Estimating Wage Volatilities for College versus High School Careers

Stacey H. Chen [1]

June 16, 2003

[1]The earlier version of this paper is entitled "Is Investing in College Education Risky?" I am indebted to Mark Bils and Eric Hanushek for helpful insights and comments. I am also grateful to Gordon Dahl, Shakeeb Khan, and Lance Lochner for useful suggestions. Christian Belzil, Moshe Buchinsky, Pedro Carneiro, Vicent Hogan, Boyan Jovanovic, Robert McMillan, Ignacio Palacios-Huerta, Michael Sattinger, Christopher Taber, two anonymous reviewers, and seminar participants at several universities and institutions made helpful comments. This paper has also benefited from discussions with Samuel Danthine, Justin Fox, Sreya Kolay, Toshihiko Mukoyama, Kharis Templeman, and Kazuhiro Yuki. All errors are mine.

Abstract

This paper examines whether income uncertainty rises with education, important for understanding why some students are hesitant to attend college. I focus on the identification of the volatility differential — the gap in the variance of the log wage between college and high school careers. I control for individual heterogeneity, including demographics, scholastic ability, parental education, and fields of major. More importantly, I address a self selection problem that arises because risk-averse agents tend to choose careers with lower volatility. I find that the volatility differential between college and high school is significantly positive, suggesting that income uncertainty can be a relevant determinant of schooling decisions especially for risk-averse high school graduates whose returns to college education are at the margin. (**JEL Classification:** J310. **Key Words:** self-selection, volatility differential, schooling choice.)

1 Introduction

A consensus to emerge from the numerous studies measuring the return to schooling is that attending college can be viewed as a profitable investment. Extensive surveys by Card (1995a, 1999) document that the returns to schooling range approximately between 8 percent and 13 percent per school year. In addition to being profitable, attending college seems to be affordable given that federal student aid programs provide guaranteed loans and tuition subsidies to needy students. Nevertheless, a number of academically talented young people do not attend a postsecondary institution. Table 1 presents statistics on college attendance according to National Longitudinal Survey of Youth (NLSY). Among the cohort of high school graduates between the ages 32 and 40 in 1997 with a scholastic ability test score in the top quartile, around 16 percent did not attend college.

Much work explaining the reluctance to attend college emphasizes the influence of family income and parental education on schooling choice.[1] In this paper, I aim to examine one alternative explanation by testing whether income uncertainty rises with educational attainment. Due to potential moral hazard problems and lack of collateral for financing investment in human capital, the options of reducing uncertainties used in financial markets, such as insurance and diversification, are often unavailable for reducing uncertainties about the payoffs to human capital investment. This makes it entirely possible that wage uncertainty will have an important effect on the decision to attend college.

A number of articles have attempted to analyze the impact of income uncertainty on schooling choice. Most of them assume full employment. Assuming saving and borrowing

[1]Kane (1994), Ellwood and Kane (2000) and others argue that borrowing constraints (or the short-run effect of family income) is an important factor contributing to the reluctance to attend college. On the other hand, several authors assert that college attendance or non-attendance is mainly explained by the long-run effect of family income. See Cameron and Heckman (1999, 2001), Cameron and Taber (2000), Carneiro and Heckman (2002), Heckman and Lochner (2000), Keane and Wolpin (2001), Shea (2000) and others. Recent studies by Acemoglu and Pischke (2000) suggest that family income has a significant effect on college enrollments.

are prohibited, Weiss (1972) shows that the coefficient of variation in earnings is a valid measure for income uncertainty only if the degree of risk aversion is moderate.[2] Allowing for saving/borrowing and a general form of risk attitudes, Levhari and Weiss (1974) approach this issue in a two-period model, in which the dynamic of income uncertainty is suppressed.[3] They find that income uncertainty discourages people from investing in human capital, provided that dispersion of earnings increases with the level of schooling. Johnson (1977) and Rosen (1986) provide valuable insights into the inverse relation between schooling and dispersion of outcomes among occupations in the context of equalizing wage differentials.

Pioneering contributions by Becker (1964, 1975, 1993) measure the extent of the additional income uncertainties caused by investing in college education using the variance of the return to schooling. However, as Becker himself realized (1975, pp. 184), the variance of the returns on college education can be attributed to both the variance of college earnings and the variance of high school earnings. If people are averse to risk, a higher variation in high school earnings should make the investment in a college education more attractive, instead of less. Consequently, the variance of the returns to schooling is not a useful estimator in measuring the extent of income uncertainties generated by additional education attainment (Eden 1980).[4]

The key component of the test developed in this paper involves the estimation of the *volatility differential* – the difference in the variance of the log wage for high school graduates who attended college versus those who did not, controlling for individual demographics, scholastic ability, parental education, and fields of major. Estimating the volatility differential is challenging because of pervasive self-selection problems. In the presence of uncertainty,

[2]The assumption specified by Weiss (1972) restricts the coefficient of relative risk aversion to range between zero and one. Hause (1974) points out that Weiss's conclusion is implied by the specifications of borrowing constraints and risk attitudes.

[3]Another paper by Olson, White and Shefrin (1979) emphasizes the dynamics of income uncertainty. They study the optimal investment in schooling in a multiple-period model. They specify an indirect utility function by assuming a particular loan/repayment plan.

[4]A recent work by Carneiro, Hansen, and Heckman (2003) estimate factor models and provide point estimates of the variances of high school and college earnings, and the variance of the difference in earnings.

an individual's educational attainment depends on his attitude toward risk, measured by the dispersion of possible future income, in addition to the return to schooling. Risk-averse agents who expect that wages will become more volatile with additional schooling are likely to be more hesitant to enroll in college. If volatility indeed rises with schooling, risk-averse agents will be less likely to invest in college. Then the observed volatility differential will be lower than if no self selection operated. Neglecting the self-selection effect of personal attitude toward risk, we often understate the gap in wage volatility. In what follows, such a downward bias in the volatility differential will be referred to as "risk-aversion bias." A simple schooling choice model will be developed to motivate an appropriate empirical strategy for selectivity correction.

The volatility differential may result from lack of knowledge about (1) individual ability, (2) the quality of each university, and (3) unanticipated changes in labor market conditions. The first two sources of uncertainty cause a *permanent* shock on future earnings, while the third source causes a *transitory* shock.[5] While wage levels cannot be certain when the schooling decision is being made, an agent predicts the dispersion of the wage distribution based on private information, such as learning ability and taste for schooling. Due to the presence of private information, the level of the permanent volatility perceived by the agent can be smaller than that estimated by econometricians. In other words, part of the permanent volatility estimated in this paper can be predictable even though actual wage levels remain uncertain to individuals. I separately estimate the permanent and transitory volatilities using a two-stage fixed-effects model, which builds upon the prior work by Becker (1964), Heckman (1979), Mincer (1974), and Willis and Rosen (1979). By applying this estimation method, the risk-aversion bias mentioned earlier can be corrected for as well.

Mincer (1974, pp.103-105) provides one of the first analyses of the patterns of income volatility for various levels of schooling. He estimates the combined variance of log annual

[5] To simplify analysis, this paper focuses on the identification of permanent and transitory components of wage volatility, and ignores the case of unemployment. Becker (1964) and Mincer (1991) have documented that attending college considerably lowers the degree of unemployment risk.

earnings using a cross-sectional data base. He finds that in the later stage of work life, there is a markedly positive gap in variance between college and high school graduates. As he implicitly suggests, this gap can be attributed to a substantial rise in ability bias with educational attainment. Olson, White, and Shefrin (1979) additionally control for IQ scores, and estimate a fixed-effects model. They stress the effects of transitory volatility on schooling decisions by assuming that permanent volatility is independent of schooling. In this paper, I control for scholastic ability and parental education, and allow both permanent and transitory volatilities to vary with levels of schooling.

Recent work by Belzil and Hansen (2002), Buchinsky and Leslie (2000), and Palacios-Huerta (2003) studies dynamic optimization models of human capital investment. A common result is that the future wage's dispersion has a significant impact on schooling decisions. Assuming that the degree of risk aversion is homogeneous, Belzil et al. (2002) suggest that a rise in risk aversion increases educational attainment, while Buchinsky et al. (2000) suggest otherwise. Palacios-Huerta (2003) asserts that the variance of the stochastic component of human capital returns is substantial and estimation results are insensitive to the adjustment of ability bias.

This paper adds to the recent literature in several ways. First, I develop a simple schooling choice model that demonstrates the problems of self selection resulting from heterogeneity in the degree of risk aversion and scholastic ability. I utilize an instrument that helps correct for the selection biases. I find that the level of permanent volatility differential, particularly for a four-year college education, is sensitive to the adjustment of risk-aversion bias in addition to ability bias. Parental education, measured ability index, and fields of major have already been controlled for in my analysis. Finally, I separately identify the permanent and transitory components of volatility and the volatility differential.

The analysis makes use of the National Longitudinal Survey for Youth: 1979-1998. As an instrument for schooling, the cost of attending a private four-year college nearby is constructed and assumed to be correlated with schooling but independent of log wage.[6] Results show that based upon the identification procedure, the estimate of the permanent volatility differential increases by approximately 30 percent for average four-year college attendees, suggesting problems of self selection may not be negligible in the estimation of wage uncertainty.

My major finding is that the permanent component of volatility differential is significantly positive for a four-year college education, indicating the wage profile of average four-year college attendees is more volatile than that of average high school graduates. This finding suggests that income uncertainty may have a deterrent effect on the decision to enroll in college particularly for an agent who is highly averse to risk and has a marginal level of the return to schooling.

The next section lays out schooling choice models and empirical specifications. Section 3 overviews the data, Section 4 summarizes the results, and Section 5 concludes.

2 The Schooling Choice Model

Consider a set of high school graduates who decide whether to attend college in the beginning of period zero. They make the decision based upon the information about their abilities, the quality of each university, and the distribution of possible future incomes. However, such information is often limited when individuals have only recently graduated from high school. Owing to lack of information, high school graduates are often uncertain about the number of years of postsecondary education they will take in the future. An individual may decide to enroll in college without knowing whether he will drop out in the middle of his undergraduate years or attend graduate school later. To derive a conservative measure of wage volatility, I control for years of schooling and focus on the endogeneity of the decision to

[6] The independence condition required here for the purpose of identification should be stronger than the mean independence; a full-independence condition is needed.

attend college, $s = 1$ or 0, in the presence of income uncertainties. Since standard problems of moral hazard would arise if individuals attempted to insure against the uncertainties of the payoffs to education attainment, the degree to which an individual is risk averse has an impact on the decision to attend college.

To test whether adding two or four years of schooling to high school education makes the income stream more or less volatile, first, I control for individual heterogeneity, including scholastic ability, family background, and fields of major. Then I separately identify the permanent and transitory components of wage volatilities for college versus high school careers, and estimate the volatility differential. This section presents a simple model to motivate this test. In the first subsection a schooling choice model is introduced to characterize the problems of self selection and unobservable heterogeneity. In the second subsection, I present an identification method of volatilities and volatility differentials based on the schooling choice model. Before introducing the method, it is useful to clarify the relation between wages, schooling, and uncertainties.

Suppose that there is neither unemployment nor a borrowing constraint. Individuals can freely borrow or lend at a constant interest rate r. An individual's *hourly wage rate* $y_{it}(s)$ is determined by the level of schooling s, individual characteristics, work experience, and unanticipated shocks. If person i decides not to attend college, he earns $y_{it}(0)$ for $t = 0, 1, 2, ..., T$. If student i decides to attend college, he pays tuition fees and stays at school full-time at $t = 0$. Upon graduating from college, he earns $y_{it}(1)$ for $t = 1, 2, ..., T$.

To identify wage volatility, I control for observed differences in individuals using Mincer's (1974) wage regression model. Given a schooling choice $s = 0$ or 1, consider a log wage equation for person i who finishes education ($t = s, s + 1, ..., T$):

$$\ln y_{it}(s) = \alpha_s n_i + x_{it}'\beta_s + z_i'\gamma_s + \sigma_a(s, z_i)\, a_{is} + \sigma_\varepsilon(s, x_{it})\, \varepsilon_{it}, \tag{1}$$

where n_i is the number of years of postsecondary education, and x_{it} is a vector of characteristics, containing the number of years of work experience, experience squared, and other time-varying variables; z_i is a vector of time-invariant explanatory variables, including race, gender, scholastic ability, parental education, marital status, cohort effects, regional dummies, and a constant.

The last two terms in the regression capture income uncertainties: The time-invariant component $\sigma_a(s, z_i) a_{is}$ represents the *permanent shock*, and the time-variant component $\sigma_\varepsilon(s, x_{it}) \varepsilon_{it}$ represents the *transitory shock* in period t. Denote a_{is} and ε_{it} as a pair of standard normal disturbances, assumed to be independent of each other and across individuals. ε_{it} is assumed to be an i.i.d. random shock. Here, the model differs from traditional selection models in that I allow the variance of log wages to depend on schooling choice by incorporating the scale functions $\sigma_a(s, z_i)$ and $\sigma_\varepsilon(s, x_{it})$. In particular, permanent volatility $\sigma_a^2(s, z_i)$ depends on schooling and individual demographics, while transitory volatility $\sigma_\varepsilon^2(s, x_{it})$ depends on schooling and work experience.[7]

Individuals have rational expectations about their future wages. Distributional parameters of shocks are common knowledge as schooling decisions are made. The impact of permanent volatility on schooling is often considered more important than that of transitory volatility, since, heuristically, permanent shocks persists for one's lifetime while transitory shocks can be "averaged out" over years. Permanent Income Hypothesis is an extreme example, in which effects of transitory shocks are almost negligible. Although this concept is intuitive, an analytical solution to a stochastic dynamic model of consumption only exists under very restrictive conditions.[8] This makes it difficult to identify the relative impact of permanent versus transitory income uncertainties on one's lifetime utility. With this complication, my empirical strategy is to separately estimate permanent and transitory components of volatilities. By doing this, I can test the statistical significance of permanent volatility differential and transitory volatility differential.

If there were no selection bias, permanent and transitory volatilities could have been

[7]Measurement errors in regressors are assumed to be equally dispersed between different levels of schooling. The effect of measurement errors can be cancelled out in estimating volatility differentials.

[8]See Blundell and Stoker (1999).

identified by a fixed-effects model. However, problems of self selection and unobserved heterogeneity make the identification challenging. In the following subsection, I demonstrate a schooling choice model that clarifies the nature of the problems and further present an empirical procedure in subsection 2.2.

2.1 Schooling Choice Rule

Three important features in my model are worth noting. First, the *cost of attendance* is formulated in a manner similar to Cameron and Heckman's (1998) model:

$$\mu_{i0}(1) = -\tau_i e^{-\eta_i},$$

where τ_i denotes tuition cost, and η_i represents a composition of *unobservable individual factors*, such as learning ability, parental inspiration, and taste for schooling, all of which cannot be observed by econometricians. Suppose that η_i is uncorrelated with transitory shocks but correlated with permanent shocks, $E[\eta_i a_{is}] = \rho_s$ for $s = 0$ or 1.

Second, log earnings are decomposed into two components — non-stochastic and stochastic — in an additively separable form, as defined in Mincer's equation (1). To disentangle income uncertainty from non-stochastic individual differences, it is useful to introduce another notation. Define $\mu_{it}(s)$ as person i's *non-stochastic income stream* in period t for a given level of schooling,

$$\mu_{it}(s) = \exp[\alpha_s n_i + x_{it}'\beta_s + z_i'\gamma_s],$$

as $t = s, s+1, ..., T$. As an example, in the absence of uncertainty, $y_{it}(s) = \mu_{it}(s)$ and foregone earnings equals $\mu_{i0}(0)$. Finally, this model highlights the importance of permanent shocks by assuming that an individual's expected lifetime utility equals the expected value of a function of lifetime income, exhibiting constant relative risk aversion (see Appendix).

In the absence of uncertainty, a commonly-used approach to describe the schooling decision begins with Becker's (1964) model, in which an individual's schooling decision entirely depends on his lifetime income. Precisely,

$$s_i = I\left\{-\tau_i e^{-\eta_i} + \sum_{t=1}^{T} R^t \mu_{it}(1) > \sum_{t=0}^{T} R^t \mu_{it}(0)\right\}, \tag{2}$$

where R denotes the discount factor and $I\{\cdot\}$ represents an indicator function. Notably, both direct cost $\tau_i e^{-\eta_i}$ and foregone earnings $\mu_{i0}(0)$ have been incorporated in Becker's model.

In the presence of uncertainty, an agent who is averse to risk will discount his expected lifetime income. Following Levhari and Weiss's (1974) contribution, I focus on the effect of permanent shocks on one's expected lifetime utility by suppressing the complication of income dynamics. An appendix shows that if the lifetime utility function exhibits constant relative risk aversion, transitory shocks can be smoothed out over one's lifetime, and log wages are normally distributed, then an individual will invest in a college education when

$$\begin{aligned} &\ln\left[-\tau_i e^{-\eta_i} + \sum_{t=1}^{T} R^t \mu_{it}(1)\right] - \ln\left[\sum_{t=0}^{T} R^t \mu_{it}(0)\right] \\ &> \frac{\kappa_i - 1}{2}\left[Var\left(\sigma_a(1, z_i)a_{1i}|\eta_i\right) - Var\left(\sigma_a(0, z_i)a_{0i}|\eta_i\right)\right], \end{aligned} \tag{3}$$

where κ_i is the coefficient of relative risk aversion, and $Var(.|\eta_i)$ is the variance of permanent shocks conditional on unobserved individual factor η_i as well as all the observed attributes (omitted for convenience of demonstration). When the utility function is logarithmic ($\kappa_i = 1$), I show in the appendix that the schooling choice rule approximately degenerates into Becker's model. Here I focus on the case where the coefficient of relative risk aversion is greater than unity, $\kappa_i > 1$.

The right-hand-side of the above inequality summarizes the cost of uncertainties resulting from permanent shocks, which is proportional to the volatility differential and the degree of risk aversion. Notably, the volatility for a given level of schooling is conditional on an information set that is available to the agent while the schooling decision is made. The information set contains the composition of unmeasured individual factors η_i as well as observed individual characteristics (z_i, x_i). Because the unmeasured individual factor is correlated with permanent shocks, the level of conditional volatility is proportionally smaller than the unconditional volatility. Precisely,

$$Var\left(\sigma_a\left(s,z_i\right)a_{is}\left|\eta_i\right.\right)=\left(1-\rho_s^2\right)\sigma_a^2\left(s,z_i\right),$$

where ρ_s is the correlation coefficient between permanent shock a_{is} and individual factor η_i for a given level of schooling. The absolute value of the correlation coefficient can be viewed as a measure of the permanent shock's *predictability.* For instance, if permanent shocks and individual factors are perfectly correlated, people can anticipate the level of the permanent shock by observing η_i. As such, the schooling choice rule degenerates into Becker's model.

Selection bias arises in estimating the wage volatilities, since the wage distribution, including both the mean and the variance, depends upon the schooling decision. In addition to "ability bias" which has been intensively discussed in the literature, "risk-aversion bias" occurs because an agent who is averse to risk tends to choose a level of schooling with a less volatile wage profile. If the college wage profile is more (less) volatile than the high-school wage profile, the estimate of the volatility differential will be biased downward (upward). The treatment of risk-aversion bias requires an empirical model that reflects the decision rule (3) of schooling, where the effects of uncertainty and risk aversion are both taken into account.

Although the decision rule (3) intuitively characterizes the nature of self selection, it is not empirically useful because the data of one's lifetime wage stream is usually unavailable. To proceed, extending Willis and Rosen's (1979) model, I substitute Mincer's regression (1) into the schooling choice rule (3) to derive a reduced from of the selection equation, which facilitates a treatment for both risk-aversion bias and ability bias :[9]

$$s_i = I\left\{\eta_i - \xi_i > \theta_1 \ln \tau_i + z_i'\theta_2\right\}, \tag{4}$$

where θ_1 and θ_2 are parameters, τ_i is the cost of attendance, and z_i contains a constant and person i's time-invariant characteristics. η_i represents a composition of personal learning ability and taste for schooling, and ξ_i depends on person i's attitude toward risk and information about the distribution of log wages.

The selection equation contains an unobservable individual effect ξ_i, which cannot be identified by a conventional discrete-choice model . Even if it could be identified, selectivity corrections would involve cumbersome computation that requires multinomial integration for each sample point (Aigner, Lovell, and Schmidt 1977; Meeusen and van den Broeck 1977). To proceed, I follow Chamberlain's (1984) identification method, in which the individual effect is decomposed into two additive components: One is a linear projection of demographics and individual average of time-varying variables, $z_i'\theta_z + x_i'\theta_x$, and the other is an exogenous disturbance v_i, following a normal distribution with zero mean and variance σ_v^2. As such, the selection equation becomes $s_i = I\left\{\eta_i - v_i > \theta_1 \ln \tau_i + z_i'(\theta_2 - \theta_z) - x_i'\theta_x\right\}$, where the coefficients remain unidentified up to scale, but selectivity adjustment terms can still be derived in a conventional manner. In what follows, η_i^* is defined as a standard normal, denoting the normalized unobserved individual factor $(\eta_i - v_i)/\sqrt{1+\sigma_v^2}$, and θ_i^* is defined as a *single index*, denoting the normalized regressors $\left[\theta_1 \ln \tau_i + z_i'(\theta_2 - \theta_z) - x_i'\theta_x\right]/\sqrt{1+\sigma_v^2}$.

Cost of attendance ($\ln \tau_i$) is defined by the total cost of attending a public four-year

[9] Let p_i denote the right-hand-side of the schooling choice rule (3). The choice rule can be rewritten as

$$\begin{aligned} s_i &= I\left\{\left(-\tau_i e^{-\eta_i} + \sum\nolimits_{t=1}^{T} R^t \mu_{it}(1)\right)\left(\sum\nolimits_{t=0}^{T} R^t \mu_{it}(0)\right)^{-1} > e^{p_i}\right\} \\ &= I\left\{-\tau_i e^{-\eta_i - z_i'\gamma} + e^{\alpha_1} A_{1i} > e^{p_i} A_{0i}\right\}, \end{aligned}$$

$A_{si} \equiv \sum_{t=s}^{T} R^t e^{x_{it}'\gamma_s}$. The second equality is appropriate if (1) the number of years of postsecondary education is zero for high school graduates (by construction, $\alpha_0 n_i = 0$ for $s_i = 0$), and (2) there is no heterogeneity in the effect of time-invariant individual characteristics, $\gamma_s = \gamma$ for $s = 0$ and 1. Then, rearranging the above equality implies equation (4), where $\xi_i = -\ln\left(A_{1i}e^{\alpha_1} - A_{0i}e^{p_i}\right)$.

college nearby in terms of local unskilled workers' average wage rate. We note that $\ln \tau_i$ appears in the selection equation but is excluded from in the log wage regression. $\ln \tau_i$ can be viewed as an instrument implied by theory, since the selection equation is derived from the schooling choice model. With the assumption of normality, although an instrument is not necessary for the purpose of identification, the inclusion of an instrument often enhances the precision of estimation. In addition, this instrument can be viewed as a refinement of college proximity. Card (1995b) and Rouse (1995) have instrumented for schooling using college proximity to estimate the return to education. As Card (1995b) and Kling (2001) pointed out, proximity to a four-year college nearby influences the decision to attend college, but proximity to a two-year college has almost no impact on educational attainment. Here, the identification condition is that cost of attendance is correlated with an individual's schooling choice but uncorrelated with future wage rates. A precise definition of this instrument can be seen in the appendix.

Selectivity adjustment can be fully specified if the joint distribution of η_i^* and a_{is} is well defined. The following subsection demonstrates an identification procedure based on the assumption of normality.

2.2 Identification

The goal of empirical analysis is to cope with selection bias and distinguish the permanent component of log wage volatility from the transitory component. I utilize a fixed-effects model to disentangle the shocks, and extend a Heckman's selection scheme to correct for selection bias. The key to fix the selection bias rests on the assumption that the selection bias is time-invariant and thus can be cancelled out in a fixed-effects model. There are three steps: First, using a fixed-effects model, the transitory component of volatility is identified. Second, the combined component of volatilities is identified by a between-effects model (see equation (6) for the definition) . Finally, the permanent component of variation is singled out by subtracting the transitory component from the combined component. In what follows, I start with the case where permanent and transitory shocks are homoscedastic within levels of schooling, and then present the identification procedure for the case of heteroscedasticity. Details about data sources and estimation results are detailed in Section 3 and Section 4.

2.2.1 Homoscedasticity

Suppose that permanent and transitory volatilities depend on schooling only, i.e. $\sigma_a^2(s)$ and $\sigma_\varepsilon^2(s)$. In the first step, Mincer's equation (1) implies that the fixed-effects model can be written as:

$$\ln y_{it} - \ln y_i = (x_{it} - x_i)' \beta_s + \sigma_\varepsilon(s)(\varepsilon_{it} - \varepsilon_i),$$

where $\ln y_i, \varepsilon_i$, and x_i denote the individual average of $\ln y_{it}, \varepsilon_{it}$, and x_{it} respectively. Since the selection bias is cancelled out in the fixed-effects model, a consistent estimator for the transitory volatility is easily derived:

$$\widehat{\sigma}_\varepsilon^2(s) = \frac{MSE_f(s)}{1 - \overline{T}^{-1}}, \tag{5}$$

where MSE_f is the mean squared error in the fixed-effects model, $\overline{T} \equiv N / \sum_i (1/T_i)$, and N is the number of respondents, and T_i is the number of observations for respondent i.

Second, consider a between-effects model, defined by averaging individual earnings over years. I illustrate the way to treat selection biases by using the subsample of college attendees ($s_i = 1$ or $\eta_i^* > \theta_i^*$) as follows. Given individual attributes and private information about taste for education and attitude toward risk, $q_i \equiv (x_{i1}, x_{i2}, \cdots, x_{iT_i}, z_i, \tau_i, \eta_i)$, the between-effects model can be written as:

$$\begin{aligned} E[\ln y_i | s_i = 1, q_i] &= \alpha_1 n_i + x_i'\beta_1 + z_i'\gamma_1 \\ &\quad + E[\sigma_a(1) a_{i1} + \sigma_\varepsilon(1) \varepsilon_i | \eta_i^* > \theta_i^*, q_i], \end{aligned} \tag{6}$$

where E is the expectation operator for a given joint normal distribution of the triple $(\eta_i^*, a_{i1}, \varepsilon_{it})$. The last term is the *selectivity correction* for the mean. Under the normality condition, the selectivity correction can be fully specified by taking the conditional expectation on combined error terms:

$$E\left[\sigma_a(1)\, a_{i1} + \sigma_\varepsilon(1)\, \varepsilon_i \,|\eta_i^* > \theta_i^*, q_i\right] = \beta_\lambda(1)\, \lambda_{1i}, \tag{7}$$

where λ_{1i} is often called the inverse Mills ratio, and $\beta_\lambda(1) \equiv \sigma_a(1)\, \rho_1$ is the coefficient of the inverse Mills ratio.[10] Similarly, under the normality condition, the variation in combined error terms is fully identified:[11]

$$Var\left[\sigma_a(1)\, a_{i1} + \sigma_\varepsilon(1)\, \varepsilon_i \,|\eta_i^* > \theta_i^*, q_i\right] = \sigma_a^2(1) + \sigma_\varepsilon^2(1)/\overline{T} - \beta_\lambda^2(1)\, \delta_{1i}, \tag{8}$$

where $\delta_{1i} \equiv \lambda_{1i}(\lambda_{1i} - \theta_i^*)$ is between zero and one.[12] The last term $\beta_\lambda^2(1)\, \delta_{1i}$ is the *selectivity correction* for the variance of the combined error terms. Because this selectivity correction for the variance is positive according to theory, ordinary least squares often *understate* the true variance of the combined residuals.[13] Notably, the mean squared error MSE_b of the between-effects model is a consistent estimator for the variance of the combined residuals, which is the left-hand-side of the above equality. This yields the identification of the true variance of the combined residuals:

$$\widehat{\sigma}_a^2(1) + \widehat{\sigma}_\varepsilon^2(1)/\overline{T} = MSE_b + \widehat{\beta}_\lambda^2(1)\, \overline{\delta}_1, \tag{9}$$

where $\overline{\delta}_1$ is the average of δ_{1i} for the college subsample.

Finally, the permanent volatility $\sigma_a^2(1)$ is identified by subtracting the transitory component $\widehat{\sigma}_\varepsilon^2(1)/\overline{T}$ from the unbiased combined component (i.e. the right-hand-side of the

[10]Let f and F denote the normal density and distribution respectively. The inverse Mills ratio is given by $\lambda_{0i} = -f(\theta_i^*)/F(\theta_i^*)$ for high school graduates and $\lambda_{1i} = f(\theta_i^*)/(1 - F(\theta_i^*))$ for college attendees.

[11]Identification of the volatility differential relies on the normality assumption in this paper. Chen and Khan (2001) show that without the normality assumption the volatility differential, scaled by the variance of high-school log wage rates, can still be identified.

[12]For proofs, see Theorem 20.2 and Theorem 20.5 in Greene (1997).

[13]If there only exists ability bias and no risk-aversion bias, ordinary least squares still understate the true variance, but without risk-aversion bias it would be hard to explain why the volatility differential is biased downward. By the nature of selection, it is difficult to separate the ability bias from the risk-aversion bias while estimating the variance.

above equation). Notably, the transitory volatility has been consistently estimated in the first step; only the permanent component of variation has a downward bias. In addition, correlation coefficient ρ_1 can be identified by $\beta_\lambda(1)/\sigma_a(1)$. The identification procedure described above also works for the subsample of high school graduates.

Consistent estimators for permanent and transitory volatility differentials are derived by taking respective differences in variances. Bootstrapping is used to generate confidence intervals. In bootstrapping, I randomly draw 1000 resamples of size N (equal to the number of respondents) from the original sample with replacement, which generates an empirical distribution of volatility differentials. A percentile method is utilized to determine endpoints of confidence intervals.[14]

2.2.2 Heteroscedasticity

In this subsection I allow the permanent volatility to depend on individual characteristics and allow the transitory volatility to depend on work experience; i.e. $\sigma_a^2(s, z_i)$ and $\sigma_\varepsilon^2(s, x_{it})$. In the first step, Mincer's equation (1) in a fixed-effects model can now be written as

$$\ln y_{it} - \ln y_i = (x_{it} - x_i)' \beta_s + \sigma_\varepsilon(s, x_{it}) \varepsilon_{it} - T_i^{-1} \sum_t \sigma_\varepsilon(s, x_{it}) \varepsilon_{it}, \tag{10}$$

Since the OLS estimator β_s is consistent, statistics based on OLS residuals have the same asymptotic properties as those based on the true disturbance as well. In particular, the variance of OLS residuals is

$$E\left[e_{its}^2 | s, x_{it}, x_i\right] = \left(1 - 2T_i^{-1}\right) \sigma_\varepsilon^2(s, x_{it}) + T_i^{-2} \sum_t \sigma_\varepsilon^2(s, x_{it}), \tag{11}$$

where e_{its} is the OLS residual. Summing over t for both sides of the equality and rearranging, one can express the last term by $T_i^{-1}(T_i - 1)^{-1} \sum_t E[e_{its}^2 | s, x_{it}, x_i]$. Substituting it back into the above equality implies a consistent estimator of transitory volatility:

$$\widehat{\sigma}_\varepsilon^2(s, t) = \frac{1}{N} \sum_i \left(1 - 2T_i^{-1}\right)^{-1} \left[e_{its}^2 - T_i^{-1}(T_i - 1)^{-1} \sum_t e_{its}^2\right], \tag{12}$$

[14] In the percentile method, confidence intervals of volatility differentials are derived by choosing the top 2.5 (5) percentile and the bottom 2.5 (5) percentile for a 95 (90) percent significance level.

for $T_i > 2$.

In the second step, consider a between-effects model. As before, I illustrate a way to treat selection biases by using the subsample of college attendees. Under the assumption of normality, selectivity correction for the mean is

$$E\left[\sigma_a\left(1,z_i\right)a_{i1}+T_i^{-1}\sum_t\sigma_\varepsilon(1,x_{it})\varepsilon_{it}\left|\eta_i^*>\theta_i^*,q_i\right.\right]=\beta_\lambda\left(1,z_i\right)\lambda_{1i}, \tag{13}$$

where λ_{1i} is the inverse Mills ratio and $\beta_\lambda\left(1,z_i\right)\equiv\sigma_a\left(1,z_i\right)\rho_1$ is the coefficient of the inverse Mills ratio. After incorporating the selectivity correction, the OLS residual, by construction, has the same asymptotic properties as the true disturbance. The variance of the OLS residuals equals

$$Var\left[\sigma_a\left(1,z_i\right)a_{i1}+T_i^{-1}\sum_t\sigma_\varepsilon(1,x_{it})\varepsilon_{it}\left|\eta_i^*>\theta_i^*,q_i\right.\right]=\sigma_a^2\left(1,z_i\right)+T_i^{-2}\sum_t\sigma_\varepsilon^2(1,x_{it})-\beta_\lambda^2\left(1,z_i\right)\delta_{1i}, \tag{14}$$

where $\delta_{1i}\equiv\lambda_{1i}\left(\lambda_{1i}-\theta_i^*\right)$ ranges between zero and one. The last term $\beta_\lambda^2\left(1,z_i\right)\delta_{1i}$ is a *selectivity correction* for the variance of least squares residuals. Both selectivity corrections can be estimated by a standard probit model.

From the above equality, I derive $\sigma_a^2\left(1,z_i\right)$ by subtracting the last two terms from the variance of the combined error terms. While the last two terms are identified in the previous estimation steps, the variance of the combined error terms is estimated based on Amemiya's (1985, pp. 203-207) heteroscedastic regression model. Following Amemiya, I assume the variance of the combined disturbance follows a parametric regression model, $z_i'b_z+x_i'b_x+\omega_i$, where $b=(b_z,b_x)$ is a vector of parameters, z_i includes a constant and individual attributes, x_i is a vector of individual averages of time-varying attributes, and ω_i is disturbance, uncorrelated with other regressors and across individuals. The least squares estimator of b, although inefficient, is still consistent. Consequently, the statistics based upon b is a consistent estimator of the combined disturbance's variance as well. Finally, substituting the estimate of $z_i'b_z+x_i'b_x$ into (14) yields a consistent estimator for permanent volatility:

$$\widehat{\sigma}_a^2\left(1,z_i\right)=z_i'\hat{b}_z+x_i'\hat{b}_x-T_i^{-2}\sum_t\widehat{\sigma}_\varepsilon^2(1,x_{it})+\widehat{\beta}_\lambda^2\left(1,z_i\right)\widehat{\delta}_{1i}.$$

Other functional forms for the heteroscedastic regression model, including $\exp(z_i' b_z + x_i' b_x)$ and $(z_i' b_z + x_i' b_x)^2$, are also examined. I find results are robust to various forms of heteroscedastic regressions. Finally, as discussed in the case of homoscedasticity, I adopt bootstrapping to derive confidence intervals.

3 Data and Variable Definitions

Statistics and estimations are based on *National Longitudinal Survey for Youth* (NLSY): *1979-98*, merged with restricted geocode data. The sample consists of 12,686 young respondents, who were between the ages of 14 and 22 in 1979. Observations are included if (1) respondents have a high school diploma or a general equivalency diploma; (2) respondents are from a representative cross-section sample and are not in the military; (3) nominal hourly wages are between 1 and 150. The first criterion excludes 3,086 respondents who did not graduate from high school. The second criterion additionally excludes 4,363 non-representative respondents to derive a reliable estimate using a random sample. The remaining sample contains 5,237 respondents, and each respondent has 12 to 17 years of observations. The third criterion excludes 1,939 observations, where 54 respondents are entirely dropped. Nominal variables are normalized by Gross Domestic Product Implicit Price Deflator for the base year 1992. According to these criterions, there are 5,183 respondents and 67,018 observations remaining in the sample.

The series on highest grade completed in NLSY shows a number of obvious inconsistencies. For example, several respondents indicate in questionnaires that they have attended a two-year or a four-year college but the number of years of schooling is no more than nine. To resolve this problem, besides manual correction and programming, I use both college attendance and the highest grade completed to identify high school graduates and college attendees. Specifically, a respondent is identified as a *high school graduate* if his highest grade completed equals 12. A respondent is identified as a *two-year college attendee* if he attended a two-year college but never attended a four-year college, *or* his highest grade completed is greater than 12 and no more than 15. Finally, a respondent is identified as a *four-year college attendee* if he has attended a four-year college, *or* his highest grade completed exceeds 15. Among 5,183 respondents, 3,511 individuals are identified as college attendees.

The scholastic ability score is measured by the *Armed Forces Qualifying Test* (Afqt) score, a composite score consisting of four tests: a vocabulary test, a mathematics test, a reading comprehension test, and an analytical test. Since the test was conducted in 1980 for all cohorts in the sample, the original Afqt scores are not comparable across different age groups. To resolve this problem, I generate a variable representing individual deviations from average Afqt scores of the corresponding cohort, and then categorize the variable by 100 quantiles.

NLSY provides longitudinal information about work history, starting from the year 1975 when respondents were of age 10-18. A precise measure of *work experience* is constructed accordingly. Work experience is derived from the cumulative number of annual working weeks, divided by the number of total weeks. Notably, many respondents had work experience before completing education. Those work years before returning to school are counted as part of work experience. However, observations in those years are excluded from the wage equation because earnings between school years are not determined by education received during later years.

Cost of attendance is used as an instrument in the selection equation, which is determined by both college proximity and direct cost of attendance. Precisely defined, cost of attendance is the total expense of attending a *local* in-state public four-year college while the respondent was 17 years old.[15] If several four-year public colleges are located within an individual's county of residence, the average in-state tuition of those public four-year colleges is used to

[15]I measure cost of attendance using the county of residence at age 14, instead of that at age 17, since the former has twice as many sample points as the latter. In fact, both measures are highly correlated; their coefficient of correlation equals .88 for the full-sample.

define local cost of attendance for this individual. If no four-year public college exists within an individual's county of residence, local cost of attendance for this individual is defined by sum of average in-state tuition, room, and board charged by public four-year colleges located in the state.[16] Both tuition and college proximity variables are generated from the *Higher Education General Information Survey* (HEGIS): 1974-82. In HEGIS, tuition data is the tuition fees paid by a typical full-time undergraduate to an accredited college in an academic year. Room and board data are the expenses actually charged by institutions at a seven-day weekly basis in an academic year.

We note that cost of attendance may be correlated with income. If so, the cost of attendance cannot be utilized as an instrument. To resolve this problem, I deflate cost of attendance in a county by the average hourly wage of unskilled workers in that county. To generate the deflator, I use average wage rate per job in service, agriculture, wholesale and retail trade industries in the county of residence while the respondent was 17 year old. The local wage rates of unskilled workers are constructed from the *Regional Economic Information System* (REIS)*:* 1974-82. Both REIS and HEGIS are merged with NLSY in accordance with the county of residence. Notice that years 1974-82 are chosen because the respondents, born during the years of 1956 to 1965, were 17 years old in those years.

Fields of major are categorized into nine divisions, including (1) engineering and architecture, (2) law, communication, and public affairs and services, (3) business and management, (4) natural sciences, (5) mathematics and computer sciences, (6) social sciences, (7) health profession, (8) education, language, library sciences, letters and theology, and (9) fine arts, general, others.

Parental education is measured by three variables: an individual's mother's highest grade

[16]Two states are exceptional. First, Washington D.C. has six public four-year colleges, but none of them provides data about room or board. I use the average in-state tuition charged by these six colleges to define the local cost of attendance for all counties in Washington D.C. Second, Wyoming state has no public four-year college. I use the average of the total expense on (out-of-state) tuition, room and board at all four-year universities, including public and private, located in the surrounding states of Wyoming.

completed, father's highest grade completed, and their interaction. *Family income* is defined by the average of per capita family income between the ages 16 and 17, where per capita family income is measured by dividing the total family income by the number of family members. In the initial cohort of the NLSY survey, since half of respondents' age is greater than 17 years old, family income in the half of sample is not well defined, and the sample size is reduced by half. To derive a more reliable measure for family income, I add in the NLSY supplemental sample[17] to increase my sample size *only* in estimating levels of family income (see Table 2). After doing this, I have a set of 5,295 respondents for whom there is no missing observation on those relevant variables. The total sample is categorized into three classes — *low*, *middle* and *high family income* — according to the values at the 33th and 66th percentiles of the family income distribution, where each class contains one-third of the total sample.

4 Results

I provide two sets of results. First, I report earnings and characteristics of college versus high school careers as a baseline comparison. Second, I present the results based on the estimation procedure described in Section 2.2.

4.1 Characteristics and Earnings of College Educated versus High School

In this subsection, I present the composition of the college and high school variables. Tables 2 and 3 compare characteristics of college attendees versus high school graduates. Columns 1 to 3 use the full sample (defined in the Data Section) and columns 4 and 5 use the sample in the Afqt top quartile. Item 1 shows that high school graduates have a slightly higher proportion of white and female students than college attendees do. The ratio of whites in

[17] The NLSY supplement sample is designed to oversample Hispanic, black, and economically disadvantaged non-black and non-Hispanic youths.

the top Afqt quartile is greater than that in the overall sample, whereas the ratio of females in the top Afqt quartile is smaller. Item 2 reports the number of years of schooling and cost of attendance. The most talented people who attended college have .8 more years of postsecondary education than the average college attendees. In addition, a postsecondary education seems to be *less* costly for those who attended college. The cost of attending a local public four-year college is used as an instrument to correct selection biases.

Item 3 summarizes family background variables. There is an obvious gap in both family income and parental education between those who attended college and those who did not. In particular, individuals who attended college have 16 to 26 percent more family income than those who did not. In addition, of those people who attended college, 47 percent had parents who attended college; of those who did not attend college, only 15 percent had parents who attended college. Table 3 summarizes the employment status of the sample. For all levels of education, people with academic talent tend to earn more and work longer than the average.

Table 4 reports preliminary statistics about the effect of college attendance on earnings for white males in two eight-year spans, between the ages 23 and 28 during 1982-89, and between the ages 31 and 36 during 1990-97. Similar to Becker's (1964, 1993) analysis, the statistics are derived *without* controlling for interpersonal differences. Item 1 shows that earnings increase with age and schooling. To measure the risk and the volatility differential, I first estimate the variance and the difference in variance between college attendees and high school graduates. I also distinguish the variation in log wage caused by permanent and transitory shocks. For a preliminary analysis, I use the variance of deviations from the individual eight-year average to measure the risk caused by transitory shocks, and use the variance of the actual individual eight-year average to measure the risk caused by permanent shocks. Item 2 shows that from 1980s to 1990s, the variations in earnings caused by transitory shocks decreased, while the variations caused by permanent shocks increased. Item 3 shows that the differences in the variance of the individual average, which represents the riskiness due to permanent shocks, increased dramatically from 1980s to 1990s. In the meantime, the differences in the variance of deviation from individual average, which denotes the riskiness due to transitory shocks, decreased by 30 percent.

Using the *1940-1950 Census of Population and Education*, Becker (1964, 1993) used the *coefficients of variation* in earnings to measure earnings uncertainty for those who attended college versus those who did not.[18] In the 1940 survey, he found that four-year college graduates exceed high school graduates in terms of the variation in earnings by the ratio of 1.32 to 1 for the sample during the ages 25-29, and the ratio decreases to 1.05:1 for the sample during the ages 30-34. In the 1950 survey, the ratios for those two age groups increased to 1.67:1 and 1.24:1 respectively. Using NLSY, I compute the coefficient of variation to be comparable to Becker's finding. The results are similar to Becker's to a certain extent: the coefficients of variation in earnings are often higher for individuals who are college-educated relative to those who are not. During the 1980s, four-year college attendees exceeded high school graduates in terms of the variation in individuals average earnings by the ratio of 1.15:1 when respondents were 23-28 years old. During the 1990s, the ratio decreases to 1.05:1 when respondents were 31-36 years old.

Without controlling for individual differences, however, coefficients of variation do not fully measure the earnings uncertainty since individual characteristics that are unrelated to earnings uncertainty are all included in the coefficients. The coefficient of college attendees may exceed that of high school graduates, because people who attended college are more diverse than those who did not, *or* because college attendees tend to face higher volatility than high school graduates do. The analysis below controls for individual characteristics to measure the extent of volatility. In the previous calculation, selection bias problems were ignored. After controlling for individual characteristics and correcting for selection biases, more accurate estimates are reported in the next subsection.

[18]The estimates can be found in Table 10 of Chapter 5 in Becker (1993).

4.2 Estimation Results

This subsection reports the major estimation results. Volatility and the volatility differential caused by permanent and transitory shocks are separately identified based on a series of regressions: The first is a fixed-effects regression that estimates transitory volatility and transitory volatility differential, and the second is a probit regression that generates selectivity adjustments for mean and variance. Third, I adopt a between-effects regression that uses selectivity adjustments to generate consistent estimates of permanent volatility, permanent volatility differential, as well as the return to schooling. These results are summarized in Tables 5-8.

Table 5 presents the results of the fixed-effects regression, in which variations in deviations of log wages from individual averages are considered. Since time-invariant variables are cancelled out in this regression, the explanatory variables are all time-variant. The explanatory variables include the number of years of work experience, experience squared, marital status, and regional dummies. The number of years of schooling is time-invariant in my model, because I focus on the variation of the post-schooling wage stream; observations of work experience before finishing college are excluded from the sample although the pre-graduation work experience is counted as part of experience. Rows (1b) and (1c) compare the experience-wage profiles among various levels of schooling, which indicate that the experience-wage profile of a college career is slightly steeper relative to that of a high-school career. In addition, Marital status and regional dummies have significant explanatory power in the fixed-effects model, as we see in Rows (d) and (e).

Table 6 reports the estimates of transitory volatilities and transitory volatility differentials. Part (I) presents the results under the assumption that no heterogeneity in transitory volatility appears within a given level of schooling. Here, the transitory volatilities range between 8.3 percent and 9.5 percent, slightly higher than Olson, White, and Shefrin's (1979) results, where the estimate of the transitory volatility is 7.3 percent for high school graduates and 7.5 percent for college graduates. Row (I-1) shows that the four-year college's hourly earnings is 1.2 percentage point more volatile than the high school's, while the two-year college's is 0.5 percentage point less volatile. Clearly, transitory volatility does not monotonically increase with educational attainment; some extent of nonlinearity can be seen in the relation between schooling and transitory volatility.

Part (II) reports the results of heteroscedasticity, allowing the transitory volatility to vary with work experience as well as schooling. Row (II-1) shows that the transitory volatility profile monotonically increases with work experience for a four-year college career, but appears to be U-shaped for the high-school and two-year college careers. Row (II-2) presents the estimates of transitory volatility differentials for college versus high school. The volatility differential of a four-year education is negative during the first three years of experience and becomes positive in later years. During the first three years of work experience, the wage profile of a high school career is more volatile than that of two-year or four-year college career. Later on, the four-year college's transitory volatility monotonically increases and exceeds the high school's; the volatility differential of a four-year college career is approximately one to three percent. This result is consistent with Mincer's (1974, pp. 103-105) finding, where experience profiles of variance of log earnings are U-shaped for high school graduates and positively incline in the later stage of college career. His results also indicate a small negative variance gap in the first few years of work experience and a large positive gap in the later stages.

Results of the probit regression are reported in Appendix. It is worth emphasizing that coefficients of the probit model cannot be identified up to scale due to the presence of the unobservable heterogeneity in attitude toward risk and taste for schooling (as illustrated in equation (4)). Hence, the estimates reported in the table do not imply the magnitude of marginal effects. Nevertheless, this table highlights the statistical significance of the instrumental variable — cost of attendance. As Row (a) shows, the marginal effect of the instrument on college enrollment is approximately -2 percent with a relatively small standard deviation, indicating that the instrument has a significantly negative impact on educational attainment. The identification rests on the assumption that attendance cost is uncorrelated with hourly earnings. This seems to be a plausible assumption since the instrument has already been deflated by the average wage of unskilled workers in the county of residence. Based upon the probit regression, I calculate the selectivity adjustments for the mean and the variance (λ_{si} and δ_{si}) respectively to prepare for the next step of estimation.

Table 7 reports the coefficients of a between-effects model, in which the interpersonal variation in average log wage is explained by time-invariant variables. Notably, the selectivity adjustment for the mean (or the "inverse Mills ratio"), constructed from the probit model, is additionally included in the model. In the discussion that follows and for the sake of brevity, I omit the results that rely on the assumption of homoscedasticity and report the estimation results that allow for heteroscedasticity; the coefficients under the assumption of homoscedasticity are very close to those reported here.

In Rows (1l)-(1n), I interact the selectivity adjustment with both gender and scholarly ability in order to allow volatility to be heterogeneous across individuals. The interaction between ability and selectivity adjustment appears to be significant to help explain the variation in average log wages of four-year college attendees, while the interaction between gender and selectivity adjustment appears to be important to help explain the variation in average log wages of two-year college attendees. Neither selectivity adjustment nor its cross terms have an important impact on the variation of average log wage among high school graduates.

The control variables in the probit are all included in the between-effects model, except for cost of attendance and high school graduation years (hgy). Cost of attendance is excluded, for it serves as an instrumental variable; hgy is also excluded, for it can be fully determined by a linear combination of cohort effects, the number of years of schooling, the number of years of experience.

As shown in Row (a) of Table 7, the coefficient of years of schooling approximately equals 6 percent. This estimate is slightly smaller than the one in the literature, because I control for scholastic ability, parental education, and selectivity adjustment. Interestingly, I find that a four-year college education seems to be more profitable than a two-year college education. The coefficient of years of schooling is 4.0 percent for a two-year college education, compared to a 6.5 percent for a four-year college education. These estimates demonstrates that the slope of years of schooling is steeper for a higher level of college education, suggesting the log wage regression is not linear in schooling. This finding is consistent with recent work by Belzil and Hansen (2002), in which they find evidence in favor of a convex log wage regression function. In addition, Grubb (1997) has documented that the effect of postsecondary education on annual earnings is increasing in the level of schooling, and the returns to associate degrees are considerably lower than the returns to baccalaureate degrees.

In Row (b), I control for eight fields of major in the college subsample. I find that fields of major are important explanatory variables for hourly earnings particularly for individuals who attended a four-year college, but not for those who attended a two-year college. As Table 10 in the appendix suggests, engineering and architecture majors have the highest positive impact on hourly earnings, and business and management majors have the second highest impact among the other majors. In Table 7, Rows (c) and (d) report that the effect of work experience on wage differences among individuals is not as significant as the effect of schooling. Rows (e) and (f) report the effects of scholastic ability and parental education on log wage. Ability seems to be particularly important to explain high school wage rates.

Row (g) of Table 7 shows that African-Americans who have attended a four-year college earn relatively lower wages than non-African-Americans who are in the same level of schooling. Specifically, among those who enrolled in a four-year college, an African-American's hourly wage is lower than a non-African-American's by approximately 10 percent, if other attributes are held constant. Meanwhile, the sign of racial discrimination seems not to be significant for other levels of schooling. Row (h) reports that females have a considerable disadvantage in wage rates relative to males for all levels of schooling. It is interesting to compare the estimates here with Table 9, where being a female African-American seems to have a significantly positive impact in the probit model. It appears that that racial and gender biases in college enrollment take an opposite direction compared to the biases observed in the labor market.

Table 8 reports the estimates of the permanent components of volatility and volatility differential. Rows (1a) and (1b) report the results neglecting selectivity adjustments, and Rows (2a)-(2b) report the estimates that selectivity biases are corrected for. As shown in Row (1a), without treating selection bias, the estimate of permanent volatility is similar for college versus high school; the estimate ranges between 9 percent and 12 percent for college attendees, and approximately equals 10 percent for high school graduates. Row (1b) shows that permanent shocks of four-year college attendees are more volatile than those of high-school graduates by 2.6 percentage points. In contrast, permanent shocks of two-year college attendees are less volatile than those of high-school graduates by 0.7 percentage points. Volatility differentials of both levels of college education are statistically significant at the five percentage level. In addition, I find that the permanent volatility differential increases with Afqt scores, suggesting that individuals who have higher scholastic ability are more likely to confront more wage uncertainties relatively to those with lower ability. I also find that in terms of the permanent volatility differential, wage uncertainty for females tends to be slightly higher than that for males.

After correcting for both ability and risk-aversion biases, Row (2a) shows a modest increase in the permanent volatility for all levels of schooling. The central premise of the selectivity adjustments is that *if there were no ability bias*, risk-aversion bias would understate the college volatility and overstate the high-school volatility. Simultaneously, both ability bias and risk-aversion bias leads to an incidental truncation that understates both college and high-school volatilities. Given the composition effect of both selectivity corrections, it can be expected that the selection adjustment for high school volatility should be smaller than that for college volatility. Estimation results are consistent with this premise. I find that, of high-school graduates, permanent volatility slightly increases after selectivity adjustments because the adjustment for ability bias overtakes the adjustment for risk-aversion bias. In other words, the presence of risk-aversion bias reduces the downward bias of ability for high-school graduates. On the other hand, of college attendees (especially four-year college attendees), the selectivity corrections for both ability and risk-aversion biases shift permanent volatility upward. It turns out that college attendees (especially for four-year college attendees) have a larger selectivity adjustment for the permanent volatility relative to high school graduates. As Row (2b) shows, the selectivity adjustments increase permanent volatility differential to 3.4 percent from 2.6 percent for four-year college attendees, and increase to almost zero percent for two-year college attendees. Consequently, four-year college attendees have the permanent volatility differential corrected by about 30 percent, while two-year college attendees have almost no adjustment. Notably, the selectivity adjustments considerably widen the confidence intervals of permanent volatility differentials.

A major result is that the unbiased estimate of the permanent volatility differential is significantly positive for investing in a four-year college, but almost zero for investing in a two-year college, indicating that the investment in a four-year college involves more uncertainties about future wages relative to the investment in a two-year college. The gap in the degree of uncertainty between two-year and four-year college education can be associated with the difference in the timing of the payoffs to the educational investment. As Becker (1993, p.91-92) noted, "The long time required to collect the return on an investment in human capital reduces the knowledge available, for knowledge required is about the environment when the return is to be received, and the longer the average period between investment and return, the less such knowledge is available." Here, my empirical result clearly confirms his intuition.

The extent of the impact of the volatility differential on schooling decisions depends on levels of risk aversion. To understand the magnitude of the potential impact, it is useful to consider an example. Given that the annual return to a four-year college education equals

6.4 percent (as shown in Table 7), the lifetime return to a four-year college education is approximately 29.2 percent.[19] Suppose that agents cannot anticipate the future income stream, and the coefficient of relative risk aversion equals two – a relatively conservative number compared to the degree of risk aversion documented in the asset pricing literature. Then the schooling choice model (3) suggests that the cost of uncertainty can offset the lifetime return to schooling by 1.7 percentage points, which is one half of the permanent volatility differential. As such, the lifetime return to schooling declines from 29.2 percent to $29.2 - 1.7 = 27.5$ percent. The cost of uncertainty can be relatively substantial to those who are highly averse to risk. For instance, if the agent's coefficient of relative risk aversion ranges from three to five, the cost of uncertainty can offset the lifetime return to schooling by 3.4 percentage points to more than 6.8 percentage points. Consequently, for those people whose returns to schooling are at the threshold, the presence of uncertainty can alter their schooling decisions.

5 Conclusion

If future incomes are uncertain, risk-averse agents may discount the payoffs to the investment in human capital. The longer the time required to reap the payoffs, the more of the discount. Among other possible causes to the reluctance to invest in college education, the impact of income uncertainty has long been noted but has been studied less intensively. In this paper, I test whether the wage profile of the college career is more volatile than that of the high school career by estimating the volatility differential. I empirically demonstrate that the volatility differential can be understated because risk-averse agents tend to choose less volatile careers. The aim of this paper is to clarify the nature of the risk-aversion bias and provide an unbiased estimate of the volatility differential.

Using the National Longitude Survey for Youth: 1979-1998, I estimate a selection fixed-

[19] The lifetime return to a four-year college education is approximately $e^{.064\times 4} - 1 = .292$ assuming tuition is negligible.

effects model, where the variance of log wage, as well as the mean, depends on the decision to attend college. To treat risk-aversion bias and ability bias, the cost of attending a four-year college nearby is utilized as an instrument to help identify the volatility differential. Estimation results support the validity of this instrument — it exhibits a significant effect on the selection equation, and the selectivity adjustment based on this instrument considerably raises the estimate of volatility differential, particularly for the average four-year college attendees. After selectivity adjustments, I find that the permanent component of the volatility differential is approximately 3.4 percentage points for a four-year college education at the conventional significance level. A back-of-envelop calculation shows that with modest levels of risk aversion, the cost of income uncertainty may offset a portion of the lifetime return to education. In particular, this deterrent effect of income uncertainty can be substantial for individuals whose returns to schooling are at the threshold, suggesting that income uncertainty can be an important factor in making schooling decisions under certain circumstances.

It is worth noting that although this paper focuses on how wage uncertainty affects the investment in college education, unemployment is another important source of uncertainty associated with schooling decisions. According to Mincer's (1991) estimation, the probability of unemployment is 6.4 percent for high school graduates but only 3.5 percent for those who attended a four-year college. If unemployment risk is incorporated as part of income uncertainties, enrolling in college may acquire the benefit of reducing income uncertainty in addition to the return to schooling. A more thorough study on human capital investment under income uncertainty should be done in the future research.

References

[1] **Acemoglu, D. and Pischke, J.S.** "Changes in the Wage Structure, Family Income and Children's Education." *European Economic Review*, 2001, *45*(4-6), pp. 890-904.

[2] **Aigner, D.J., Lovell, C.A.K., and Schmidt, P.** "Formulation and Estimation of Stochastic Frontier Production Function Models." *Journal of Econometrics*, 1977, *6*(1), pp. 21-37.

[3] **Amemiya, T.** *Advanced Econometrics*, 1985. Cambridge, Massachusetts: Harvard University Press.

[4] **Becker, G.S.** *Human Capital.* 1964, 1975, and 1993. New York: Columbia University Press.

[5] **Belzil, C. and Hansen, J.** "Earnings Dispersion, Risk Aversion, and Education." IZA Discussion Paper No. 513, 2002.

[6] **Blundell, R. and Stoker, T.M.** "Consumption and the Timing of Income Risk." *European Economic Review*, 1999, 43(3), pp. 475-507.

[7] **Buchinsky, M. and Leslie, P.** "Educational Attainment and the Changing US Wage Structure: Dynamic Implications Without Rational Expectations." Working Paper, 2000.

[8] **Cameron, S. and Heckman, J.J.** "Life Cycle Schooling and Educational Selectivity: Models and Choice." *Journal of Political Economy*, 1998, 106(2), pp. 262-333.

[9] **Cameron, S. and Heckman, J.J.** "Should College Attendance Be Further Subsidized to Reduce Rising Wage Inequality? Does Family Income Foster Ability, or Is It an Important Cash Constraint Limiting College Attendance?" In *Financing College Tuition: Government Policies Social Priorities*, 1999, edited by Marvin Kosters. Washington, D. C.: AEI Press.

[10] **Cameron, S. and Heckman, J.J.** "The Dynamics of Education Attainment for Blacks, Whites and Hispanics." *Journal of Political Economy*, 2001, *109*(3), pp. 455-499.

[11] **Cameron, S. and Taber, C.** "Borrowing Constraints and Returns to Schooling." NBER Working Papers No. 7761, 2000.

[12] **Card, D.** "Earnings, Schooling, and Ability Revisited." In *Research in Labor Economics*, 1995a, edited by Solomon Polachek, Greenwich Connecticut: JAI Press, pp. 23-48.

[13] ——. "Using Geographic Variations in College Proximity to Estimate the Return to Schooling." In *Aspects of Labor Market Behavior: Essays in Honor of John Vanderkamp*, 1995b, edited by Louis N. Christofides, E. Kenneth Grant, and Robert Swidinsky, Toronto: Canada: University of Toronto Press, pp. 201-22.

[14] ——. "The Causal Effect of Education on Earnings." In Chapter 30, *Handbook of Labor Economics*, 1999, edited by O. Ashenfelter and D. Card, Vol. 3a. Handbooks in Economics, Vol. 5. Amsterdam; New York and Oxford: Elsevier Science, North-Holland, pp. 1801-63.

[15] **Carneiro, P., Hansen, K.T., and Heckman, J.J.** "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*, forthcoming.

[16] **Carneiro, P. and Heckman, J.J.** "The Evidence on Credit Constraints in Post-Secondary Schooling." *Economic Journal*, 2002, 112(482), pp. 705-34.

[17] **Chamberlain, G.** "Panel Data." In *Handbook of Econometrics*, 1984, 2nd ed., Z. Grilliches and M. Intriligator. Amsterdam: North-Holland.

[18] **Chen, S.H. and Khan, S.** "Nonparametric Estimation of Average Volatility Differentials in Selection Models with an Application to Returns to Schooling." working paper, University of Rochester, 2003.

[19] **Eden, B.** "Stochastic Dominance in Human Capital." *Journal of Political Economy*, 1980, 88(1), pp. 135-45.

[20] **Ellwood, D.T. and Kane, T.J.** "Who Is Getting a College Education? Family Background and the Growing Gaps in Education." In *Securing the Future: Investing in Children from Birth to College*, edited by Sheldon Danziger and Jane Waldfogel, 2000, New York: Ford Foundation Series on Asset Building.

[21] **Greene, W.H.** *Econometric Analysis.* 1997. Saddle River, NJ: Prentice Hall.

[22] **Grubb, W.N.** "The Returns to Education in the Sub-Baccalaureate Labor Market, 1984-1991." *Economics of Education Review*, 1997, *16*(3), pp. 231-245.

[23] **Hause, J.C.** "The Risk Element in Occupational and Educational Choices: Comment." *Journal of Political Economy*, 1974, *82*(4), pp. 803-7.

[24] **Heckman, J.J.** "Sample Selection Bias as a Specification Error." *Econometrica*, 1979, *47*(1), pp. 153-61.

[25] **Heckman, J.J. and Lochner, L.** "Rethinking Education and Training Policy: Understanding the Sources of Skill Formation in a Modern Economy." In *Securing the Future: Investing in Children From Birth to College*, edited by Sheldon Danziger and Jane Waldfogel, 2000, New York: Ford Foundation Series on Asset Building.

[26] **Johnson, W.** "Uncertainty and the Distribution of Earnings." In *The Distribution of Economic Well-being*, edited by Juster, F. Thomas, 1977, National Bureau of Economic Research.

[27] **Kane, T.J.** "College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education." *Journal of Political Economy*, 1994, *102*(5), pp. 878-911.

[28] **Keane, M.P. and Wolpin, K.I.** "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *International Economic Review*, 2001, 42(4), pp. 1051-1103.

[29] **Kling, J.** "Interpreting Instrumental Variable Estimates of the Returns to Schooling." *Journal of Business and Economic Statistics*, 2001, 19(3), pp. 358-64.

[30] **Levhari, D. and Weiss, Y.** "The Effect of Risk on the Investment in Human Capital." *American Economics Review*, 1974, 64(6), pp. 950-63.

[31] **Meeusen, W. and van den Broeck, J.** "Efficiency Estimation from Cobb-Douglas Producation Functions with Composed Error." *International Economic Review,* 1977, 18(2), pp. 435-44.

[32] **Mincer, J.** *Schooling, Experience and Earnings.* 1974, New York: Columbia University Press.

[33] ——. "Education and Unemployment." National Bureau of Economic Research (Cambridge, MA), Working Paper No. 3838, 1991.

[34] **Olson, L., White, H., and Shefrin, H.M.** "Optimal Investment in Schooling When Incomes Are Risky." *Journal of Political Economy*, 1979, 87(3), pp. 522-39.

[35] **Palacios-Huerta, I.** "An Empirical Analysis of the Properties of Human Capital Returns." forthcoming in the *Aemerican Economic Review.*

[36] **Rosen, S. H.** "The Theory of Equalizing Differential," in *Handbook of Labor Economics*, 1986, Volumn 1, ed by O. Ashenfelter and R. Layard. Elsevier Science Publishers, pp. 641-692.

[37] **Rouse, C.E.** "Democratization or Diversion? The Effect of Community Colleges on Educational Attainment." *Journal of Business and Economic Statistics,* 1995, *13*(2), pp. 217-24.

[38] **Shea, J.** "Does Parents' Money Matter?" *Journal of Public Economics*, 2000, *77*(2), pp. 155-84.

[39] **Weiss, Y.** "The Risk Element in Occupational and Educational Choices." *Journal of Political Economy*, 1972, *86*(6), pp. 1203-13.

[40] **Willis, R.J. and Rosen, S.** "Education and Self-Selection." *Journal of Political Economy*, 1979, *87*(5), pp. S7-36.

Table 1: College enrollment rates of high school graduates, age 32-40 in 1997, with AFQT in the top quartile, by family income and parental education.

	College enrollment and types of colleges			
	No enroll-ment	**2-year college**	**4-year college**	**Sample sizes**
1. All high school graduates	15.8	15.3	68.9	2265
2. By family income:				
Low	17.9	19.3	62.9	140
Middle	12.7	18.3	69.0	394
High	10.5	12.9	76.7	630
3. By parental education:				
High school or below	21.9	19.6	58.6	1537
College attendees	4.6	11.1	84.3	1519

Note: : (i) Afqt (Armed Force Qualification Test) scores are deflated by the corresponding cohort average. (ii) Disadvantage groups, represented by the supplemental sample, are included in Items 2-3. In Item 1, I use random sample. (iii) Family income is defined in the paper. Notice that family income variables are available only for those cohorts between the ages between 32-35 in 1997. (iv) Parental education is defined by the highest education level of parents.
Source: National Longitudinal Survey of Youth of 1979-98.

Table 2: Comparison of Baseline Characteristics of Individuals At Age 32-40 in 1997, by College Attendance and Scholastic Ability (AFQT) Scores. (Continued)

		Overall sample		**Afqt Top quartile**	
		High		**High**	
	Total	**School**	**College**	**School**	**College**
	(1)	(2)	(3)	(4)	(5)
1. Demographics					
Percent whites	87.6	85.8	88.5	96.7	96.7
	(33.0)	(34.9)	(32.0)	(18.0)	(17.8)
Percent blacks	10.3	11.3	9.7	1.6	2.2
	(30.3)	(31.6)	(29.7)	(12.8)	(14.5)
Percent males	48.9	50.9	47.9	56.0	53.4
	(50.0)	(50.0)	(50.0)	(49.8)	(49.9)
2. Education					
Years of schooling	13.9	12	14.8	12	15.6
	(2.2)	(0.0)	(2.2)	(0.0)	(2.1)
Average log cost of	6.2	6.4	6.1	6.5	6.2
attending local public	(1.0)	(1.0)	(1.0)	(.9)	(.9)
4-year college at age 17					
3. Family Background					
Average family income	11778.4	10064.2	12631.0	12025.6	13923.6
per person	(7519.4)	(6520.5)	(7833.9)	(6509.4)	(8041.5)
Percent a parent	36.1	14.6	46.6	21.7	58.3
attended college	(48.0)	(35.3)	(49.9)	(41.3)	(49.3)
Sample size	3216	1091	2125	182	1112

Note: (i) Standard deviations are in parentheses. (ii) Cost of attending local public four-year college is defined in the Data Section. (iii) Parents are recognized as have attended college if the years of schooling is equal to or greater than 13. (iv) Sample sizes of the family income variable are about half of the full-sample. The sample sizes listed in the bottom row are the number of respondents for all variables, except for family income.

Table 3: Comparison of Baseline Characteristics of Individuals At Age 32-40 in 1997, by College Attendance and Scholastic Ability (AFQT) Scores.

		Overall sample		**Top quartile**	
	Total	**Attended College**	**High School**	**Attended College**	**High School**
4. Employment status					
Average nominal	16.7	18.6	12.7	21.8	14.6
hourly wage	(13.4)	(14.5)	(9.6)	(16.3)	(10.6)
Average log real	2.5	2.6	2.2	2.7	2.4
hourly wage	(6.9)	(.7)	(.6)	(.7)	(.6)
Average annual	2116.8	2128.6	2091.4	2175.1	2132.2
working hours	(823.4)	(827.9)	(814.0)	(857.4)	(792.2)
Average working	14.8	14.9	14.5	15.4	16.1
experience, years	(4.0)	(3.8)	(4.4)	(3.4)	(3.8)
Sample size	3500	2353	1147	1224	226

Note: (v) Nominal hourly wage is calculated from annual wages and earnings divided by total working hours in the year. (vi) Real wage is derived by normalizing the nominal wage by the Gross Domestic Product Implicit Price Deflator for the base year 1992. (vii) Years of work experience is defined by the accumulated annual ratio of the number of working weeks to the number of total weeks since 1975. Note that in 1975 respondents were age 10 to 18. (viii) Standard errors are in parentheses.

Source: National Longitudinal Survey for Youth of 1979-1998.

Table 4: Mean and Variance of Log Wage: 1982-97, White Males.

	High School	College Attendees		
Log of Earnings:	**Graduates**	**Total**	**2-year**	**4-year**
1. Cross-sectional mean of individual average:				
1982-1989, age 23-28	2.201	2.370	2.337	2.383
	(.007)	(.005)	(.010)	(.006)
1990-1997, age 31-36	2.390	2.692	2.577	2.733
	(.008)	(.006)	(.011)	(.007)
2. Variance of				
a. individual average:				
1982-1989, age 23-28	.258	.247	.227	.253
1990-1997, age 31-36	.303	.390	.367	.387
b. deviations from individual average:				
1982-1989, age 23-28	.107	.112	.101	.116
1990-1997, age 31-36	.078	.083	.073	.086
3. Difference in variance of				
a. individual average:				
1982-1989, age 23-28	-	-.010**	-.031**	-.004**
(F-test; p-value)	-	(.000^{+})	(.000^{+})	(.000^{+})
1990-1997, age 31-36	-	.086^{+}	.064	.083
(F-test; p-value)	-	(.000^{+})	(.000^{+})	(.000^{+})
b. deviations from individual average:				
1982-1989, age 23-28	-	.005**	-.006**	.009**
(F-test; p-value)	-	(.000^{+})	(.000^{+})	(.000^{+})
1990-1997, age 31-36	-	.005**	-.005**	.008**
(F-test; p-value)	-	(.000^{+})	(.000^{+})	(.000^{+})
4. Number of respondents:				
1982-1989, age 23-28	661	1348	351	997
1990-1997, age 31-36	637	1244	322	922
5. Number of Observations				
1982-1989, age 23-28	2660	5529	1478	4051
1990-1997, age 31-36	1758	3544	924	2620

Note: (i) In the first item, standard errors are in parentheses. (ii) ** and * indicate the 5% and 10% significance levels respectively. (iii) Note that NLSY has no observations of wage in 1994 and 1996. (iv) '+' indicates the corresponding is close to zero but not equal to zero

Table 5: Fixed-Effects Wage Regressions.
Dependent Variable: $\ln y_{it} - \ln y_i$

Estimates and Tests	**High School Graduates**	**College Attendees**		
		Total	**2-year**	**4–year**
1. Log wage equation (fixed-effects):				
a. Years of schooling	-	-	-	-
b. Experience	.065**	.072**	.074**	.073**
	(.003)	(.003)	(.004)	(.003)
c. Experience squared	-.002**	-.002**	-.002**	-.002**
	(.000+)	(.000+)	(.000+)	(.000+)
d. Marital status	.038**	.079**	.065**	.083**
	(.013)	(.010)	(.020)	(.013)
e. Regional dummies	yes**	yes**	yes**	yes**
(F-test; p-value)	(.006)	(.000+)	(.000+)	(.000+)
2. R^2	.103	.155	.152	.158
6. Number of respondents	1551	2828	843	1985
7. Number of observations	16952	25098	8566	16532

Note: (i) Standard errors are in parenthesis. (ii) The ** and * indicate the 5 percent and 10 percent significance level respectively. (iii) "Reginal dummies" include urban, south, north east, and west. (iv) 'Marital Status' stands for a dummy variable denoting whether the respondence has ever married before.

Table 6: Transitory Volatilities and Volatility Differentials.

Estimates and Tests	**High School Graduates**	**College Attendees Total**	**2-year**	**4-year**
I. Assuming Homoscedasticity:				
1. Transitory Volatilities	.083	.094	.077	.095
2. Transitory Volatility Differentials	-	.011**	-.005**	.012**
(F-test; p-value)		(.000+)	(.000+)	(.000+)
II. Allowing for Heteroscedasticity:				
1. Transitory Volatility by Work Experience:				
Experience 0-3 years	.088	.080	.071	.080
Experience 3-6 years	.075	.083	.066	.086
Experience 6-9 years	.080	.089	.074	.088
Experience 9-12 years	.086	.100	.082	.095
Experience 12-15 years	.087	.106	.091	.106
Experience more than 15 years	.089	.112	.091	.122
2. Transitory Volatility Differential by Work Experience				
Experience 0-3 years	-	-.008**	-.017**	-.008**
(F-test; p-value)		(.000+)	(.000+)	(.001)
Experience 3-6 years	-	.008**	-.009**	.011**
(F-test; p-value)		(.000+)	(.000+)	(.000+)
Experience 6-9 years	-	.008**	-.007**	.008**
(F-test; p-value)		(.005)	(.000+)	(.000+)
Experience 9-12 years	-	.014**	-.004**	.009**
(F-test; p-value)		(.000+)	(.000+)	(.021)
Experience 12-15 years	-	.019**	.004*	.019**
(F-test; p-value)		(.000+)	(.060)	(.000+)
Experience more than 15 years	-	.022**	.002	.033**
(F-test; p-value)		(.000+)	(.256)	(.000+)

Note: (i) and (ii): same as Table 4. (iii) '-' indicates 'not available'.
(v) I obtain the volatility estimates using a function of ordinary least squares residuals. The results reported here are based on a regression of the absolute values of least squares residuals on experience, experience squared, Afqt, parental education, race, and gender. See Identification Section for more details.

Table 7: Between-Effects Models.
Dependent Variable: Individual Average of Log Wage ($\ln y_i$)

Estimates and Tests	**High School Graduates**	**College Attendees**		
		Total	**2-year**	**4–year**
1. Log wage equation (between-effects):				
a. Years of schooling	-	.061**	.040**	.065**
		(.005)	(.015)	(.006)
b. Field of major	-	yes**	yes	yes**
(F-test; p-value)		(.000^{+})	(.306)	(.000^{+})
c. Experience	.030	.034**	.020	.033*
	(.021)	(.015)	(.028)	(.019)
d. Experience squared	.002	.000^{+}	.002	-.000^{+}
	(.001)	(.001)	(.001)	(.001)
e. Scholastic Ability (Afqt)	.003**	-.001	.001	-.001
	(.001)	(.001)	(.002)	(.002)
f. Parental education	yes	yes	yes	yes
(F-test; p-value)	(.982)	(.487)	(.487)	(.783)
g. Black	-.008	-.049	.022	-.096*
	(.049)	(.045)	(.075)	(.057)
h. Male	.210**	.330**	.389**	.292**
	(.046)	(.041)	(.073)	(.052)
i. Current marital status	.073*	.075**	.033	.105**
	(.041)	(.032)	(.059)	(.038)
j. Regional dummies	yes**	yes**	yes**	yes**
(F-test; p-value)	(.000^{+})	(.000^{+})	(.000^{+})	(.000^{+})
k. Cohort effects	yes*	yes	yes	yes**
(F-test; p-value)	(.072)	(.295)	(.771)	(.019)
l. Inverse Mills ratio	.087	.003	-.063	.014
	(.092)	(.110)	(.193)	(.136)
m. Inverse Mills ratio*Afqt	-.001	.001*	.002	.002*
	(.001)	(.001)	(.001)	(.001)
n. Inverse Mills ratio*Male	.029	-.068**	-.133**	-.045
	(.048)	(.023)	(.052)	(.028)
2. R^2	.375	.332	.366	.364
3. Number of respondents	1210	2359	662	1697
4. Number of observations	13785	21227	6899	14328

Note: (i) and (ii): same as Table 4. (iii) A Hispanic dummy is included.

(iv) Parental education includes the mother's and father's years of education

Table 8: Average Permanent Volatility and Average Volatility Differentials.

Estimates and Test Statistics	**High School Graduates**	**College Attendees**		
		Total	**2-year**	**4–year**
1. Before Correcting for Selection Bias:				
a. Permanent Volatility	.099	.118	.091	.124
b. Permanent Volatility Differential	-	.019**	-.008**	.026**
(F-test; p-value)		(.000+)	(.000+)	(.000+)
2. After Correcting for Selection Bias				
a. Permanent Volatility	.102	.121	.095	.136
b. Permanent Volatility Differential	-	.020**	-.006	.034**
(Bootstrapping; confidence intervals)		(.002,.040)	(-.024,.029)	(.001,.079)

Note: (i) and (ii): same as Table 4. (iii) The confidence intervals of the volatility differentials in Row (b) are based on bootstrapping with 1000 replications. The size for each replication is the number of observations. (iv) To allow for heterogeneity in permanent volatility, I use general least squares to estimate the between-effects model. First, I regress the absolute values of the least squares residuals on Afqt, parental education, race, gender, and cohort effects. Second, the fitted values of variances are used to conduct a weighted-least-squares regression. Notably, gender, Afqt scores, and cohort effects are often important to explain the magnitude of permanent volatility differentials.

A The Schooling Choice Rule

The schooling choice rule (3) can be derived as follows. Consider a multiple-period model where an individual makes a schooling decision in the first period and obtains income in each period afterwards. His income is determined by Mincer's equation conditional on his characteristics and information about permanent and transitory shocks. He makes his schooling decision by calculating his expected lifetime utility. The calculation is based on three assumptions. First, the utility function exhibits constant relative risk aversion. Second, permanent shocks, transitory shocks, and unobservable individual effects are normally distributed. Permanent shocks are independent across individuals but are correlated with unobservable individual effects, whereas transitory shocks are identically and independently distributed over periods and across individuals. Third, the effect of transitory shocks on lifetime earnings is smoothed out over one's lifetime. I express the expect lifetime utility by a certainty equivalent, which leads to the schooling choice rule (3). The details follow.

Person i makes a schooling decision by comparing incomes streams between attending and not attending college. Suppose that there is neither unemployment nor a borrowing constraint. The schooling decision is based on individual characteristics, private information associated with schooling η_i, and distributional parameters of permanent shocks and transitory shocks. Assume that the private information is correlated with permanent shocks but uncorrelated with transitory shocks. Notably, since only person i can observe his own individual effect η_i, his perceived earnings stream is slightly different from what econometricians observe in Mincer's equation,

$$\begin{aligned}\ln y_{it}(s) = {} & \alpha_s n_i + x_{it}'\beta_s + z_i'\gamma_s \\ & + \sigma_a(s, z_i)\sqrt{1-\rho_s^2}a_{is} + \sigma_\varepsilon(s, x_{it})\varepsilon_{it}, \end{aligned} \tag{15}$$

for $t = s, s+1, ..., T$, where ρ_s is the correlation coefficient between a_{is} and η_i for a level of schooling, σ_a is the variance of permanent shock (unconditional on η_i), and σ_ε is the variance of transitory shocks. Conditional on individual factor η_i, the variance of permanent shocks is

$$Var\left[\sigma_a\left(s,z_i\right)a_i\left|\eta_i\right.\right]=\sigma_a^2\left(s,z_i\right)\left(1-\rho_s^2\right). \tag{16}$$

Note that (15) and (16) do not require the assumption of normality. Let μ_{it} denote the deterministic component of income, $\alpha_s n_i + x_{it}'\beta_s + z_i'\gamma_s$. Summing over his income stream y_{it}, person i's lifetime earnings can be written as:

$$y_i^p\left(s\right)\equiv\sum_{t=s}^{T}R^t\mu_{it}(s)\exp\left[\sigma_a\left(s,z_i\right)\sqrt{1-\rho_s^2}a_{is}+\sigma_\varepsilon\left(s,x_{it}\right)\varepsilon_{it}\right],$$

where R is a discount factor. Notably, $\mu_{i0}(0)$ indicates the *foregone earnings*, and $\mu_{i0}(1)$ represents the *cost of attendance*. It is convenient to define the non-stochastic component of lifetime earnings as

$$\mu_i^p\left(s\right)=\sum_{t=s}^{T}R^t\mu_{it}(s).$$

Assume that individual utility is a function of lifetime earnings, exhibiting constant relative risk aversion,

$$u\left(y_i^p\right)=\frac{\left(y_i^p\right)^{1-\kappa_i}}{1-\kappa_i},$$

where κ_i denotes the *Arrow-Pratt coefficient of relative risk aversion.*

If κ_i is not unity, the expected lifetime utility can be derived as follows,

$$\begin{aligned}E_0\left[u\left(y_i^p\left(s\right)\right)\right]&=E_0\left[\left(1-\kappa_i\right)^{-1}\left(y_i^p\left(s\right)\right)^{1-\kappa}\right]\\&\approx\left(1-\kappa_i\right)^{-1}\left(\mu_i^p\left(s\right)\right)^{1-\kappa_i}\exp\left[\left(1-\kappa_i\right)^2\left(1-\rho_s^2\right)\sigma_a^2\left(s,z_i\right)/2\right],\end{aligned}$$

where E_0 is the expectation operator based on the information in the beginning of period one. The second line is appropriate only if (i) the permanent shock a_{is} is normally distributed so that $E_0\left[e^{\sigma a_{is}}\right]=\exp\left[\sigma^2/2\right]$, and (ii) the expected lifetime utility can be approximated by the first-order Taylor expansion around $\varepsilon_{it}=E\left[\varepsilon_{it}\right]=0$. Then, the expected lifetime utility can be expressed by a *certainty equivalent*, defined by $E_0\left[u\left(y_i^p(s)\right)\right]=u\left(ce_i\left(s\right)\right)$, such that

$$ce_i\left(s\right)=\exp\left[\ln\mu_i^p\left(s\right)-\frac{\kappa_i-1}{2}\left(1-\rho_s^2\right)\sigma_a^2\left(s,z_i\right)\right].$$

Clearly, a comparison between expected lifetime utility is equivalent to a comparison between certainty equivalents. The schooling choice rule (3) then follows.

As an example, consider the case where κ_i equals unity (i.e. the utility function is logarithmic). Based on the first-order Taylor expansion around $\varepsilon_{it} = E[\varepsilon_{it}] = 0$, the expected lifetime utility can be approximated by the log of lifetime income. Precisely,

$$E_0[\ln y_i^p] \approx \ln \mu_i^p(s) + E[\sigma_a(s, z_i)\sqrt{1-\rho_s^2}a_{is}] = \ln \mu_i^p(s).$$

Table 9: Probit Models.
Dependent Variable: College enrollment (s_i)

Estimates and Tests	**Coefficients**	**Changes in Probability**
1. Independent variables:		
a. Log cost of attendance	-.063**	-.020**
	(.027)	(.008)
b. High school graduation year	yes**	yes**
(Chi square (2); p-value)	(.000+)	(.000+)
c. Afqt scores	.024**	.007**
	(.001)	(.000+)
d. Parental education	yes**	yes**
(Chi square (3); p-value)	(.000+)	(.000+)
e. Black	.502**	.156**
	(.087)	(.027)
f. Male	-.204**	-.063**
	(.049)	(.015)
g. Regional dummies at age 17	yes**	yes**
(Chi square (4); p-value)	(.437)	(.437)
h. Work experience	-.046**	-.014**
	.010	.003
i. Marriage status	-.457**	-.142**
	.073	(.022)
j. Regional dummies	yes**	yes**
	(.000+)	(.000+)
2. LR test, chi-squared	1312.20	
3. Number of respondents	4131	

Note: (i) and (ii): same as Table 4. (iii) A Hispanic dummy is included. (iv) Parental education includes the mother's and father's years of education and their cross term. (v) Cost of attendance has deflated by the average hourly wage of local unskilled worker (vi) Row (b) contains quadratic terms of high school graduation years. (vii) Regional dummies include whether one lived in an urban area, northeastern, west, and south at age 14. (viii) The changes in probability are evaluated at sample means of other variables.

Table 10: The Effects of the Field of Major in Undergraduate Education on Log Wage, Compared to the Fine Arts and General Fields.

Coefficients	**College Attendees**		
	Total	**2-year**	**4–year**
1. Engineering and architecture	.206**	.048	.280**
	(.032)	(.059)	(.038)
2. Law, communication, and public	.126**	-.033	.200**
affairs and services	(.025)	(.050)	(.030)
3. Business and management	.195**	.036	.281**
	(.034)	(.061)	(.042)
4. Natural sciences	.137**	.046	.169**
	(.032)	(.058)	(.040)
5. Mathematics and computer sciences	.090**	-.055	.146**
	(.034)	(.068)	(.040)
6. Social Sciences	.027	-.052	.069
	(.039)	(.090)	(.043)
7. Health profession	.009	-.063	.055
	(.040)	(.072)	(.048)
8. Education, language, library sciences,	.043	-.045	.067*
letters and theology	(.030	(.055)	(.037)

Note: Same as Table 4.