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

Flavio Cunha† James Heckman‡

Federal Reserve Bank of Minneapolis

Micro Conference

First draft, May 2005

This draft, October 11, 2006

1 Introduction

A vast literature documents an increase in wage inequality in the American economy 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.

Increased variability is not the same as increased uncertainty. The goal of this paper is to understand how much of the recent increase in inequality is due to an increase in heterogeneity or sorting that is predictable to the agents but not known to the observing economist and how much is due to uncertainty. We establish that any explanation for the recent rise in wage inequality has to recognize that individuals possess an array of abilities and not just a single skill, that the joint distribution of these abilities has changed over time along with the prices of the vector of abilities, and that increasing uncertainty at the agent level is an important fact of recent economic life.

The Gorman-Lancaster model of earnings is a useful framework for thinking about these issues.[1] It

*This research was supported by NIH R01-HD043411. Cunha is grateful to Rob Dugger, the Committee on Economic Development, PAES project, America's Promise and the Claudio Haddad Dissertation Fund at the University of Chicago. This research is an outgrowth of the work reported in Cunha, Heckman, and Navarro (2005).

†University of Chicago, Department of Economics.

‡University of Chicago, Department of Economics; University College Dublin; and the American Bar Foundation.

[1]See Heckman and Scheinkman (1987).

writes earnings at schooling level s at time t, $Y_{s,t}$ as

$$Y_{s,t} = X\beta_{s,t} + \theta\alpha_{s,t} + \varepsilon_{s,t}.$$

We suppress the subscripts for individual i. Let X and θ be the observed and unobserved skills of an agent (X and θ may be vectors); $\beta_{s,t}$, $\alpha_{s,t}$ the prices of these skills in market s; and $\varepsilon_{s,t}$ unmeasured factors, known by the individual, or some forms of productivity shocks, unknown by the individual before its realization. In versions of the Gorman-Lancaster model of earnings, agents pick their sector s based on utility maximization.[2]

The analysis of Juhn, Murphy, and Pierce (1993) can be fit in this framework. It links the increase in wage inequality to technological progress. The authors claim that as technology advances, the demand for skill has grown at a faster pace than its supply, causing skill prices to increase and the wage gap between skilled and unskilled workers to widen.[3] To arrive at this conclusion, they interpret quantiles of the distribution of the residuals in log wage equations as quantiles of the distribution of ability and interpret the rise in inequality as a rise in unobserved skill prices. This is possible in the Gorman-Lancaster model if the following conditions hold for $Y_{s,t}$ interpreted as log earnings: 1) there is one, and only one, type of unobserved ability (θ is a scalar); 2) the distribution of this ability across individuals is invariant over time (θ is a time-invariant random variable); 3) there are no unanticipated shocks in earnings so that the economy is described by a deterministic environment ($\varepsilon_{s,t} = 0$); and 4) the price of θ is the same for all schooling levels ($\alpha_{s,t} = \alpha_t$). The error term is $\theta\alpha_t$, and since θ is invariant, α_t can be identified up to scale. If θ varies with t, it is not possible to distinguish changing variances of θ from changing α_t without using further information.

In this paper, we use further information to distinguish changes in α_t from changes in the variance of θ. We compare the life cycle earnings of individuals born between 1941 and 1952 to those born between 1957 and 1964. We test and reject the hypothesis that there is only one component of unobserved ability. In the recent period, θ is a vector of dimension six. This finding is consistent with the empirical evidence presented in a large body of literature that shows that abilities are multiple in nature.[4] These abilities

[2]See for example Heckman and Sedlacek (1985) or Taber (2001).

[3]See Tinbergen (1975) for an early analysis of the race between demand and supply.

[4]See, for example, Heckman and Rubinstein (2001), Carneiro and Heckman (2003), Heckman, Stixrud, and Urzua (2006) and Gould (2002).

may be combined in different quantities to produce the same wage outcome. For example, an individual with low stocks of cognitive ability but a persistent personality can be more successful than a smart person with no motivation. Hence, the assumption of a one-to-one mapping from quantiles of the distribution of wage residuals to quantiles of the distribution of abilities is not supported by the data.

We reject the hypothesis that the distribution of unobserved abilities has remained fixed over time. Our findings are consistent with evidence by Gottschalk and Moffitt (1994), who show a substantial increase in the variance of the θ when they compare earnings in the period 1970–1978 with earnings in the period 1979–1987. This could be due to an increase in the price of the skill or an increase in the variance of θ. We show how to separate the increase in $\alpha_{s,t}$ from increases in the variance of θ over time.

We demonstrate that an increase in microeconomic uncertainty plays an important role in explaining the increase in wage inequality. Again, our findings are consistent with the analysis of Gottschalk and Moffitt (1994), who document an increase in "earnings instability" (the $\varepsilon_{s,t}$), demonstrating that the variance of temporary shocks rose considerably from the period 1970–1978 to the period 1979–1987. We use choices to estimate the information set of the agents at the age college enrollment decisions are made. As they age, agents learn about components of θ and the $\varepsilon_{s,t}$. We show that the stochastic process that fits the unforecastable components in labor income has changed across cohorts and that, as a result, uncertainty, or earnings instability, or turbulence, have increased substantially.[5]

We model schooling and earnings equations jointly to identify the information set of agents at the time college going decisions are made. Modelling schooling choices is not just an econometric exercise to correct for selection in earnings. Schooling choices are the source of information that allows us to separate what is known and acted on by individuals at the time schooling choices are made — which we call heterogeneity — from what is not known — which we call uncertainty.

This paper is in six parts. Part 2 presents the model. Part 3 presents the econometrics and the empirical results. Part 4 relates our framework to previous models in the literature. Part 5 discusses a more general framework. Part 6 concludes.

[5]See Ljungqvist and Sargent (2004), who discuss the rise of turbulence in the recent economy.

2 The Model

We estimate the information set of the agents. We identify this information set by analyzing both the choices and the outcomes associated with choices made by the individuals.

2.1 Earnings Equations

To motivate our econometric procedures, we start by describing the earnings equations for $t = 1, \ldots, T$, which are life cycle outcomes over horizon T. We assume that $(Y_{0,t}, Y_{1,t})$, $t = 1, \ldots, T$, have finite means and can be expressed in terms of conditioning variables X in the following manner:

$$Y_{0,t} = X\beta_{0,t} + U_{0,t} \tag{1}$$

$$Y_{1,t} = X\beta_{1,t} + U_{1,t}, \qquad t = 1, \ldots, T. \tag{2}$$

The error terms $U_{s,t}$ are assured to satisfy $E(U_{s,t} \mid X) = 0$, $s = 0, 1$.

2.2 Choice Equations

We assume that agents make schooling choices based on expected present value income maximization given information set $\mathcal{I}$. We discuss this assumption in section 5. Write the index I of present values as

$$I = E\left[\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}(Y_{1,t} - Y_{0,t}) - C \middle| \mathcal{I}\right], \tag{3}$$

where C is the cost of attending college. We denote by Z and U_C the observable and unobservable determinants of costs, respectively. We assume that costs can be written as

$$C = Z\gamma + U_C. \tag{4}$$

If we define $\mu_I(X, Z) = \sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1} X\left(\beta_{1,t} - \beta_{0,t}\right) - Z\gamma$ and $U_I = \sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}(U_{1,t} - U_{0,t}) - U_C$, and substitute (1), (2), and (4) into (3) we obtain

$$I = E\left[\mu_I(X, Z) + U_I \middle| \mathcal{I}\right]. \tag{5}$$

More generally, we define U_I as the error in the choice equation and it may or may not include $U_{1,t}$, $U_{0,t}$, or U_C, depending on what is in the agent's information set. Similarly, $\mu_I(X,Z)$ may only be based on expectations of future X and Z at the time schooling decisions are made. The schooling decision of agents is

$$S = \mathbf{1}\left[I \geq 0\right]. \tag{6}$$

2.3 Test Score Equations

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

$$M_k = X^M \beta_k^M + U_k^M. \tag{7}$$

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

2.4 Heterogeneity and Uncertainty

Consider college earnings in period t, $Y_{1,t}$. It is only observed for the agents who choose to attend college $(S = 1)$. Consequently, from a standard selection argument, from observational data we can only identify the cross-sectional mean college earnings conditional on explanatory variables X and $S = 1$:

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

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

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

Consequently, because $E[Y_{1,t}| X, S=1] = E[Y_{1,t}| X, E[U_I|\mathcal{I}] \geq -\mu_I(X,Z)]$, it follows that

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

To extract the information known to agents, we separate U_I into two components. The first component, $E[U_I|\mathcal{I}]$, is used by the agent to make schooling choices. This expectation is determined by the elements in the information set of the agent which influence their schooling decision. The second component, $U_I - E[U_I|\mathcal{I}]$, does not affect selection into schooling because it is not known to the agent at the time schooling decisions are made.

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

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

Clearly $(U_{s,t} - E[U_{s,t}|\mathcal{I}])$ affects realized earnings. To determine the unobservable components that are in the information set of the agent we need to determine which specification of the information set $\mathcal{I}$ better characterizes the dependence between schooling choices and future earnings. We can determine the components that are not in the information set of the agent by varying the specification of $(U_{s,t} - E[U_{s,t}|\mathcal{I}])$ while keeping $\mathcal{I}$ fixed, 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 $E[U_{s,t}|\mathcal{I}]$ and $(U_{s,t} - E[U_{s,t}|\mathcal{I}])$ in a framework convenient for testing.

2.5 Factor Models

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 the test score equations into 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. In this notation, we can write equation (7) as

$$M_k = X^M\beta_k^M + \alpha_1^M\theta_1 + \varepsilon_1^M. \tag{8}$$

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 of X^M and ε_k^M. The ε_k^M are mutually independent and independent of θ. Modelling test scores in this fashion allows them to be noisy measures of cognitive skill.

2.5.1 Earnings and Choice Equations

We decompose the error terms in the earnings equations into three components. The first component is the cognitive factor θ_1. The second component is a "productivity" factor θ_2 which affects earnings and schooling choices, but not test scores. In our empirical work, we fit models with as many as six factors, but for expositional simplicity, in this section we use a two factor model. The third component is the idiosyncratic error term which affects only the period-t, schooling-s earnings equation, $\varepsilon_{s,t}$. We rewrite equations (1) and (2) as

$$Y_{0,t} = X\beta_{0,t} + \alpha_{1,0,t}\theta_1 + \alpha_{2,0,t}\theta_2 + \varepsilon_{0,t} \tag{9}$$

and

$$Y_{1,t} = X\beta_{1,t} + \alpha_{1,1,t}\theta_1 + \alpha_{2,1,t}\theta_2 + \varepsilon_{1,t}. \tag{10}$$

We assume that factor θ_j is independent from X, $\varepsilon_{s,t}$, and θ_l for $l \neq j$ and for all s, t. The $\varepsilon_{\ell,t}$, $\ell = 0,1$ and $t = 1, \ldots, T$, are mutually independent.

The cost equation is decomposed like the earnings equations, so that (4) can be rewritten as

$$C = Z\gamma + \alpha_{1,C}\theta_1 + \alpha_{2,C}\theta_2 + \varepsilon_C. \tag{11}$$

Given the specifications with the factors in (9), (10), and (11), we can rewrite the schooling choice equation as

$$I = E\left[\begin{array}{c|c} \sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)_i^{t-1} X\left(\beta_{1,t}-\beta_{0,t}\right) - Z\gamma + \theta_1\left[\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}\left(\alpha_{1,1,t}-\alpha_{1,0,t}\right)-\alpha_{1,C}\right] & \\ +\theta_2\left[\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}\left(\alpha_{2,1,t}-\alpha_{2,0,t}\right)-\alpha_{2,C}\right] + \sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}\left(\varepsilon_{1,t}-\varepsilon_{0,t}\right)-\varepsilon_C & \mathcal{I} \end{array}\right]. \tag{12}$$

We assume that for all subscripts the ε's are mutually independent and independent of the X, Z, and (θ_1, θ_2).

2.6 The Estimation of the Components in the Information Set

We now show how to determine the unobservable components of the information set $\mathcal{I}$ of the agent at the age schooling choices are made by exploiting the structure of factor models. Assume that X, Z, and ε_C are in the information set $\mathcal{I}$. To economize notation, define

$$\alpha_{k,I} = \sum_{t=1}^{T} \left(\frac{1}{1+\rho}\right)^{t-1} (\alpha_{k,1,t} - \alpha_{k,0,t}) - \alpha_{k,C} \text{ for } k = 1, 2. \tag{13}$$

Suppose that we propose that $\{\theta_1, \theta_2\} \subset \mathcal{I}$, but $\varepsilon_{s,t} \notin \mathcal{I}$. Given the definitions of $\alpha_{1,I}, \alpha_{2,I}$ and $\mu_I(X, Z)$, if the null hypothesis is true, the index governing schooling choices is

$$I = \mu_I(X, Z) + \alpha_{1,I}\theta_1 + \alpha_{2,I}\theta_2 + \varepsilon_C. \tag{14}$$

Suppose for the sake of argument that we know both $\mu_I(X, Z)$ and $\beta_{s,t}$ for all s and t. From discrete choice analysis we can essentially determine I up to scale.[6] The econometric literature shows that for purposes of studying identification, it does no harm to assume we know I. Given observations on X and Z we can obtain from the data the covariance between the terms $I - \mu_I(X, Z)$ and $Y_{1,1} - X\beta_{1,1}$. Under the null hypothesis $\{\theta_1, \theta_2\} \subset \mathcal{I}$, this covariance is

$$Cov\left(I - \mu_I(X, Z), Y_{1,1} - X\beta_{1,1})\right) = \alpha_{1,I}\alpha_{1,1,1}\sigma^2_{\theta_1} + \alpha_{2,I}\alpha_{2,1,1}\sigma^2_{\theta_2}. \tag{15}$$

We can test the null $\{\theta_1, \theta_2\} \subset \mathcal{I}$ against many different alternative hypotheses. 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 expected present value of schooling (12) can be written as

$$I = \mu_I(X, Z) + \alpha_{1,I}\theta_1 + \varepsilon_C. \tag{16}$$

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

$$Cov\left(I - \mu_I(X, Z), Y_{1,1} - X\beta_{1,1})\right) = \alpha_{1,I}\alpha_{1,1,1}\sigma^2_{\theta_1}, \tag{17}$$

[6] See, e.g., Matzkin (1992).

and the difference between the school index generated by the null and the alternative hypothesis is the term $\alpha_{2,I}\alpha_{2,1,1}\sigma^2_{\theta_2}$ that appears in (15) but not in (17). We can characterize these tests by defining parameters Δ_{θ_1} and Δ_{θ_2} such that

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

Agents know and act on the information contained in factors 1 and 2, so that $\{\theta_1, \theta_2\} \subset \mathcal{I}$, if we reject the hypothesis that 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 $\mu_I(X,Z)$, the parameters β and α in the test and earnings equations, the distribution of the factors, F_θ, as well as the distribution of idiosyncratic components F_ε in the test, earnings and cost equations. Carneiro, Hansen, and Heckman (2003) present formal proofs of semi-parametric identification of this approach. An appendix discusses an intuitive explanation of identification using normality. Normality is *not* required to secure identification, and our estimates are *not* based on normality assumptions.

3 Empirical Results

In order to study the evolution of labor earnings risk in the U.S. economy we compare two distinct samples. The first sample consists of white males born between 1957 and 1964. We obtain information on them from NLSY/1979 data pooled from their birth cohort counterparts from the PSID data. The second sample consists of white males born between 1941 and 1952 who are surveyed from the NLS/1966 combined with their birth cohort counterparts from the PSID data. We pool the surveys to increase sample sizes. In what follows, we refer to the samples as NLSY/1979 and NLS/1966, respectively.

Following our theoretical analysis, we consider only two schooling choices: high school and college graduation. We use $s = 0$ to denote those who stop their schooling at high school and $s = 1$ to denote those who go to college.

The Web Data Appendix Tables 1 and 2 present descriptive statistics of the NLS/1966 and NLSY/1979 samples, respectively. In both samples, college graduates have higher test scores, fewer siblings and parents with higher levels of education. In the NLSY/1979, college graduates are more likely to live in locations where the tuition for four-year college is lower. This is not true for the college graduates in NLS/1966[7].

In our empirical analysis we consider labor income from ages 22 to 41. Web Data Appendix Tables 3 and 4 show mean and standard deviations for earnings in high school and college for NLSY/1979 and NLS/1966, respectively[8]. In both data sets, college graduates start off with lower mean labor income than high-school graduates. The overtaking age in both data sets is 26. The standard error of earnings tends to increase with age for high school and college graduates in both data sets.

In both data sets, we observe cognitive test scores, which are the left-hand side variables in the measurement system for cognitive ability (M in the notation of section 2). For the NLSY/1979 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 others from it. This justifies our interpretation of θ_1 as ability.

In the NLS/1966, there are many different achievement tests, but we use the two most commonly reported ones: the OTIS/BETA/GAMMA and the California Test of Mental Maturity (CTMM). One problem with the NLS/1966 sample is that there are no respondents for whom we observe scores from at least two distinct tests. That is, for each respondent we observe at most one test score. We supplement the information from these test scores by considering other proxies for cognitive achievement. These are the tests on "knowledge of the world of work."

There are three different tests. The first is a question regarding occupation. The respondent is asked about the duties of a given profession, say draftsman. For this specific example, there are three possible answers: (a) makes scale drawings of products or equipment for engineering or manufacturing purposes, (b) mixes and serves drinks in a bar or tavern, (c) pushes or pulls a cart in a factory or warehouse. The second test asks the level of education associated with each occupation mentioned in the first test. The third test is an earnings comparison test. Specifically, it asks the respondent who he/she believes makes more in a year, for pairs of occupations.

In Web Data Appendix Table 5 we show that even after controlling for parental education, number of siblings, urban residence at age 14, and dummies for year of birth, the "knowledge of the world of work" test scores are correlated with the cognitive test scores. The correlation with OTIS/BETA/GAMMA and

[7]See Web Data Appendix for details on the construction of the tuition variables used in this paper.

[8]Earnings figures are adjusted for inflation using the CPI and we take the year 2000 as the base year.

CTMM is stronger for the occupation and education tests than for the earnings-comparison test.

We model the test score j, M_j, as

$$M_j = X^M \beta_j^M + \theta_1 \alpha_j^M + \varepsilon_j^M. \tag{18}$$

The covariates X^M include family background variables, year of birth dummies, and characteristics of the individuals at the time of the test.[9] To set the scale of θ_1, we normalize $\alpha_1^M = 1$.

One of the advantages of using factor models instead of the test score itself is that factor models allow for test scores to be noisy measures of cognitive skills. Another advantage is that this method does not require the observation of test scores for all individuals. This is important because full samples exhibit different earnings characteristics than incomplete samples. Web Data Appendix Table 6 and Web Data Appendix Figures 1 through 4 compare the time series of the means and standard errors of earnings in the full NLSY/1979 sample and the NLSY/1979 subsample with observed test scores. While mean earnings are the same in both samples, the standard error seems to be more volatile in the subsample with observed test scores than in the full sample. Web Data Appendix Table 7 and Web Data Appendix Figures 5 through 9 show the same comparison for the NLS/1966, but now the conclusion is different. Mean high-school earnings from age 35 to 41 tend to be higher in the subsample with observed test scores than in the full sample. The same seems to be true for the time series for the standard error of college earnings. Although there are no differences in mean college earnings, the standard errors diverge in the distinct samples, and they are much higher in the full-sample than in the subsample with observed test scores (see Web Data Appendix Figure 8). Web Data Appendix Tables 8-10 compare the serial correlation matrices for NLSY/1979 in high-school, college and overall sample, respectively. Parallel information for NLS/1966 survey is reported in Web Data Appendix Tables 11-13. Although there are few differences in the pattern of serial correlation when one compares the full sample with the subsample with observed test scores, the information contained in the subsample with observed test scores alone would not suffice to compute all the cells in the correlation matrix.

[9] In both NLSY/1979 and NLS/1966 we include mother's education, father's education, number of siblings, urban residence at age 14, dummies for year of birth of the individuals, and an intercept. In the NLSY/1979 sample we also control for the fact that the test taker is enrolled at school and the highest grade completed at the time of the test. In the NLS/1966 all of the respondents were enrolled at school at the time of the test (in fact, the test score is obtained in a survey from schools). We don't know the highest grade completed at the time of the test for the NLS/1966 sample. See Herriott and Kohen (undated, found in our Web Data Appendix) for an analysis of the test scores in NLS/1966.

For the NLSY/1979, a six factor model fits the data best:

$$Y_{s,t} = X\beta_{s,t} + \theta_1\alpha_{1,s,t} + \theta_2\alpha_{2,s,t} + \theta_3\alpha_{3,s,t} + \theta_4\alpha_{4,s,t} + \theta_5\alpha_{5,s,t} + \theta_6\alpha_{6,s,t} + \varepsilon_{s,t},\ t = 1, \ldots, T^*,\ s = 0, 1, \quad (19)$$

where $t = 1$ corresponds to age 22 and T^* is age 41. For the NLS/1966, only a five factor model is required to fit the data. The identification of the model requires the normalization of some of the factor loadings. Table 1A shows the factor loading normalizations imposed in the NLSY/1979; the same information for the NLS/1966 is found in Table 1B. In both samples, the covariates X are urban residence at age 14, dummies for year of birth of the individual, and an intercept.

The cost function C is

$$C = Z\gamma + \theta_1\alpha_{1,C} + \theta_2\alpha_{2,C} + \theta_3\alpha_{3,C} + \theta_4\alpha_{4,C} + \theta_5\alpha_{5,C} + \theta_6\alpha_{6,C} + \varepsilon_C. \quad (20)$$

The covariates Z are urban residence at age 14; dummies for year of birth; an intercept; and variables that affect the costs of going to college but do not affect outcomes $Y_{s,t}$, such as mother's education, father's education, number of siblings, and local tuition. Because we only have earnings data into the early 40's for both samples, the truncated discounted earnings after the 40's are absorbed into the definition of C.

Each factor θ_k, is generated by a mixture of J_k normal distributions,

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

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 $p_{k,j} > 0$. Ferguson (1983) shows that 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 Carneiro, Hansen, and Heckman (2003). For all factors, a three-component model ($J_k = 3$, $k = 1, \ldots, 6$) is adequate. For all $\varepsilon_{s,t}$ we use a four-component model.[10]

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

3.1 How the model fits the data

The model fits the data well. Figures 1A and 1B plot the densities of earnings in the overall sample for the NLSY/1979 and NLS/1966, respectively. In Figures 2A and 2B we show the comparison of actual versus model prediction for high school earnings at age 31, for both NLSY/1979 and NLS/1966 samples. Similarly, in Figures 3A and 3B we plot the actual versus the predicted densities of college earnings at age 31.[11] When we perform formal tests of equality of predicted versus actual densities, we pass these tests for most of the ages. The model fits the NLS/1966 data marginally better than it fits the NLSY/1979 data.

We also conduct χ^2 goodness-of-fit tests for the earnings correlation matrices. In Table 1, we show that the six factor model can fit the correlation matrix for the NLSY/1979 sample. We can not reject the equality of actual and predicted correlation matrix for the NLS/1966 model when we use our five factor model. However, a five factor model would not be able to replicate the earnings correlation matrix for the NLSY/1979. Consequently, in what follows, we work with a six factor model for the NLSY/1979 and a five factor model for NLS/1966.

3.2 The Evolution of Joint Distributions and Returns to College

In estimating the distribution of earnings in counterfactual schooling states within a policy regime (e.g., the distributions of college earnings for people who actually choose to be high school graduates under a particular tuition policy), one standard approach is to assume that both college and high school distributions are the same except for an additive constant—the coefficient of a schooling dummy in an earnings regression possibly conditioned on the covariates. We relax this assumption and identify the joint distribution of counterfactuals without imposing this condition or related rank invariance conditions.[12]

We identify both *ex ante* and *ex post* joint distributions. Let $E\left(Y_s|\mathcal{I}\right)$ denote the *ex ante* present value of lifetime earnings at schooling level s. Suppose that we want to compute the means and covariances between *ex ante* college and *ex ante* high-school earnings conditional on the information set, which we estimate to be based on three factors as documented below. For this case, the mean present value of

[11] In the website for the paper we show the same figures for all ages, for the overall, high-school, and college earnings, for both the NLSY/1979 and NLS/1966. We abstain from reporting them in the paper because there are 120 such figures.

[12] Abbring and Heckman (2007) discuss alternative assumptions used to identify joint counterfactual distributions.

earnings is

$$E\left(Y_s \middle| \mathcal{I}\right) = \sum_{t=1}^{T^*} \frac{X\beta_{s,t} + \theta_1\alpha_{1,s,t} + \theta_2\alpha_{2,s,t} + \theta_3\alpha_{3,s,t}}{(1+\rho)^{t-1}},$$

where T^* is the maximum age at which we observe earnings. To simplify notation, the first age we consider (22) is denoted $t = 1$ and the last age we consider (41) is denoted T^*. Conditional on covariates X, the covariance between $E\left(Y_1 \middle| \mathcal{I}\right)$ and $E\left(Y_0 \middle| \mathcal{I}\right)$ is

$$\begin{aligned} Cov\left(E\left(Y_1 \middle| \mathcal{I}\right), E\left(Y_0 \middle| \mathcal{I}\right)\right) &= Var\left(\theta_1\right)\left(\sum_{t=1}^{T^*} \frac{\alpha_{1,1,t}}{(1+\rho)^{t-1}}\right)\left(\sum_{t=1}^{T^*} \frac{\alpha_{1,0,t}}{(1+\rho)^{t-1}}\right) \\ &\quad + \cdots + Var\left(\theta_3\right)\left(\sum_{t=1}^{T^*} \frac{\alpha_{3,1,t}}{(1+\rho)^{t-1}}\right)\left(\sum_{t=1}^{T^*} \frac{\alpha_{3,0,t}}{(1+\rho)^{t-1}}\right). \end{aligned}$$

Tables 2A and 2B present the conditional distribution of the present values of *ex ante* college earnings given *ex ante* high school earnings decile by decile for the NLSY/1979 and NLS/1966 samples. If the dependence across outcomes were perfect and positive, as postulated by Juhn, Murphy, and Pierce (1993), the diagonal elements would be '1' and the off diagonal elements would be '0.' We estimate positive dependence between the relative positions of individuals in the two distributions, but the dependence is not perfect. For example, for the NLSY/1979 sample, 29.95% of the individuals who are in the first decile of the high school present value of earnings distribution would be in the first decile of the college present value of earnings distribution. For the NLS/1966 sample, this figure is 70.36%. The comparison of tables 2A and 2B shows that the correlation between *ex ante* high school and *ex ante* college present value of lifetime earnings has become weaker for recent cohorts.

We can also compute the covariance between the present value of *ex post* college and high-school earnings conditional on X. This is

$$\begin{aligned} Cov\left(Y_1, Y_0 \middle| X\right) &= Var\left(\theta_1\right)\left(\sum_{t=1}^{T^*} \frac{\alpha_{1,1,t}}{(1+\rho)^{t-1}}\right)\left(\sum_{t=1}^{T^*} \frac{\alpha_{1,0,t}}{(1+\rho)^{t-1}}\right) \\ &\quad + \cdots + Var\left(\theta_6\right)\left(\sum_{t=1}^{T^*} \frac{\alpha_{6,1,t}}{(1+\rho)^{t-1}}\right)\left(\sum_{t=1}^{T^*} \frac{\alpha_{6,0,t}}{(1+\rho)^{t-1}}\right). \end{aligned}$$

In Tables 3A and 3B we show the conditional distribution of the present values of *ex post* college earnings given *ex post* high school earnings. In NLSY/1979, *ex post* present values of earnings exhibit greater correlation than present values of *ex ante* earnings (the correlation is 0.16 for *ex ante* and 0.28 for *ex post*). On the other hand, in the NLS/1966 sample, *ex post* earnings exhibit lower correlation than *ex ante* earnings (the correlation is 0.91 for ex-ante and 0.62 for *ex post*). This trend in the correlations of *ex post* present values of earnings across schooling states is consistent with the analysis in Gould (2002). We find no evidence supporting perfect positive dependence or independence in *ex ante* earnings, nor *ex post* earnings.

Knowledge of the joint distributions allow us to compare factual with counterfactual distributions. Take agents who choose to be high-school graduates. We can compare the density of the present value of *ex post* earnings in the high-school sector with those in the college sector for the people who are high-school graduates. This information is plotted in Figures 4A and 4B for the NLSY/1979 and NLS/1966, respectively. In both data sets we see that the high-school agents would have higher earnings if they had chosen to be college graduates. Similarly, for the college graduates, we can compare the actual density of present value of earnings in the college sector with that in the high-school sector. We display these densities in Figures 5A and 5B for the NLSY/1979 and NLS/1966, respectively. Again, in both data sets the densities of high-school present value of earnings is to the left of the college density.

From such distributions we can generate the distribution of rates of return to college. The *ex post* gross rate of return R (excluding cost) is

$$R = \frac{Y_1 - Y_0}{Y_0}.$$

The typical high school student would have returns around 29% to a college education over the whole life cycle for the NLS/1966 sample and around 31% for the NLSY/1979 sample. For the typical college graduate this return is around 33% for the NLS/1966 sample and 40% for the NLSY/1979 sample. For the individuals at the margin, these figures are 31% and 35% for the NLS/1966 and NLSY/1979 samples, respectively. Returns to college have increased for college graduates and individuals at the margin, but not so much for the high school graduates.

Another interesting calculation one can perform given the knowledge of the joint distribution is the percentage of the individuals who regret their schooling choice, which we report in Table 5. Perhaps not surprisingly, a higher fraction of the high-school individuals regret not graduating from college (7.5% in NLSY/1979 and 9.7% in NLS/1966) than the other way around (around 3% of the individuals regret not stopping their schooling upon high-school graduation, for both the NLSY/1979 and NLS/1966).

3.3 The Evolution of Uncertainty and Heterogeneity

The valuation or net utility function for schooling choice is

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

Individuals go to college if $I > 0$. As explained in section 2.6, the correlation between schooling choices and future information allows us to disentangle heterogeneity from uncertainty. In the NLSY/1979, we test and do not reject the hypothesis that individuals, at the time they make college going decisions, know their Z and the factors θ_1, θ_2, and θ_3. However, they do not know the cohort dummies in X and the factors θ_4, θ_5, θ_6, or $\varepsilon_{s,t}$, $s = 0, 1$, $t = 1, \ldots, T^*$, at the time they make their educational choices.

For the NLS/1966 we test and do not reject that hypothesis that the individuals know their Z, X, and the factors θ_1, θ_2, and θ_3. However, they do not know the cohort dummies in X and the factors θ_4, θ_5, or $\varepsilon_{s,t}$, $s = 0, 1$, $t = 1, \ldots, T^*$, at the time they make their educational choices.

Therefore, our first result is that components not in the information sets of the agents at age 18 that we estimate from schooling choices in the NLSY/1979 and NLS/1966 are different. Next, we explore the implication of this difference for the increase in uncertainty.

3.3.1 Total Residual Variance and Variance of Unforecastable Component

For the NLSY/1979 the present value of lifetime (i.e., from age 22 ($t = 1$) to age 41 (T^*)) realized earnings in school level s can be written as

$$Y_s = \sum_{t=1}^{T^*} \frac{Y_{s,t}}{(1+\rho)^{t-1}} = \sum_{t=1}^{T^*} \frac{X\beta_{s,t} + \theta_1\alpha_{1,s,t} + \theta_2\alpha_{2,s,t} + \theta_3\alpha_{3,s,t} + \theta_4\alpha_{4,s,t} + \theta_5\alpha_{5,s,t} + \theta_6\alpha_{6,s,t} + \varepsilon_{s,t}}{(1+\rho)^{t-1}}.$$

We define total residual as the sum of the unobserved (to the econometrician) components,[13]

$$Q_s = \sum_{t=1}^{T^*} \frac{\theta_1\alpha_{1,s,t} + \theta_2\alpha_{2,s,t} + \theta_3\alpha_{3,s,t} + \theta_4\alpha_{4,s,t} + \theta_5\alpha_{5,s,t} + \theta_6\alpha_{6,s,t} + \varepsilon_{s,t}}{(1+\rho)^{t-1}}, \tag{21}$$

[13] In our empirical analysis we fix $\rho = 0.05$.

and note that Q_s combines terms that are known and unknown by the agent at the time of the schooling choice. The total residual variance in schooling level s is by definition $Var\left(Q_s\right)$.

The unforecastable component is the sum of the components that are not in the information set of the agent at the time schooling choices are made. For the NLSY/1979, the unforecastable component is

$$P_s = \sum_{t=1}^{T^*} \frac{\theta_4 \alpha_{4,s,t} + \theta_5 \alpha_{5,s,t} + \theta_6 \alpha_{6,s,t} + \varepsilon_{s,t}}{(1+\rho)^{t-1}}. \tag{22}$$

The variance of the unforecastable component in schooling level s is, by definition, $Var\left(P_s\right)$. Note that $Var\left(P_s\right) \leq Var\left(Q_s\right)$.

Table 6A displays the total residual variance and the variance of unforecastable component for each schooling level for both NLS/1966 (Panel A) and NLSY/1979 (Panel B). Total residual variance in present value of lifetime college earnings increase from 460.6260 (NLS/1966) to 709.7487 (NLSY/1979). This implies an increase of almost 55% in the total residual variance. The increase is larger for the present value of high school earnings: it goes from 284.8089 in NLS/1966 to 507.2910, corresponding to an increase of almost 80%.

The variance of the unforecastable component has also increased. For college earnings, it is 181.3712 for the NLS/1966 and it becomes 372.3509 for the NLSY/1979. For high school earnings, it is 128.4315 for the NLS/1966 and it becomes 272.3596 for the NLSY/1979. In percentage terms, this implies that the variance of the unforecastable component increased 105% for college and 112% for high school.

We can also make the same comparisons for the gross returns to college:

$$R = \sum_{t=1}^{T^*} \frac{Y_{1,t} - Y_{0,t}}{(1+\rho)^{t-1}}.$$

The total residual in the gross returns to college can be defined as $\Delta Q = Q_1 - Q_0$,

$$\Delta Q = \sum_{t=1}^{T^*} \frac{\theta_1 \Delta\alpha_{1,t} + \theta_2 \Delta\alpha_{2,t} + \theta_3 \Delta\alpha_{3,t} + \theta_4 \Delta\alpha_{4,t} + \theta_5 \Delta\alpha_{5,t} + \theta_6 \Delta\alpha_{6,t} + \Delta\varepsilon_t}{(1+\rho)^{t-1}},$$

and the unforecastable component in the gross returns to college is defined as $\Delta P = P_1 - P_0$,

$$\Delta P = \sum_{t=1}^{T^*} \frac{\theta_4 \Delta\alpha_{4,t} + \theta_5 \Delta\alpha_{5,t} + \theta_6 \Delta\alpha_{6,t} + \Delta\varepsilon_t}{(1+\rho)^{t-1}}.$$

From Table 6A we see that total residual variance in gross returns to college increased from 351 in NLS/1966 to 906 in NLSY/1979, which implies an increase of around 160%. The variance of the unforecastable component increased from 327 to 432, or roughly 32%.

These figures show that the increase in the variance of unforecastable components of earnings is a key element in explaining the increase in total residual variance in high school and college earnings. Furthermore, both the total residual variance and the variance of unforecastable components have increased more for low-skill workers (i.e., the high-school graduates) than high-skill workers (i.e., college graduates).

The same exercise can be repeated for the evolution of unobserved heterogeneity. For both the NLSY/1979 and NLS/1966 the unobserved heterogeneity component is

$$H_s = \sum_{t=1}^{T^*} \frac{\theta_1 \alpha_{1,s,t} + \theta_2 \alpha_{2,s,t} + \theta_3 \alpha_{3,s,t}}{(1+\rho)^{t-1}}.$$

We illustrate these findings in Figures 7A and 7B, which plot the density of total residual versus the density of unforecastable components for high-school earnings. Note that the latter has a much less dispersed density than the former. Figures 8A and 8B make the same comparison for college earnings. Finally, Figures 9A and 9B show the corresponding figures for returns. Figure 9B shows that all the densities of total residual and unforecastable components are very similar.

Table 6B displays the total residual variance and the variance of heterogeneity (or forecastable) component for each schooling level for both NLS/1966 (Panel A) and NLSY/1979 (Panel B). Individuals have become more diverse. For college earnings, the variance of forecastable components is 279 for the NLS/1966 and it is 337 for the NLSY/1979, which corresponds to roughly 21% increase. For high school earnings, it is 156 for the NLS/1966 and it becomes 234 for the NLSY/1979, which implies an increase of more than 50%. As we document above, there is no selection in returns in the NLS/1966. This can only happen if agents could not forecast returns well in the 1966 cohort. As a result, most of the variance of unobservable component in returns for that cohort is due to uncertainty and not forecastable heterogeneity. This explains the substantial increase in the variance of heterogeneity in the variance of returns to college.

3.3.2 The Variance of the Unforecastable Component by Age

We have shown that the variance of the unforecastable component in the present value of lifetime earnings has increased for both college and high school graduates. In this section, we show that the increase is not uniform across all ages. For every age t and schooling level s let $P_{s,t}$ denote the unforecastable component in school level s age t earnings. Our estimated information, together with the identifying normalizations displayed in Table 1, imply that the unforecastable components for ages 22 through 25 are given by

$$P_{s,t} = \frac{\varepsilon_{s,t}}{(1-\rho)^{t-1}} \text{ for } t = 1, \ldots, 4, \tag{23}$$

and for ages 26 through 41,

$$P_{s,t} = \frac{\theta_4 \alpha_{4,s,t} + \theta_5 \alpha_{5,s,t} + \theta_6 \alpha_{6,s,t} + \varepsilon_{s,t}}{(1+\rho)^{t-1}} \text{ for } t = 5, \ldots, T^*. \tag{24}$$

Figure 10 shows that the variance of unforecastable components in high school earnings in NLS/1966 and NLSY/1979 are about the same until age 27/28. From age 29 on, the variances diverge. They both increase with age, but the cohort in NLSY/1979 experiences a faster increase than in NLS/1966. At age 41, the variance of the unforecastable component in high school earnings NLSY/1979 is almost three times larger than its counterpart in the NLS/1966 sample.

A similar pattern arises in the variance of the unforecastable component in college earnings. Figure 11 shows that until around age 30, the profiles of the variances are roughly the same for NLSY/1979 and NLS/1966. From age 31 on, the series diverge, and the variances in the NLSY/1979 sample increase at a faster rate. At at age 37, the variances of the unforecastable component in NLSY/1979 are more than twice those in NLS/1966 sample.

We conclude that the variance of unforecastable components began to increase at earlier ages for high-school graduates. Furthermore, the gradient of the increase was sharper for high-school graduates as well.

3.3.3 Persistence of Shocks

Note that given our estimated information set, from (23) and (24) it follows that

$$Cov\left(P_{s,\tau}, P_{s,t}\right) = 0 \text{ if } \tau \text{ or } t = 1, \ldots, 4.$$

However, at other ages,

$$Cov\left(P_{s,\tau}, P_{s,t}\right) \neq 0 \text{ if } \tau \text{ and } t = 5, \ldots, T^*.$$

In Web Appendix Tables 1 through 3 we plot the correlation of unforecastable components across ages 26 ($t = 5$) through 41 (T^*) for high-school, college, and overall sample, respectively. The correlations of unforecastable components have increased from the NLS/1966 to NLSY/1979 sample. For example, in the overall sample, 92 out of 120 total off-diagonal lower triangular elements are greater in the NLSY/1979 matrix than its NLS/1966 counterpart. For the college sample, 101 out of 120 are greater in NLSY/1979 than in the NLS/1966 matrix. In the high-school sample, only 42 out of 120 are greater in the NLSY/1979 than in NLS/1966 sample.

A clearer way to see these results is by plotting the correlogram of these components. Let $\phi(s, t, k)$ denote the correlogram of the unforecastable components of earnings at school level s and age t with respect to ages $k = t + 1, t + 2, \ldots, T^*$:

$$\phi(s, t, k) = \frac{Cov\left(P_{s,t}, P_{s,k}\right)}{Var\left(P_{s,t}\right)}.$$

In Figure 12 we compare $\phi(s, t, k)$ for the overall sample from NLS/1966 (gray column) against that from NLSY/1979 (black column). There is no clear pattern. We arrive at a different conclusion by looking at schooling-specific correlograms. In Figure 13 we plot the correlogram at age 31 for the high-school sector. Shocks are more persistent in the NLSY/1979 (black column) than the NLS/1966 sample (gray column). In Figure 14 we see that for the college sample, again, no clear pattern arises from the data.

3.3.4 The Evolution of Skill Prices

The Gorman-Lancaster model relates realized earnings with characteristics and the price of characteristics. For schooling level s at age t, we model earnings according to the Gorman-Lancaster model:

$$Y_{s,t} = X\beta_{s,t} + \theta_1\alpha_{1,s,t} + \theta_2\alpha_{2,s,t} + \theta_3\alpha_{3,s,t} + \theta_4\alpha_{4,s,t} + \theta_5\alpha_{5,s,t} + \theta_6\alpha_{6,s,t} + \varepsilon_{s,t}.$$

We have established that θ_4, θ_5, and θ_6 are unforecastable components at the time schooling choice. We have shown how the factor loadings $\alpha_{4,s,t}$, $\alpha_{5,s,t}$, and $\alpha_{6,s,t}$ have contributed for the serial correlation matrix of the uncertain components within each schooling group.[14] The factor θ_1 is, in fact, known by the agent at the time of the schooling choice. We identify factor θ_1 directly from the cognitive test scores, so we take this as measures of cognitive skills of the agents. In the Gorman-Lancaster model, the factor loadings $\alpha_{1,s,t}$, are in fact cognitive skill prices at schooling level s and age t. In Figure 15 we show that the prices of the factor θ_1 in the high school sector have increased from NLS/1966 to the NLSY/1979. In Figure 16 we show that the prices have increased even more sharply in the college sector. This evidence is consistent with higher demand for cognitive skills in more recent years.

3.3.5 Accounting for Macro Uncertainty

Our estimates of uncertainty are microeconomic in character. A large literature in macroeconomics documents that aggregate instability has decreased over the past 30 years. To capture this phenomenon, one could introduce time dummies into the earnings equation. Given the standard problem of the lack of identification of individual age, period, and cohort effects, we work with cohort dummies when we form estimates. We find that variables that capture macro uncertainty (cohort dummies) do not enter the schooling choice equation. Thus, macro uncertainty is not forecastable by agents at the time schooling choices are being made. However, realized macro shocks affect earnings outcomes. Macro uncertainty decreased for later cohorts by 90% (see Table 7). These estimates are consistent with evidence by Gordon (2005) and others that US business cycle volatility has decreased in recent years.

[14] Note that under our assumptions we could also calculate the serial correlation matrix across schooling choices.

4 Models with Sequential Updating of Information

In this paper, we have analyzed one shot models of schooling choice. In truth, schooling is a sequential decision process made with increasingly richer information sets at later stages of the choice process.

Keane and Wolpin (1997) and Eckstein and Wolpin (1999) pioneered the estimation of dynamic discrete choice models for analyzing schooling choices. They assume expected income maximization and do not entertain a range of alternative market structures facing agents.[15] In the notation of this paper, they assume a one (discrete) factor model with factor loadings that are different across different counterfactual states, but are constant over time ($\alpha_{s,t} = \alpha_s$, $s = 1, \ldots, \bar{S}$, where there are $\bar{S}$ schooling states).[16] At a point in time, t, $\varepsilon_{s,t}$, $s = 1, \ldots, \bar{S}$ are assumed to be multivariate normal random variables. Over time the $\varepsilon_t = (\varepsilon_{1,t}, \ldots, \varepsilon_{s,t})$ are assumed to be independent and identically distributed. They *assume* that agents know θ but not the ε_t, $t = 1, \ldots, T$. The unobservables are thus equicorrelated over time (age) because the factor loadings are assumed equal over time and ε_t is independent and identically distributed over time. They make parametric normality assumptions in estimating their models. Keane and Wolpin (1997) impose their discrete factor in the schooling choice and outcome equations rather than testing for whether or not the factor appears in both sets of equations in all time periods as we do in this paper. In their model, about 90% of the variance in lifetime returns is predictable at age 16. Our setup allows for more factors and for testing which factors enter different equations at different time periods. We allow for temporally persistent shocks.

In an important paper, Taber (2001) estimates a dynamic programming model of schooling choice with a Gorman-Lancaster model for *log* earnings and interprets his estimated increasing factor loadings over time on unobserved ability as evidence for rising skill prices on unobserved factors. His model is not directly comparable to our model because of his use of logs–rather than levels. The interpretation of the factor loadings as skill prices is strained with a log dependent variable. He allows for sequential updating of information, but does not estimate the initial information set of the agents, but rather imposes it like Keane and Wolpin (1997). Unlike Keane and Wolpin, he tests whether additional serially correlated information enters the agent baseline information set as the agent ages.

[15] But see Keane and Wolpin (2001), where credit constraints are explicitly modelled.

[16] Thus instead of assuming that θ is continuous, as we do, they impose that θ is a discrete-valued random variable that assumes a finite, known number of values.

Heckman and Navarro (2006) formulate and identify semiparametric sequential schooling models based on the factor structures of the sort used in this paper. They present a new semiparametric identification analysis for this class of models. See Abbring and Heckman (2007) for a discussion of this literature.

5 More General Preferences and Market Settings

To focus on the main ideas, we have used the simple market structure of complete contingent claims markets, or alternatively, we have assumed risk neutrality. An alternative interpretation is that we use expected present value income maximization as our schooling choice criterion. What can be identified in more general environments? In the absence of perfect certainty or perfect risk sharing, preferences and market environments also determine schooling choices. The separation theorem allowing consumption and schooling decisions to be analyzed in isolation of each other that we have used in this paper breaks down.

If we postulate information processes *a priori*, and assume that preferences are known up to some unknown parameters as in Flavin (1981), Blundell and Preston (1998) and Blundell, Pistaferri, and Preston (2004), we can identify departures from specified market structures. In work in progress, Cunha and Heckman (2006) postulate an Aiyagari (1994) – Laitner (1992) economy with one asset and parametric preferences to identify the information processes in the agent's information set. They take a parametric position on preferences and a nonparametric position on the economic environment and the information set.

An open question, not yet fully resolved in the literature, is how far one can go in nonparametrically jointly identifying preferences, market structures and information sets.[17] Navarro (2005) adds consumption data to the schooling choice and earnings data to secure identification of risk preference parameters (within a parametric family) and information sets, and to test among alternative models for market environments. Alternative assumptions about what analysts know produce different interpretations of the same evidence. The lack of full insurance interpretation given to the empirical results by Flavin (1981) and Blundell, Pistaferri, and Preston (2004) may be a consequence of their misspecification of the generating process of agent information sets.

[17] This point was first made in the Hicks Lecture at Oxford, April 2004, and is published in Cunha, Heckman, and Navarro (2005).

6 Summary and Conclusion

This paper investigates the sources of rising wage inequality the US labor market. We find that increasing inequality arises from both increasing micro uncertainty and increasing heterogeneity predictable by agents. The latter could arise from increased sorting. While both components have increased since the late 1960s but the fraction of the variability due to micro uncertainty has increased. Aggregate uncertainty has decreased. Thus the recent increase of uncertainty is microeconomic in character.

A Identification of the Model

We provide an intuitive discussion of identification based on the normal case, because together with the assumption of complete Arrow Debreu markets it allows for closed form solutions. See Carneiro, Hansen, and Heckman (2003) for proofs of semi-parametric identification of the distributions of the factors θ and uniquenesses ε.

A.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 X^M are independent from both θ_1 and ε_k^M, for $k = 1, \ldots, K$. Second, the factor θ_1 is independent from ε_k^M, for $k = 1, \ldots, K$. Third, the uniqueness ε_k^M is independent from ε_l^M for any $k \neq l$, for $k, l = 1, \ldots, K$. The first assumption allows us to conclude that β_k^M can be consistently estimated from a simple OLS regression of M_k against X^M. Given knowledge of these parameters we can construct differences $M_k - X^M \beta_k^M$ and compute the covariances:

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

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

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

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 $M_k - X^M\beta_k^M$. Because the factor θ_1 and uniquenesses ε_k are independently normally distributed random variables, we have identified their distribution.

A.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 X and 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 $\alpha_{1,s,t}, \alpha_{2,s,t}, \alpha_{1,C}$, and $\alpha_{2,C}$ the factors θ can affect $U_{0,t}, U_{1,t}$, and U_C differently. Therefore, by adopting 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 X is:

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

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

$$E\left(Y_{1,t}\middle| X, S=1\right) = X\beta_{1,t} + \alpha_{1,1,t}E\left[\theta_1\middle| X, S=1\right] + \alpha_{2,1,t}E\left[\theta_2\middle| X, S=1\right] + E\left[\varepsilon_{1,t}\middle| X, S=1\right]. \quad (29)$$

The event $S = 1$ corresponds to the event $I = E\left(\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}(Y_{1,t} - Y_{0,t}) - C\middle|\mathcal{I}\right) \geq 0$. 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 $E\left(U_{s,t}\middle|\mathcal{I}\right)$. 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 $U_{s,t} - E\left(U_{s,t}\middle|\mathcal{I}\right)$. 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

$$E\left(\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}(Y_{1,t} - Y_{0,t}) - C\middle|\mathcal{I}\right) = \mu_I(X,Z) + \alpha_{1,I}\theta_1 + \alpha_{2,I}\theta_2 - \varepsilon_C.$$

Let η be the linear combination of three independent normal random variables: $\eta = \alpha_{1,I}\theta_1 + \alpha_{2,I}\theta_2 - \varepsilon_C$. Then, $\eta \sim N\left(0, \sigma_\eta^2\right)$, with $\sigma_\eta^2 = \alpha_{1,I}^2\sigma_{\theta_1}^2 + \alpha_{2,I}^2\sigma_{\theta_2}^2 + \sigma_{\varepsilon_c}^2$ and

$$S = 1 \Leftrightarrow \eta > -\mu_I(X,Z). \quad (30)$$

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

$$E\left(Y_{1,t}\middle| X, S=1\right) = X\beta_1 + \alpha_{1,1,t}E\left[\theta_1\middle| X, \eta > -\mu_I(X,Z)\right] + \alpha_{2,1,t}E\left[\theta_2\middle| X, \eta > -\mu_I(X,Z)\right]. \quad (31)$$

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 $Cov\left(\theta_1, \eta\right) = \sigma^2_{\theta_1}\alpha_{1,I}$ and $Cov\left(\theta_2, \eta\right) = \sigma^2_{\theta_2}\alpha_{2,I}$ it follows that

$$E\left[\theta_1 | X, \eta > -\mu_I(X, Z)\right] = \frac{\sigma^2_{\theta_1}\alpha_{1,I}}{\sigma^2_\eta} E\left[\eta | \eta > -\mu_I(X, Z)\right],$$

$$E\left[\theta_2 | X, \eta > -\mu_I(X, Z)\right] = \frac{\sigma^2_{\theta_2}\alpha_{2,I}}{\sigma^2_\eta} E\left[\eta | \eta > -\mu_I(X, Z)\right].$$

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

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

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

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

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 $Y_{s,t} - X\beta_{s,t}$ and compute the covariances

$$Cov\left(M_1 - X^M\beta_1^M, Y_{0,t} - X\beta_{0,t}\right) = \alpha_{1,0,t}\sigma^2_{\theta_1}, \tag{35}$$

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

The left-hand side of (35) 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 (35) and

(36) we can recover $\alpha_{1,0,t}$ and $\alpha_{1,1,t}$ for all t. Note that we can also identify the $\frac{\alpha_{1,C}}{\sigma_\eta}$ by computing the covariance:

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

The argument why $\frac{\alpha_{1,C}}{\sigma_\eta}$ can be recovered is simple: Using (35) and (36) we can identify $\alpha_{1,1,t}$ and $\alpha_{1,0,t}$ for all t.The only remaining term to be identified is the ratio $\frac{\alpha_{1,C}}{\sigma_\eta}$, which we can from the covariance equation (37).

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

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

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

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

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 $\alpha_{2,0,2}$. From (38) we can recover $\sigma_{\theta_2}^2$. We now add the covariances from the college earnings:

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

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

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

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

References

Abbring, J. H. and J. J. Heckman (2007). Econometric evaluation of social programs, part III: Social experiments, distributional treatment effects and general equilibrium policy evaluation. In J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics, Volume 6*. Amsterdam: Elsevier. Forthcoming.

Aiyagari, S. R. (1994, August). Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics 109*(3), 659–684.

Blundell, R., L. Pistaferri, and I. Preston (2004, October). Consumption inequality and partial insurance. Technical Report WP04/28, Institute for Fiscal Studies.

Blundell, R. and I. Preston (1998, May). Consumption inequality and income uncertainty. *Quarterly Journal of Economics 113*(2), 603–640.

Carneiro, P., K. Hansen, and J. J. Heckman (2003, May). 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. 2001 Lawrence R. Klein Lecture.

Carneiro, P. and J. J. Heckman (2003). Human capital policy. In J. J. Heckman, A. B. Krueger, and B. M. Friedman (Eds.), *Inequality in America: What Role for Human Capital Policies?* Cambridge, MA: MIT Press.

Cunha, F. and J. J. Heckman (2006). The evolution of earnings risk in the us economy. Presented at the 9th World Congress of the Econometric Society, London.

Cunha, F., J. J. Heckman, and S. Navarro (2005, April). Separating uncertainty from heterogeneity in life cycle earnings, the 2004 hicks lecture. *Oxford Economic Papers 57*(2), 191–261.

Eckstein, Z. and K. I. Wolpin (1999, November). Why youths drop out of high school: The impact of preferences, opportunities, and abilities. *Econometrica 67*(6), 1295–1339.

Ferguson, T. S. (1983). Bayesian density estimation by mixtures of normal distributions. In H. Chernoff, M. Rizvi, J. Rustagi, and D. Siegmund (Eds.), *Recent Advances in Statistics: Papers in Honor of Herman Chernoff on his Sixtieth Birthday*, pp. 287–302. New York: Academic Press.

Flavin, M. A. (1981, October). The adjustment of consumption to changing expectations about future income. *Journal of Political Economy 89*(5), 974–1009.

Gordon, R. J. (2005). What caused the decline in U. S. business cycle volatility? In C. Kent and D. Norman (Eds.), *The Changing Nature of the Business Cycle*, pp. 61–104. Sydney, Australia: Economics Group, Reserve Bank of Australia. Proceedings of a conference held at the H.C. Coombs Centre for Financial Studies, Kirribilli, Australia on 11-12 July 2005.

Gottschalk, P. and R. Moffitt (1994). The growth of earnings instability in the U.S. labor market. *Brookings Papers on Economic Activity 2*, 217–254.

Gould, E. D. (2002, January). Rising wage inequality, comparative advantage, and the growing importance of general skills in the United States. *Journal of Labor Economics 20*(1), 105–147.

Heckman, J. J. (1976, December). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement 5*(4), 475–492.

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

Heckman, J. J. and Y. Rubinstein (2001, May). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review 91*(2), 145–149.

Heckman, J. J. and J. Scheinkman (1987, April). The importance of bundling in a gorman-lancaster model of earnings. *Review of Economic Studies 54*(2), 243–355.

Heckman, J. J. and G. L. Sedlacek (1985, December). Heterogeneity, aggregation, and market wage functions: An empirical model of self-selection in the labor market. *Journal of Political Economy 93*(6), 1077–1125.

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

Juhn, C., K. M. Murphy, and B. Pierce (1993, June). Wage inequality and the rise in returns to skill. *Journal of Political Economy 101*(3), 410–442.

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

Keane, M. P. and K. I. Wolpin (2001, November). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review 42*(4), 1051–1103.

Laitner, J. (1992, December). Random earnings differences, lifetime liquidity constraints, and altruistic intergenerational transfers. *Journal of Economic Theory 58*(2), 135–170.

Levy, F. and R. J. Murnane (1992, September). U.S. earnings levels and earnings inequality: A review of recent trends and proposed explanations. *Journal of Economic Literature 30*(3), 1333–1381.

Ljungqvist, L. and T. J. Sargent (2004, April-May). European unemployment and turbulence revisited in a matching model. *Journal of the European Economic Association 2*(2-3), 456–468.

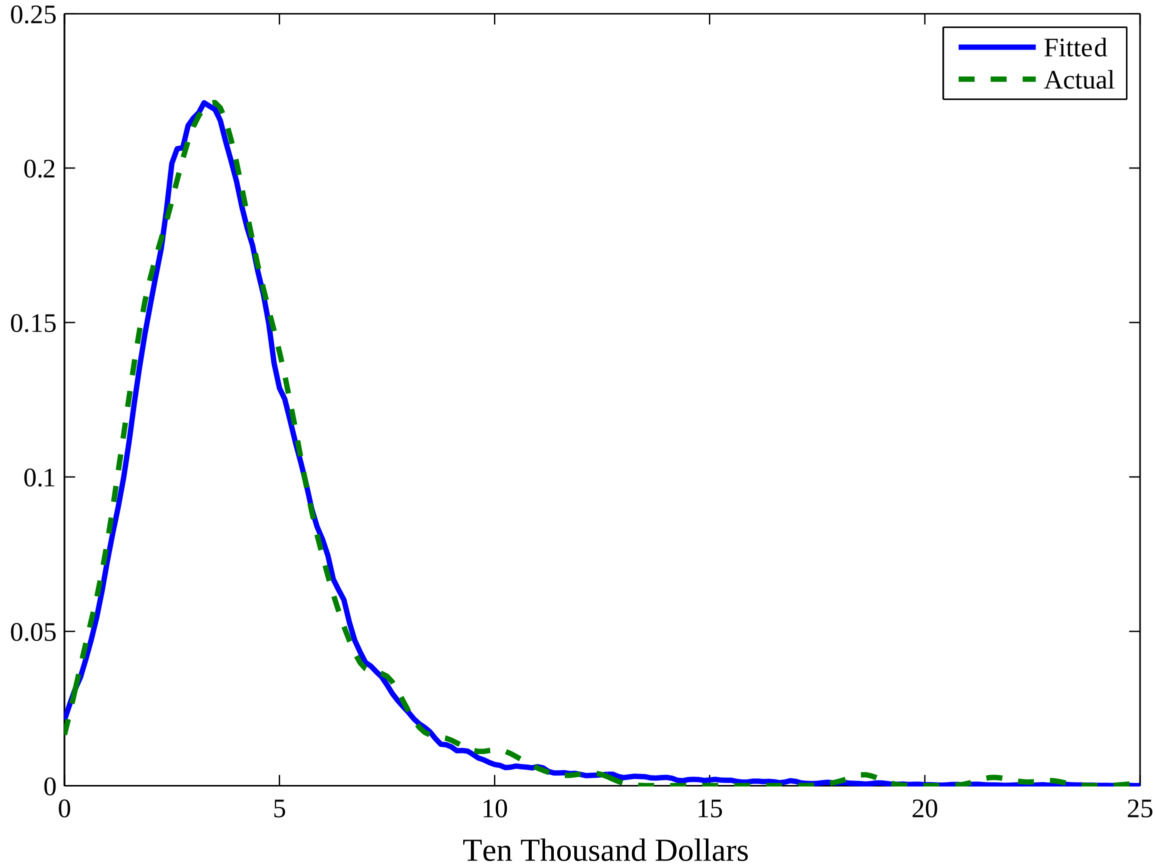
Matzkin, R. L. (1992, March). Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models. *Econometrica 60*(2), 239–270.

Navarro, S. (2005). *Understanding Schooling: Using Observed Choices to Infer Agent's Information in a Dynamic Model of Schooling Choice When Consumption Allocation is Subject to Borrowing Constraints.* Ph.D. Dissertation, University of Chicago, Chicago, IL.

Taber, C. R. (2001, July). The rising college premium in the eighties: Return to college or return to unobserved ability? *Review of Economic Studies 68*(3), 665–691.

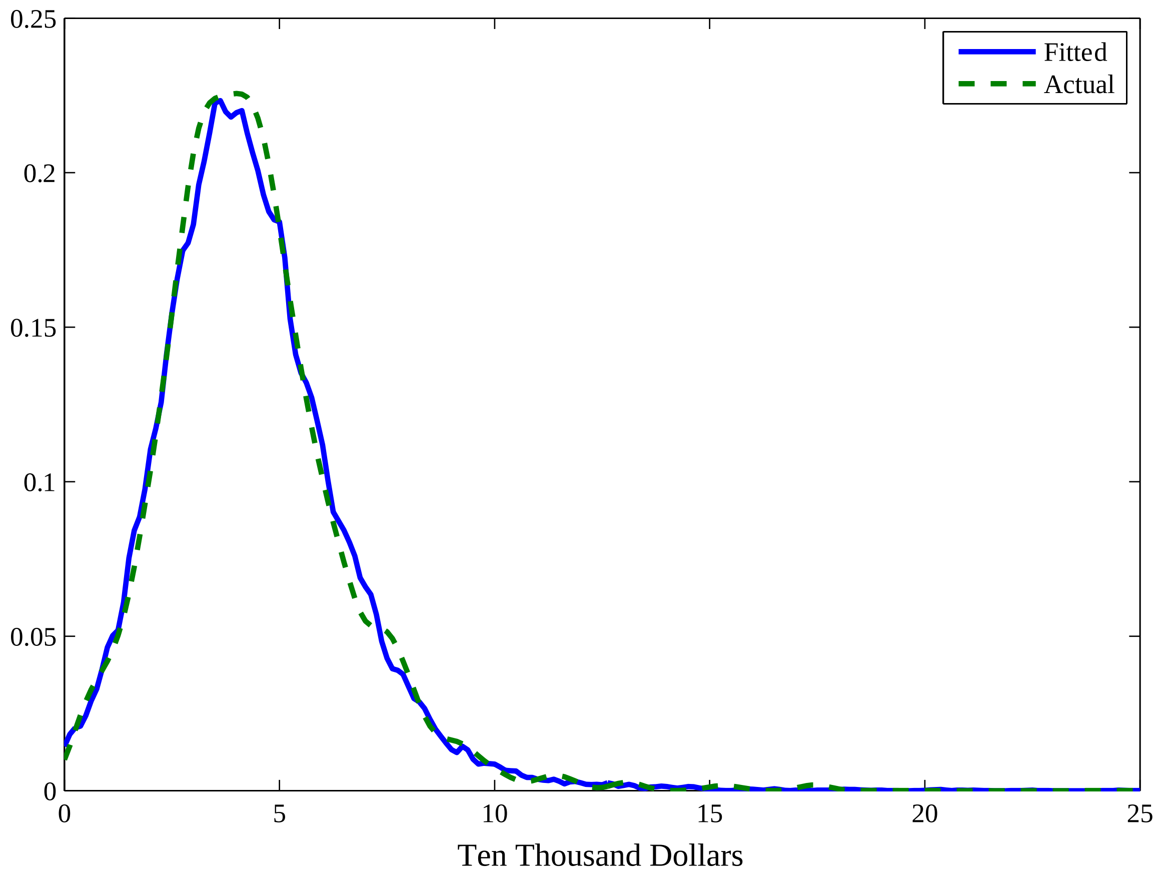
Tinbergen, J. (1975). *Income Distribution: Analysis and Policies.* New York: North-Holland Publishing Company.

Figure 1A
Densities of earnings at age 31
Overall Sample NLSY/1979



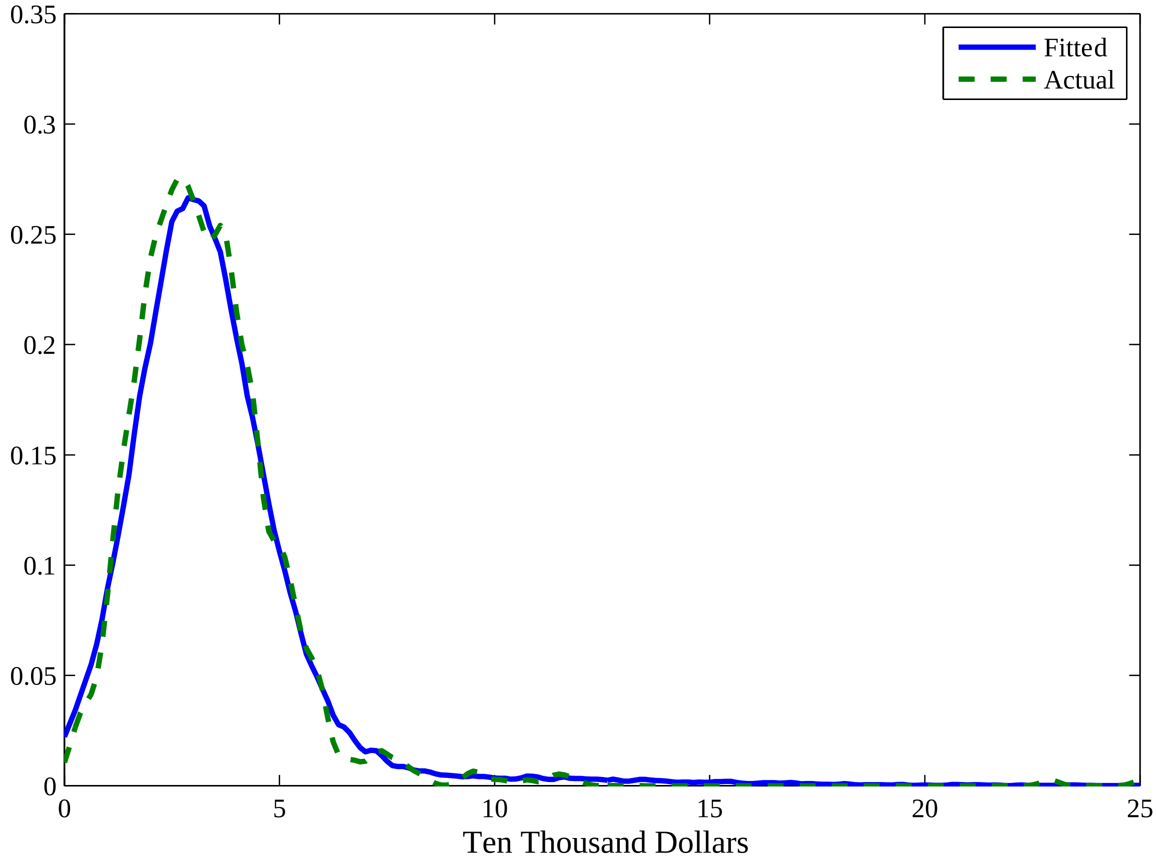
Let Y denote earnings at age 31 in the overall sample. Here we plot the density functions f(y) generated from the data (the solid curve), against that predicted by the model (the dashed line).

Figure 1B
Densities of earnings at age 31
Overall Sample NLS/1966



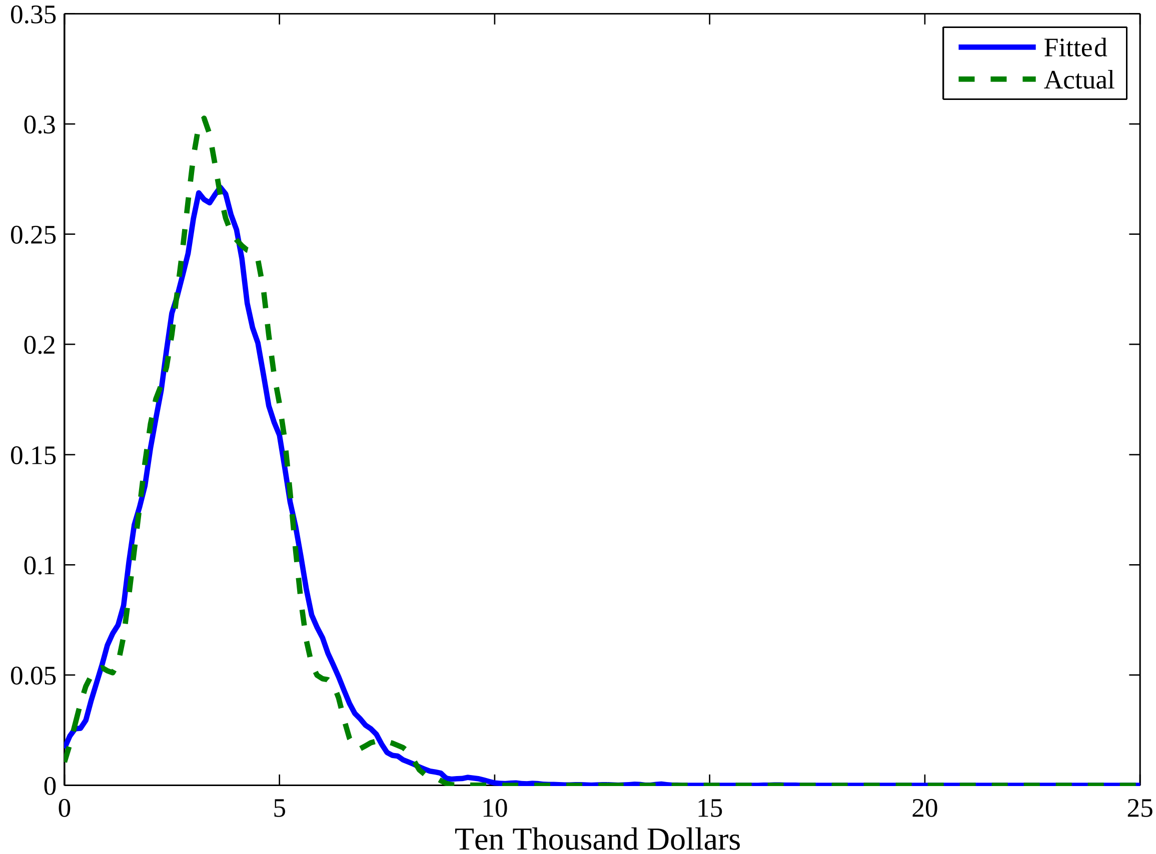
Let Y denote earnings at age 31 in the overall sample. Here we plot the density functions f(y) generated from the data (the solid curve), against that predicted by the model (the dashed line).

Figure 2A
Densities of earnings at age 31
High School Sample NLSY/1979



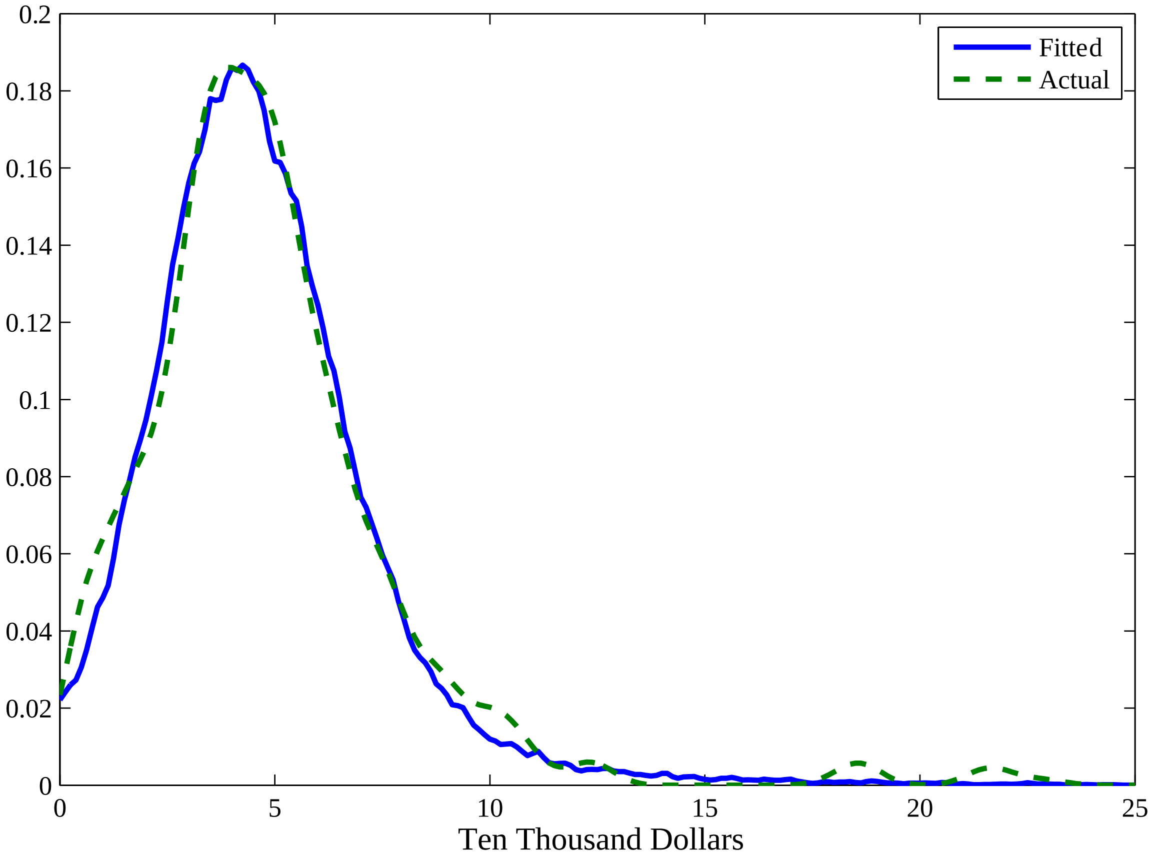
Let Y_0 denote earnings in high-school sector (S = 0) at age 31. Here we plot the density function $f(y_0|S=0)$ generated from the data (the solid curve) against that predicted by the model (the dashed line).

Figure 2B
Densities of earnings at age 31
High School Sample NLS/1966



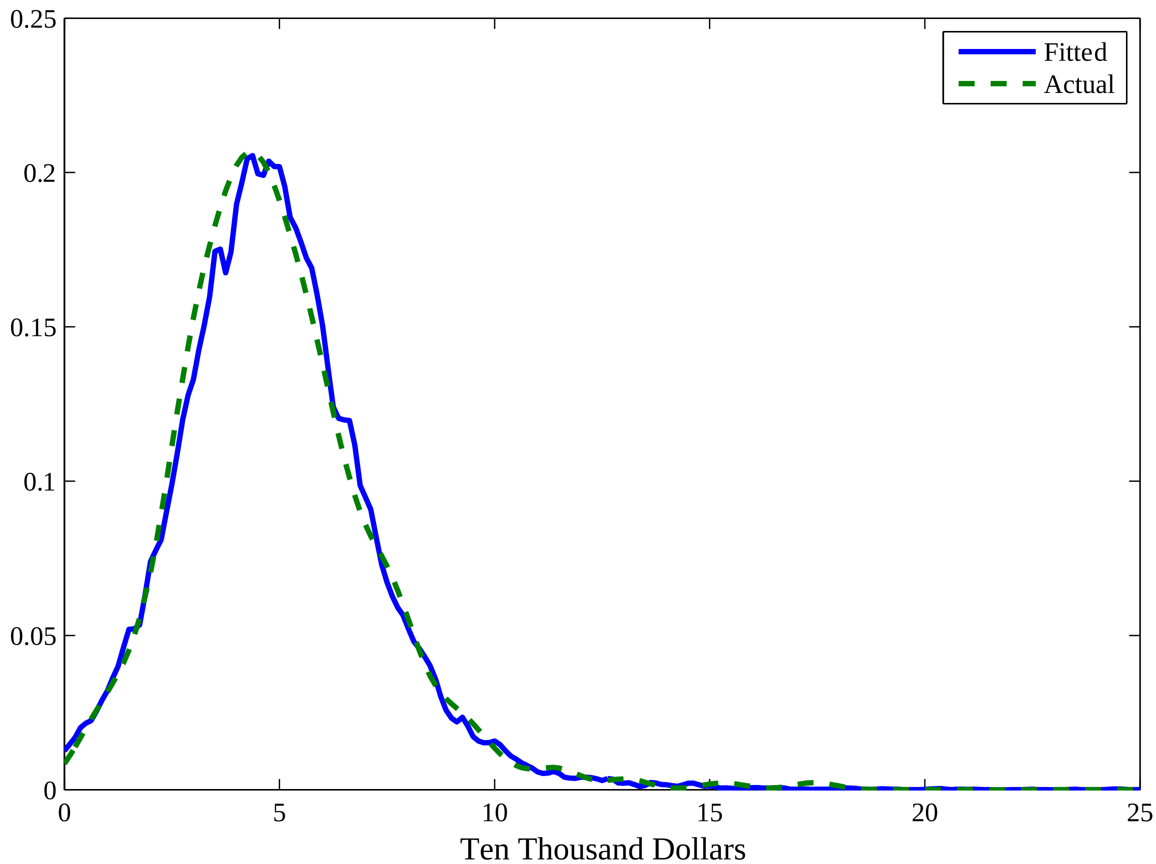
Let Y_0 denote earnings in high-school sector (S = 0) at age 31. Here we plot the density function $f(y_0|S=0)$ generated from the data (the solid curve) against that predicted by the model (the dashed line).

Figure 3A
Densities of earnings at age 31
College Sample NLSY/1979



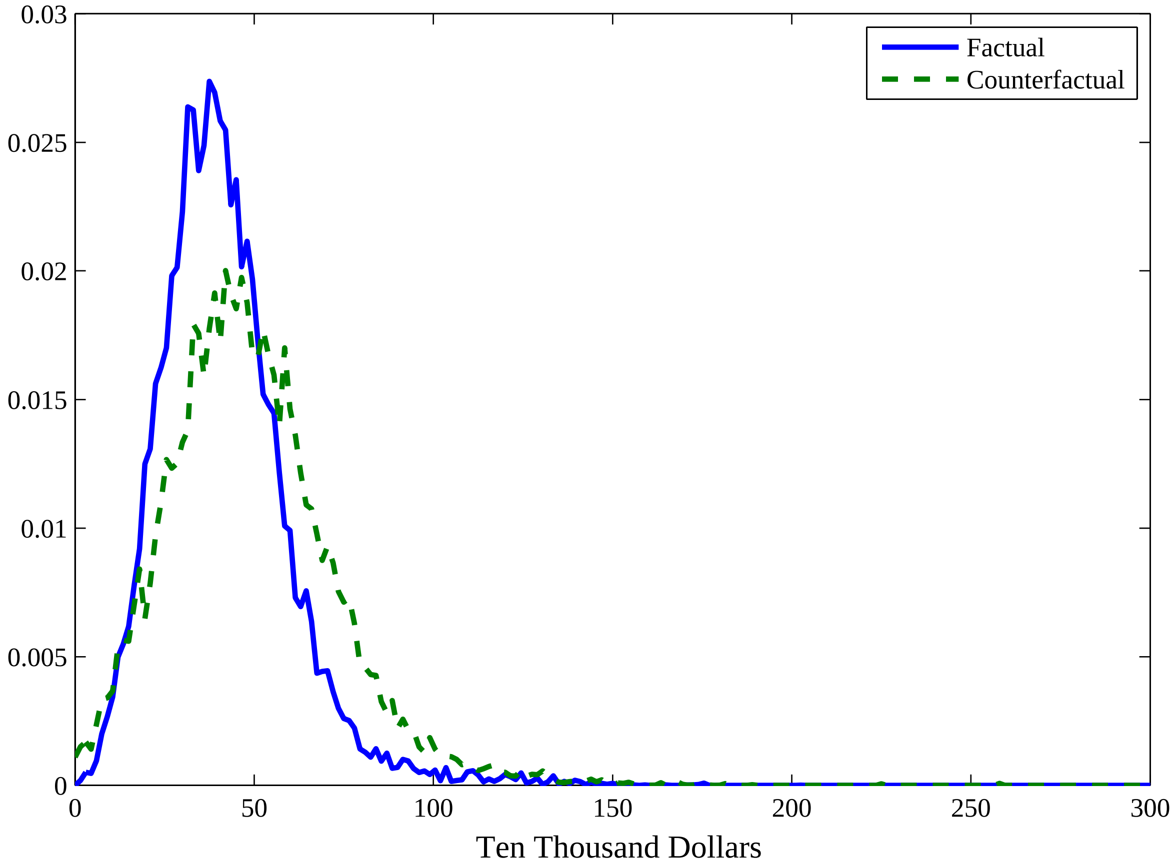
Let Y_1 denote earnings in the college sector (S = 1) at age 31. Here we plot the density function $f(y_1|S=1)$ generated from the data (the solid curve) against that predicted by the model (the dashed line).

Figure 3B
Densities of earnings at age 31
College Sample NLS/1966



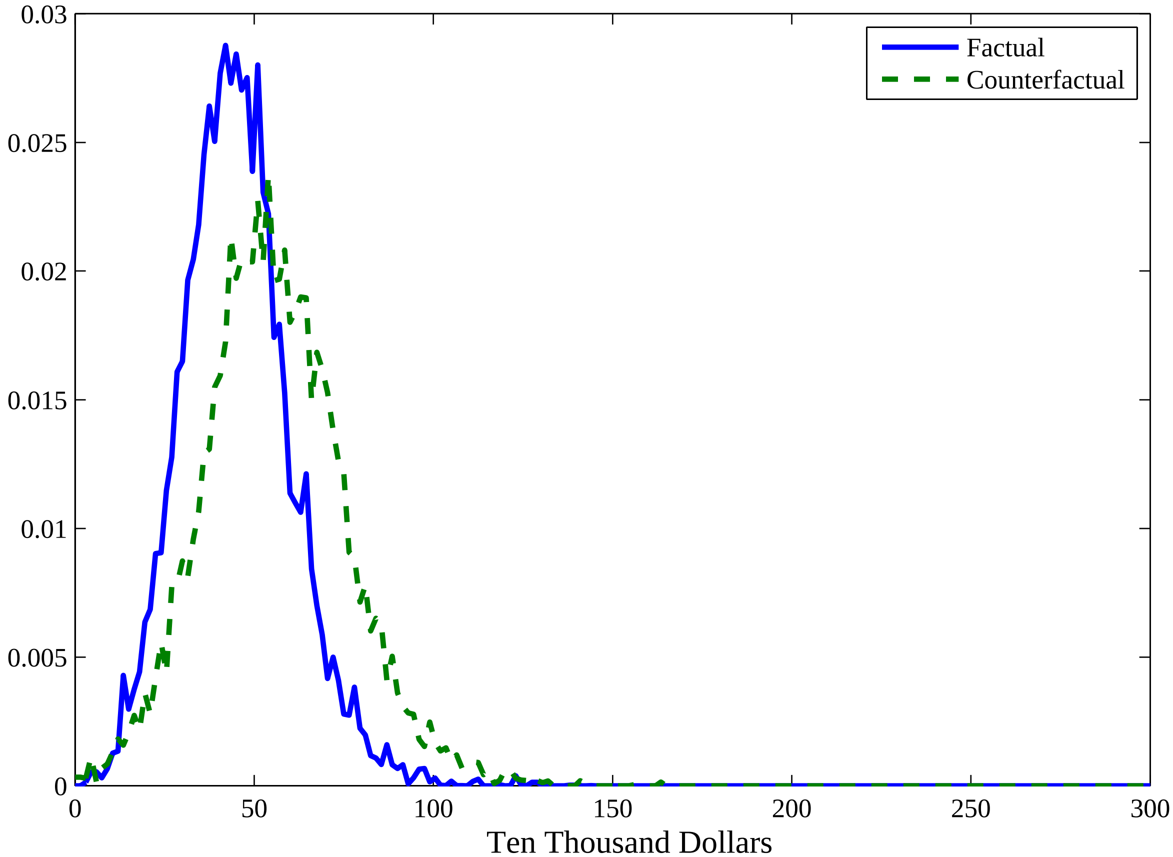
Let Y_1 denote earnings in the college sector (S = 1) at age 31. Here we plot the density function $f(y_1|S=1)$ generated from the data (the solid curve) against that predicted by the model (the dashed line).

Figure 4A
Densities of present value of earnings
High School Sample NLSY/1979



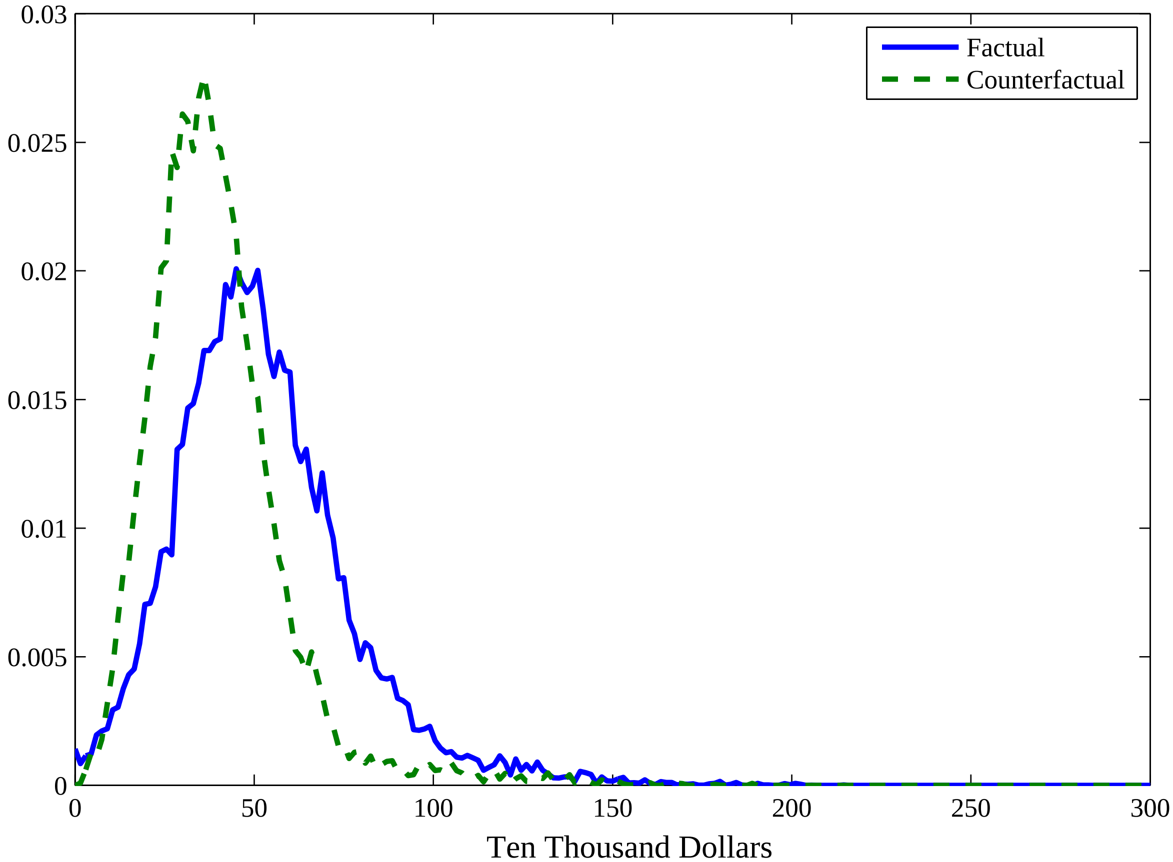
Let Y_0 denote the present value of earnings from age 22 to 41 in the High School sector (S = 0). Let Y_1 denote the present value of earnings from age 22 to 41 in the college sector (S = 1). Here we plot the factual density function $f(y_0|S=0)$ (the solid curve) against the counterfactual density function $f(y_1|S=0)$ (the dashed curve). We use a discount rate of 5%.

Figure 4B
Densities of present value of earnings
High School Sample NLS/1966



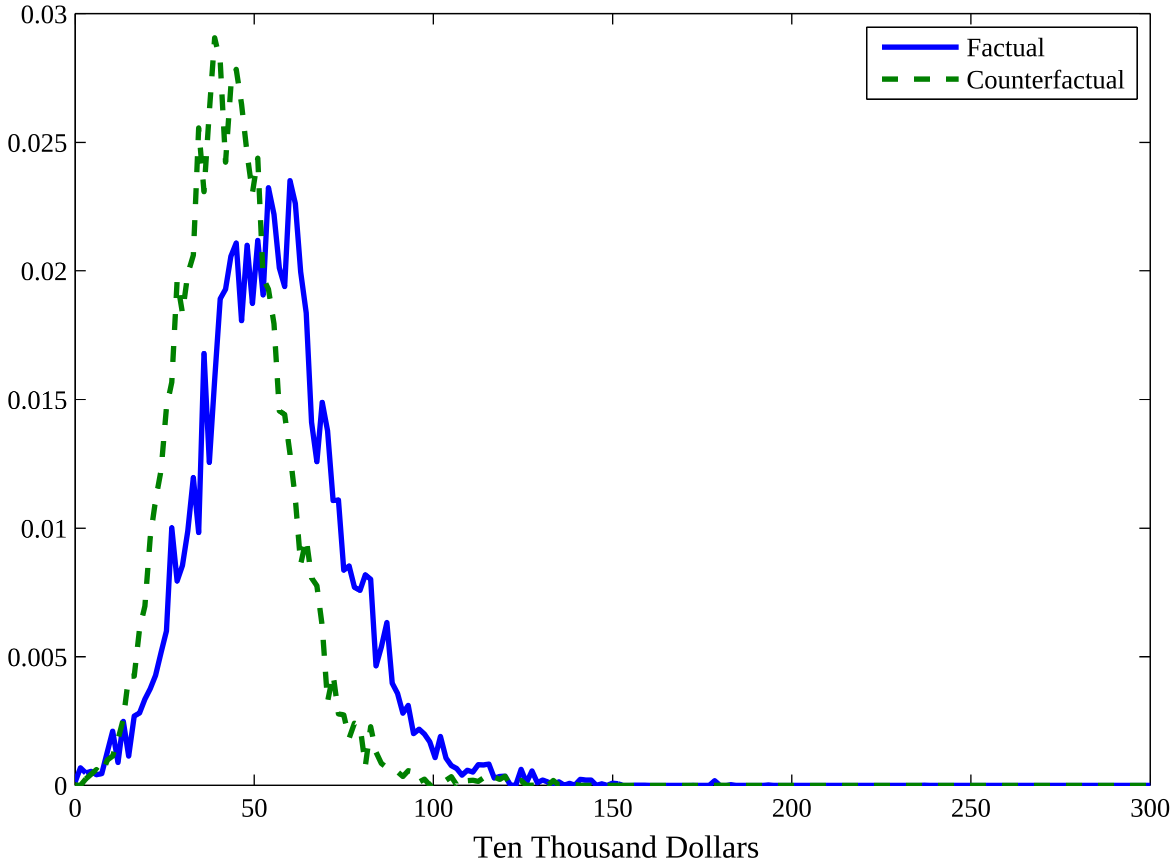
Let Y_0 denote the present value of earnings from age 22 to 41 in the High School sector (S = 0). Let Y_1 denote the present value of earnings from age 22 to 41 in the college sector (S = 1). Here we plot the factual density function $f(y_0|S=0)$ (the solid curve) against the counterfactual density function $f(y_1|S=0)$ (the dashed curve). We use a discount rate of 5%.

Figure 5A
Densities of present value of earnings
College Sample NLSY/1979



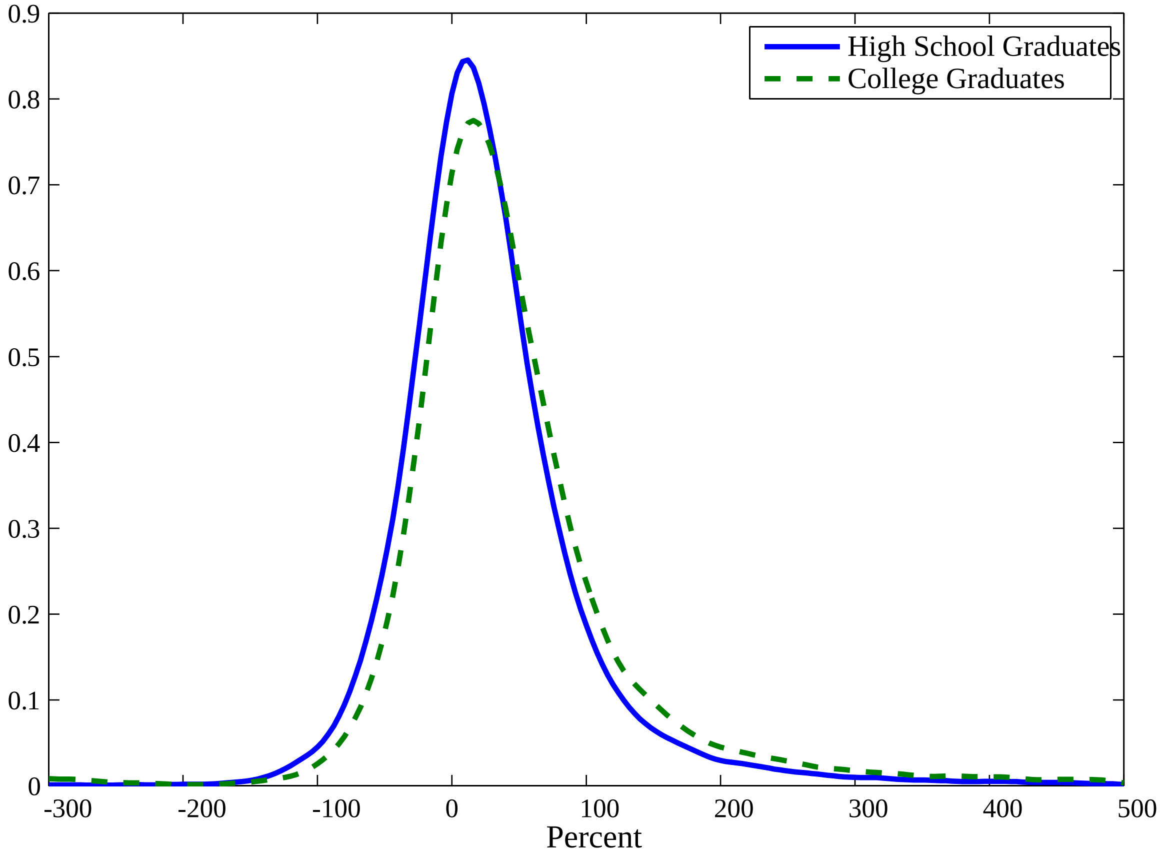
Let Y_0 denote the present value of earnings from age 22 to 41 in the High School sector (S = 0). Let Y_1 denote the present value of earnings from age 22 to 41 in the college sector (S = 1). Here we plot the factual density function $f(y_1|S=1)$ (the solid curve) against the counterfactual density function $f(y_0|S=1)$ (the dashed curve). We use a discount rate of 5%.

Figure 5B
Densities of present value of earnings
College Sample NLS/1966



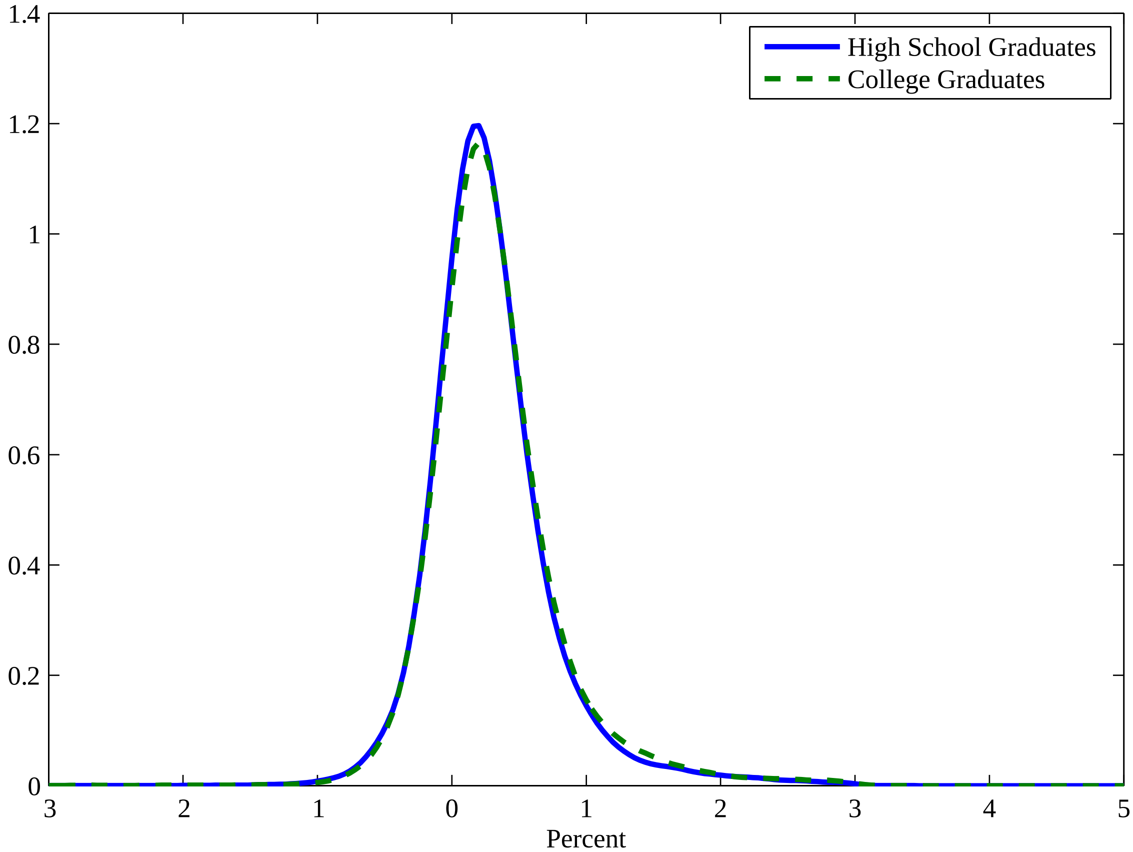
Let Y_0 denote the present value of earnings from age 22 to 41 in the High School sector (S = 0). Let Y_1 denote the present value of earnings from age 22 to 41 in the college sector (S = 1). Here we plot the factual density function $f(y_1|S=1)$ (the solid curve) against the counterfactual density function $f(y_0|S=1)$ (the dashed curve). We use a discount rate of 5%.

Figure 6A
Densities of Returns to College
NLSY/1979 Sample



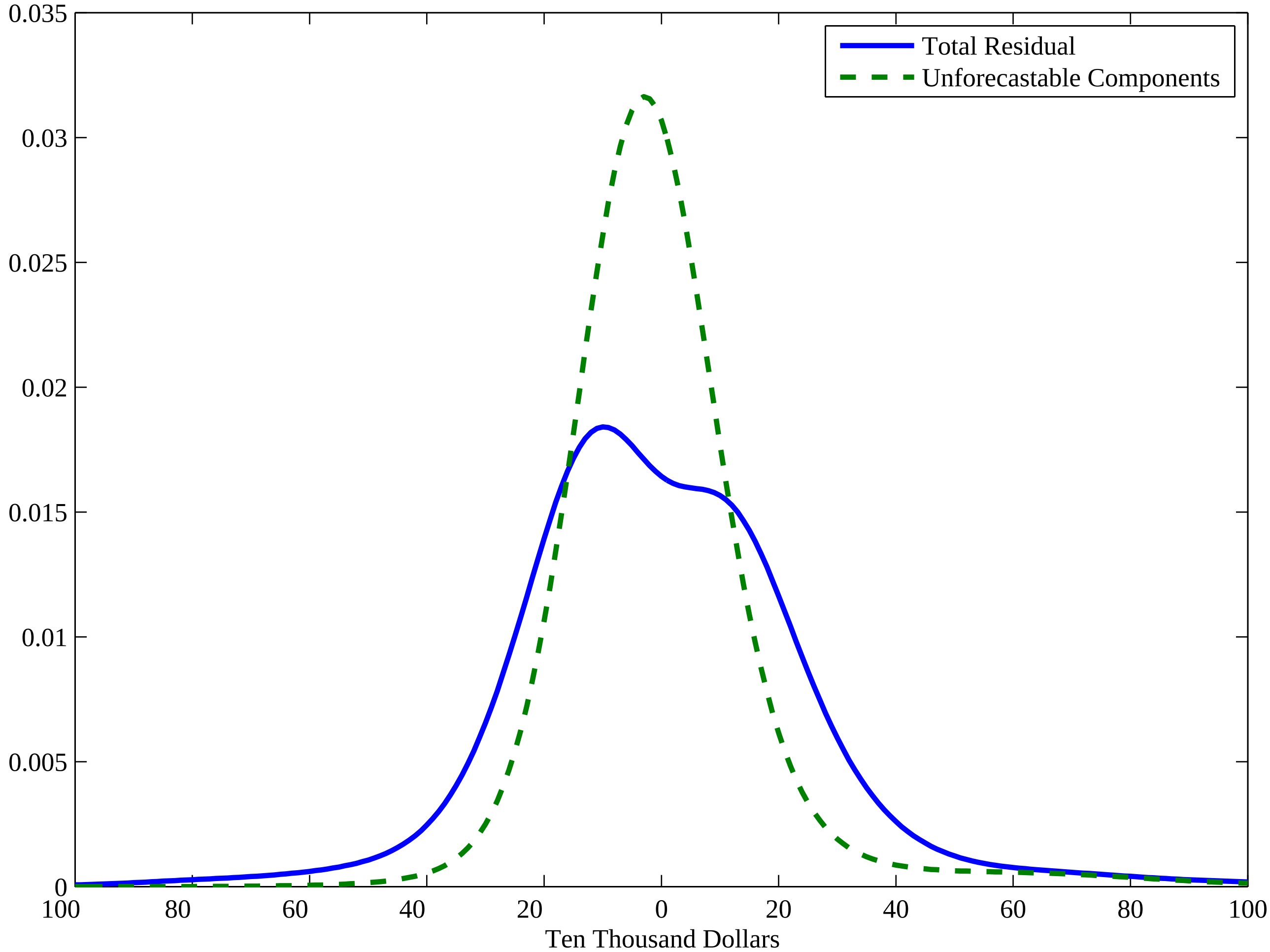
Let Y_0, Y_1 denote the present value of earnings from age 22 to age 41 in the high school and college sectors, respectively. Define ex post returns to college as the ratio $R=(Y_1-Y_0)/Y_0$. Let f(r) denote the density function of the ex post returns to college R. The solid line is the density of ex post returns to colege for high school graduates, that is, f(r|S=0). The dashed line is the density of ex post returns to college for college graduates, that is, f(r|S=1).

Figure 6B
Densities of Returns to College
NLS/1966 Sample



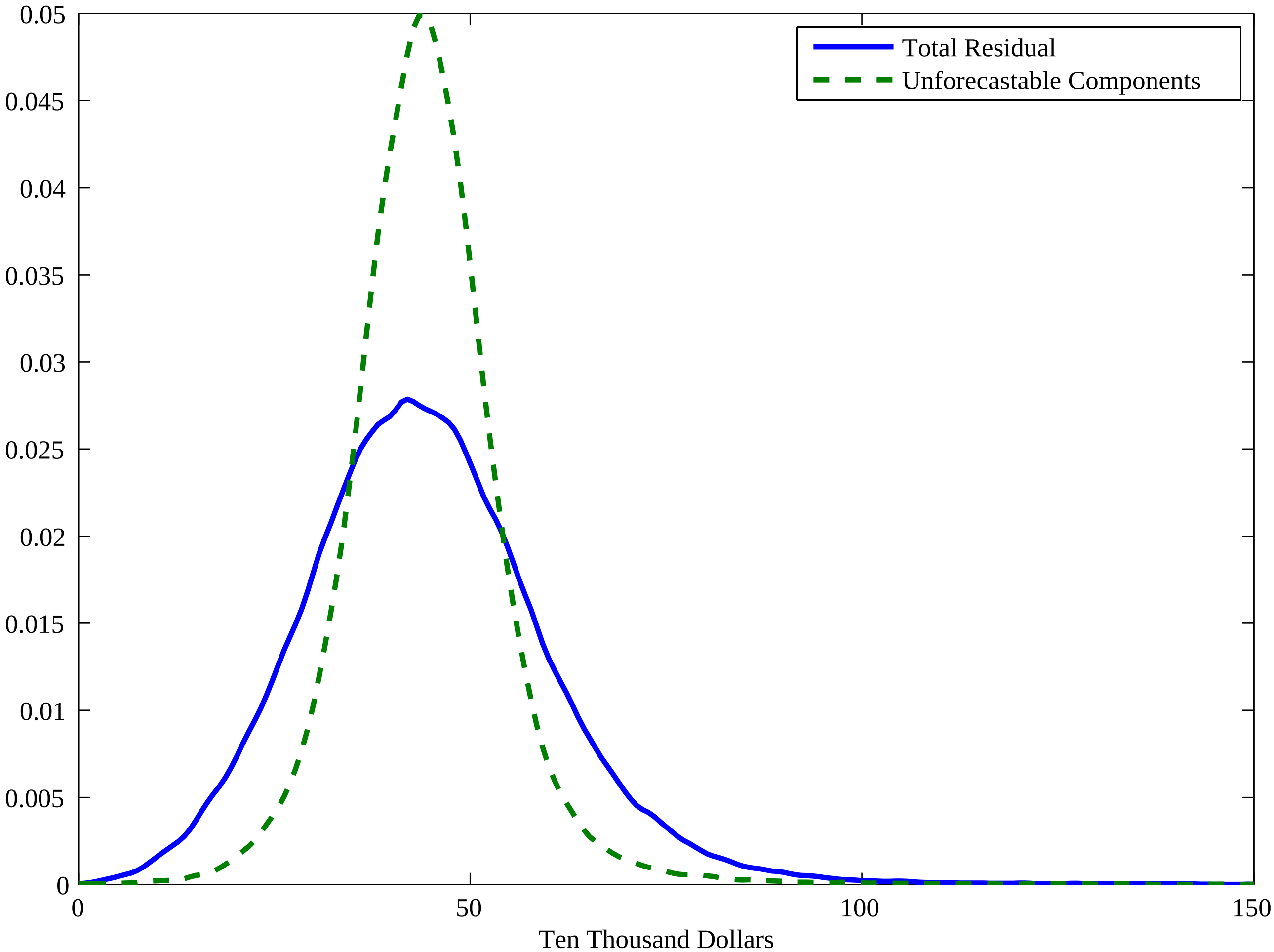
Let Y_0, Y_1 denote the present value of earnings from age 22 to age 41 in the high school and college sectors, respectively. Define ex post returns to college as the ratio $R=(Y_1-Y_0)/Y_0$. Let f(r) denote the density function of the ex post returns to college R. The solid line is the density of ex post returns to colege for high school graduates, that is, f(r|S=0). The dashed line is the density of ex post returns to college for college graduates, that is, f(r|S=1).

Figure 7A
The densities of total residual vs unforecastable components
in present value of high school earnings for the NLSY/1979 sample



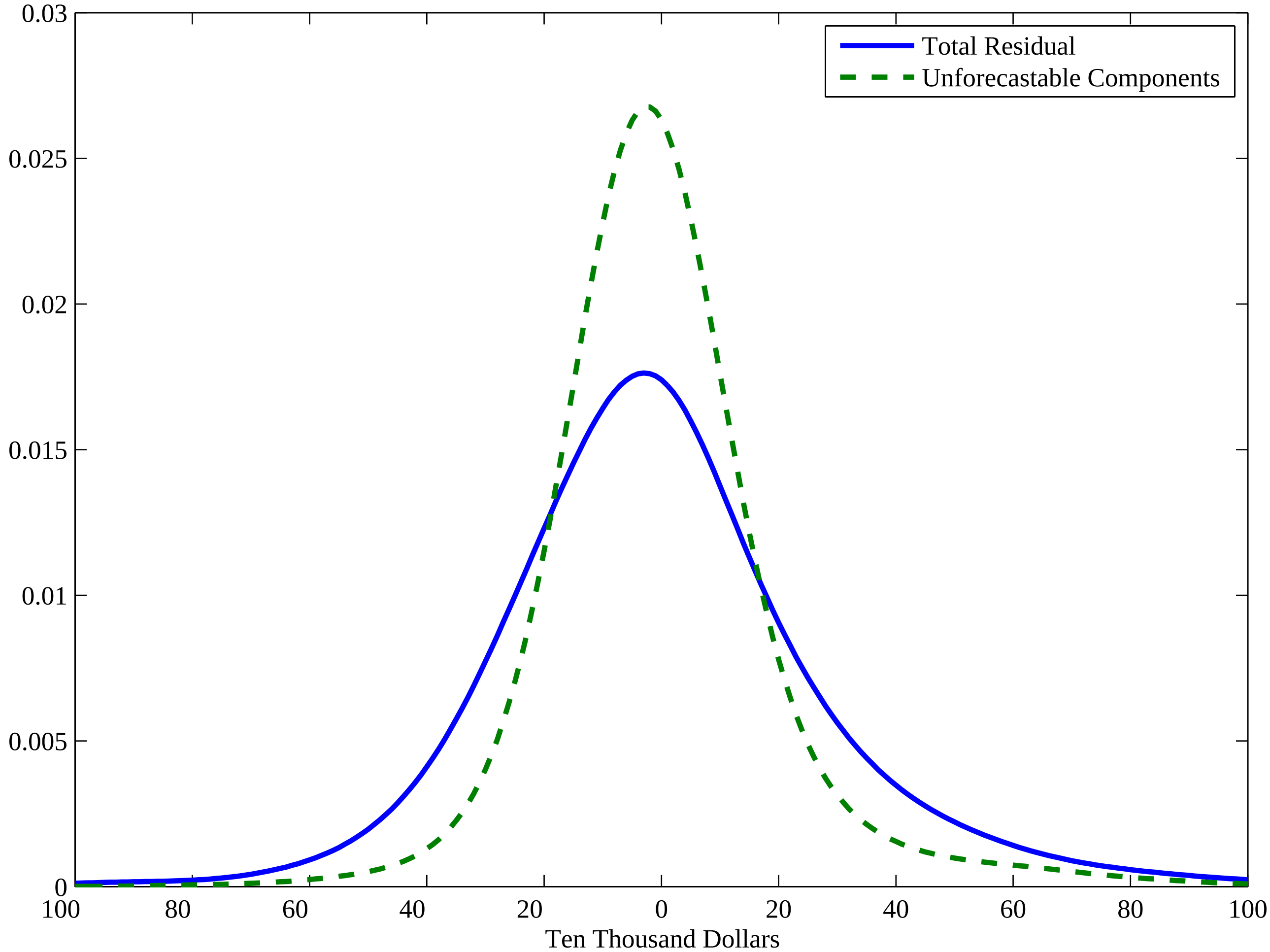
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of high-school earnings from ages 22 to 41 for the NLSY/1979 sample of white males. The present value of earnings is calculated using a 5% interest rate.

Figure 7B
The densities of total residual vs unforecastable components
in present value of high school earnings for the NLS/1966 sample



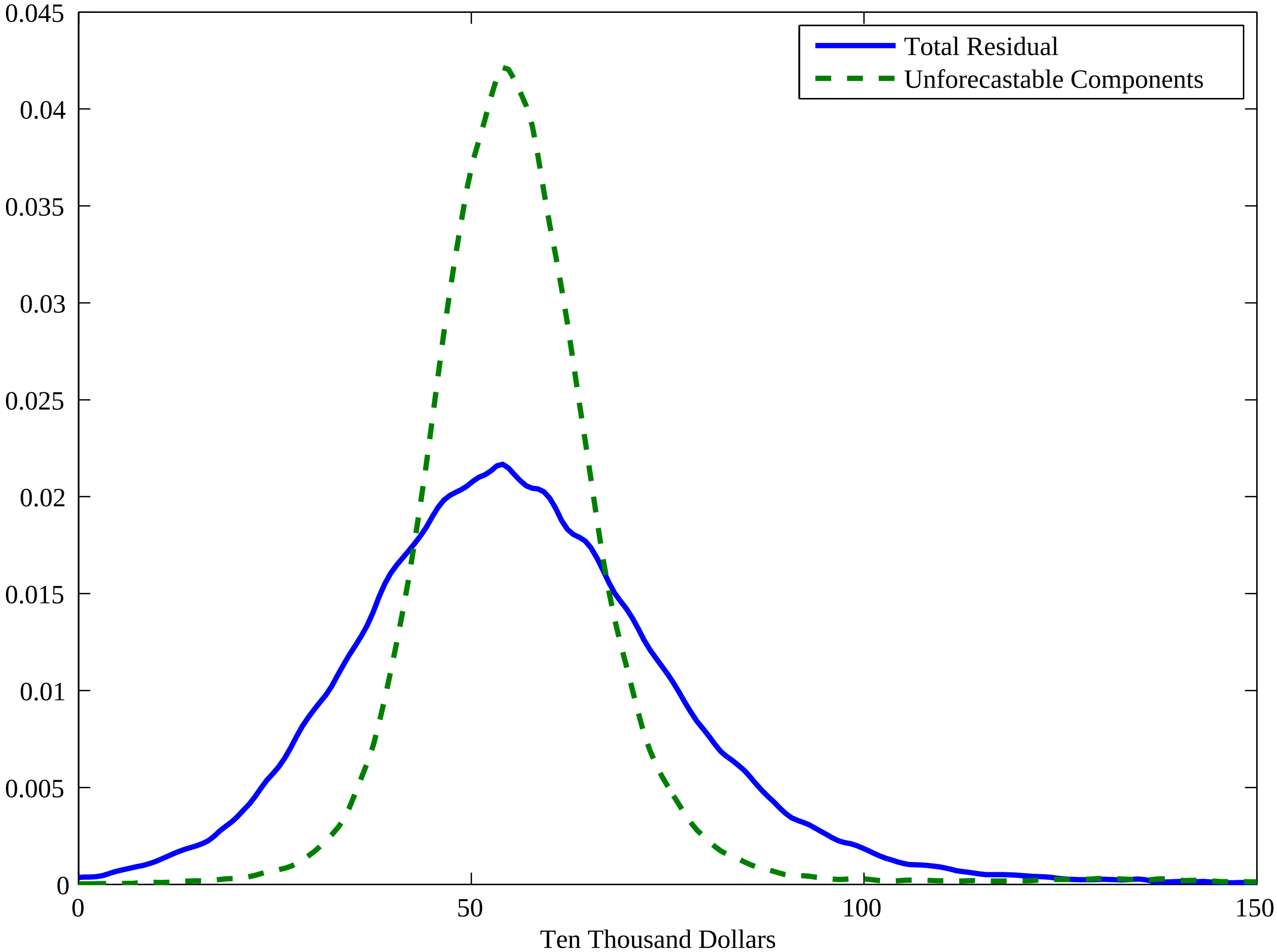
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of high-school earnings from ages 22 to 41 for the NLS/1966 sample of white males. The present value of earnings is calculated using a 5% interest rate.

Figure 8A
The densities of total residual vs unforecastable components
in present value of college earnings for the NLSY/1979 sample



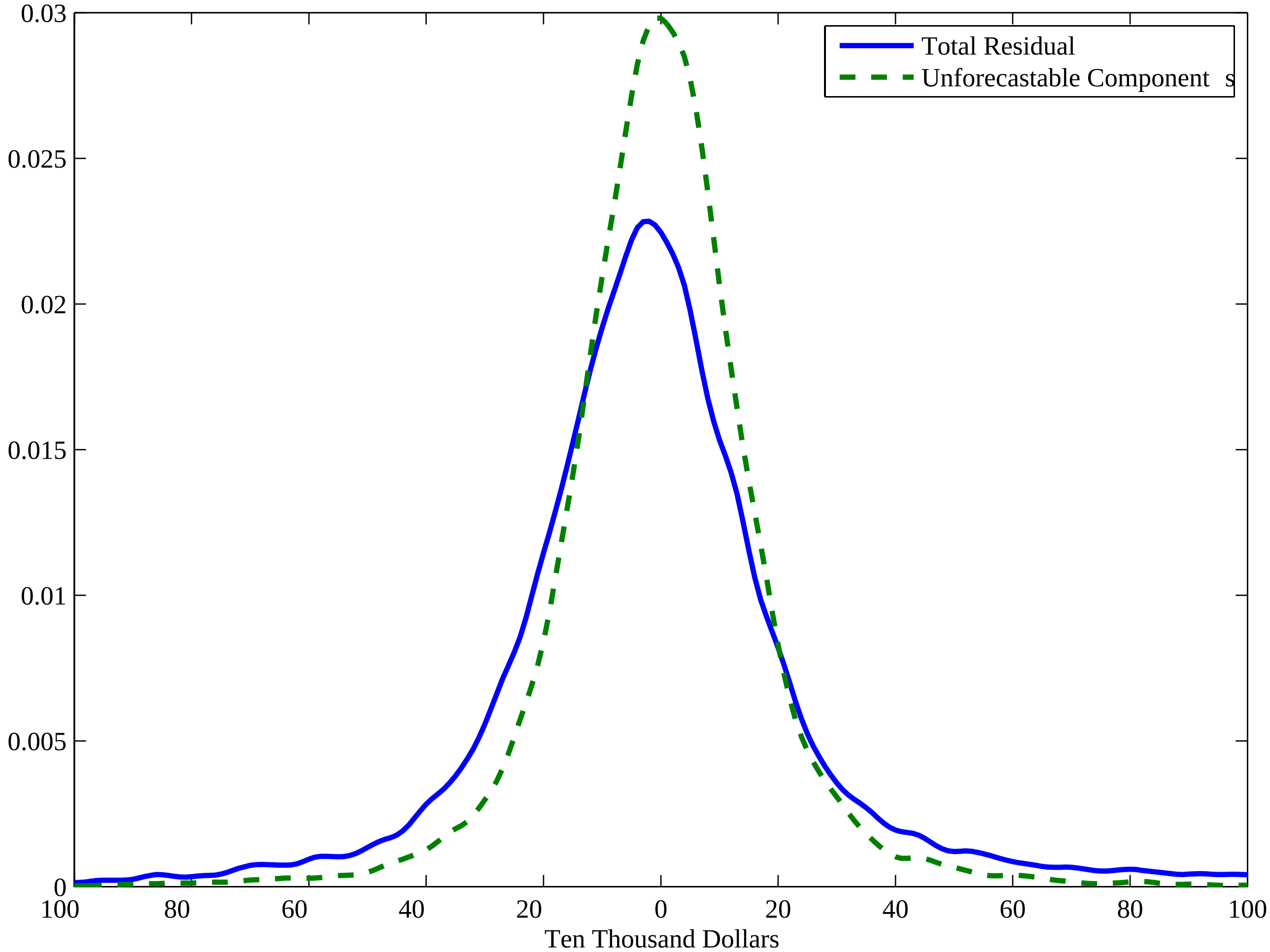
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of college earnings from ages 22 to 41 for the NLSY/1979 sample of white males. The present value of earnings is calculated using a 5% interest rate.

Figure 8B
The densities of total residual vs unforecastable components
in present value of college earnings for the NLS/1966 sample



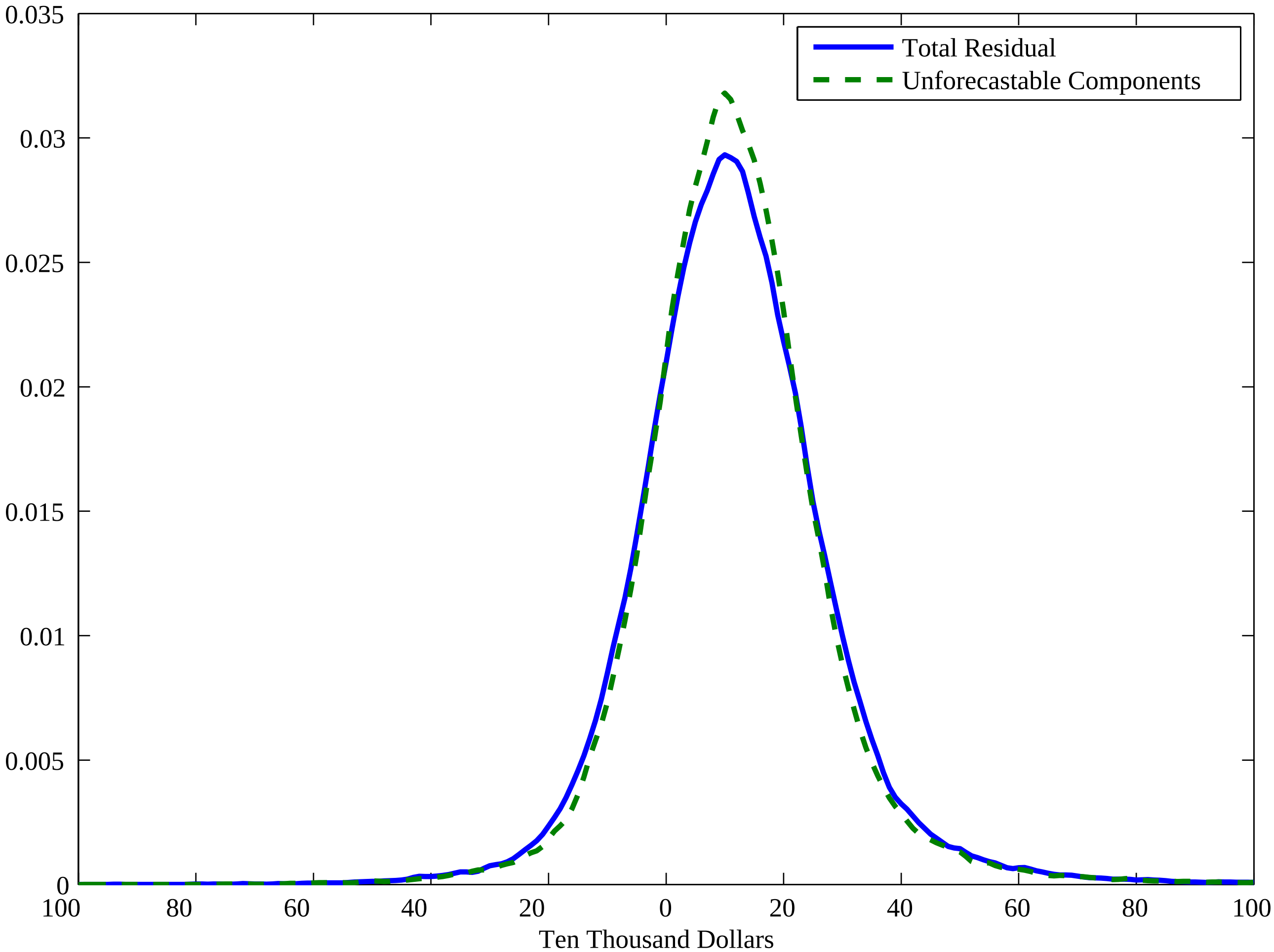
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of college earnings from ages 22 to 41 for the NLS/1966 sample of white males. The present value of earnings is calculated using a 5% interest rate.

Figure 9A
Densities of total residual vs unforecastable components
returns college vs high school for the NLSY/1979 sample



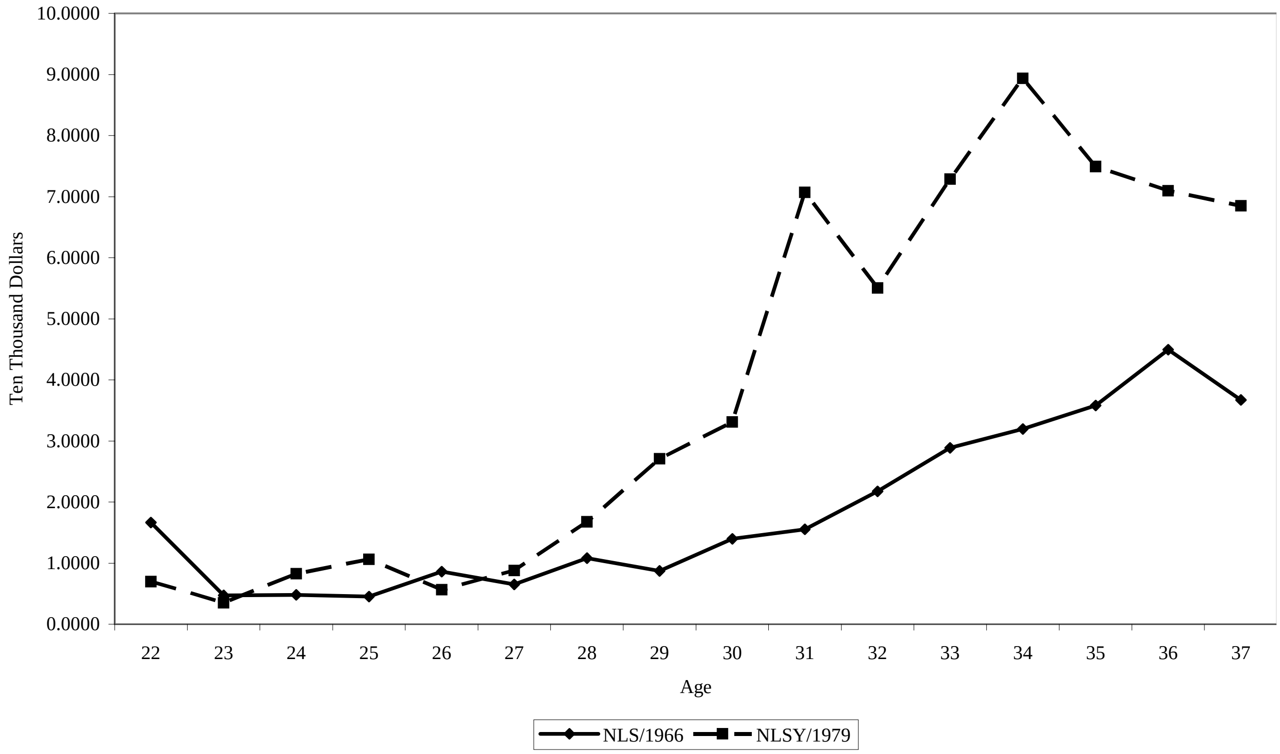
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of earnings differences (or returns to college) for the white males sample of the NLSY/1979 from ages 22 to 41. The present value of returns to college is calculated using a 5% interest rate.

Figure 9B
Densities of total residual vs unforecastable components
returns college vs high school for the NLS/1966 sample



In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of earnings differences (or returns to college) for the white males sample of the NLSY/1979 from ages 22 to 41. The present value of returns to college is calculated using a 5% interest rate.

Figure 10
Evolution of Variance of Unforecastable Components - High School Sector



For each schooling level s, at each age t, we model earnings $Y_{s,t}$ according to:

$$Y_{s,t} = X\beta_{s,t} + \theta\alpha_{s,t} + \varepsilon_{s,t}$$

For the NLS/1966 data set, the vector θ contains 5 elements. We test and cannot reject that the agents know the factors θ_1, θ_2, and θ_3 but they don't know factors θ_4, θ_5, and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. For the NLSY/1979 data set, the vector θ contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors θ_1, θ_2, and θ_3 but they don't know factors $\theta_4, \theta_5, \theta_6$ and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. Let $P_{s,t}$ denote the unforecastable components at the time of the schooling choice. For the NLS/1966, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \varepsilon_{s,t}$. For the NLSY/1979, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \alpha_{6,s,t}\theta_6 + \varepsilon_{s,t}$. In Figure 10, we compare the variance of $P_{s,t}$ from NLS/1966 (the solid curve) with the one from NLSY/1979 (the dashed curve) at different ages of the individuals who are high-school graduates. We see that until age 27, the estimated variance of $P_{s,t}$ from NLS/1966 and NLSY/1979 are very similar, but from age 28 on, the variance of $P_{s,t}$ from NLSY/1979 is much larger than the counterpart from NLS/1966.

Figure 11
Evolution of Variance of Unforecastable Components - College Sector

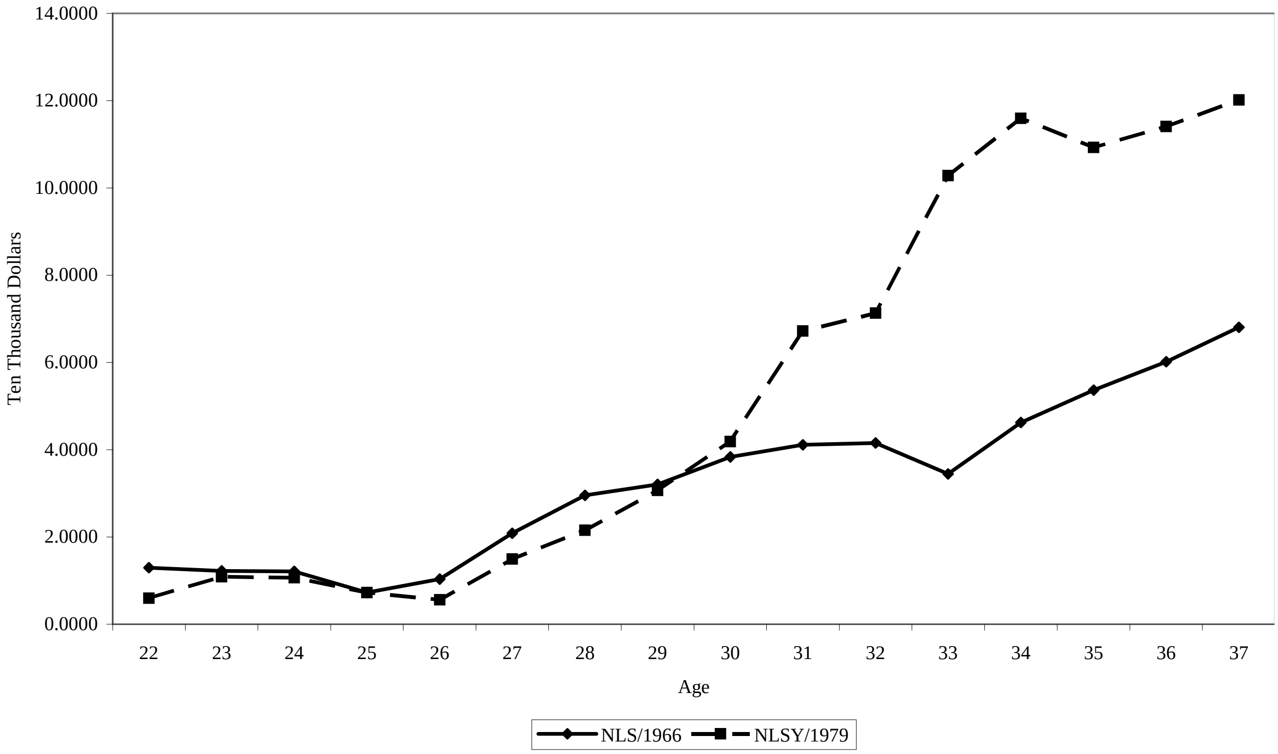


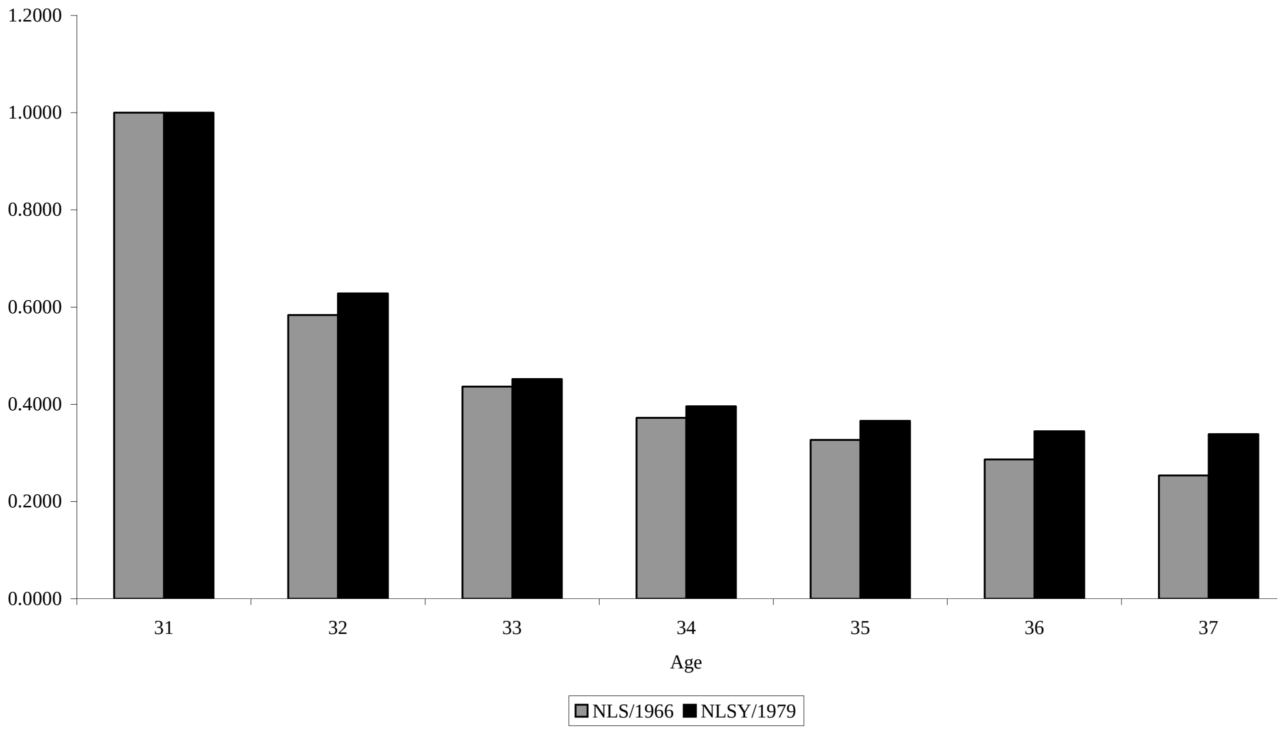
Figure 1:

For each schooling level s, at each age t, we model earnings $Y_{s,t}$ according to:

$$Y_{s,t} = X\beta_{s,t} + \theta\alpha_{s,t} + \varepsilon_{s,t}$$

For the NLS/1966 data set, the vector θ contains 5 elements. We test and cannot reject that the agents know the factors θ_1, θ_2, and θ_3 but they don't know factors θ_4, θ_5, and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. For the NLSY/1979 data set, the vector θ contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors θ_1, θ_2, and θ_3 but they don't know factors $\theta_4, \theta_5, \theta_6$ and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. Let $P_{s,t}$ denote the unforecastable components at the time of the schooling choice. For the NLS/1966, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \varepsilon_{s,t}$. For the NLSY/1979, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \alpha_{6,s,t}\theta_6 + \varepsilon_{s,t}$. In Figure 11, we compare the variance of $P_{s,t}$ from NLS/1966 (the solid curve) with the one from NLSY/1979 (the dashed curve) at different ages of the individuals who are college graduates. We see that until age 30, the estimated variance of $P_{s,t}$ from NLS/1966 and NLSY/1979 are very similar, but from age 31 on, the variance of $P_{s,t}$ from NLSY/1979 is much larger than the counterpart from NLS/1966.

Figure 12
Correlogram at Age 31: NLS/1966 vs NLSY/1979
Overall Sample



For each schooling level s, at each age t, we model earnings $Y_{s,t}$ according to:

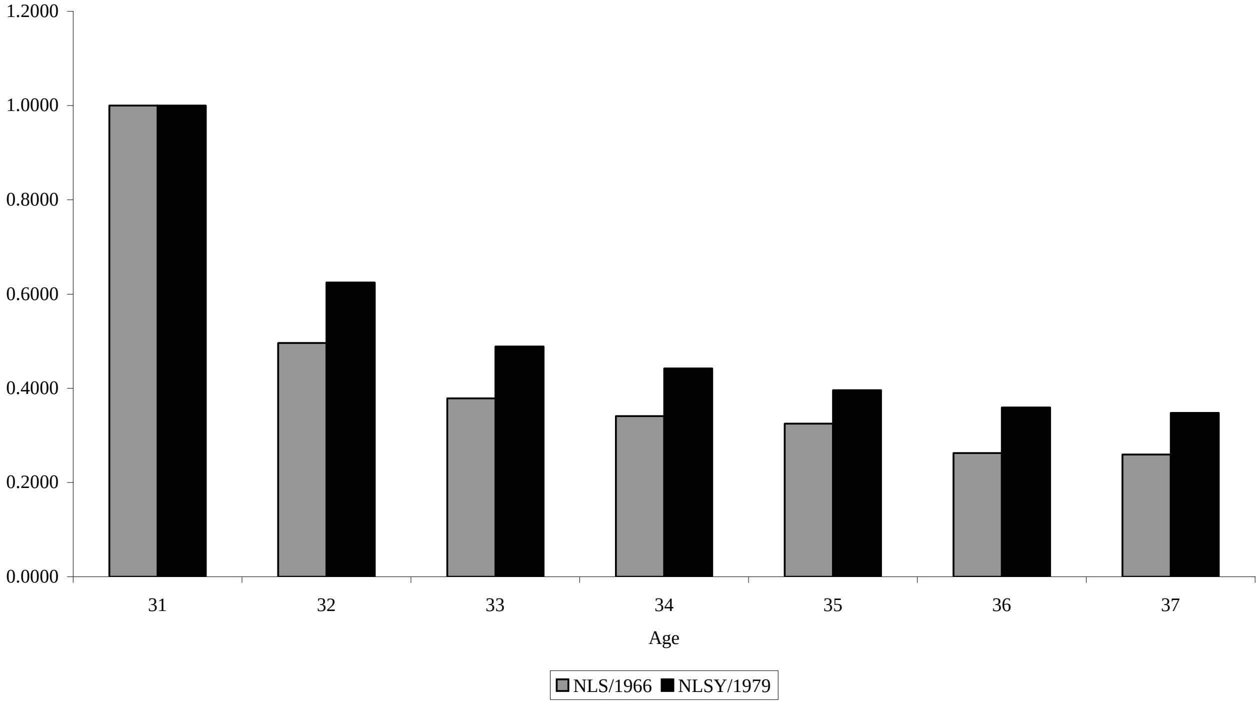
$$Y_{s,t} = X\beta_{s,t} + \theta\alpha_{s,t} + \varepsilon_{s,t}$$

For the NLS/1966 data set, the vector θ contains 5 elements. We test and cannot reject that the agents know the factors θ_1, θ_2, and θ_3 but they don't know factors θ_4, θ_5, and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. For the NLSY/1979 data set, the vector θ contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors θ_1, θ_2, and θ_3 but they don't know factors $\theta_4, \theta_5, \theta_6$ and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. Let $P_{s,t}$ denote the unforecastable components in sector s and age t at the time of the schooling choice. For the NLS/1966, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \varepsilon_{s,t}$. For the NLSY/1979, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \alpha_{6,s,t}\theta_6 + \varepsilon_{s,t}$. Let $\phi(s,t,k)$ denote the correlogram at age t :

$$\phi(s,t,k) = \frac{Cov(P_{s,t}, P_{s,k})}{Var(P_{s,t})} \text{ for } k = t, t+1, t+2, ..., T$$

In Figure 12, we plot $\phi(s,t,k)$ from NLS/1966 (the gray columns) with the one from NLSY/1979 (the black columns) when the agents are 31 years-ol~~d ($t = 31$) for t~~he overall sample.

Figure 13
Correlogram at Age 31: NLS/1966 vs NLSY/1979
High School Sample



For each schooling level s, at each age t, we model earnings $Y_{s,t}$ according to:

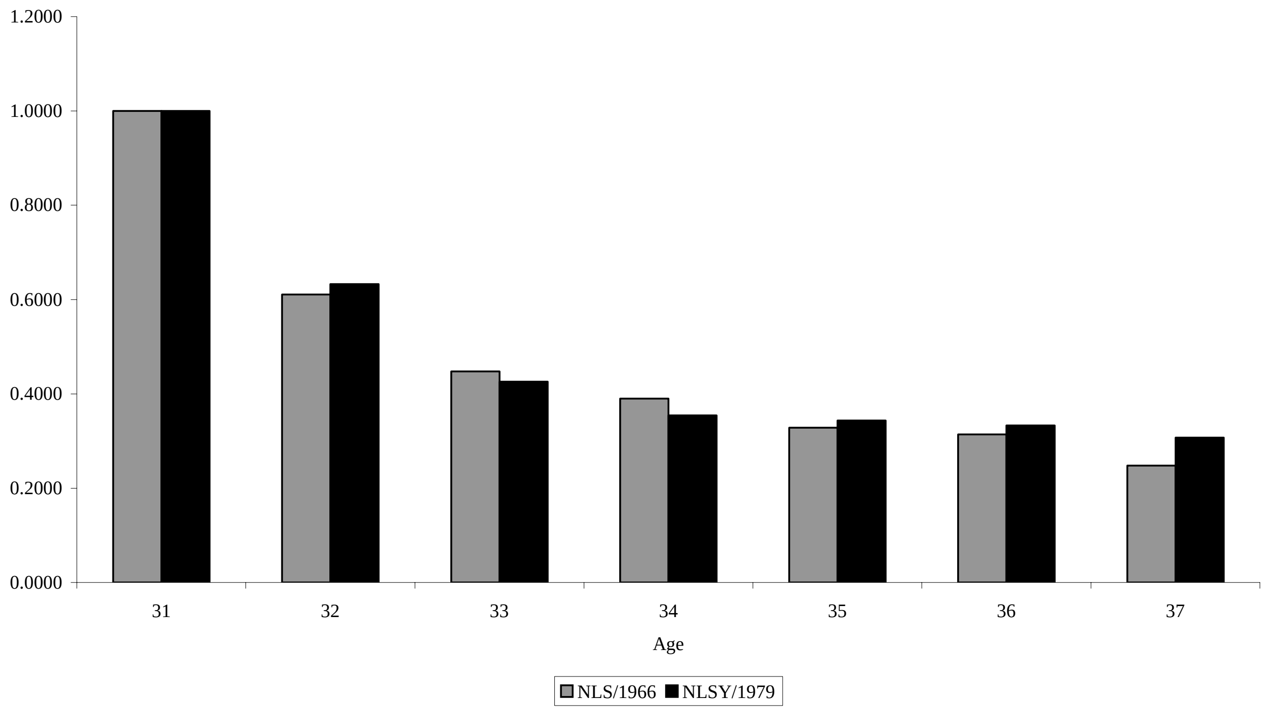
$$Y_{s,t} = X\beta_{s,t} + \theta\alpha_{s,t} + \varepsilon_{s,t}$$

For the NLS/1966 data set, the vector θ contains 5 elements. We test and cannot reject that the agents know the factors θ_1, θ_2, and θ_3 but they don't know factors θ_4, θ_5, and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. For the NLSY/1979 data set, the vector θ contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors θ_1, θ_2, and θ_3 but they don't know factors $\theta_4, \theta_5, \theta_6$ and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. Let $P_{s,t}$ denote the unforecastable components in sector s and age t at the time of the schooling choice. For the NLS/1966, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \varepsilon_{s,t}$. For the NLSY/1979, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \alpha_{6,s,t}\theta_6 + \varepsilon_{s,t}$. Let $\phi(s,t,k)$ denote the correlogram at age t :

$$\phi(s,t,k) = \frac{Cov(P_{s,t}, P_{s,k})}{Var(P_{s,t})} \text{ for } k = t, t+1, t+2, ..., T$$

In Figure 13, we plot $\phi(s,t,k)$ from NLS/1966 (the gray columns) with the one from NLSY/1979 (the black columns) when the agents are 31 years-ol~~d $(t = 31)$ for t~~he high school sample.

Figure 14
Correlogram at Age 31: NLS/1966 vs NLSY/1979
College Sample



For each schooling level s, at each age t, we model earnings $Y_{s,t}$ according to:

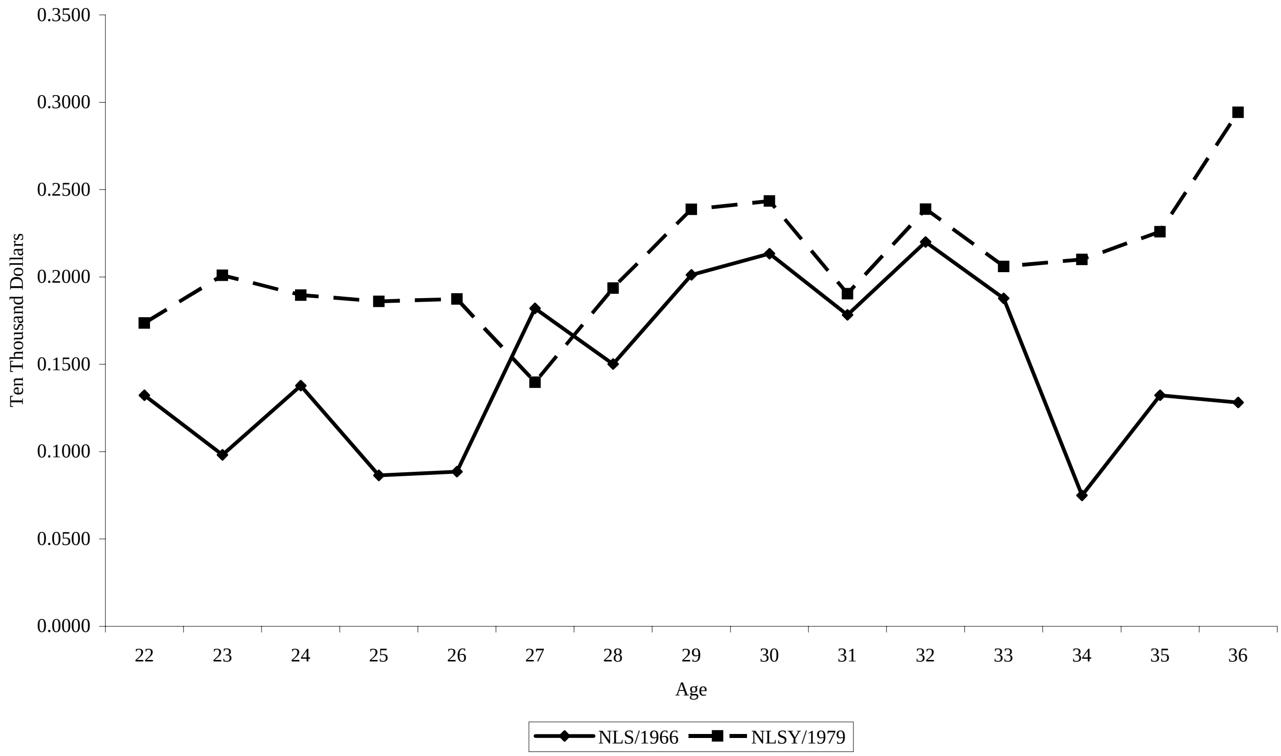
$$Y_{s,t} = X\beta_{s,t} + \theta\alpha_{s,t} + \varepsilon_{s,t}$$

For the NLS/1966 data set, the vector θ contains 5 elements. We test and cannot reject that the agents know the factors θ_1, θ_2, and θ_3 but they don't know factors θ_4, θ_5, and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. For the NLSY/1979 data set, the vector θ contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors θ_1, θ_2, and θ_3 but they don't know factors $\theta_4, \theta_5, \theta_6$ and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. Let $P_{s,t}$ denote the unforecastable components in sector s and age t at the time of the schooling choice. For the NLS/1966, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \varepsilon_{s,t}$. For the NLSY/1979, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \alpha_{6,s,t}\theta_6 + \varepsilon_{s,t}$. Let $\phi(s, t, k)$ denote the correlogram at age t :

$$\phi(s, t, k) = \frac{Cov(P_{s,t}, P_{s,k})}{Var(P_{s,t})} \text{ for } k = t, t+1, t+2, ..., T$$

In Figure 14, we plot $\phi(s, t, k)$ from NLS/1966 (the gray columns) with the one from NLSY/1979 (the black columns) when the agents are 31 years-ol~~d ($t = 31$) for t~~he college sample.

Figure 15
Evolution of Cognitive Skill Prices - High School Sector

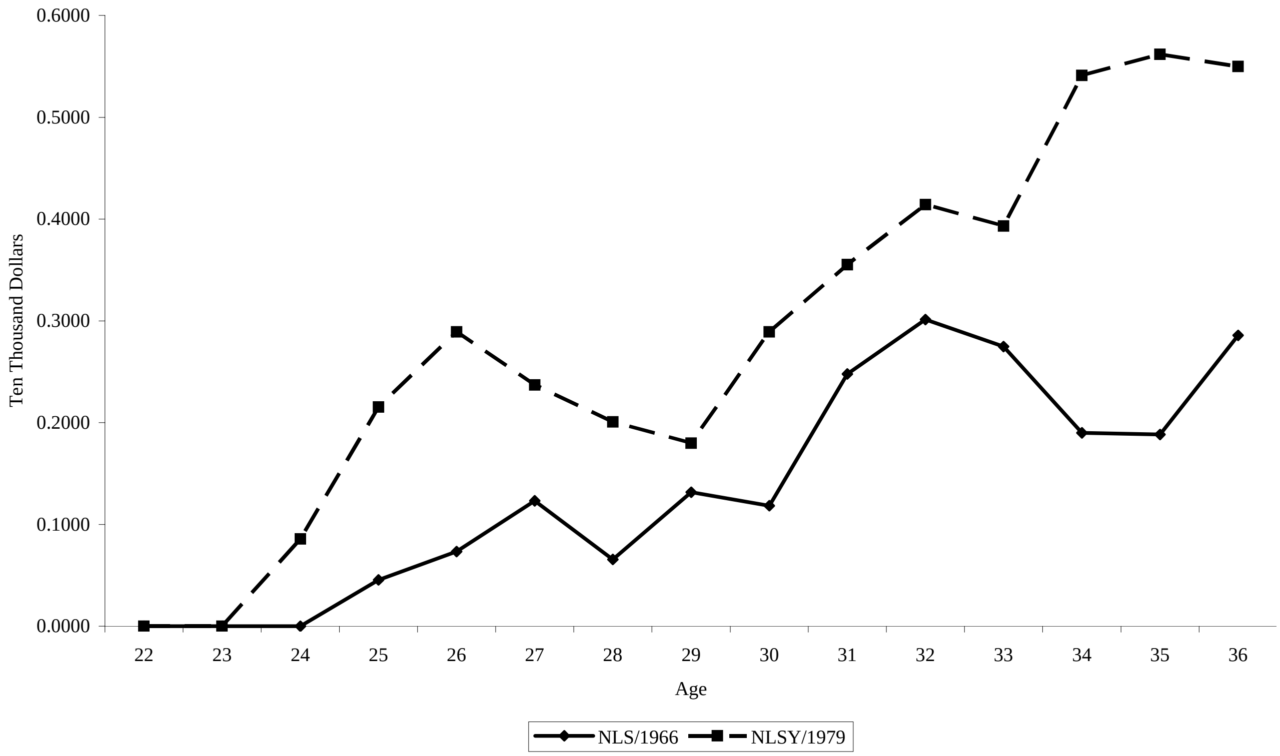


For each schooling level s, at each age t, we model earnings $Y_{s,t}$ according to:

$$Y_{s,t} = X\beta_{s,t} + \theta\alpha_{s,t} + \varepsilon_{s,t}$$

For the NLS/1966 data set, the vector θ contains 5 elements. For the NLSY/1979 data set, the vector θ contains 6 elements. In both cases, the first factor, θ_1, is identified from cognitive tests. According to the Gorman-Lancaster model of earnings, the loading on the first factor in the earnings equations is the price of cognitive skills. In Figure 15 we plot the loading on factor θ_1 in the high-school earnings equation from ages 22 to 36 for both the NLS/1966 and NLSY/1979.

Figure 16
Evolution of Cognitive Skill Prices - College Sector



For each schooling level s, at each age t, we model earnings $Y_{s,t}$ according to:

$$Y_{s,t} = X\beta_{s,t} + \theta\alpha_{s,t} + \varepsilon_{s,t}$$

For the NLS/1966 data set, the vector θ contains 5 elements. For the NLSY/1979 data set, the vector θ contains 6 elements. In both cases, the first factor, θ_1, is identified from cognitive tests. According to the Gorman-Lancaster model of earnings, the loading on the first factor in the earnings equations is the price of cognitive skills. In Figure 16 we plot the loading on factor θ_1 in the college earnings equation from ages 22 to 36 for both the NLS/1966 and NLSY/1979.

Table 1

Test of Equality of Predicted versus Actual Correlation

Matrices of Earnings (from ages 22 to 41)

NLSY/1979 and NLS/1966

	High School	**College**	**Overall**
NLS/1966 - 5 Factors	15.6968	210.4133	114.8754
NLS/1979 - 6 Factors	70.6451	156.5446	187.5425
NLS/1979 - 5 Factors	64.2682	309.2815	226.2401
Critical Value*	222.0741	222.0741	222.0741

* 95% Confidence

Table 2A: 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 unknown θ at their means, so Information Set=$\{\theta_1,\theta_2,\theta_3\}$
Corrrelation(Y_C,Y_H) = 0.1666

	College									
High School	1	2	3	4	5	6	7	8	9	10
1	0.2995	0.1685	0.1114	0.0789	0.0570	0.0413	0.0393	0.0431	0.0471	0.1137
2	0.2273	0.2119	0.1597	0.1271	0.0907	0.0678	0.0450	0.0288	0.0180	0.0236
3	0.1532	0.1840	0.1656	0.1472	0.1146	0.0914	0.0642	0.0434	0.0230	0.0132
4	0.1110	0.1368	0.1492	0.1474	0.1418	0.1184	0.0882	0.0588	0.0334	0.0148
5	0.0748	0.1100	0.1244	0.1413	0.1459	0.1403	0.1172	0.0836	0.0462	0.0162
6	0.0494	0.0866	0.1146	0.1204	0.1371	0.1399	0.1283	0.1242	0.0736	0.0258
7	0.0306	0.0582	0.0904	0.1094	0.1264	0.1436	0.1506	0.1430	0.1064	0.0414
8	0.0236	0.0348	0.0531	0.0769	0.0989	0.1252	0.1638	0.1799	0.1676	0.0761
9	0.0264	0.0262	0.0316	0.0459	0.0651	0.0929	0.1308	0.1784	0.2431	0.1594
10	0.0457	0.0182	0.0214	0.0216	0.0321	0.0446	0.0772	0.1176	0.2291	0.3925

Table 2B: Ex-Ante 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 unknown θ at their means, so Information Set=$\{\theta_1,\theta_2,\theta_3\}$
Corrrelation(Y_C,Y_H) = 0.9174

	College									
High School	1	2	3	4	5	6	7	8	9	10
1	0.7036	0.2155	0.0622	0.0137	0.0035	0.0015	0.0000	0.0000	0.0000	0.0000
2	0.2225	0.3780	0.2475	0.1085	0.0285	0.0110	0.0035	0.0000	0.0005	0.0000
3	0.0500	0.2505	0.2960	0.2320	0.1090	0.0455	0.0120	0.0035	0.0015	0.0000
4	0.0145	0.1005	0.2250	0.2585	0.2150	0.1135	0.0545	0.0135	0.0045	0.0005
5	0.0045	0.0435	0.1055	0.1945	0.2545	0.2135	0.1265	0.0460	0.0105	0.0010
6	0.0010	0.0115	0.0435	0.1190	0.2035	0.2455	0.2100	0.1335	0.0295	0.0030
7	0.0000	0.0030	0.0150	0.0500	0.1190	0.2185	0.2705	0.2095	0.1040	0.0105
8	0.0005	0.0000	0.0055	0.0200	0.0555	0.1085	0.2080	0.3125	0.2460	0.0435
9	0.0000	0.0000	0.0005	0.0035	0.0105	0.0380	0.1045	0.2390	0.3920	0.2120
10	0.0000	0.0000	0.0000	0.0005	0.0010	0.0045	0.0105	0.0425	0.2115	0.7295

Table 3A: Ex-Post Conditional Distributions for the NLSY/1979 (College Earnings Conditional on High School Earnings)
Pr(d_i<Yc<d_i+1 | 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
Information Set={$\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6$}
Corrrelation(Y_C,Y_H) = 0.2842

	College									
High School	1	2	3	4	5	6	7	8	9	10
1	0.2118	0.1614	0.1188	0.0932	0.0782	0.0654	0.0532	0.0554	0.0651	0.0974
2	0.1684	0.1777	0.1557	0.1213	0.1038	0.0862	0.0640	0.0516	0.0417	0.0296
3	0.1374	0.1676	0.1464	0.1390	0.1244	0.0954	0.0754	0.0577	0.0333	0.0234
4	0.1080	0.1336	0.1433	0.1378	0.1213	0.1115	0.0980	0.0746	0.0475	0.0243
5	0.0787	0.1105	0.1232	0.1335	0.1345	0.1291	0.1144	0.0862	0.0614	0.0286
6	0.0656	0.1028	0.1149	0.1201	0.1276	0.1330	0.1250	0.0998	0.0823	0.0288
7	0.0548	0.0779	0.0842	0.1097	0.1196	0.1224	0.1410	0.1331	0.1132	0.0441
8	0.0428	0.0507	0.0741	0.0880	0.0994	0.1224	0.1410	0.1585	0.1539	0.0693
9	0.0416	0.0436	0.0474	0.0577	0.0803	0.1001	0.1277	0.1728	0.1939	0.1348
10	0.0386	0.0204	0.0269	0.0292	0.0339	0.0520	0.0704	0.1155	0.1945	0.4186

Table 3B: Ex-Post Conditional Distributions for the NLS/1966 (College Earnings Conditional on High School Earnings)
Pr(d_i<Yc<d_i+1 | 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
Information Set={$\theta_1,\theta_2,\theta_3,\theta_4,\theta_5$}
Corrrelation(Y_C,Y_H) = 0.6226

	College									
High School	1	2	3	4	5	6	7	8	9	10
1	0.4001	0.1813	0.1023	0.0717	0.0611	0.0406	0.0422	0.0306	0.0337	0.0364
2	0.2144	0.2239	0.1663	0.1207	0.0862	0.0676	0.0486	0.0261	0.0256	0.0205
3	0.1286	0.1716	0.1591	0.1496	0.1181	0.0960	0.0695	0.0515	0.0340	0.0220
4	0.0870	0.1426	0.1551	0.1576	0.1386	0.1131	0.0810	0.0650	0.0365	0.0235
5	0.0450	0.0905	0.1390	0.1400	0.1405	0.1395	0.1165	0.0960	0.0625	0.0305
6	0.0350	0.0720	0.1126	0.1196	0.1456	0.1416	0.1306	0.1211	0.0900	0.0320
7	0.0210	0.0600	0.0710	0.1046	0.1201	0.1521	0.1466	0.1531	0.1126	0.0590
8	0.0205	0.0320	0.0455	0.0816	0.0951	0.1261	0.1562	0.1797	0.1667	0.0966
9	0.0180	0.0205	0.0305	0.0430	0.0755	0.0830	0.1476	0.1741	0.2316	0.1761
10	0.0125	0.0115	0.0235	0.0135	0.0225	0.0415	0.0611	0.1041	0.2077	0.5020

Table 4

Mean Rates of Return to College by Schooling Group

	NLS/1966		NLSY/1979	
Schooling Group	Mean Returns	Standard Error	Mean Returns	Standard Error
High School Graduates	0.2937	0.0083	0.3095	0.0113
College Graduates	0.3107	0.0114	0.3994	0.0129
Individuals at the Margin	0.3081	0.0446	0.3511	0.0535

Table 5

Percentage that Regret Schooling Choices

Schooling Group	NLS/1966	NLSY/1979
Percentage of High School Graduates who Regret Not Graduating from College	0.0966	0.0749
Percentage of College Graduates who Regret Graduating from College	0.0337	0.0311

Table 6A
Evolution of Uncertainty

Panel A: NLS/1966

	College	High School	Returns
Total Residual Variance	460.6260	284.8089	351.4026
Variance of Unforecastable Components	181.3712	128.4315	327.3480

Panel B: NLSY/1979

	College	High School	Returns
Total Residual Variance	709.7487	507.2910	906.0066
Variance of Unforecastable Components	372.3509	272.3596	432.8733

Panel C: Percentage Increase

	College	High School	Returns
Percentage Increase in Total Residual Variance	54.083%	78.116%	157.826%
Percentage Increase in Variance of Unforecastable Components	105.298%	112.066%	32.236%

Table 6B
Evolution of Heterogeneity

Panel A: NLS/1966

	College	High School	Returns
Total Residual Variance	460.6260	284.8089	351.4026
Variance of Forecastable Components (Heterogeneity)	279.2549	156.3774	24.0546
Panel B: NLSY/1979			
	College	High School	Returns
Total Residual Variance	709.7487	507.2910	906.0066
Variance of Forecastable Components (Heterogeneity)	337.3978	234.9314	473.1333
Panel C: Percentage Increase			
	College	High School	Returns
Percentage Increase in Total Residual Variance	54.083%	78.116%	157.826%
Percentage Increase in Variance of Forecastable Components	20.821%	50.234%	1866.914%

Table 7

Share of Variance of Business Cycle in Total Variance of Unforecastable Components

	NLS/1966		**NLSY/1979**	
	Point Estimate	Standard Error	Point Estimate	Standard Error
High School	0.0586	0.0060	0.0069	0.0009
College	0.1193	0.0126	0.0158	0.0021

Let $Y_{s,t}$ denote the labor income at schooling sector s and age t. Let d_k denote the cohort dummy that takes the value one if the agent was born in year k and zero otherwise. Let X denote the vector of variables containing a dummy indicating whether the agent lived in the South Region at age 14 and a constant term. Let θ_j denote the factor j and $\alpha_{s,t,j}$ denote its factor loading at schooling sector s and age t. Let $\varepsilon_{s,t}$ denote the uniqueness. The model is:

$$Y_{s,t} = X\beta_{s,t} + \sum_{k=\tau_0}^{\tau_1} \gamma_{k,s,t} d_k + \theta_1\alpha_{s,t,1} + \theta_2\alpha_{s,t,2} + \theta_3\alpha_{s,t,3} + \theta_4\alpha_{s,t,4} + \theta_5\alpha_{s,t,5} + \theta_6\alpha_{s,t,6} + \varepsilon_{s,t}$$

The cohort dummies can capture aggregate shocks. Under this interpretation, we test and reject that the agents know the aggregate shocks at the time of the schooling choice. We test and reject that the agent knows the uniqueness $\varepsilon_{s,t}$ and factors $\theta_{4,}\theta_5$, and θ_6 at the time of the schooling choice. Consequently, the total unforecastable component (aggregate and idiosyncratic components) is given by:

$$\tilde{P}_{s,t} = \sum_{k=\tau_0}^{\tau_1} \gamma_{k,s,t} d_k + \theta_4\alpha_{s,t,4} + \theta_5\alpha_{s,t,5} + \theta_6\alpha_{s,t,6} + \varepsilon_{s,t}$$

In school sector s lifetime earnings, this component is given by the discounted summation:

$$\tilde{Q}_s = \sum_{t=22}^{41} \left[\frac{\sum_{k=\tau_0}^{\tau_1} \gamma_{k,s,t} d_k}{(1+\rho)^{t-22}} \right] + \sum_{t=22}^{41} \left[\frac{\theta_4\alpha_{s,t,4} + \theta_5\alpha_{s,t,5} + \theta_6\alpha_{s,t,6} + \varepsilon_{s,t}}{(1+\rho)^{t-22}} \right]$$

The variance of the total unforecastable component (aggregate plus idiosyncratic uncertainty) is:

$$Var\left(\tilde{Q}_s\right) = Var\left(\sum_{t=22}^{41} \left[\frac{\sum_{k=\tau_0}^{\tau_1} \gamma_{k,s,t} d_k}{(1+\rho)^{t-22}} \right] \right) + Var\left(\sum_{t=22}^{41} \left[\frac{\theta_4\alpha_{s,t,4} + \theta_5\alpha_{s,t,5} + \theta_6\alpha_{s,t,6} + \varepsilon_{s,t}}{(1+\rho)^{t-22}} \right] \right)$$

The share of aggregate uncertainty in the total variance of the unforecastable component, m_s, is:

$$m_s = \frac{Var\left(\sum_{t=22}^{41} \left[\frac{\sum_{k=\tau_0}^{\tau_1} \gamma_{k,s,t} d_k}{(1+\rho)^{t-22}} \right] \right)}{Var\left(\tilde{Q}_s\right)}$$

In the table, we plot m_s for s = high school, college, for both the NLSY/1979 and NLS/1966. For example, 5.86% of the total variance of unforecastable components in high-school lifetime earnings is due to the aggregate uncertainty in the NLS/1966 sample and 0.7% in the NLSY/1979 sample.