293.6141	290.8390	290.5154
High School	89.4809	96.8433	83.8973	104.1163	103.1273	102.8044
Returns	202.6363	221.5384	198.0082	340.0533	334.0230	331.0939
	Share of Total Variance due to Forecastable Component			**Share of Total Variance due to Forecastable Component**		
Schooling Group	U$ 500.00	U$ 2,500.00	U$ 5,000.00	U$ 500.00	U$ 2,500.00	U$ 5,000.00
College	0.6284	0.6699	0.6316	0.6586	0.6513	0.6472
High School	0.7574	0.7847	0.7540	0.5637	0.5594	0.5572
Returns	0.4690	0.5193	0.4670	0.5716	0.5604	0.5530
	Share of Total Variance due to Unforecastable Component			**Share of Total Variance due to Unforecastable Component**		
Schooling Group	U$ 500.00	U$ 2,500.00	U$ 5,000.00	U$ 500.00	U$ 2,500.00	U$ 5,000.00
College	0.3716	0.3301	0.3684	0.3414	0.3487	0.3528
High School	0.2426	0.2153	0.2460	0.4363	0.4406	0.4428
Returns	0.5310	0.4807	0.5330	0.4284	0.4396	0.4470

[1] Let $\tilde{Y}_{s,a}$ denote the (cpi-adjusted, base year 2000) earnings of a person with schooling s and age a. The index s takes on two values: $s \in \{0,1\}$ where $s=0$ denotes a high-school graduate and $s=1$ denotes a college graduate agent. The index t takes on values in the set $t \in \{1,2,...,24\}$. Let ρ denote the rate of intertemporal preference. When $t=1$ we define $Y_{s,1}$ as:

$$Y_{s,1} = \tilde{Y}_{s,18} + \frac{\tilde{Y}_{s,19}}{(1+\rho)}$$

When $t=2$ we define $Y_{s,2}$ as:

$$Y_{s,2} = \frac{\tilde{Y}_{s,20}}{(1+\rho)^2} + \frac{\tilde{Y}_{s,21}}{(1+\rho)^3}$$

In general, we define $Y_{s,t}$ as:

$$Y_{s,t} = \frac{\tilde{Y}_{s,18+2(t-1)}}{(1+\rho)^{2(t-1)}} + \frac{\tilde{Y}_{s,19+2(t-1)}}{(1+\rho)^{2(t-1)+1}}$$

In our empirical work we drop observations for which $Y_{s,t}$ is less than the earnings of a person who works part time for 50 weeks and receives a wage rate which corresponds to half minimum wage. This value is around $U\$2,500.00$ in dollars of 2000. In this table, we check the robustness of our results by considering two alternative values for lower bound earnings: $U\$500.00$ and $U\$5,000.00$.

Table

Share of Total Variance due to the Unforecastable Component According to Different Methodologies

	NLS/1966		NLSY/1979	
	College	High School	College	High School
Two-Factor Model	0.3401	0.2154	0.3487	0.4406
Fixed Effect Model - Earnings	0.3326	0.3466	0.3951	0.4422
Random Effect - Earnings	0.3996	0.4513	0.4633	0.5183
Fixed Effect Model - Log Earnings	0.3349	0.3483	0.3845	0.3917
Random Effect - Log Earnings	0.4048	0.4402	0.4502	0.4701

Let $Y_{s,t}$ denote present value at age 18 of real earnings in schooling level s at period t. Let $\mu_{s,t}(X)$ denote the mean of $Y_{s,t}$. We model $Y_{s,t}$ as:

$$Y_{s,t} = \mu_{s,t}(X) + \alpha_{s,t}\theta_1 + \delta_{s,t}\theta_2 + \varepsilon_{s,t}$$

Consequently, the total variance of present value of lifetime earnings in schooling level s, V_s, is:

$$V_s = Var\left(\sum_{t=1}^{24} Y_{s,t} - \mu_{s,t}(X)\right) = \sigma_{\theta_1}^2\left(\sum_{t=1}^{24}\alpha_{s,t}\right)^2 + \sigma_{\theta_2}^2\left(\sum_{t=1}^{24}\delta_{s,t}\right)^2 + \sum_{t=1}^{24}\sigma_{\varepsilon_{s,t}}^2$$

Because we estimate that the agent knows both factors θ_1 and θ_2 at the time of the schooling choice, the share of the total variance that is due to the unforecastable component is x_2 where:

$$x_2 = \frac{\sum_{t=1}^{24}\sigma_{\varepsilon_{s,t}}^2}{\sigma_{\theta_1}^2\left(\sum_{t=1}^{24}\alpha_{s,t}\right)^2 + \sigma_{\theta_2}^2\left(\sum_{t=1}^{24}\delta_{s,t}\right)^2 + \sum_{t=1}^{24}\sigma_{\varepsilon_{s,t}}^2}$$

In row one, column one, we report x_2 for college for the two-factor model that we estimate in this paper.

For the fixed effect models we assume that the fixed effect is known at the time of the schooling choice and that the rest of the residuals are unknown by the agent at the time of the schooling choice. So we model earnings (or log earnings) as:

$$Y_{t,i} = \mu + \beta_1 age + \beta_2 age^2 + \beta_3 age^3 + \beta_4 age^4 + \nu_i + u_{i,t}$$

where $p(age)$ is a polynomial of degree four in a. Let σ_ν^2 denote the variance of the fixed effect, let $\sigma_{u_t}^2$ denote the variance of the innovations in income, which we also assume is $i.i.d$ over time and across agents. The total variance is, consequently, $V_{FE} = \sigma_v^2 + \sigma_{u_t}^2$. The share of total variance due to unforecastable component is, consequently,

$$x_3 = \frac{\sigma_{u_t}^2}{\sigma_v^2 + \sigma_{u_t}^2}$$

We proceed in the same way for the random effects models.