# Is Investing in College Education Risky?

Stacey H. Chen[1]
University of Rochester and State University of New York at Albany

October 8, 2001

[1]Department of Economics, State University of New York at Albany, 1400 Washington Avenue, Albany, NY 12222 (email: schen@albany.edu). This paper is based on chapter 1 of my dissertation at the University of Rochester. I am deeply indebted to my advisors, Mark Bils and Eric Hanushek, for their guidance. I am also grateful to Gordon Dahl, Shakeeb Khan, and Lance Lochner for their valuable suggestions. Thanks to Traci Bach, Boyan Jovanovic, and Daniel Sullivan for their helpful comments. Thanks to Eric Hanushek and Thomas Kane for kindly offering their databases. Thanks also to the Friends of the Library at the University of Rochester to sponsor the access to the Higher Education General Information Survey. This paper also benefits from the discussion with Samuel Danthine, Pohan Fong, Justin Fox, Sreya Kolay, Yoonsoo Lee, Toshihiko Mukoyama, Kharis Templeman, and Kazuhiro Yuki. Thanks to participants at the presentations at Academia Sinica, National Taiwan University, National University of Singapore, SUNY Albany, University of Pennsylvania, and University of Rochester. All errors are mine.

**Abstract**

Attending college seems to be a profitable and affordable investment in the US. Nevertheless, a number of academically talented young people still hesitate to attend college. This puzzle motivates this paper to test for whether college education is a risky investment. To measure the riskiness of college attendance, I estimate the risk differential in earnings between college attendees and high school graduates. This paper copes with selection bias problems and distinguishes permanent earnings risk from transitory earnings risks. Evidence indicates that investing in a four-year college education is indeed risky, suggesting that, under certain circumstances, the riskiness of college attendance is an important factor in the schooling choice (JEL D81, C25, I2. Key words: schooling, risk differential, risk premium, selection bias).

# 1 Introduction

While numerous studies have measured the return to schooling using different methodologies, a consensus has emerged that going to college is a profitable investment. Extensive surveys by Card (1995a, 1999) report that the estimated returns to schooling are approximately between 8 percent and 13 percent per school year, which compares well with the high rate of returns on stocks. Going to college seems to be not only profitable, but also 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 the statistics on college attendance according to National Longitudinal Survey of Youth. 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, about 16 percent did not attend college.

Much recent work explaining the reluctance to attend college emphasizes the importance of family income and parental education to the schooling decision.[1] The aim of this paper is to examine one alternative explanation by testing whether investing in a college education is risky. If there is uncertainty in returns to education, measuring the risk premium on schooling is as important as estimating the mean return. As future payoffs to college education are uninsurable, an individual's return to schooling is a random variable. Provided that individual returns are normally distributed, both the mean and variance will govern the schooling decision. To measure the riskiness of college attendance, this paper estimates the *average risk differential* — the difference in variance of log earnings between college attendees and high school graduates, conditional on predetermined individual differences. I find that the risk differential is statistically significant and the magnitude can be substantial, suggesting that the riskiness of college attendance is important under certain circumstances.

[1] Kane (1994), Ellwood and Kane (2000) and others argue that borrowing constraints (or the short-run effect of family income) may be a 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 (1998, 1999a, 1999b), Cameron and Taber (2000), Heckman and Lochner (2000), Keane and Wolpin (1999), Shea (1998) and others. Recent studies by Acemoglu and Pischke (2000) suggest that family income has a significant effect on college enrollments.

The riskiness of attending college may result from lack of knowledge about (1) individual ability, (2) the quality of each university, and (3) unanticipated changes in market conditions. The first two sources of risk cause a *permanent* shock on future earnings, while the third source causes a *transitory* shock. Mincer (1991) has documented that the risk of unemployment is lower for more educated people. To simplify the analysis, this paper concentrates attention on earnings risk and ignores the case of unemployment. Since no insurance market exists to avoid uncertainties of future earnings, the analysis focusing on the earnings risk is not trivial.

A number of early articles have attempted to measure the riskiness of schooling in a no-unemployment context. Assuming that individuals cannot borrow, Weiss (1972) shows that the coefficient of variation in earnings is a valid measure for the risk of schooling only if risk aversion is moderate.[2] Allowing for free borrowing and general risk attitudes, Levhari and Weiss (1974) approach this issue in a two-period model, in which the dynamics of income risk is suppressed.[3] They find that earnings uncertainty discourages investing in human capital if the dispersion of earnings increases with the level of investment in human capital.

Pioneering estimations by Becker (1963, 1993) and Weiss (1972) use the coefficient of variation in earnings to estimate the risk of schooling without controlling for individual differences. Estimations by Mincer (1974) and Olson, White and Shefrin (1979) control for individual heterogeneity and use the residual variance of an ordinary least squares estimation to measure earnings risk. In particular, Olson, White and Shefrin identify transitory shocks as the only source of earnings uncertainty and use a fixed-effects model to estimate the transitory earnings uncertainties.[4] Besides the problem of omitting permanent earnings risk, potential problems of *selection bias* are ignored as well. Selection biases arise if the distribution of individual returns to schooling correlates with the schooling decision. As a result, the ordinary least

[2]The imposed condition restricts the coefficient of relative risk aversion to be between zero and one.

[3]Another paper by Olson, White and Shefrin (1979) emphasizes the dynamics of income risk. 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]Recent literature on the equity premium puzzle considers transitory earnings uncertainty to be the only source of earnings risk. See Campbell (1996), Buckinsky and Leslie (1997), and Palacio-Huerta (1999). Little attention is paid to permanent earnings uncertainty.

squares estimation often *understates* the true variance, and understate or overstate the differences in variances. In the existing literature, the problem of selection bias has long been neglected in estimating the earnings risk for different levels of schooling.

In this paper I use Heckman's two-step scheme to correct for selection biases. To characterize the problem of selection bias, I develop a simple schooling choice model built on the work by Levhari and Weiss (1974) and Willis and Rosen (1979). Based on the schooling choice model, a measure for risk premium on schooling is defined, and the self-selection problem is characterized. The model implies that under certain circumstances the *permanent* risk differential is a major determinant of risk premium on schooling, while the transitory risk differential is not. The aim of my empirical studies is to disentangle these two risk differentials and treat the selectivity biases. After treating selection biases, estimates of permanent risk differentials dramatically increase by 77-163 percent. Using a fixed-effects model to disentangle permanent and transitory earnings uncertainties, I find that the permanent risk differential of a four-year college education is significantly positive, suggesting that a four-year college education is a risky investment. In addition, I find that the transitory risk differential of a four-year college education is significantly negative, implying that attending a four-year college may reduce transitory earnings uncertainties.

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 must decide whether or not to attend college in the beginning of period zero. They make this decision based upon the information about their abilities, the quality of each university, and the distribution of random shocks. However, this kind of information may be very limited when individuals have only recently graduated from high school. Owing to lack of information, high school graduates may be uncertain about the number of years of postsecondary education they will take in the future. An individual may decide to attend college without knowing whether he will drop out in the middle of his undergraduate years or attend graduate school later. My analysis focuses on the endogeneity of the decision to attend college, $s = 1$ or 0, given uncertainty in returns to schooling. Without an insurance market to secure the returns to schooling, such uncertainty is inevitable.

The uncertainty in the return to schooling can be measured by the variance of log earnings, conditional on individual characteristics. This section presents a simple model to estimate this measure. An individual's decision to attend college is essentially like choosing a distribution of future earnings. Both the mean and variance are involved in determining his schooling decision. Individuals attend college only if the lifetime benefit from a college education exceeds the opportunity cost, taking both the returns and the spread of the returns into account. Hence, the distribution of earnings observed by econometricians is incidentally truncated, but the true distribution of earnings is not. Ignoring this truncation problem, an ordinary least squares estimation understates the true variance of earnings for each level of schooling.

In the first subsection I use a schooling choice rule to characterize this truncation problem. In the second subsection I show how the truncation problem can be solved so that the true risk and risk differential of schooling can be identified. Before these two subsections, it is useful to clarify the link between schooling and earnings.

Suppose there is neither unemployment nor a borrowing constraint. Individuals can freely borrow or lend at a constant interest rate $r$. *Annual earnings* $y_{it}(s_i)$ is determined by the level of schooling $s_i$, individual characteristics, work experience $(t-s_i)$ at period $t$, and unanticipated shocks. If person $i$ decides not to attend college, he can earn $y_{it}(0)$ in period $t = 0, 1, ..., T$, where $T$ is the *total number of work years* during one's lifetime. 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 can earn $y_{it}(1)$ in period $t = 1, 2, ..., T$.

To identify the risk of schooling, I control for individual differences using Mincer's (1974) earnings function. Given a schooling choice $s = 0$ or 1, consider a log annual earnings equation for person $i$ who finishes his education $(t = s, s+1, ..., T)$:

$$\ln y_{it}(s) = \alpha_s n_i + x_{it}'\beta_{s1} + z_i'\beta_{s2} + \sigma_a(s)\, a_i + \sigma_\varepsilon(s)\, \varepsilon_{it}, \qquad (1)$$

where $n_i$ is the number of schooling years, $x_{it}$ is a vector of the number of years of work experience $(t - n_i)$ and experience squared, and $z_i$ is a vector of individual $i's$ characteristics. $z_i$ consists of $i$'s scholastic ability test score (Afqt), $i$'s parental education, whether $i$ lived with parents at age 14, $i$'s gender, race, marital status, and $i$'s cohort effects, and regional dummies.[5] The time-invariant individual effect $\sigma_a(s)\, a_i$ represents a *permanent shock,* while the transitory component $\sigma_\varepsilon(s)\, \varepsilon_{it}$ represents a *transitory shock*. Although both shocks are unobserved by an individual while his schooling choice is made, parameters about the distributions of both shocks are common knowledge. Suppose $a_i$ and $\varepsilon_{it}$ are standard normal, $N(0,1)$, and independent of each other and across individuals. The transitory shocks are assumed to be identically and independently distributed. Supposed also that measurement errors in regressors are equally dispersed between different levels of schooling. Hence, the effect of measurement errors is cancelled out in estimating risk differentials.

The *risk* for a given schooling choice $s$ is defined by the variance of permanent and transitory shocks, conditional on $n_i, x_{it}$ and $z_i$. Transitory shocks are considered less important relative to permanent shocks in measuring the risk because, heuristically, transitory shocks can be "averaged out" over one's lifetime, while permanent shocks persist for each period. The Permanent Income Hypothesis is an extreme example, in which the risk caused by transitory shocks is almost negligible. Although this concept is intuitively clear, an analytical solution for a stochastic dynamic model of consumption only exists under very restrictive conditions.[6] Consequently, it is difficult to determine the relative importance of both permanent and transitory earnings uncertainties. With this complication, my empirical strategy is to *separately* estimate the variances of permanent and transitory shocks. By doing this, I test whether the average risk differential is significantly positive for permanent shocks and for transitory shocks.[7]

If there is no selection bias, both variances can be easily estimated by a fixed-effects model. However, problems of selection bias arise because the distribution of earnings is dependent upon the schooling decision. To understand the nature of the problems, it is useful to consider how a risk-averse

[5]The series on marital status and the regional dummies actually vary over time. They are treated as time-invariant variables for convenience of demonstration.

[6]See Blundell and Stoker (1999).

[7]Heterogeneity in the risk and the risk differential can be also considered in this framework. For instance, the variance may differ in parental education; $p = 1$ or $0$ indicates whether or not parents attended college. Then, the variance can be written as $\sigma_a^2(s,p)$ and $\sigma_\varepsilon^2(s,p)$, and the hypothesis testing is conducted conditional on $p$.

individual decides to invest in a college education in the presence of uncertainties. The following subsection presents a simple schooling choice rule to facilitate my empirical analysis.

## 2.1 Schooling Choice Rule

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

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

where $\tau_i$ is tuition fees and $\eta_i$ is an *unobservable individual effect* which is unobservable to econometricians but is observed by individual $i$. $\eta_i$ indicates the effect of parental taste or family support on the decision to attend college. For exposition, I assume that the $\eta_i$ is standard normal although this is not necessary (see below). Note that permanent shocks may be correlated with $\eta_i$, $E[\eta_i a_i] = \rho$, while the transitory shocks are identically distributed and independent, particularly of $\eta_i$.

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 earnings uncertainty from non-stochastic individual differences, it is useful to define another notation. Define $\mu_{it}(s)$ as person $i$'s *non-stochastic earnings stream* in period $t$ for a given level of schooling $s$,

$$\mu_{it}(s) = \exp\left[\alpha_s n_i + x'_{it}\beta_{s1} + z'_i\beta_{s2}\right],$$

as $t = s, s+1, ..., T$. As an example, if there is *no* uncertainty, $y_{it}(s) = \mu_{it}(s)$. Third, the model emphasizes the importance of permanent shocks by assuming an individual's expected lifetime utility is a function of lifetime earnings, exhibiting constant relative risk aversion (see Appendix).

If there is *no* uncertainty, a commonly-used approach to describe the schooling decision begins with Becker's (1967) model. In Becker's model, an individual makes a schooling decision by calculating his lifetime earnings. The schooling choice rule satisfies:

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

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

If there is uncertainty, an individual who is risk-averse will discount his payoff to schooling. To approach the schooling choice problem, one may consider a stochastic dynamic model of consumption. Unfortunately, such a model is rarely solvable. Instead, it is useful to begin with a simple model with *no* transitory shocks. Built on Levhari and Weiss's (1974) two-period model, the model suppresses the dynamics of earnings risk and emphasizes the importance of permanent shocks.

An appendix shows that if the lifetime utility function exhibits constant relative risk aversion, individuals are averse to risk, transitory shocks can be smoothed out over a lifetime, and the distribution of earnings is log-normal, then the schooling choice rule satisfies: $s_i = 1$ only if

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

where $\gamma$ is the *coefficient of relative risk aversion* and $\rho$ is the correlation coefficient between permanent shocks $a_i$ and the individual effect $\eta_i$. If individuals are risk neutral (i.e. $\gamma = 1$) or permanent shocks can be fully anticipated by observing $\eta_i$ (i.e. $E(a\eta_i) = 0$), the choice rule degenerates to the case where there is no uncertainty. In this paper, the gap in permanent earnings risk $(\sigma_a^2(1) - \sigma_a^2(0))$ is referred to as the *average risk differential* caused by permanent shocks.

The right-hand-side of the above inequality, denoting the cost resulting from earnings uncertainty, defines the *risk premium* on schooling investment:

$$risk_\gamma \equiv \frac{\gamma-1}{2}\left(1-\rho^2\right)\left[\sigma_a^2(1) - \sigma_a^2(0)\right], \tag{4}$$

Earnings uncertainty can influence individuals' schooling decisions if schooling is risky and individuals are averse to risk. It could happen that schooling is beneficial to individuals' future earnings but is unattractive to those who are highly averse to risk because going to college is risky. The goal of my empirical analysis is to examine the significance of riskiness of college attendance. To do so, I estimate the average risk differentials caused by permanent shocks and transitory shocks, and treat potential selection biases.

The schooling choice rule motivates an *empirical selection equation,* which facilitates a treatment for the selection bias. The empirical selection equation can be derived by substituting Mincer's regression (1) into the schooling choice rule (3),[8]

$$s_i = I\left\{\eta_i > \theta_0 + \theta_1 \ln \tau_i + z_i'\theta_2\right\}, \tag{5}$$

where $\theta_k$ are parameters, $\tau_i$ is person $i$'s cost of attendance, $z_i$ is person $i$'s time-invariant characteristics, and $\eta_i$ represents the unobservable individual effect associated with parental taste or support for schooling. $\eta_i$ can be normalized to be a random variable with zero mean and unit variance, given parameters $\theta_0, \theta_1$ and $\theta_2$. I assume that $\eta_i$ follows a standard normal distribution. The regressors $z_i$ include the year in which person $i$ graduated from high school (hgy),[9] hgy squared, Afqt, parental education, gender, race, and regional dummies. The log of cost of attendance $\ln \tau_i$ is an instrumental variable implied by the schooling choice rule. I denote the *single index function* by $w_i'\theta = \theta_0 + \theta_1 \ln \tau_i + z_i'\theta_2$. Notice that the risk premium (including the risk differential as well as the correlation and preference parameters) is hidden in the constant term $\theta_0$ and cannot be singled out. Details on data sources and variable definitions can be seen in the Data section.

The empirical selection equation characterizes the selection problems in estimating risk differentials. In the case where individual effect $\eta_i$ is correlated with permanent earnings uncertainty $a_i$, the distribution of observed

[8]The schooling choice rule (3) can be rewritten as

$$\begin{aligned} s_i &= I\left\{\left(-\tau_i e^{-\eta_i} + \sum\nolimits_{t=1}^{T} R^t \mu_{it}(1)\right)\left(\sum\nolimits_{t=0}^{T} R^t \mu_{it}(0)\right)^{-1} > e^{risk_\gamma}\right\} \\ &= I\left\{-\tau_i e^{-\eta_i - z_i'\beta_2} + e^{\alpha} A_1 > e^{risk_\gamma} A_0\right\}, \end{aligned}$$

where $A_s \equiv \sum_{t=s}^{T} R^t e^{x_{it}'\beta_{s1}}$. The second equality is appropriate only if the following three conditions are satisfied: (1) the years of schooling are constant within both of the schooling groups (so that $\alpha_1 n_i = \alpha$ for $s_i = 1$, and $\alpha_0 n_i = 0$ for $s_i = 0$), (2) there is no heterogeneity for the effect of time-invariant individual characteristics, $\beta_{s2} = \beta_2$, and (3) there is no transition from work to school, such that years of work experience $x_{it} = t - s_i$ for all $i$, where $t$ refers to the number of periods from the high school graduation year to the current year. Hence, $A_s$ only depends on schooling level $s$ but does not vary across individuals. Then, it implies that

$$s_i = I\left\{\eta_i > \theta_0 + \theta_1 \ln \tau_i + z_i'\theta_2\right\},$$

where $\theta_0 = -\ln\left(A_1 e^{\alpha} - A_0 e^{risk_\gamma}\right), \theta_1 = 1$, and $\theta_2 = -\beta_2$.

[9]Notice that number of years of schooling and cohort effects *cannot* fully determine the high school graduation year because the number of years of schooling at college is not constant across individuals.

earnings will be more concentrated for a given level of schooling. Hence, the estimated variance of $a_i$ using OLS is often biased downward. The extent of this bias can be fully specified if the $\eta_i$ and $a_i$ are normally distributed. The next subsection presents a full parametric approach to identify earnings risk and risk differentials.

## 2.2 Identifying Average Risk Differentials

Two major tasks are necessary to identify both permanent earnings risk and transitory earnings risk, and, meanwhile, cope with potential selection biases. The first task is to utilize a fixed-effects model to disentangle permanent shocks from transitory shocks. The second is to use Heckman's scheme to correct selection bias, if any. More specifically, the transitory component of variations in log earnings can be identified by a fixed-effects model since time-invariant selection biases are entirely cancelled out. Meanwhile, the combined component of variations can be identified by a between-effects model, in which selection biases arise but can be corrected by Heckman's scheme. Finally, the permanent component of variation is singled out by subtracting the transitory component from the combined component.

The key to fix the selection bias is to exploit the fact that selection biases are time-invariant. The selection bias is simply subtracted out in the fixed-effects model. From Mincer's equation (1), the fixed-effects model can be written as

$$\ln y_{it} - \ln y_i = (x_{it} - x_i)' \alpha + \sigma_\varepsilon (s) (\varepsilon_{it} - \varepsilon_i), \tag{6}$$

where $\ln y_i, \varepsilon_i$, and $x_i$ denote the individual average of $\ln y_{it}, \varepsilon_{it}$, and $x_{it}$ respectively. A consistent estimator for the variance of transitory shocks can be derived:

$$\widehat{\sigma}_\varepsilon^2 (s) = \frac{MSE_f (s)}{1 - \overline{T}^{-1}}, \tag{7}$$

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 $\eta_i > \theta' w_i$) as follows. Given the information about individual attributes, $q_i \equiv (x_{i1}, x_{i2}, \cdots, x_{iT}, w_i, \eta_i)$, the between-effects model can be written as:

$$E\left[\ln y_i\left|\eta_i > w_i'\theta, q_i\right.\right] = \alpha_1 n_i + x_i'\beta_1 + z_i'\beta_2 + E\left[\sigma_a(1)\,a_i + \sigma_\varepsilon(1)\,\varepsilon_i\left|\eta_i > w_i'\theta, q_i\right.\right], \quad (8)$$

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

$$E\left[\sigma_a(1)\,a_i + \sigma_\varepsilon(1)\,\varepsilon_i\left|\eta_i > w_i'\theta, q_i\right.\right] = \beta_\lambda(1)\,\lambda_{1i},$$

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

$$Var\left[\sigma_a(1)\,a_i + \sigma_\varepsilon(1)\,\varepsilon_i\left|\eta_i > w_i'\theta, q_i\right.\right] = \sigma_a^2(1) + \sigma_\varepsilon^2(1)/\overline{T} - \beta_\lambda^2(1)\,\delta_{1i}, \quad (9)$$

where $\delta_{1i} \equiv \lambda_{1i}(\lambda_{1i} - w_i'\theta)$ is between zero and one.[12] The last term $\beta_\lambda^2(1)\,\delta_{1i}$ is a *selectivity correction* for the variance of the combined error terms. Because this selectivity correction is positive according to theory, the ordinary least squares estimation often *understates* the true variance of the combined residuals. 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:

$$\widehat{\sigma}_a^2(1) + \widehat{\sigma}_\varepsilon^2(1)/\overline{T} = MSE_b + \widehat{\beta}_\lambda^2(1)\,\overline{\delta}_1, \quad (10)$$

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

Finally, the variance $\sigma_a^2(1)$ of permanent shock is identified by subtracting the variance of transitory shocks from the unbiased combined component, i.e.

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

[11]Identification of the average risk differential relies on the normality assumption in this paper. Another working paper by Chen and Khan (2001) shows that without the normality assumption the average risk differential, scaled by the variance of earnings of high school graduates, can still be identified.

[12]For proofs, see Theorem 20.2 and Theorem 20.4 in Greene (1997).

the right-hand-side of the above equation. Notably, the variance of transitory shocks has been consistently estimated in the first step. It is the permanent component of variation that has selection biases and is understated by the ordinary least squares. This identification procedure also works for the subsample of high school graduates.

The average risk differentials for permanent and transitory shocks can be identified by taking respective differences in the variances. Since a limiting distribution of an estimated risk differential is not available, bootstrapping is used to generate confidence intervals. In bootstrapping, I randomly draw 1001 resamples of size $N$ from the original sample with replacement, which generates an empirical distribution of average risk differentials. A percentile method is utilized to determine the endpoints of confidence intervals.[13] The next section provides details about data. Section 4 reports empirical results.

# 3 Data

Statistics and estimations are based on the *National Longitudinal Survey for Youth* (NLSY): *1979-98.* 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 3086 respondents who did not graduate from high school. The second criterion additionally excludes 4363 non-representative respondents to derive a reliable estimate using a random sample. The remaining sample contains 5237 respondents, and each respondent has 12 to 17 years of observations. The third criterion excludes 1939 observations, where 54 respondents are entirely dropped. Nominal variables are normalized by the Gross Domestic Product Implicit Price Deflator for the base year 1992. According to these criterions, there are 5183 respondents and 67018 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

[13]For a 95 percent significance level, the confidence interval of the risk differential is derived by choosing the top 2.5 percentile and the bottom 2.5 percentile.

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 the 5183 respondents, 3511 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 for all cohorts in the sample in 1980, the original Afqt scores are not comparable across different age groups. To resolve this problem, I generate a variable representing the individual deviations from the average Afqt for 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 the 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 cost of attending a *local* in-state public four-year college while the respondent was 17 years old.[14] 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 define the local cost of attendance for this individual. If no four-year

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

public college exists within an individual's county of residence, local cost of attendance for this individual is defined by the sum of average in-state tuition, room, and board charged by the public four-year colleges located in the state.[15] 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 at age 17. 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 at age 17. 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.

*Parental education* is measured by three variables: an individual's mother's highest grade 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 statistics on family income, I add in the NLSY supplemental sample[16] to increase my sample size *only* in estimating levels of family income (see Table 2). After

[15]Two states are exceptional. First, Washington D.C. have 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.

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

doing this, I have a set of 5295 respondents for whom there are 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 three sets of results. First, I examine the characteristics and earnings of college-educated versus high school graduates. These results are given as a baseline comparison. They also demonstrate the importance of controlling individual characteristics and treating selection biases. Next, I present results from the estimation procedure described in Section 2.2. Finally, I interpret the results by approximating the risk premium of college attendance based on a specific risk attitude.

## 4.1 Characteristics and Earnings of College Trained versus High School

In this subsection, I present the composition of the college and high school variables. I also compare their means and variances of log earnings, and their coefficients of variation in earnings. These results are use to examine the importance of controlling for individual characteristics and treating selection biases.

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 schools have a slightly higher proportion of white and female students than colleges do. The ratio of whites in 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 risk 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.[17] 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 1:1.05 for the sample during the ages 30-34. In the 1950 survey, the ratios for those two age groups increased to 1:1.67 and 1:1.24 respectively. Using NLSY, I compute

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

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 individual average earnings by the ratio of 1:1.15 when respondents were 23-28 years old. During the 1990s, the ratio decreases to 1:1.05 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 tent to face higher earnings risk than high school graduates do. The analysis below controls for individual characteristics to measure the extent of earnings risk. 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.

## 4.2 Estimation Results

This subsection reports the empirical results after controlling for individual heterogeneity and treating selection biases. The earnings risk and the risk differentials caused by permanent and transitory shocks are identified using the method introduced in Section 2.2. There are three steps. First, I use a *fixed-effects* model to estimate the risk and risk differential of transitory shocks. Note that the selection bias is time-invariant. The nuisance term of selectivity correction is "differenced" out in the fixed-effects model. Hence, the estimated risk differential of transitory shocks is consistent and has no selection bias. The second and third steps correspond to Heckman's two-stage scheme, in which the selection bias associated with permanent shocks is treated. In the second step, I estimate a *probit* model to derive a selectivity correction term. Using this selection correction, in the third step I identify the combined variance of the shocks using a *between-effects* model, as shown in equation (8). Finally, the permanent earnings risk is singled out from the combined component by simple subtraction. The results of these three steps are summarized in Tables 5-8.

Table 5 presents the results of the fixed-effects model, in which deviations of log wage from the individual average are considered for all survey years (see equation (6)). The explanatory variables are all time-variant, including the number of years of work experience, experience squared, marital status and regional dummies. Notably, the observations of work experience before finishing college are excluded from the sample; therefore, the number of years of schooling is time-invariant and is excluded from the fixed-effects model. Items 1b and 1c compare the experience-earnings profiles among different levels of schooling, which indicate that college attendees tend to have a steeper experience-earnings profile relative to high school graduates.

Item 3 in Table 5 presents the risk and risk differentials that result from transitory shocks. In Item 3a it appears that transitory earnings risk is slightly smaller for higher school graduates than the college attendees. To estimate the risk differentials, I take differences in variance between college attendees and high school graduates. Item 3b shows that the risk differentials of transitory shocks are significantly negative and approximately 17 percent to 19 percent. These suggest that college attendees, either two-year or four-year, may face *less* transitory earnings uncertainties than high school graduates do. Notice that the estimates of risk differentials caused by transitory shocks are consistent, with no selection bias.

Similar to Olson, White, and Shefrin (1979), the transitory earnings uncertainty is estimated using a fixed-effects model. Olson et al presumes that attending college causes additional transitory earnings uncertainty but does not cause any additional permanent earnings uncertainty. They find that the variance of permanent shocks (or the variance of individual specific effects) is a constant .0743, and the variance of transitory shocks *increases* in levels of schooling. For instance, the variance of transitory shocks is .0727 for high school graduates and .1204 for individuals with one year of college education. However, Table 5 indicates that the variance of transitory shocks *decreases* by levels of schooling if permanent earnings risk can vary with these levels.

Item 4 of Table 5 reports the risk and risk differentials caused by permanent shocks. The estimates in this item are *biased* because the self-selection problem has not been treated yet. Comparing Item 2a to 3a, permanent earnings uncertainties are smaller than transitory earnings uncertainties for high school graduates, but larger for college attendees. In fact, Item 5 shows that the permanent shocks account for 54 to 58 percent of the total unexplained variation in earnings of college attendees. In Item 4b, permanent risk differentials are all significantly positive, compared to Item 3b where transitory risk differentials are all significantly negative. These results seem to imply that investing in a college education may *decrease* the transitory earnings risk but *increase* the permanent earnings risk.

Table 6 reports the determinants of college attendance, according to a probit estimation. In row (a), the cost of attendance has a strong negative effect on schooling decisions. If the cost of attendance increases one percent, the college enrollment rate may decrease by 3.9 percent, with as low as .008 of the standard error. This variable is chosen as an instrumental variable to treat selection biases because it is directly correlated with the schooling choice but not correlated with individual earnings. Note that to remove the local income effect on the instrument, I have deflated the instrument by the average wage of local unskilled workers. See the Data Section for details.

The high school graduation year, as shown in row (b), has significant explanatory power for college enrollment. Rows (c) and (d) reveal that both scholastic ability and parental education are influential in the decision to enroll in college. Rows (e) and (f) present evidence that both blacks and females are more likely to enroll in college than non-African American males. To be more specific, other things being equal, blacks have a 22 percent higher probability of attending college than other races, while females have a 6.2 higher probability of attending college than males. As an extreme example, a high-ability black female who lived with both parents during youth in a low-tuition county is more likely to attend college than any other type of individual.

Table 7 reports the results of the between-effects model, in which the variation in individual average earnings over survey years is explained by several time-invariant variables. The selectivity correction term (or the inverse Mills ratio), constructed from the probit model, is included as one of the regressors. Row (k) reports that, except for the two-year college attendees, the inverse Mills ratio is significant for all levels of schooling. In particular, for four-year college attendees, the coefficient for the inverse Mills ratio in the log wage equation is .336, with a small .088 of the standard error. This suggests that the selectivity bias may not be negligible and the instrumental variable effectively corrects the selectivity bias in most of the cases.

The controls of the probit are included in the between-effects model, but both cost of attendance and high school graduation years (hgy) are excluded. Cost of attendance is excluded since it serves as an instrumental variable; hgy is also excluded since it can be fully determined by years of schooling, years of experience, and cohort effects.

As shown in row (a) of Table 7, the marginal returns to years of schooling are 4.4 percent and 6.5 percent for a two-year and four-year college education, respectively. The estimates in the literature range from 8 to 13 percent (see Card [1995a, 1999]), larger than my estimates because I control for both scholastic ability scores and parental education. Rows (b) and (c) suggest that the effects of work experience on earnings are not as significant as the effects of schooling. Rows (d) and (e) report the effects of scholastic ability and parental education on earnings. We note that an individual's ability only plays a minor role in explaining earnings, whereas parental education is often important to explain the levels of earnings, especially for those who enrolled in a four-year college. In contrast, both scholastic ability and parental education are influential in shaping the schooling decision, as shown in rows (c) and (d) in Table 6.

Rows (f) and (g) of Table 7 report that female African-Americans have a significant disadvantage in earnings relative to males or non-African-Americans. In particular, holding other individual attributes constant, African Americans earn 21.6 percent less than non-African-Americans, and females earn 26.5 percent less than males. This is a puzzle given that female African-Americans are more likely to attend college than males or non-African-Americans, as shown in Table 6. This suggests that racial and gender biases in college enrollment are less severe than the racial and gender biases in labor markets.

Item 2 of Table 7 shows that interpersonal differences account for 33 to 40 percent of variation in individual average earnings over the years. The rest of the variation may be due to earnings uncertainty or measurement errors. If the variance of measurement errors are independent of schooling choices, the variation in earnings caused by measurement errors can be subtracted out in estimating risk differentials. In this case, the difference in variances of earnings is still a reasonable measure for the riskiness of college attendance even with the presence of measurement errors.

Table 8 reports the *consistent* estimates of risk and risk differentials. For convenience of comparison, in item 1 I duplicate the estimated transitory risk and risk differentials that are derived from item 3 of Table 5. Item 2 of Table 8 presents the estimates of permanent earnings risk and risk differentials. Compared to the estimates with selection biases in item 4 of Table 5, the variances of permanent shocks adjust upward for all levels of schooling after correcting for selection biases, as equation (9) suggests. This selectivity correction adjustment ranges between 11 percent and 41 percent.

Taking differences in these unbiased variances, the consistent estimates of permanent risk differentials increase dramatically. The selectivity correction adjustment on the permanent risk differentials ranges between 53 percent and 147 percent. In particular, the risk differential for a four-year college education increases from .051 to .126, while the risk differential for a two-year college education increases from .015 to .023. Confidence intervals, generated by bootstrapping, indicate that the permanent risk differential of a four-year college education is significantly positive, while that of a two-year college education is insignificant and approaches zero. Interpretations of my findings follow.

## 4.3 Discussion

To understand how risk differentials influence schooling decisions, recall the schooling choice rule defined in equation (3). There are three components that determine the magnitude of risk premium: the risk differential, the degree of risk aversion, and the correlation coefficient between unobserved individual effects and earnings uncertainty. As an example, suppose that the coefficient of relative risk aversion equals two, and the correlation between unobserved individual effects and earnings uncertainty is negligible. The schooling choice model in Section 2 shows that the risk premium of schooling is approximately equal to one half of the risk differential caused by permanent shocks.

Given that the annual return to a four-year college education is 6.5 percent, according to row (a) of Table 7, the lifetime return to a four-year college education is about 29.7 percent.[18] Since the risk differential caused by permanent shocks is .126 for a four-year college attendee, the risk premium on a four-year college education is approximately equal to $12.6/2 = 6.3$ percent. Hence, the net lifetime returns to a four-year college education is reduced to $29.7-6.3 = 23.4$ percent. Notice that the risk premium on a four-year college education offsets almost a quarter of the return to schooling. In contrast, for a two-year college education, the risk differential is approximately zero. These results suggest that investing in a four-year college education is indeed risky, while a two-year college education may not be.[19] To interpret this difference, one may consider that a four-year college education is a long-term

[18] $e^{.065\times4} - 1 = .297$.

[19] Note that if fields of major at college are controlled in the regression, the estimated risk differential will increase by a small extent.

investment relative to a two-year college education. The postponement of payoffs from a four-year college education may generate a degree of earnings uncertainty.

Finally, it is worth noting that although this paper focuses on the riskiness of schooling caused by earnings uncertainty, unemployment is also an important source of risk associated with college attendance, as pointed out by Mincer (1991). According to Mincer's estimation, the probability of unemployment is 6.4 percent for high school graduates, but only 3.5 percent for four-year college graduates. Of course, if the risk of unemployment is incorporated as a part of earnings uncertainties, the risk differential will be smaller than the one that I estimate above.

# 5 Conclusion

Is investing in a college education risky? This paper has separately estimated the risk differentials caused by permanent and transitory shocks to answer this question. I find there is substantial and statistically significant risk associated with an investment in a college education.

After correcting for the selection bias problems neglected in the literature, the estimates of permanent risk differentials dramatically increase. A back-of-the-envelope calculation implies that the risk premium on a four-year college education may offset a large portion of the returns to schooling. A numerical example shows that the risk premium may reduce the rate of the return from 29.7 percent to 23.4 percent. This finding suggests that the riskiness of college attendance can potentially be a factor in shaping the schooling choice in certain circumstances.

It is worth noting that although this paper focuses on the riskiness of schooling caused by earnings uncertainties, unemployment is also an important source of the risk associated with college attendance. If the risk of unemployment is incorporated as a part of earnings uncertainties, the estimated risk differential will be smaller. On the other hand, the estimated risk differentials control for the number of years of college education. Nevertheless, many students enroll in college not knowing whether they will drop out during school years, nor whether they will attend a graduate school in the future. If the uncertainty in the number of years of schooling is taken into account, the risk differential will be much larger and offset a greater portion of the return to schooling than that estimated in this paper.

# References


[1] **Acemoglu, Daron 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] **Altonji, Joseph G.** "The Demand for and Return to Education When Education Outcomes Are Uncertain." *Journal of Labor Economics*, 1993, 11(1), 48-83.

[3] **Angrist, Joshua D. and Newey, Whitney K.** "Over-Identification Tests in Earnings Functions with Fixed Effects." *Journal of Business and Economics Statistics*, 9(3), 1991, pp. 317-23.

[4] **Attanasio, Orazio P. and Davis, Steven K.** "Relative Wage Movements and the Distribution of Consumption." *Journal of Political Economy*, 104(6), 1996, pp. 1227-62.

[5] **Becker, Gary S.** *Human Capital.* 1964; 1993. New York: Columbia University Press.

[6] ——. *Human Capital and the Personal Distribution of Income*. 1967, Ann Arbor: The University of Michigan Press.

[7] **Blundell, Richard and Preston, Ian.** "Consumption Inequality and Income Uncertainty." *Quarterly Journal of Economics*, 1998, pp. 603-40.

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

[9] **Buchinsky, Moshe and Leslie, Phillip.** "Educational Attainment and the Changing US Structure: Some Dynamic Implications." Working Paper 97/13, 1997, Department of Economics, Brown University.

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

[11] ——. "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*, 1999a, edited by Marvin Kosters. Washington, D. C.: AEI Press.

[12] ——. "The Dynamics of Education Attainment for Blacks, Whites and Hispanics." Working paper W7249, 1999b, Cambridge, Mass: National Bureau of Economic Research.

[13] **Cameron, Stephen and Taber, Christopher.** "Borrowing Constraints and Returns to Schooling." 2000, Manuscript.

[14] **Campbell, J.Y.** "Understanding Risk and Return." *Journal of Political Economy*, 104(2), 1996, pp. 298-345.

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

[16] ——. "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.

[17] ——. "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.

[18] **Deaton, Angus.** *Understanding Consumption.* 1992. Clarendon Lectures in Economics. Oxford: Oxford University Press.

[19] **Efron, Bradley and Tibshirani, Robert.** "Bootstrap Method for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy." *Statistics Science*, 1(1), 1986, pp. 54-75.

[20] **Ellwood, David T. and Kane, Thomas 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, Wiliam H.** *Econometric Analysis.* 1997, Saddle River, NJ: Prentice Hall.

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

[23] —— **and Lochner, Lance.** "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.

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

[25] ——. "Rising Public College Tuition and College Entry: How Well Do Public Subsidies Promote Access to College." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5164, 1995.

[26] **Keane, Michaek P. and Wolpin, Kenneth I.** "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." Unpublished manuscript, 1999.

[27] **Kyriazidou, Ekaterini.** "Estimation of a Panel Data Sample Selection Model." *Econometrica*, 65(6), 1997, pp. 1335-64.

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

[29] **Meghir, C. and Palme, M.** "Assessing the Effect of Schooling on Earnings Using a Social Experiment." Institute for Fiscal Studies Working Paper: 99(10), 1999.

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

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

[32] **Mooney, Christopher Z. and Duval, Robert D.** "Bootstrapping: a Nonparametric Approach to Statistical Inference." Newbury Park CA: Sage Publication, 1993.

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

[34] **Palacio-Huerta,I.** "An Empirical Analysis of the Properties of Human Capital Returns." Working Paper, Brown University and University of Chicago, 1999.

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

[36] **Taber, Christopher R.** "Semiparametric Identification and Heterogeneity in Discrete Choice Dynamic Programming Models." *Journal of Econometrics*, 96(2), 2000, pp. 201-229.

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

[38] **Willis, Robert J. and Rosen, Sherwin.** "Education and Self-Selection." *Journal of Political Economy*, 87(5), 1979, 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 institutes** | | | |
|---|---|---|---|---|
| | **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** | |
|---|---|---|---|---|---|
| | **Total** | **High School** | **College** | **High 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 attending local public 4-year college at age 17[2] | 6.2 | 6.4 | 6.1 | 6.5 | 6.2 |
| | (1.0) | (1.0) | (1.0) | (.9) | (.9) |
| 3. Family Background | | | | | |
| Average family income per person | 11778.4 | 10064.2 | 12631.0 | 12025.6 | 13923.6 |
| | (7519.4) | (6520.5) | (7833.9) | (6509.4) | (8041.5) |
| Percent a parent attended college[3] | 36.1 | 14.6 | 46.6 | 21.7 | 58.3 |
| | (48.0) | (35.3) | (49.9) | (41.3) | (49.3) |
| Sample size[4] | 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: College Attendance and the Mean and Variance of Log Earnings: 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) The ** and * indicate the 5% and 10% significance levels respectively. (iii) Note that NLSY has no observations for wage and earnings in 1994 and 1996.

Table 5: Earnings, Earnings Risk and Risk Differentials — Fixed Effects.

| **Estimates and Tests** | **High School Graduates** | **College Attendees Total** | **College Attendees 2-year** | **College Attendees 4–year** |
|---|---|---|---|---|
| 1. Log wage equation (fixed-effects): | | | | |
| a. Years of schooling | - | - | - | - |
| b. Experience | .065** | .072** | .074** | .073** |
| | (.003) | (.002) | (.004) | (.003) |
| c. Experience squared | -.002** | -.002** | -.002** | -.002** |
| | (.001) | (.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 |
| 3. Transitory shocks: | | | | |
| a. Variance | .188 | .166 | .167 | .165 |
| b. Risk differentials | - | -.022** | -.021** | -.023** |
| (F-test; p-value) | - | (.000+) | (.000+) | (.000+) |
| 4. Permanent shocks (with selection bias): | | | | |
| a. Variance | .177 | .230 | .192 | .228 |
| b. Risk differentials | - | .052** | .015** | .051** |
| (F-test; p-value) | - | (.000+) | (.000+) | (.000+) |
| 5. Percentage of unexplained variation due to permanent shocks (with selection bias) | 48.4 | 58.1 | 53.5 | 53.9 |
| 6. Number of respondents | 1551 | 2828 | 843 | 1985 |
| 7. Number of observations | 16952 | 25098 | 8566 | 16532 |

Note: (i) and (ii): same as Table 4. (iii) '-' indicates 'not available'. (iv) "Reginal dummies" include urban, south, north east, and west.

Table 6: Explaining College Attendance – the Probit Model.

| **College enrollment** | **Coefficients** | **Changes in Probability** |
|---|---|---|
| 1. Explanatory variables: | | |
| a. Log cost of attendance | -.125** | -.039** |
| | (.025) | (.008) |
| b. High school graduation year | yes** | yes** |
| (Chi square (2); p-value) | (.000+) | (.000+) |
| c. Afqt scores | .023** | .007** |
| | (.001) | (.000+) |
| d. Parental education | yes** | yes** |
| (Chi square (4); p-value) | (.000+) | (.000+) |
| e. Black | .691** | .218** |
| | (.084) | (.027) |
| f. Male | -.196** | -.062** |
| | (.046) | (.015) |
| g. Regional dummies at age 17 | yes** | yes** |
| (Chi square (4); p-value) | (.011) | (.011) |
| 2. LR test, chi-squared | 1191.80 | |
| 3. Number of respondents | 4138 | |

Note: (i) and (ii): same as Table 4. (iii) The instrument in row a is deflated by the average hourly wage of local unskilled workers. (iv) Row b contains quadratic terms of high school graduation years. (v) Parental education includes the number of years of education of mother, the number of years of education of father, and their cross terms. (vi) Regional dummies include whether one lived in an urban area, northeastern, west, and south at age 14. (vii) A dummy variable for races other than black and white is also included in the model. (viii) The changes in probability are evaluated at sample means of other variables.

Table 7: Between-Effects Model.

| Estimates and Tests | High School Graduates | College Attendees | | |
|---|---|---|---|---|
| | | Total | 2-year | 4–year |
| 1. Log wage equation (between-effects): | | | | |
| a. Years of schooling | - | .061** | .044** | .065** |
| | - | (.004) | (.015) | (.006) |
| b. Experience | .029 | .035** | .026 | .032* |
| | (.021) | (.016) | (.028) | (.019) |
| c. Experience squared | .002 | .000+ | .002 | -.000+ |
| | (.001) | (.001) | (.001) | (.001) |
| d. Afqt | .001 | -.001 | -.000+ | -.003* |
| | (.001) | (.001) | (.002) | (.002) |
| e. Parental education | yes | yes** | yes | yes** |
| (Chi squared; p-value) | (.852) | (.012) | (.432) | (.005) |
| f. Black | -.067** | -.155** | -.083 | -.216** |
| | (.049) | (.048) | (.077) | (.063) |
| g. Male | .274** | .264** | .270** | .265** |
| | (.023) | (.018) | (.031) | (.023) |
| h. Current marital status | .064** | .074** | .053** | .091** |
| | (.028) | (.022) | (.039) | (.026) |
| i. Regional dummies | yes** | yes** | yes** | yes** |
| (F-test; p-value) | (.000+) | (.000+) | (.000+) | (.000+) |
| j. Cohort effects | yes | yes | yes | yes** |
| (F-test; p-value) | (.146) | (.277) | (.715) | (.047) |
| k. Inverse Mills ratio | .181** | .264** | .196 | .336** |
| | (.078) | (.069) | (.114) | (.088) |
| 2. $R^2$ | .379 | .379 | .402 | .334 |
| 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) Parental education include years of schooling of mother and father. (iv) Regional dummies include whether currently lived in an urban area, northeast, west, and south.

Table 8: Earnings, Earnings Risk and Risk Differentials.

| **Estimates and Test Statistics** | **High School Graduates** | **College Attendees** | | |
|---|---|---|---|---|
| | | **Total** | **2-year** | **4–year** |
| 1. Earnings risk and risk differentials due to: | | | | |
| A. Transitory shocks: | | | | |
| a. Variance | .188 | .166 | .167 | .165 |
| b. Risk differentials | - | -.022** | -.021 | -.023** |
| (p-value) | - | (.000+) | (.168) | (.000+) |
| B. Permanent shocks: | | | | |
| a. Variance | .197 | .283 | .220 | .322 |
| b. Risk differentials | - | .086** | .023 | .126** |
| (confidence intervals) | - | (.022, .172) | (-.015, .084) | (.037, .266) |
| | - | | | |
| 2. Percentage of unexplained variation due to permanent shocks: | | | | |
| | 51.2 | 63.0 | 56.8 | 66.1 |

Note: (i) and (ii): same as Table 4. (iii) The confidence intervals of the risk differentials in row b are based on bootstrapping with 1001 replications. The size for each replication is the number of observations.

# 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 college attendance decision in the first period and obtains earnings in each period afterwards. His earnings are determined by Mincer's earnings 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 his earnings streams between attending and not attending college. His decision is based on the following information: personal characteristics $(x_{i1}, x_{i2}, \cdots, x_{iT}, z_i)$, the effect of family support for schooling $\eta_i$, and the perceived distributions of permanent shocks $a_i$ and transitory shocks $\varepsilon_{it}$. Notably, since only person $i$ can observe his individual effect of family support for schooling $\eta_i$, his perceived earnings stream is slightly different from what econometricians observe in Mincer's equation (1),

$$\begin{aligned}\ln y_{it}(s) = {} & x_{it}'\alpha + z_i'\beta \\ & + \sigma_a(s)\sqrt{1-\rho_a^2}a_i + \sigma_\varepsilon(s)\,\varepsilon_{it}.\end{aligned}$$

This equation is appropriate because the triple $(a_i, \varepsilon_{it}, \eta_i)$ is normally distributed; $\rho_a$ is the correlation coefficient between $a_i$ and $\eta_i$, $\sigma_a$ is the unconditional variance of permanent shock, and $\sigma_\varepsilon$ is the unconditional variance of transitory shocks. The conditional variances of permanent and transitory shocks are

$$\begin{aligned}Var\left[\sigma_a(s)\,a_i\,|\eta_i\right] &= \sigma_a^2(s)\left(1-\rho_a^2\right)\\ Var\left[\sigma_\varepsilon(s)\,\varepsilon_{it}\,|\eta_i\right] &= \sigma_\varepsilon^2(s)\,.\end{aligned}$$

Let $\mu_{it}$ denote the deterministic component of earnings, $x'_{it}\alpha + z'_i\beta$. Summing up his earnings stream $y_{it}$, person $i$'s lifetime earnings can be written as:

$$y_i^p(s) \equiv \sum_{t=0}^{T} R^t \mu_{it}(s) \exp\left[\sigma_a(s)\sqrt{1-\rho_a^2}a_i + \sigma_\varepsilon(s)\varepsilon_{it}\right],$$

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

$$\mu_i^p(s) = \sum_{t=0}^{T} R^t \mu_{it}(s).$$

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

$$u(y_i^p) = \frac{(y_i^p)^{1-\gamma}}{1-\gamma},$$

where $\gamma$ denotes the *Arrow-Pratt coefficient of relative risk aversion.* If $\gamma$ equals unity, the preference degenerates into a logarithmic function. The *expected* lifetime utility can be derived as follows,

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

where $E_0$ is the expectation operator based on the information in the beginning of period 1. The second line is appropriate only if (i) the permanent shock $a_i$ is normally distributed so that $E_0[e^{ta_i}] = \exp[t^2\sigma_a^2(s)/2]$, and (ii) the expected lifetime utility can be approximated by the first-order Taylor expansion around $\varepsilon_{it} = E[\varepsilon_{it}] = 0$. Then, the expected lifetime utility can be expressed by a *certainty equivalent*, defined by $E_0[u(y_i^p(s))] = u(ce_i(s))$, such that

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

The comparison between expected lifetime earnings is equivalent to the comparison between certainty equivalents. The schooling choice rule 3 then follows:

$$s_i = I\left\{\ln \mu_i^p(1) - \ln \mu_i^p(0) > \frac{\gamma-1}{2}\left(1-\rho_a^2\right)\left[\sigma_a^2(1) - \sigma_a^2(0)\right]\right\},$$

where $I\{\cdot\}$ is an indicator function.