Fifty Years of Mincer Earnings Regressions

James J. Heckman
Lance J. Lochner

and

Petra E. Todd
First draft June, 1998
Revised March 19, 2003

Abstract

The Mincer earnings function is the cornerstone of a large literature in empirical economics. This paper discusses the theoretical foundations of the Mincer model and examines the empirical support for it using data from Decennial Censuses and Current Population surveys. While data from the 1940 and 1950 Censuses provide some support for Mincer's model, data from later decades are inconsistent with it. We examine the importance of relaxing functional form assumptions in estimating internal rates of return to schooling and in accounting for taxes, tuition, nonlinearity in schooling, and nonseparability between schooling and work experience. Inferences about trends in rates of return to high school and college obtained from our more general model differ substantially from inferences drawn from estimates based on a Mincer earnings regression. Important differences also arise between cohort-based and cross-sectional estimates of the rate of return to school. In the recent period of rapid technological change, widely used cross-sectional applications of the Mincer model produce dramatically biased estimates of cohort returns to schooling. We also examine the implications of accounting for uncertainty and agent expectation formation. Even when the static framework of Mincer is maintained, accounting for uncertainty substantially affects rate of return estimates. Considering the sequential resolution of uncertainty over time in a dynamic setting gives rise to option values, which fundamentally changes the analysis of schooling decisions. In the presence of sequential resolution of uncertainty and option values, the internal rate of return - a cornerstone of classical human capital theory - is not a useful guide to policy analysis.

JEL Code: C31

James J. Heckman
Department of Economics
University of Chicago
1126 East 59th Street, Chicago, IL 60637
Telephone: (773) 702-0634
Fax: (773) 702-8490
E-mail: jjh@uchicago.edu

Lance Lochner
Department of Economics
University of Rochester
Rochester, NY 14627
Telephone: (585) 275-4239
Fax: (585) 256-2309
E-mail: lance@troi.cc.rochester.edu

Petra Todd
Department of Economics
University of Pennsylvania
20 McNeil, 3718 Locust Walk, Philadelphia, PA.
Telephone: (215)898-4084
Fax: (215)573-2057
E-mail: petra@athena.sas.upenn.edu

1 Introduction

Jacob Mincer's model of earnings (1974) is a cornerstone of empirical economics. It is the framework used to estimate returns to schooling,[1] returns to schooling quality,[2] and to measure the impact of work experience on male-female wage gaps.[3] It is the basis for economic studies of education in developing countries[4] and has been estimated using data from a variety of countries and time periods. Recent studies in economic growth use the Mincer model to analyze the relationship between growth and average schooling levels across countries.[5]

In one equation, Mincer's framework captures two distinct economic concepts: (a) a pricing equation or hedonic wage function revealing how the labor market rewards productive attributes like schooling and work experience and (b) the rate of return to schooling which can be compared with the interest rate to determine optimality of human capital investments. Assuming stationarity of the economic environment, the analyst can use the Mincer model to identify both skill prices and rates of return to investment. This happy coincidence only occurs under special conditions, which were approximately valid in the 1960 Census data used by Mincer (1974). Unfortunately, these conditions have been at odds with data ever since. As a result, the widely used Mincer model applied to more recent data does not provide valid estimates of returns to schooling, nor do related studies that associate a rising college - high school wage differential with an increase in the return to schooling. (See, *e.g.* Murphy and Welch, 1992, Katz and Murphy, 1992, Katz and Autor, 1999.)

A large literature refers to the coefficient on schooling in an earnings regression as a rate of return to schooling without stating the conditions under which this interpretation is valid. This approach to estimating returns has been a main vehicle used to document the rise in returns to schooling over the past twenty years. Yet, it neglects major determinants of actual returns, such as the direct and indirect costs of schooling, taxes, length of working life, and uncertainty about future returns at the time schooling decisions are made.

---

[1]See, e.g., Psachoropoulus (1981), Willis (1986), Ashenfelter and Krueger (1994), Ashenfelter and Rouse (1998), Smith and Welch (1989), Krueger (1993).

[2]See Behrman and Birdsall (1983) and Card and Krueger (1992).

[3]See Mincer and Polachek (1974).

[4]See Glewwe (2002).

[5]See Bils and Klenow (2000).

Additionally, while some widely cited studies point out that educational wage differentials vary over the lifecycle and that the pattern for earnings-experience-schooling relationships has changed over time (*e.g.* Murphy and Welch, 1992, Katz and Murphy, 1992, Katz and Autor, 1999), these studies offer little guidance in mapping those differentials into a rate of return measure that can be used to study educational decisions or policy.

This paper makes the following points. (1) Building on the analysis of Willis (1986), we present conditions under which the coefficient on schooling in a Mincer earnings function estimates the rate of return to schooling, assuming stationarity of the economic environment and perfect certainty. (2) Using Census data for the years 1940 - 1990, we test these conditions and reject them, even in the 1960 Census data used in the original Mincer analysis. (3) We develop an alternative nonparametric method to estimate rates of return to schooling that does not rely on the Mincer model. (4) Using our method, we estimate internal rates of return to school (i.e. the discount rate that equates the present value of two earnings streams associated with different schooling levels) that differ substantially in both levels and time trends from estimates based on the Mincer earnings equation. Although the empirical literature has focused on neglect of higher order terms in experience as a major source of misspecification in the Mincer model (see, *e.g.* Murphy and Welch, 1990), we find that this neglect has only minor consequences for estimated rates of return. Far more important is relaxing Mincer's assumptions of linearity in schooling and separability between schooling and experience. An interesting by-product of our analysis is the discovery that the real story of educational returns in the 1980s is not the increase in the returns to college as emphasized by Katz and Murphy (1992) and others, but rather the increase in the return to graduating from high school. The floor fell out from the wages of the unskilled.

(5) We also explore the importance of Mincer's stationarity assumptions about the economic environment, and allow lifecycle earnings-education-experience profiles to differ across cohorts. In this case, cross sections are no longer useful guides to the lifecycle earnings or schooling returns of any particular individual. Accounting for the nonstationarity of earnings over time has empirically important effects on estimated rates of return to schooling.

(6) We relax the implicit assumption of perfect certainty about future earnings streams associated with different schooling levels that underlies Mincer's model. We first consider

a model of uncertainty in a static setup without any updating of information. Accounting for uncertainty in this way substantially reduces estimated internal rates of return to more plausible levels. The resulting estimates are consistent with the qualitative conclusions of a model that ignores uncertainty.

We then propose a substantial break from Mincer's approach by allowing for the sequential resolution of uncertainty. That is, with each additional year of schooling, information about the value of different schooling choices and opportunities becomes available generating an option value of schooling.[6] Completing high school generates the option to attend college and attending college generates the option to complete college. Our findings suggest that part of the economic return to finishing high school or attending college includes the potential for completing college and securing the high rewards associated with a college degree. Both the sequential resolution of uncertainty and non-linearity in returns to schooling contribute to sizeable option values.

Accounting for option values challenges the validity of a major empirical tool used in human capital theory since the seminal work of Becker (1964) – the internal rate of return. When the schooling decision is made at the beginning of life and age-earnings streams across schooling levels are known and cross only once, then the internal rate of return (IRR) can be compared with the interest rate as a valid rule for making education decisions (Hirschleifer, 1970). When schooling decisions are made sequentially as information is revealed, a number of problems arise that invalidate this rule. We examine these problems and the empirical role that option values play in determining rates of return to schooling. Our analysis points to a need for more empirical studies that incorporate the sequential nature of individual schooling decisions and uncertainty about education costs and future earnings to help determine their importance.

This paper does not examine the implicit assumption of the Mincer model that schooling is exogenous. This assumption has been challenged elsewhere. See Griliches (1977), Willis and Rosen (1979), Willis (1986), Card (1995, 1999), Heckman and Vytlacil (1998, 2003), and Carneiro, *et al.* (2001). Unfortunately, the current empirical debate on the importance of accounting for the endogeneity of schooling is far from settled. The instruments used in this literature have been seriously challenged (Carneiro and Heckman, 2002), and

---

[6]Weisbrod (1962) developed the concept of the option value of education. For one formalization of his analysis, see Comay, Melnik and Pollatschek (1973).

the Census data used in this paper yield large samples but few instruments. This paper uses these data to examine other, neglected, aspects of the Mincer model. Assumptions about the functional form of the earnings function, the consequences of tuition and taxes, uncertainty, and stability of the economic environment have been largely neglected in the empirical literature. This paper fills that void by systematically analyzing these issues, maintaining the exogeneity of schooling like most of the literature following Mincer (See, e.g., Katz and Murphy, 1992, and Katz and Autor, 1999).

This paper proceeds in the following way. Section 3 reviews two distinct theoretical foundations for the Mincer model that are often confused. Section three presents empirical evidence on the validity of the Mincer specification. Using nonparametric estimation techniques, we formally test (and reject) key assumptions of Mincer's model. In Section four, we develop an alternative nonparametric approach that allows for income taxes, college tuition, and length of working life that may depend on the amount of schooling. We explore the empirical importance of assumptions that are needed to equate the Mincer schooling coefficient with the internal rate of return to schooling, and provide estimates of the return that take into account more general earnings functions, taxes, tuition, and a varying length of working life. We also consider the impact of allowing for uncertainty in a static decision framework.

Section five considers the interpretation of Mincer regression estimates based on cross-section data in a changing economy. We contrast cross-sectional estimates with those based on repeated cross-sections drawn from the CPS that follow cohorts over time.

In Section six, we introduce a framework with sequential resolution of uncertainty and an option value of schooling. We discuss why the internal rate of return is no longer a valid guide to schooling investments in this environment and argue that another measure of the rate of return used in modern capital theory is more appropriate. Section seven concludes.

2   The Theoretical Foundations of Mincer's Earnings Regression

The Mincer (1958, 1974) model specifies

$$\ln[w(s,x)] = \alpha_0 + \rho_s s + \beta_0 x + \beta_1 x^2 + \varepsilon \tag{1}$$

where $w(s, x)$ is wage at schooling level s and work experience x, ρ_s is the "rate of return to schooling" (assumed to be the same for all schooling levels) and ε is a mean zero residual with $E(\varepsilon|s, x) = 0$. This model is motivated by two conceptually different theoretical frameworks, which we briefly review in this section.

2.1 The compensating differences model of Mincer (1958)

The first Mincer model (1958) uses the principle of compensating differences to explain why persons with different levels of schooling receive different earnings over their lifetimes. This model assumes that individuals have identical abilities and opportunities, that there is perfect certainty, that credit markets are perfect, that the environment is perfectly certain, but that occupations differ in the amount of training required. Schooling is costly because individuals forego earnings while in school, but it entails no direct costs. Because individuals are assumed to be *ex ante* identical, they require a compensating differential to work in occupations that require a longer training period. The size of the compensating differential is determined by equating the present value of earnings streams net of costs associated with different levels of investment.

Let $w(s)$ represent the annual earnings of an individual with s years of education, assumed to be constant over his lifetime. Let r be an externally determined interest rate and T the length of working life, which is assumed not to depend on s. The present value of earnings associated with schooling level s is

$$V(s) = w(s) \int_s^T e^{-rt} dt = \frac{w(s)}{r}(e^{-rs} - e^{-rT}).$$

An equilibrium characterized by heterogeneous schooling choices requires that individuals be indifferent between schooling levels. Allocations of people to different schooling levels are driven by demand conditions. Equating the earnings streams associated with different schooling levels and taking logs yields

$$\ln w(s) = \ln w(0) + \ln((1 - e^{-rt})/(1 - e^{-r(T-s)})) + rs.$$

The second term on the right-hand-side is an adjustment for finite life, which converges to zero as T gets large.[7]

---

[7]This term also disappears if the retirement age, T, is allowed to increase one-for-one with s.

Mincer (1958) observed that this simple framework yields a number of interesting implications: (i) For large T, the coefficient on years of schooling in a Mincer regression equals the interest rate, r, (ii) people with more education receive higher earnings, (iii) the difference between earnings levels of people with different years of schooling is increasing in the interest rate and age of retirement, and (iv) the ratio of earnings for persons with education levels differing by a fixed number of years is roughly constant across schooling levels.

If we define the internal rate of return to schooling as the discount rate that equates the lifetime earnings streams for different education choices, then the internal rate of return equals the interest rate, r. Combined with implication (i), the coefficient on years of schooling in a Mincer regression yields an estimate of the internal rate of return. This coefficient also reflects the percentage increase in lifetime earnings associated with an additional year of school when T is large.

2.2 Mincer's (1974) accounting-identity model

Mincer's (1974) second model is motivated by entirely different assumptions from his earlier model, but it yields an earnings specification similar to that of the first. The second model builds on an accounting identity model developed in Becker (1964) and Becker-Chiswick (1966). Unlike the first model, the second model focuses on the life-cycle dynamics of earnings and on the relationship between observed earnings, potential earnings, and human capital investment, both in terms of formal schooling and on-the-job investment. At the same time, no explicit assumptions are made about the background economic environment.

Mincer (1974) writes observed earnings as a function of potential earnings net of human capital investment costs, where potential earnings in any time period depend on investments in previous time periods. Let E_t be potential earnings at time t. Investments in training can be expressed as a fraction of potential earnings invested, i.e. $C_t = k_t E_t$, where k_t is the fraction invested at time t. Let ρ_t be the return to training investments made at time t. Then,

$$E_{t+1} = E_t + C_t \rho_t = E_t(1 + k_t \rho_t).$$

Repeated substitution yields $E_t = \prod_{j=0}^{t-1}(1 + \rho_j k_j)E_0$.

Formal schooling is defined as years spent in full-time investment ($k_t = 1$). Assume that the rate of return on formal schooling is constant for all years of schooling ($\rho_t = \rho_s$) and that formal schooling takes place at the beginning of life. Also assume the rate of return to post-school investment, ρ_t, is constant over time and equals ρ_0. Then, we can write

$$\ln E_t = \ln E_0 + s \ln(1 + \rho_s) + \sum_{j=s}^{t-1} \ln(1 + \rho_0 k_j),$$

which yields the approximate relationship (for small ρ_s and ρ_0)

$$\ln E_t \approx \ln E_0 + s\rho_s + \rho_0 \sum_{j=s}^{t-1} k_j.$$

To establish a relationship between potential earnings and years of labor market experience, Mincer (1974) approximates the Ben Porath (1967) model and further assumes a linearly declining rate of post-school investment:

$$k_{s+x} = \kappa \left(1 - \frac{x}{T} \right) \tag{2}$$

where $x = t - s \geq 0$ is the amount of work experience as of age t. The length of working life, T, is assumed to be independent of years of schooling. Under these assumptions, the relationship between potential earnings, schooling and experience is given by:

$$\ln E_{x+s} \approx [\ln E_0 - \kappa\rho_0] + \rho_s s + \left(\rho_0\kappa + \frac{\rho_0\kappa}{2T} \right) x - \frac{\rho_0\kappa}{2T}x^2.$$

Observed earnings equal potential earnings less investment costs, producing the following relationship for observed earnings:

$$
\begin{aligned}
\ln w(s, x) & \approx \ln E_{x+s} - \kappa \left(1 - \frac{x}{T} \right) \\
& = [\ln E_0 - \kappa\rho_0 - \kappa] + \rho_s s + \left(\rho_0\kappa + \frac{\rho_0\kappa}{2T} + \frac{\kappa}{T} \right) x - \frac{\rho_0\kappa}{2T}x^2. \\
& = \alpha_0 + \rho_s s + \beta_0 x + \beta_1 x^2.
\end{aligned}
$$

Thus, we arrive at the standard form of the Mincer earnings model (equation (1)) that regresses log earnings on a constant term, a linear term in years of schooling, and linear and quadratic terms in years of labor market experience.

In most applications of the Mincer model, it is assumed that the intercept and slope coefficients in equation (1) are identical across persons. This implicitly assumes that E_0, κ,

ρ_0 and ρ_s are the same across persons and do not depend on the schooling level. However, Mincer formulates a more general model that allows for the possibility that κ and ρ_s differ across persons, which produces a random coefficient model

$$\ln w(s_i, x_i) = \alpha_{0i} + \rho_{si} s_i + \beta_{0i} x_i + \beta_{1i} x_i^2 + \varepsilon_i$$

Letting $\alpha_0 = E(\alpha_{0i})$, $\rho_s = E(\rho_{si})$, $\beta_0 = E(\beta_{0i})$, $E(\beta_{1i}) = \beta_1$, we may write this expression as

$$\ln w(s, x) = \alpha_0 + \rho_s s + \beta_0 x + \beta_1 x^2 + [(\alpha_{0i} - \alpha_0) + (\rho_{si} - \rho_s)s + (\beta_{0i} - \beta_0)x + (\beta_{1i} - \beta_1)x^2],$$

where the terms in brackets are part of the error.[8] Mincer initially assumes that $(\alpha_{0i} - \alpha_0), (\rho_{si} - \rho_s), (\beta_{0i} - \beta_0), (\beta_{1i} - \beta_1)$ are independent of (s, x); although he relaxes this assumption in later work (Mincer, 1997).

Implications for log earnings-age and log earnings-experience profiles and for the interpersonal distribution of life-cycle earnings

Mincer derives several implications from the accounting identity model under different assumptions about the relationship between formal schooling and post-school investment patterns. Under the assumption that post-school investment patterns are identical across persons and do not depend on the schooling level, he shows that $\frac{\partial \ln w(s,x)}{\partial s \partial x} = 0$ and $\frac{\partial \ln w(s,x)}{\partial s \partial t} = \frac{\rho_0 \kappa}{T} > 0$. These two conditions imply:

(i) log-earnings experience profiles are parallel across schooling levels, and

(ii) log-earnings age profiles diverge with age across schooling levels.

Mincer (1974) presents informal empirical support for both of these implications of the model using cross-sectional data from the 1960 Decennial Census. In Section 3, we extend his analysis to more Census cross sections and show that the data from the 1940-1950 Censuses provide some empirical support for patterns *(i)* and *(ii)*. The 1960 and 1970 data are roughly consistent with the model, but pattern *(i)* does not pass conventional statistical tests. Data from the more recent Census years are much less supportive of Mincer's model.

---

[8]In the random coefficients model, the error term of the derived regression equation is heteroskedastic.

The framework described above also has important implications for understanding how individual earnings patterns vary with population averages at each age in the life-cycle. One implication is that for each schooling class, there is an age in the life-cycle at which the interpersonal variance in earnings is minimized. Consider the accounting identity for observed earnings at experience level x and schooling level s, which we can write as

$$w(s,x) = E_s + \rho_0 \sum_{j=s}^{s+x-1} C_j - C_{s+x}.$$

In logs

$$\ln w(s,x) \cong \ln E_s + C_0 \sum_{j=0}^{x-1} k_{s+j} - k_{s+x}.$$

Interpersonal differences in observed earnings of individuals with the same E_0 and ρ_s arise because of differences in $\ln E_s$ and in post-school investment patterns as determined by k_j. When $\ln E_s$ and κ (from equation 2) are uncorrelated, it can be shown that the variance of log earnings is minimized when experience is approximately equal to $1/\rho_0$. (See the derivation in Appendix A.) At this experience level, variance in earnings is solely a consequence of differences in schooling levels or ability and is unrelated to differences in post-school investment behavior. Prior to and after this time period (often referred to as the 'overtaking age'), there is an additional source of variance due to differences in post-school investment. As discussed by Mincer (1974), this yields another important implication that can be examined in the data, namely:

(iii) the variance of earnings over the life-cycle has a U-shaped pattern

Below, we show that this prediction of the model is supported in Census data from both early and recent decades.[9]

3   Empirical Evidence for the Mincer Model

We now examine the empirical support for three key implications of Mincer's accounting identity model given above by *(i)*, *(ii)*, and *(iii)*. We extend Mincer's (1974) analysis

---

[9]In addition to Mincer (1974), studies by Schultz (1975), Smith and Welch (1979), Hause (1980), and Dooley and Gottschalk (1984) also provide evidence of this pattern for wages and earnings.

of subsamples of white males from the 1960 decennial U.S. census to include both white and black males from the 1940-1990 decennial Censuses. Earnings correspond to annual earnings, which includes both wage and salary income and business income.[10]

Figure 1 presents nonparametric estimates of the experience - log earnings profiles for each of the Census years for white and black males. Analogous estimates of the age - log earnings profiles are shown for 1940, 1960, and 1980 in Figure 2. Nonparametric local linear regression is used to generate the estimates.[11] The estimated profiles for white males from the 1940-1970 Censuses generally support the fanning-out by age and the parallelism by experience patterns (implications *(i)* and *(ii)* above) predicted by Mincer's accounting identity model. For black males, the patterns are less clear, partly due to the order of magnitude smaller sample sizes which result in less precise estimates. For 1960 and 1970, when the sample sizes of black males are much larger relative to earlier years, experience - log earnings profiles for black males show convergence across education levels over the life-cycle.

Earnings-experience profiles for the 1980-1990 Censuses show convergence for both white and black males. Thus, while data from the 1940-1950 Censuses provide support for implications *(i)* and *(ii)* of Mincer's model, the evidence for implication *(i)* is weaker for 1960 and 1970. The data from 1980 and 1990 do not support the model.[12] Formal statistical tests, reported in Table 1, reject the hypothesis of parallel experience - log earnings profiles for whites during all years except 1940 and 1950. Thus, even in the 1960 data used by Mincer, we reject parallelism. For black males, parallelism is only rejected in 1990, although the samples are much smaller. (The formulae for the test statistics are given in Appendix C.)

Figure 3 examines the support for implication *(iii)*–a U-shaped variance in earnings– for three different schooling completion levels: eighth grade, 12th grade, and college (16 years of school). For the 1940 Census year, the variance of log-earnings over the life-cycle is relatively flat for whites. It is similarly flat in 1950, with the exception of increasing

---

[10]Business income is not available in the 1940 Census. Appendix B provides detailed information on the construction of our data subsamples and variables.

[11]Details about the nonparametric estimation procedure are given in Appendix C. The bandwidth parameter is equal to 5 years. Estimates are not very sensitive to changes in the bandwidth parameter in the range of 3-10 years.

[12]Murphy and Welch (1992) also document differences in earnings-experience profiles across education levels using data from the 1964-1990 Current Population Surveys.

variance at the tails. However, data for black and white men from the 1960-1990 Censuses clearly exhibit the U-shaped pattern predicted by Mincer's accounting-identity model.[13]

Table 2 reports standard cross-section regression estimates of the Mincer return to schooling for all Census years derived from earnings specification (1). The estimates indicate a rate of return to schooling of around 10-13% for white men and 9-15% for black men over the 1940-90 period. While estimated coefficients on schooling tend to be lower for blacks than whites in the early decades, they are higher in 1980 and 1990. The estimates suggest that the rate of return to schooling for blacks increased substantially over the 50 year period, while it first declined and then rose for whites. The coefficient on experience rose for both whites and blacks over the five decades. At the same time, earnings profiles have become more concave as reflected in the increasingly more negative estimated coefficients for experience squared.

4   Estimating Rates of Return

Under the assumptions invoked in the compensating differentials model described in Section 2, the coefficient on schooling equals both the real interest rate and the internal rate of return to schooling. The coefficient on schooling in an accounting identity model can also be interpreted as an average rate of return. These observations have led many economists to label that coefficient the "Mincer rate of return," and a large empirical literature focuses on its estimation.

In this section, we explore what earnings equations estimate within a simple income maximizing framework under perfect certainty developed in Rosen (1977) and Willis (1986). We assume that individuals choose education levels to maximize their present value of lifetime earnings, as in Mincer's compensating differences model, taking as given a post-school earnings profile, which may be determined through on-the-job investment as in the accounting-identity model. The model analyzed in this section relaxes many of the assumptions that were imposed in the models of Section 2, such as the restriction that log earnings increase linearly with schooling and the restriction that log earnings-experience profiles are parallel across schooling classes. We also incorporate additional features, such as

---

[13]For the sake of brevity, only a subset of years are shown in the figures. Figures for 1950, 1970, and 1990 are available from the authors upon request.

school tuition and nonpecuniary costs of schooling, income taxes, and a length of working life that may depend on the schooling level. When these features are incorporated, the coefficient from a Mincer regression need no longer equal the real interest rate (the rate of return on capital). It also loses its interpretation as the internal rate of return to schooling. Therefore, instead of fitting Mincer equations, we estimate rates of returns by a procedure applied in Hanoch (1967), which is further described below.

Let $w(s, x)$ be wage income at experience level x for schooling level s; $T(s)$, the last age of earnings, which may depend on the schooling level; v, private tuition and non-pecuniary costs of schooling; τ, a proportional income tax rate; and r, the before-tax interest rate.[14] Individuals are assumed to choose s to maximize the present discounted value of lifetime earnings[15]

$$V(s) = \int_0^{[T(s)-s]} (1 - \tau)e^{-(1-\tau)r(x+s)}w(s, x)dx - \int_0^s ve^{-(1-\tau)rz}dz. \tag{3}$$

The first order condition for a maximum yields

$$[T'(s) - 1]e^{-(1-\tau)r(T(s)-s)}w(s, T(s) - s) - (1 - \tau)r \int_0^{T(s)-s} e^{-(1-\tau)rx}w(s, x)dx$$

$$+ \int_0^{T(s)-s} e^{-(1-\tau)rx}\frac{\partial w(s, x)}{\partial s}dx - v/(1 - \tau) = 0. \tag{4}$$

---

[14]The standard framework implicitly assumes that individuals know these functional relationships, credit markets are perfect, education does not enter preferences, and there is no uncertainty.

[15]This expression embodies an institutional feature of the U.S. economy where income from all sources is taxed but one cannot write-off tuition and non-pecuniary costs of education. However, we assume that agents can write-off interest on their loans. This assumption is consistent with the institutional feature that persons can deduct mortgage interest, that 70% of American families own their own homes, and that mortgage loans can be used to finance college education.

Defining $\tilde{r} = (1 - \tau)r$ (the after-tax interest rate) and re-arranging terms yields

$$\tilde{r} = \underbrace{\frac{[T'(s) - 1]e^{-\tilde{r}(T(s)-s)}w(s, T(s) - s)}{\int_0^{T(s)-s} e^{-\tilde{r}x}w(s, x)dx}}_{(\text{Term } 1)}$$

$$+ \underbrace{\frac{\int_0^{T(s)-s} e^{-\tilde{r}x}\left[\frac{\partial \log w(s,x)}{\partial s}\right]w(s, x)dx}{\int_0^{T(s)-s} e^{-\tilde{r}x}w(s, x)dx}}_{(\text{Term } 2)} - \underbrace{\frac{v/(1 - \tau)}{\int_0^{T(s)-s} e^{-\tilde{r}x}w(s, x)dx}}_{(\text{Term } 3)}. \tag{5}$$

Term 1 represents a life-earnings effect – the change in the present value of earnings due to a change in working-life associated with additional schooling (expressed as a fraction of the present value of earnings measured at age s). Term 2 is the weighted effect of schooling on log earnings by experience, and Term 3 is the cost of tuition expressed as a fraction of lifetime income measured at age s.

The special case assumed by Mincer (and most labor economists) writes $v = 0$ (or assumes that the third term is negligible) and $T'(s) = 1$ (no private tuition costs and no loss of work life from schooling). This simplifies the first order condition to

$$\tilde{r} \int_0^{T(s)-s} e^{-\tilde{r}x}w(s, x)dx = \int_0^{T(s)-s} e^{-\tilde{r}x}\frac{\partial w(s, x)}{\partial s}dx.$$

As described in Section two, Mincer's model further imposes multiplicative separability between the schooling and experience components of earnings, so $w(s, x) = \mu(s)\varphi(x)$ (i.e. log earnings profiles are parallel in experience across schooling levels). In this special case, $\tilde{r} = \mu'(s)/\mu(s)$. If this holds for all s, then wage growth must be log linear in schooling and $\mu(s) = \mu(0)e^{\rho_s s}$. If all of these assumptions hold, then the coefficient on schooling in a Mincer equation (ρ_s) estimates the internal rate of return to schooling, which should equal the after-tax interest rate.

From equation (5) we observe, more generally, that the difference between after-tax interest rates and the Mincer coefficient can be composed of three parts: a life-earnings part (Term 1), a second part which depends on the structure of the schooling return over the lifecycle, and a tuition cost part (Term 3). The second part is the difference between Term 2 averaged over all schooling and experience categories and the Mincer rate of return

estimated from equation (1). It reflects deviations from linearity of log earnings in schooling and parallelism in experience profiles across education levels.

The evidence for 1980 and 1990 described in Section 3 argues strongly against the assumption of multiplicative separability of log earnings in schooling and experience. In recent decades, log earnings-experience profiles differ across schooling groups. In addition, college tuition costs are nontrivial and are not offset by work in school for most college students. These factors account for some of the observed disparities between the after-tax interest rate and the steady-state Mincer coefficient. Finally, the least squares estimate obtained from a standard Mincer regression does not control for variation in the ability of persons attending college, so classical ability bias could also partly account for the disparity.[16]

One can view $\tilde{r}$ as a marginal internal rate of return to schooling after incorporating tuition costs, earnings increases, and changes in the retirement age. That is, $\tilde{r}$ is the discount rate that equates the net lifetime earnings for marginally different schooling levels at an optimum. As in the model of Mincer (1958), this internal rate of return should equal the interest rate in a world with perfect credit markets, once all costs and benefits from schooling are considered.

After allowing for taxes, tuition, variable length of working life, and a flexible relationship between earnings, schooling and experience, the coefficient on years of schooling in a log earnings regression no longer equals the internal rate of return. However, it is still possible to calculate the internal rate of return using the observation that it is the discount rate that equates lifetime earnings streams for two different schooling levels (Becker, 1964, states this logic. Hanoch, 1967, applies it). Typically, internal rates of return are based on non-marginal differences in schooling. Incorporating tuition and taxes, the internal rate of

---

[16]The evidence on the importance of ability bias is mixed. See, *e.g.* Griliches (1977), Card (1995), Heckman and Vytlacil (2001) and Carneiro, *et.al* (2001) and Carneiro (2002). The evidence reported in Cawley, *et.al* (2000) demonstrates that fundamental identification problems plague studies of the effect of ability on earnings.

return for schooling level s_1 versus s_2, $r_I(s_1, s_2)$, solves

$$\int_0^{[T(s_1)-s_1]} (1-\tau)e^{-r_I(x+s_1)}w(s_1, x)dx - \int_0^{s_1} ve^{-r_I z}dz$$

$$= \int_0^{[T(s_2)-s_2]} (1-\tau)e^{-r_I(x+s_2)}w(s_2, x)dx - \int_0^{s_2} ve^{-r_I z}dz. \qquad (6)$$

As with $\tilde{r}$ above, r_I will equal the Mincer coefficient on schooling under the assumptions of parallelism over experience across schooling categories (i.e. $w(s, x) = \mu(s)\varphi(x)$), linearity of log earnings in schooling ($\mu(s) = \mu(0)e^{\rho_s s}$), no tuition costs ($v = 0$), no taxes ($\tau = 0$), and equal work-lives irrespective of years of schooling ($T'(s) = 1$).[17] In the next section, we compare rate of return estimates based on specification (1) to those obtained by directly solving for $r_I(s_1, s_2)$ in equation (6).

4.1 How model specifications and accounting for taxes and tuition affect internal rate of return (IRR) estimates

Using data for white and black men from 1940-90 decennial Censuses, we examine how internal rate of return (IRR) estimates change when different assumptions about the model are relaxed. Tables 3a and 3b report internal rates of return to schooling for each Census year and for a variety of pairwise schooling level comparisons for white and black men, respectively.[18] These estimates assume that workers spend 47 years working irrespective of their educational choice (i.e. a high school graduate works until age 65 and a college graduate until 69). Initially, the only assumptions we relax are functional form assumptions on the earnings equation, and we ignore taxes and tuition. To calculate each of the IRR estimates, we first estimate a log wage equation under the assumptions indicated in the tables. Then, we predict earnings under this specification for the first 47 years of experience, and the IRR is taken to be the root of equation (6).[19] As a benchmark, the first row for each year reports the IRR estimate obtained from the Mincer specification for log

---

[17]When tuition costs are negligible, proportional taxes on earnings will have no effect on estimated internal rates of return, because they reduce earnings at the same rate regardless of educational choices.

[18]As lower schooling levels are reported only in broader intervals in the 1990 Census, we can only compare 6 years against 10 years and cannot compare 6 years against 8 years or 8 against 10 years as we do for the earlier Census years. We assume the private cost to elementary and high school is zero in all the calculations.

[19]Strictly speaking, we solve for the root of the discrete time analog of equation (6).

wages (equation (1)). The IRR could equivalently be obtained from a Mincer regression coefficient.[20]

Relative to the Mincer specification, row 2 relaxes the assumption of linearity in schooling by including indicator variables for each year of schooling. This modification leads to substantial differences in the estimated rate of return to schooling, especially for schooling levels associated with degree completion years (12 and 16) which now show much larger returns than other schooling years. For example, the IRR to finishing high school is 30% for white men in 1970, while the rate of return to finishing 10 rather than 8 years of school is only 3%. In general, imposing linearity in schooling leads to upward biased estimates of the rate of return to grades that do not produce a degree, while it leads to downward biased estimates of the degree completion years (high school or college). Sheepskin effects are an important feature of the data.[21] There is a considerable body of evidence against linearity. (See *e.g.* Bound, Jaeger and Baker 1995, Heckman, Layne-Farrar and Todd, 1996, Jaeger and Page, 1996, Solon and Hungerford, 1987.) Row 3 relaxes both linearity in schooling and the quadratic specification for experience, which produces similar estimates. The assumption that earnings are quadratic in experience is empirically innocuous for estimating returns to schooling once linearity and separability are relaxed.

Finally, row 4 fully relaxes all three Mincer assumptions (i.e. earnings are non-parametrically estimated as a function of experience, separately within each schooling class, which does not impose any assumption other than continuity on the functional earnings-experience relationship). Comparing these results with those of row three provides a measure of the bias induced by assuming separability of earnings in schooling and experience. In many cases, especially in recent decades, there are large differences. This finding is consistent with the results reported in Section 3, which showed that earnings profiles in recent decades are no longer parallel in experience across schooling categories.

The estimates in Table 3a show a large increase in the return to completing high school for whites, which goes from 24% in 1940 to 50% in 1990, and even more dramatic increases for blacks (Table 3b). It is possible that these increases partially reflect a selection effect,

---

[20]They would be identically equal if our internal rate of return calculations were computed in continuous time. Because we use discrete time to calculate internal rates of return, $r_I = e^{\rho_s} - 1$, which is approximately equal to ρ_s when it is small.

[21]We use the term "sheepskin effects" to refer to exceptionally large rates of return at degree granting years of schooling. We cannot, however, distinguish in the Census data which individuals receive a diploma among individuals reporting 12 or 16 years of completed schooling.

stemming from a decrease in the average quality of workers over time who drop out of high school.[22] There is also a significant increase over time in the marginal internal rate of return to 14 years and 16 years of school, consistent with changes in the demand for labor favoring skilled workers. The Mincer coefficient implies a much lower return to schooling than do the nonparametric estimates, with an especially large disparity for the return to high school completion. For whites, the return to a 4-year college degree is similar under the Mincer and nonparametric models, but for blacks the Mincer coefficient understates the return by about 10%. While the recent literature has focused on the rising returns to college, the increase in returns to completing high school has been substantially greater.

A comparison of the IRR estimates based on the most flexible model for black males and white males shows that for all years except 1940, the return to high school completion is higher for black males, reaching a peak of 58% in 1990 (compared with 50% for whites in 1990). The internal rate of return to completing 16 years is also higher for blacks, by about 10% in 1990.

Estimated internal rates of return clearly differ depending on the set of assumptions imposed by the earnings model. While the assumption that log earnings profiles are quadratic is fairly innocuous, the assumptions of linearity in schooling and separability in schooling and experience are not. Comparing the unrestricted estimates in row 4 with the Mincer-based estimates in row 1 reveal substantial differences for nearly all grade progressions and all years.

Table 4 examines how the IRR estimates change when we account for income taxes (both flat and progressive) and college tuition.[23] For ease of comparison, the first row for each year reports estimates of the IRR for the most flexible earnings specification, not accounting for tuition and taxes. (These estimates are identical to the fourth row in Tables 3a and 3b.) All other rows account for private tuition costs for college (v) assumed equal to the average college tuition paid in the U.S. that year. The average college tuition paid by

---

[22]Though, it is worth noting that the fraction of white men completing high school is relatively stable after 1970. Among black men, high school graduation rates continued to increase until the early 1980s.

[23]Because we assume that schooling is free (direct schooling costs are zero) through high school and because internal rates of return are independent of flat taxes when direct costs of schooling are zero, internal rates of return to primary and secondary school are identical across the first three specifications in the table. Empirically, taking into account progressive tax rates has little impact on the estimates for these school completion levels. (Tables are available upon request.) For these reasons, we only report in Table 4 the IRR estimates for comparisons of school completion levels 12 and 14, 12 and 16, and 14 and 16.

students increased steadily since 1950 as shown in Figure 4a. In 1990, it stood at roughly \$3,500 (in 2000 dollars).[24] Row three accounts for flat wage taxes using estimates of average marginal tax rates (τ) from Barro and Sahasakul (1983) and Mulligan and Marion (2000), which are plotted for each of the years in Figure 4b. Average marginal tax rates increased from a low of 5.6% in 1940 to a high of 30.4% in 1980 before falling to 23.3% in 1990. The final row accounts for the progressive nature of our tax system using federal income tax schedules (Form 1040) for single adults with no dependents and no unearned income. (See Appendix B for details.)

When costs of schooling alone are taken into account (comparing row 2 with row 1), the return to college generally falls by a few percentage points. Because the earnings of blacks are typically lower than for whites but tuition payments are assumed here to be the same, accounting for tuition costs has a bigger effect on the estimates for the black samples. For example, internal rates of return to the final two years of college decline by about one-fourth for whites and one-third for blacks. Further accounting for taxes on earnings (rows 3 and 4) has little additional impact on the estimates. Interestingly, the progressive nature of the tax system typically reduces rates of return by less than a percentage point. Overall, failure to account for tuition and taxes leads to an overstatement of the return to college. However, the time trends in the return are fairly similar whether or not one adjusts for taxes and tuition.

Figure 5 graphs the time trend in the IRR to high school completion for white and black males, comparing estimates based on (i) the Mincer model and (ii) the flexible nonparametric earnings model accounting for progressive taxes and tuition. Estimates based on the Mincer specification tend to understate returns to high school completion and also fail to capture the substantial rise in returns to schooling that has taken place since 1970. Furthermore, the sizeable disparity in returns by race is not captured by the Mincer equation estimates.

Figure 6 presents similar estimates for college completion. Again, the Mincer model yields much lower estimates of the IRR in comparison with the more flexible model that also takes into account taxes and tuition. Nonparametric estimates of the return to college

---

[24]Average college tuition was computed by dividing the total tuition and fees revenue in the U.S. by total college enrollment that year. Federal and state support are not included in these figures. See Appendix A for further details on the time series we used for both tuition and taxes.

completion are generally 5-10% higher than the corresponding Mincer-based estimates even after accounting for taxes and tuition. Additionally, the more general specification reveals a substantial decline in the IRR to college between 1950 and 1960 for blacks that is not reflected in the Mincer-based estimates.

Using the flexible earnings specification, we also examine how estimates depend on assumptions about the length of working life, comparing two extreme cases. Previous estimates assume that individuals work for 47 years regardless of their schooling (i.e. $T'(s) = 1$). An alternative assumption posits that workers retire at age 65 regardless of their education (i.e. $T'(s) = 0$). We find virtually identical results for all years and schooling comparisons for both assumptions about the schooling - worklife relationship.[25] Because earnings at the end of the life-cycle are heavily discounted, they have little impact on the total value of lifetime earnings and, therefore, have little effect on internal rate of return estimates.

4.2   Accounting for Uncertainty in a Static Version of the Model

We have, thus far, computed internal rates of return using fitted values from earnings specifications. Under Mincer's assumptions about the earnings process, when tuition and taxes are negligible, and the working life is the same across schooling levels, these estimates correspond directly to the coefficient on schooling in a Mincer regression. This subsection discusses the interpretation of estimates generated by Mincer's strategy and demonstrates that it makes an implicit assumption about how individuals forecast their earnings. We suggest other ways to estimate the IRR used by agents in making their schooling choices that are based on more plausible expectation formation mechanisms.

Full earnings profiles for all schooling choices are not known by individuals making decisions about schooling, so individuals must use some method of predicting their future earnings. Of course, the same is true for the econometrician calculating internal rates of return to schooling. As previously discussed, it is common in the literature to use log specifications for earnings. Thus, it is common to assume $\ln w = Z\gamma + \varepsilon$, so $w = e^{Z\gamma}e^{\varepsilon}$ and

$$E(w|Z) = e^{Z\gamma}E(e^{\varepsilon}).$$

Assume for the moment that Mincer's assumptions about earnings are correct, so that

---

[25]Results available from authors upon request.

equation (1) describes the true earnings process and that $E(\varepsilon|x,s) = 0$. So far, we have estimated internal rates of return using fitted values for w in place of the true values. That is, we use the following estimate for log earnings: $\hat{w}(s,x) = exp(\hat{\alpha}_0 + \hat{\rho}_s s + \hat{\beta}_0 x + \hat{\beta}_1 x^2)$, where $\hat{\alpha}_0$, $\hat{\rho}_s$, $\hat{\beta}_0$, and $\hat{\beta}_1$ are the regression estimates. This procedure implicitly assumes that when making their schooling choices, individuals take fitted earnings profiles as their prediction of their own future earnings, ignoring any potential person-specific deviations. In other words, we calculate the IRR for an individual at the mean value for ε (zero) at all experience and schooling classifications. Thus our IRR estimator $\hat{r}_I$ solves

$$\sum_{x=0}^{\infty} \frac{\hat{w}(s+j,x)}{(1+\hat{r}_I)^{s+j+x}} - \sum_{x=0}^{\infty} \frac{\hat{w}(s,x)}{(1+\hat{r}_I)^{s+x}} - v \sum_{x=1}^{j} \frac{1}{(1+\hat{r}_I)^{s+x}} = 0,$$

which is the discrete time analogue to the model of equation (3) for two schooling levels s and $s+j$, assuming an infinite horizon. When $v = 0$ (no tuition costs), or if tuition costs are negligible,

$$\text{plim } \hat{r}_I = e^{\rho_s} - 1 \approx \rho_s.$$

This is an ex ante rate of return.

Suppose instead that agents base their expectations of future earnings at different schooling levels on the mean earnings profiles for each schooling level, or on $E(w|s,x)$. In this case, the estimator of the rate of return is given by the root of

$$\sum_{x=0}^{\infty} \frac{E(w(s+j,x)|s,x)}{(1+\hat{r}_I)^{s+j+x}} - \sum_{x=0}^{\infty} \frac{E(w(s,x)|s,x)}{(1+\hat{r}_I)^{s+x}} - \sum_{x=1}^{j} \frac{v}{(1+\hat{r}_I)^{s+x}} = 0 \qquad (7)$$

If $v = 0$ and Mincer's assumptions hold,

$$\frac{e^{\rho_s j}}{(1+\hat{r}_I)^j} \sum_{x=0}^{\infty} \frac{e^{\beta_0 x + \beta_1 x^2} E(e^{\varepsilon(s+j,x)}|s,x)}{(1+\hat{r}_I)^x} = \sum_{x=0}^{\infty} \frac{e^{\beta_0 x + \beta_1 x^2} E(e^{\varepsilon(s,x)}|s,x)}{(1+\hat{r}_I)^x}.$$

If $E[e^{\varepsilon(s,x)}|s,x] = E[e^{\varepsilon(s+j,x)}|s,x]$ for all x, then the two sums are equal and plim $\hat{r}_I = e^{\rho_s} - 1$ as before. In this special case, using $\hat{w}(s,x) = exp(\hat{\alpha}_0 + \hat{\rho}_s s + \hat{\beta}_0 x + \hat{\beta}_1 x^2)$ or $E(w(s,x)|s,x)$ will yield estimates of the internal rate of return that are asymptotically equivalent. However, if $E(e^{\varepsilon(s+j,x)}|s,x)$ is a more general function of s and x, the estimators of the ex ante return will differ.

In the more general case, using estimates of $E(w(s,x)|s,x)$ yields an estimated rate of return with a probability limit

$$\text{plim } \hat{r}_I = e^{\rho_s}\left[M(s,j)\right]^{1/j} - 1 \approx \rho_s + \frac{1}{j}(\ell n\ M(s,j)),$$

where

$$M(s,j) = \frac{\sum\limits_{x=0}^{\infty} e^{\beta_0 x + \beta_1 x^2} E(e^{\varepsilon(s+j,x)}|s,x)(1+r_I)^{-x}}{\sum\limits_{x=0}^{\infty} e^{\beta_0 x + \beta_1 x^2} E(e^{\varepsilon(s,x)}|s,x)(1+r_I)^{-x}}. \tag{8}$$

This estimator will be larger than ρ_s if the variability in earnings is greater for more educated workers (i.e. $M(s,j) > 1$) and smaller if the variability is greater for less educated workers (i.e. $M(s,j) < 1$). If individuals use mean earnings at given schooling levels in forming the expectations that govern their schooling decisions, this estimator is more appropriate. Inspection of Figure 3 reveals that, at young ages, the variability in earnings for low education groups is the highest among all groups. If discounting dominates wage growth with experience, we would expect that $M(s,j) < 1$.[26]

These calculations assume that agents are forecasting the unknown $\varepsilon(s,x)$ using (s,x). If they also use another set of variables q, then these calculations are all conditional on q $(\hat{r}_I = \hat{r}_I(q))$ and we would have to average over q to obtain the average *ex ante* rate of return. If agents know $\varepsilon(s,x)$ at the time they make their schooling decisions, then the *ex ante* return and the *ex post* return are the same, and $\hat{r}_I$ now depends on the full vector of "shocks" confronting agents. Returns would then be averaged over the distribution of all "shocks" to calculate an expected return. Due to the nonlinearity of the equation used to calculate the internal rate of return, the rate of return based on an average earnings profile is not the same as the mean rate of return. Thus, *ex ante* and *ex post* mean rates of return are certain to disagree.

When ρ_s varies in the population, these results must be further modified. Assume ρ_s varies across individuals, that $E(\rho_s) = \bar{\rho}_s$, and that ρ_s is independent of x and $\varepsilon(s+j,x)$ for all x, j. Also, assume $v = 0$ for expositional purposes. Using fitted earnings, $\hat{w}(s,x)$, to calculate internal rates of return yields an estimator, $\hat{r}_I$, that satisfies

$$\text{plim } \hat{r}_I = e^{\bar{\rho}_s} - 1 \approx \bar{\rho}_s.$$

---

[26]More generally if $v \neq 0$, then $\hat{r}_I$ converges to the root of equation (7). Neglecting this term leads to an upward bias, as previously discussed.

This estimator calculates the *ex ante* internal rate of return for someone with the mean increase in annual log earnings $\rho_s = \bar{\rho}_s$ and with the mean deviation from the overall average $\varepsilon(s, x) = \varepsilon(s + j, x) = 0$ for all x.

On the other hand, assuming agents cannot forecast ρ_s, using estimates of mean earnings $E(w(s, x)|s, x)$ will yield an estimator for r with

$$\text{plim } \hat{r}_I = e^{\bar{\rho}_s} \left[kM(s, j) \right]^{1/j} - 1 \approx \bar{\rho}_s + \frac{1}{j}[\ell n \ k + \ell n \ (s, j)],$$

where $k = \frac{E(e^{(s+j)(\rho_s - \bar{\rho}_s)}|s,x)}{E(e^{s(\rho_s - \bar{\rho}_s)}|s,x)}$ and $M(s, j)$ is defined in equation (8).

For $\bar{\rho}_s > 0$, it is straightforward to show that $k > 1$, which implies that everything else the same, the estimator, $\hat{r}_I$, based on mean earnings will be larger when there is variation in the return to schooling than when there is not. Furthermore, the internal rate of return is larger for someone with the mean earnings profile than it is for an individual with the mean value of ρ_s. Again, if agents know ρ_s, we should compute $\hat{r}_I$ conditioning on ρ_s and construct the mean rate of return from the average of those $\hat{r}_I$. Again, the mean *ex post* and *ex ante* rates of return are certain to differ unless there is perfect foresight.

Table 5 reports estimates of the *ex ante* IRR based on the earlier estimation strategy as well as adjusted estimates that use mean earnings within each education and experience category rather than predicted earnings at $\varepsilon = 0$ (both the adjusted and unadjusted estimates account for tuition and progressive taxes). The adjusted estimates generate much lower (and more reasonable) IRR estimates than the unadjusted ones.[27]

Using mean earnings rather than earnings for someone with the mean residual generally leads to lower estimated internal rates of return for most schooling comparisons. Thus, even if the Mincer specification for log earnings is correct, the internal rate of return guiding individual decisions is lower than the Mincer estimated rate of return when individuals base their schooling decisions on average earnings levels within schooling and experience categories. In other words, predicted earnings obtained using the coefficients from a log earnings regression evaluated where $\varepsilon = 0$ is an inaccurate measure of the average earnings within each schooling and experience category.

The adjustment for uncertainty reported in this section makes the strong assumption that all variation is unforecastable at the time schooling decisions are made. A better approach would be to extract components of variation that are forecastable at the time

---

[27]We lack the required panel data on individuals to compute *ex post* rates of return.

schooling decisions are being made (heterogeneity) from components that are unforecastable (true uncertainty). Only the latter components should be used to compute $M(s, j)$. Methods for extracting heterogeneity from uncertainty are available (Carneiro, Hansen, and Heckman, 2003) but require panel data and cannot be applied to Census cross sections. We consider sequential uncertainty in section 6, but first we consider cohort bias within the Mincer framework.

5 How do Cross-sectional IRR Estimates Compare with Cohort-based Estimates?

Thus far, following Mincer and an entire literature, we have estimated returns to schooling using cross-section data, which takes the standard synthetic cohort approach assuming that younger workers base their earnings expectations on the current experiences of older workers. In this case, cross-section and cohort earnings-education-experience profiles are the same. However, if skill prices are changing over time and workers are able to at least partially anticipate these changes, then estimates of the return to different schooling levels based on cross-sectional data may not represent the ex ante rates of return governing human capital investment decisions. While estimates based on cross-section data reflect current price differentials and opportunity costs, they do not capture future skill price differentials that forward-looking individuals would take into account. Consider, for example, a cohort of individuals deciding whether to attend college just prior to a permanent increase in the relative price of college educated workers. Those cohorts will experience higher returns to college than earlier cohorts, which would be reflected in cohort-based estimates but not in cross-section estimates. If cohorts anticipate the rise in the skill premium, they will base their schooling decisions on their true cohort-specific rate of return and not the rate of return estimated from a cross-section of workers. However, if individuals do not anticipate the price change, cross-section estimates may better represent the expected return from attending college that guides their decisions. Thus, expectations about the future play a crucial role in determining whether cross-section or cohort-based estimates influence schooling decisions.

Another possible source of discrepancy between cross-section and cohort-based rate of return estimates is a change in cohort quality. Consider an increase in school quality for

24

cohorts entering the market after some date. If relative skills for some schooling classes increase permanently, then cohort rates of return jump up with the first 'new' cohort and remain higher for all succeeding cohorts. Cross-section estimates only reflect the changes slowly as more and more high quality cohorts enter the sample each year. As a result, they under-estimate true rates of return for all cohorts entering the labor market after the change in school quality, with the bias slowly disappearing as time progresses.

Mincer (1974) explicitly addressed the distinction between cross-section and cohort-based lifecycle earnings patterns. However, he found that patterns for wage growth in a 1956 cross-section of male workers were quite similar to the 1956 to 1966 growth in wages for individual cohorts. At the time he was writing, the empirical discrepancy between cross-section and cohort-based estimates was relatively small, and the data required to compute full life-cycle earnings profiles did not exist. More recently, however, collections of micro data over many years have made cohort analyses possible, and these analyses reveal that wage patterns have changed dramatically across cohorts and that cross-sections no longer approximate cohort or life cycle change (MaCurdy and Mroz, 1995, and Card and Lemieux, 2000). While these studies question whether or not these changes are due to changes in relative skill prices or cohort quality, there is little question that life-cycle earnings profiles based on a cross-section of workers no longer accurately reflect the true earnings patterns for any given cohort. As a result, the rates of return to schooling estimated from cross-sections of workers reported in the previous section are likely to differ from the rates of return faced by cohorts making their schooling decisions.

In our cohort analysis, we focus on the actual returns earned by each cohort without regard for whether changes in those returns over time are due to changes in cohort quality or skill prices. We simply ask how the actual ex post returns earned by individual cohorts compare with returns estimated from a cross-section of individuals at the time those cohorts made their schooling decisions. We use repeated cross-section data from the 1964-2000 Current Population Survey (CPS) March Supplements, comparing cross-section estimates of the return to schooling with estimates that combine all years of the CPS to follow cohorts over their lifecycles. Given the sensitivity noted in the previous sections to changes in functional form specification, we adopt a flexible earnings specification and compute internal rates of return to high school completion (12 vs. 10 years of schooling) and college completion (16 vs. 12 years of schooling) that relax the assumptions that log earnings

are parallel in experience and linear in schooling. Our estimates also take into account average marginal tax rates and tuition costs using the time series generated from CPS data.[28] Because earnings are not observed at every experience level for any cohort in the sample, a fully non-parametric approach is infeasible, and we require a way of extrapolating the earnings function to work experience levels not observed in the data. We assume that log earnings profiles are quadratic in experience for each education classification in a specification that allows the intercept and coefficients on experience and experience-squared to vary by schooling class and year or cohort of data. That is, we estimate log earnings for each year or for each cohort using regressions of the following form given by[29]

$$log(w(s,x)) = \alpha_s + \beta_{0s}x + \beta_{1s}x^2 + \varepsilon_s,$$

where the regression coefficients are allowed to vary by schooling group. Two sets of estimates are generated: (i) regressions are estimated separately for each year of CPS data (to produce a set of cross-section estimates), and (ii) all CPS cross-sections are combined and separate regressions are estimated for each cohort by following them over their lifecycles (to produce a set of cohort-based estimates). Both sets of estimates are used to generate predicted lifecycle earnings profiles for each cohort or cross-section of individuals, which are then used to compute internal rates of return to high school and college by the method described in the previous section.[30]

Figures 7a and 7b show cohort and cross-section high school and college completion IRR estimates for white men, corresponding to CPS estimates in Table 6a. Cross-section estimates are shown for each year of the sample from 1964-1995, and cohort-based estimates are shown for cohorts turning age 18 in 1950 through 1983.[31] The cohort-based estimates reported in Figure 7a reveal relative stability in the return to high school for cohorts making

---

[28]An average marginal tax rate of 25% is assumed for all years after 1994, the final year of tax rates reported in Mulligan and Marion (2000). This corresponds to the average of all rates since 1950, after which rates changed very little from year to year.

[29]In estimating earnings profiles for those with 10 years of education, we combine individuals with 9-11 years, with separate intercept terms for each of the education levels. This is done to increase precision in estimation. See Appendix A for additional details on the coding of the education variables.

[30]In addition to the quadratic specification, we also tried using a cubic and quartic in experience to extrapolate for the missing experience levels. For cohorts with 25 or fewer years of data, extrapolations based on higher order polynomial specifications were unreliable, so we adopted the more parsimonious quadratic specification.

[31]We do not estimate returns for cohorts beyond 1983, since there are too few years of earnings observations for those cohorts to produces stable and reliable estimates.

their high school completion decisions prior to 1960, followed by a large increase in the IRR for cohorts making their decisions over the first half of the 1960s, followed by another period of relative stability. Returns increased from around 10% among 1950-60 cohorts to around 40% for post-1965 cohorts. Cross-section based estimates increase consistently over most of the 1964-1995 period. In general, cross-section estimated rates of return under-estimate the true rates of return earned by cohorts of white men making their schooling decisions in the late 1960s and 1970s. Dramatic differences are also observed for the college-going decision of white men as shown in Figure 7b. While cross-section estimates show declining returns to college over the 1970s (from 12% down to 8%), cohort-based estimates show increasing returns over that period. After declining over time for cohorts making their college-going decisions in the 1950s, the cohort-based rates of return to college increase sharply in the 1960s, stabilize (or even fall) briefly in the early 1970s, then continue on a sharp upward trend through the early 1980s. The rate of return estimated from cross-sections of individuals does not begin to increase until much later, in 1980, rising quickly until the mid 1980s. Cross-section estimates over-estimate the rate of return faced by cohorts making their college attendance decisions around 1965 by as much as 4 percentage points, while estimates in the early 1980s under-estimate the return by nearly the same amount. Table 6b reports comparable numbers for black men. Again, in recent years, cohort rates of return exceed cross sectionally estimated rates.

If the observed discrepancies between cross-section and cohort-based estimated rates of return are due to price changes over time that could be at least partly anticipated or are due to changing cohort quality, then cross-section estimates would not reflect the rates of return that govern schooling decisions. On the other hand, if changes in skill prices were entirely unanticipated, then cross-section estimates may provide a better indication of the returns governing schooling decisions than would the actual returns experienced by each cohort. A better understanding of the underlying causes for such dramatic changes in wages and of individual expectations are needed.

In summary, cross-section estimates of the rate of return to schooling should be cautiously interpreted, particularly when skill prices are changing over time or when cohort quality is changing. If one is interested in empirically estimating historical rates of return, a cohort analysis is clearly preferable. Data from 1964-2000 March CPS suggest that returns estimated from a cross-section of workers are not only biased in levels, but they

also suggest time patterns that sometimes differ from those obtained using a cohort-based estimation strategy. If one is interested in estimating the rates of return governing school investment decisions, then whether to use cross-section or cohort-based estimates depends on the extent to which individuals are able to forecast future changes in wages and skill prices.

We next turn to considering the impact of sequential resolution of uncertainty on conventional estimates of returns to schooling.

6 The Internal Rate of Return and The Sequential Resolution of Uncertainty

Human capital theory was developed in an era before the modern tools of dynamic decision making under uncertainty were fully developed. Concepts central to human capital theory like the internal rate of return are not generally appropriate to the evaluation of investment programs under sequential resolution of uncertainty. A more general analysis is required.

For two reasons, the dynamic nature of schooling suggests that the returns to education may include an option value (Weisbrod, 1962). First, the return to one year of school may include the potential for greater returns associated with higher levels of education when the returns to school are not constant across all schooling levels. For example, finishing high school provides access to college, and attending college is a necessary first step to obtaining a college degree. Given the large increase in earnings associated with college completion, the total return to high school or college attendance may include the potential for even greater returns associated with finishing college. Mincer's assumption that earnings are log linear in schooling implicitly rules out this type of option value.

Second, when there is uncertainty about college costs or future earnings and when each additional year of schooling reveals new information about those costs or earnings, the full returns to schooling will include the expected value of newly revealed information. Consider the following example. Finishing high school opens the possibility of attending college if tuition costs and opportunity costs turn out to be low. The returns to high school completion, therefore, include both the expected increase in earnings associated with completing high school and the *ex ante* expected value of the information learned about college costs. The value of this information depends on the probability that the individual

decides to continue on to college and the expected return if he does so. Failing to finish high school precludes an individual from learning about these costs and eliminates the college option entirely. Earnings each period may also be uncertain, and the decision to continue on in school may depend on both current and expected future labor market conditions. By ignoring uncertainty, the literature based on the Mincer earnings equation neglects this source of option value as well. Both sources of option values to schooling suggest that education decisions are made sequentially and should not be treated as a static discrete choice problem made once in a lifetime by individuals – the traditional approach used in human capital theory. (See, *e.g.*, Mincer, 1958, Willis and Rosen, 1979, or Willis, 1986).

The empirical evidence presented earlier (also see Bound, Jaeger and Baker 1995, Heckman, Layne-Farrar and Todd, 1996, Solon and Hungerford, 1987) strongly rejects Mincer's (1958) implicit assumption that internal rates of return to each year of schooling are identical and equal to a common interest rate. This alone undermines the interpretation of the coefficient on schooling in a log earnings regression as a rate of return. But, this non-linearity, combined with the sequential resolution of uncertainty, creates additional problems for estimating rates of returns using Mincer regressions. Because the returns to college completion are high, it may be worthwhile to finish high school to keep the option of college open. The total return to high school and earlier schooling choices may, therefore, include a non-trivial option value. To analyze this option value, we present two simple dynamic models with uncertainty about the value of future schooling choices given an individual's current education. Following most of the literature, we assume that individuals maximize the expected value of lifetime earnings given the available information.

To gain some understanding about the separate roles of nonlinearity and uncertainty in generating option values, first consider the option value framework of Comay, Melnik, and Pollatschek (1973), which assumes that there is no uncertainty about earnings conditional on final schooling attainment but that individuals face some exogenously specified probability ($\pi_{s+1,s}$) of being accepted into grade $s+1$ if they choose to apply after finishing grade s.[32] They face a lottery where the chance of being admitted to the next round of schooling does not depend on earnings values. For someone attending exactly s years of school, define

---

[32]They also consider the probability of failing conditional on attending the next grade. The results from such an analysis are quite similar to those discussed here.

the discounted present value of lifetime earnings as of the schooling completion date as:

$$W_s = \sum_{x=0}^{T} (1+r)^{-x} w(s,x).$$

The interest rate, r, is exogenously specified. If an individual that chooses to apply for grade $s+1$ is rejected, he begins working immediately, earning W_s. In this environment, the total expected value of attaining $s \in \{1, 2, ..., \bar{S}\}$ years of school, given the information available at $s-1$, is

$$E_{s-1}(V_s) = (1 - \pi_{s+1,s})W_s + \pi_{s+1,s}E_{s-1}max\left\{W_s, \frac{E_s(V_{s+1})}{1+r}\right\}$$

for $s < \bar{S}$ and $E_{\bar{S}-1}(V_{\bar{S}}) = W_{\bar{S}}$. This assumes that each grade of school takes one period and that direct costs of schooling are negligible.

The *ex ante* option value of grade s as perceived at $s-1$ is defined as the difference between the total expected value of that opportunity, $E_{s-1}(V_s)$, and the present discounted value of earnings if the person does not continue in school, W_s:

$$
\begin{aligned}
O_{s,s-1} &= E_{s-1}[V_s - W_s] \\
&= E_{s-1}max\left\{0, \pi_{s+1,s}\left(\frac{E_s(V_{s+1})}{1+r} - W_s\right)\right\} \\
&= max\left\{0, \pi_{s+1,s}\left(\frac{E_{s-1}(V_{s+1})}{1+r} - W_s\right)\right\},
\end{aligned}
$$

where the final equality follows from the fact that there is no uncertainty about earnings conditional on the final schooling outcome. Notice, that when earnings grow with an additional year of schooling at the same rate as the interest rate, as is assumed by Mincer (1958), or if the growth in earnings is at the same rate as the individual-specific interest rate in the accounting identity model, then $W_s = \frac{W_{s+1}}{1+r}$ for each individual and all s. Mincer's assumption of linearity of log earnings in schooling implicitly rules out any option value of schooling in the present context.[33] Intuitively, if the earnings profiles associated with all

---

[33]Proof: $V_{\bar{S}} = W_{\bar{S}}$ at $\bar{S}$, so

$$E_{\bar{S}-2}\left(V_{\bar{S}-1}\right) = \left(1 - \pi_{\bar{S},\bar{S}-1}\right)W_{\bar{S}-1} + \pi_{\bar{S},\bar{S}-1}\max\left\{W_{\bar{S}-1}, \frac{W_{\bar{S}}}{1+r}\right\},$$

since there is no uncertainty about earnings conditional on final schooling levels. For proportional earnings growth at rate r, both versions of the Mincer model imply that $W_s = \frac{1}{1+r}W_{s+1}$ for all s. Thus, people may differ in their earnings levels and face different individual specific interest rates as in the accounting

schooling choices provide the same present value when discounted back to the same date, then there is no value attached to the possibility of continuation. Thus linearity of log wages in years of schooling with a growth rate equal to the interest rate implies no option value of education in the Comay, Melnik, and Pollatschek (1973) framework.

More generally, this model does generate option values when future wage growth is greater than $1 + r$. For example, if college graduation offers large returns, finishing high school will carry an option value since there is some probability that an individual will be accepted into college. In this case, the total value of a high school degree includes the value of a lottery ticket that pays the rewards of a college degree to 'winners'. The option value of high school represents the value of this lottery ticket.

The Comay, Melnik, and Pollatschek (1973) model assumes that the probability of transiting to higher grades (conditional on the desire to do so) is exogenous. Schooling is a sequence of lotteries. Because there is no uncertainty about future earnings paths conditional on schooling or about the future costs of or returns to schooling, their model isolates the role played by a non-linear log earnings - schooling relationship in determining option values.

We next present an economically more interesting model of the schooling choice problem that incorporates uncertainty in future earnings (or school costs) and sheds light on the impact of that uncertainty on the option value of education. Suppose that there is uncertainty about net earnings conditional on s, so that actual lifetime earnings for someone with s years of school are

$$W_s = \left[\sum_{x=0}^{T} (1+r)^{-x} w(s, x) \right] \epsilon_s.$$

This form of uncertainty is a one time, schooling specific shock. We assume that $E_{s-1}(\epsilon_s) = 1$ and define expected earnings associated with schooling s conditional on current schooling $s - 1$,

$$\bar{W}_s = E_{s-1}(W_s).$$

---

identity model. They may also face different $\pi_{s+1,s}$. For any sequence of $\pi_{s+1,s}$ and r, we obtain

$$E_{\bar{S}-2}\left(V_{\bar{S}-1}\right) = W_{\bar{S}-1} = \frac{W_{\bar{S}}}{1+r}.$$

Backward induction produces $E_{s-2}\left(V_{s-1}\right) = W_{s-1} = \frac{W_s}{1+r}$ for all s, which implies no option value for any schooling level.∎

The disturbance, ϵ_s, may represent a shock to additional schooling costs or to current earnings that is revealed after the decision to attend grade s is made but prior to any future schooling decisions. Individuals with s years of schooling must decide whether to quit school, receiving lifetime earnings of W_s, or to continue on in school for an additional year and receive an expected lifetime earnings of $E_s(V_{s+1})$.

The decision problem for a person with s years of schooling given the sequential revelation of information is to go to another year of school if

$$W_s \leq \frac{E_s(V_{s+1})}{1+r},$$

so

$$V_s = \max \left\{ W_s, \frac{E_s(V_{s+1})}{1+r} \right\}$$

for $s < \bar{S}$. At the maximum schooling level, $\bar{S}$, after information is revealed, we have $V_{\bar{S}} = W_{\bar{S}} = \bar{W}_{\bar{S}} \epsilon_{\bar{S}}$.

Notice, that the endogenous probability of going on from school level s to $s+1$ is

$$p_{s+1,s} = Pr\left(\epsilon_s \leq \frac{E_s(V_{s+1})}{(1+r)\bar{W}_s} \right),$$

where $E_s(V_{s+1})$ may depend on ϵ_s, and the average earnings of a person who stays at schooling level s is

$$\bar{W}_s E_{s-1}\left(\epsilon_s | \epsilon_s > \frac{E_s(V_{s+1})}{(1+r)\bar{W}_s} \right). \tag{9}$$

Thus, the expected value of schooling level s as of current schooling $s-1$ is:

$$E_{s-1}(V_s) = (1 - p_{s+1,s})\bar{W}_s E_{s-1}\left(\epsilon_s | \epsilon_s > \frac{E_s(V_{s+1})}{(1+r)\bar{W}_s} \right) + p_{s+1,s}\left(\frac{E_{s-1}(V_{s+1})}{1+r} \right).$$

The *option value* of schooling s, given that the agent has the information about $s-1$ outcomes, is the difference between the expected value of the earnings associated with schooling s and the corresponding value function:

$$O_{s,s-1} = E_{s-1}\left[V_s - W_s \right].$$

We can define sequential *option values* for all levels of s. Clearly, option values are non-negative for all schooling levels, since $V_s \geq W_s$ for all s. The option value for the highest schooling level is zero, since $V_{\bar{S}} = W_{\bar{S}}$.

The ex ante rate of return to schooling s at level $s-1$ is

$$R_{s,s-1} = \frac{E_{s-1}(V_s) - W_{s-1}}{W_{s-1}}.$$

Accounting for direct costs of schooling C_s, we may write this as

$$\widetilde{R}_{s,s-1} = \frac{E_{s-1}(V_s) - (W_{s-1} + C_{s-1})}{W_{s-1} + C_{s-1}}$$

This assumes that tuition or direct costs are incurred up front and returns are revealed one period later.

This is an appropriate *ex ante* rate of return concept because if

$$W_{s-1} + C_{s-1} \leq \frac{E_{s-1}(V_s)}{1+r},$$

i.e.

$$r \leq \frac{E_{s-1}(V_s) - (W_{s-1} + C_{s-1})}{W_{s-1} + C_{s-1}} = \widetilde{R}_{s,s-1},$$

then it would be optimal to advance one more year of schooling (from $s-1$ to s) given the return on physical capital r.

This analysis highlights the sequential nature of the schooling choice problem under uncertainty. The schooling allocations that arise out of this framework will differ from those implied by the standard Mincer approach, which uses a static decision rule based on expected earnings profiles as of some initial period. The approach taken here recognizes that individuals face uncertainty at the time they make their schooling decisions and that some of that uncertainty is resolved after each decision is made. After completing a schooling level, individuals observe the shock associated with that level and can base their decision to continue in school on its realization. This creates an option value of attending school. If the shock is bad, one can always continue to the next higher schooling level.

It is interesting to note that even when $\bar{W}_s = \frac{\bar{W}_{s+1}}{1+r}$ as assumed by Mincer's models, there is still an option value in this framework. This is because completing $s+1$ reveals new information about the actual returns associated with that choice and offers the option of continuing on to level $s+2$ with fresh draws of the ϵ. In contrast to its role in the simple Comay, Melnik, and Pollatschek (1973) model, Mincer's assumption that log earnings are linear in schooling does not rule out option values once we introduce shocks to schooling costs or earnings. More generally, when future earnings choices (W_{s+2} vs. W_{s+1} in this

example) offer very large expected returns, the option value might be quite substantial – both sources for option values are operating.

Conventional rate of return calculations for comparing schooling levels s and $s+1$ base the calculation only on the earnings streams associated with s and $s+1$. Taking into account the option value also requires consideration of the earnings stream associated with higher schooling levels. That is, the value of graduating from high school instead of dropping out is affected by the expected earnings associated with graduating from college. Keane and Wolpin (1997) and Eckstein and Wolpin (1999) develop sequential models of schooling that are more general than the model presented here. Their econometric procedures implicitly incorporate the option value of schooling, but they do not discuss it, nor do they present numerical estimates of its importance.[34]

To clarify the role of uncertainty and non-linearity of log earnings in terms of schooling, we present simulations of a five schooling-level version of our model with uncertainty in Tables 7a and 7b. In both tables, we assume an interest rate of $r = 0.1$ and that ϵ_s is independent and identically distributed log-normal: $log(\epsilon_s) \sim N(0, \sigma)$ for all s.[35] We assume that $\sigma = 0.1$ in the results presented in the tables. Table 7a reports various outcomes related to the returns to schooling when we assume log earnings are linear in schooling (i.e. $\bar{W}_{s-1} = \bar{W}_s/(1+r)$). Schooling continuation probabilities ($p_{s,s-1}$) and the proportional increase in $\bar{W}$ associated with an increase in schooling from $s-1$ to s are shown. By assumption, the latter equals $r = 0.1$ for all education levels. Column 4 displays the proportional increases in observed earnings (where observed earnings are measured by equation 9) from period $s-1$ to s, which are always less than r. In the presence of

---

[34]In the ordered choice model of Cameron and Heckman (1998) and Hansen, Heckman and Mullen (2003), there is no option value arising from sequential resolution of uncertainty, because of the assumed one sided nature of the information revelation process. But, there may be option value arising from the nonlinearity of the model. Is is interesting to note that schooling choice models that assume no information updating fit the data on schooling choices as well as models that incorporate such updating.

[35]We also considered models with an AR(1) process for the shocks: $log(\epsilon_s) = \rho log(\epsilon_{s-1}) + v_s$ where $v_s \sim N(0, \sigma)$. The case where $\rho = 0$ corresponds to Tables 7a and 7b. For $\rho = 1$, $E(\epsilon_{s+1}|\epsilon_s) = \epsilon_s$ and a good or bad shock affects expected future outcomes in the same proportion as current outcomes. Thus, the outcome of ϵ_s has no effect on schooling decisions. In the linear case corresponding to Table 7a, expected rates of return as measured by $E_{s-1}(R_{s,s-1})$ range between those reported in the table (when ρ is near zero) and the linear increase in earnings, $r = 0.10$ (when ρ is near one). Expected returns for the more general non-linear case differ little from those shown in Table 7b, since nearly everyone chooses to attend the highest level of schooling regardless of the value for ρ. This implies that returns always reflect the expected increase in earnings between the current schooling level and the highest possible schooling level, which is, on average, independent of ρ.

uncertainty, self-selection leads to a substantial downward bias in the observed returns to schooling, especially for higher schooling transitions. Option values as a fraction of the total expected value of a schooling level $(O_{s,s-1}/E_{s-1}(V_s))$ are reported in column 5. They show the expected decline with schooling levels attained. The final three columns report average measures of the return to schooling for different sets of individuals. Column 6 reports the average return for the entire population $(E_{s-1}[R_{s,s-1}])$, while column 7 reports estimates of the return for those who choose to continue on to grade s ("treatment on the treated") and column 8 reports the expected return that would be received by those who choose not to continue in school ("treatment on the untreated"). Comparing average returns with the proportional increase in $\bar{W}$ or in observed earnings, we see that total rates of return to schooling are substantially higher for all but the final schooling transition due to the added option value of school and self-selection that takes place. When log earnings are linear in schooling, real returns are actually declining in accumulated schooling since option values are decreasing in s.[36] Returns for those who choose to continue are noticeably larger than average returns, while returns for those who choose not to continue are all less than r. The least squares estimate of the rate of return to school (i.e. the coefficient on schooling in a log earnings regression) is only 0.063, far below the true average rate of return or treatment on the treated estimates. It also under-estimates the rate of increase in expected earnings, $\bar{W}_s$, and does not accurately reflect the pricing relationship for wages and schooling. Even with linearity of mean log earnings in schooling, Mincer-based estimates of the return are substantially downward biased in the presence of uncertainty. Not surprisingly, this bias (along with option values) disappears as the variance in ϵ_s goes to zero. However, we find a bias as large as -0.01, roughly 10% of the true return, when σ is as low as 0.01.[37]

Table 7b adds nonlinearity to this model to demonstrate its added effect on rates of return and option values. The simulation reported in this table assumes that increases in population mean log earnings from the first to the second and third to fourth levels of school are both 0.1, but the increase associated with going from level two to three

---

[36]We have assumed that individuals cannot choose to recall wage profiles associated with earlier schooling choices (i.e. someone with s years of school cannot choose to work for $W_{s-1}, W_{s-2},...$) if they receive a low realization for W_s. Allowing for recall of earlier schooling opportunities would provide a force offsetting the tendency for option values to decline with schooling if we continued to define the option value as $E_{s-1}(V_s - W_s)$. On the other hand, if option values were re-defined as $E_{s-1}(V_s) - max\{W_1, W_2, ..., E_{s-1}(W_s)\}$ with the introduction of recall, they would tend to decline even faster with schooling.

[37]Results available from the authors upon request .

is 0.3 and from four to five is 0.2. This roughly mimics the patterns observed in the later Census years with schooling levels three and five representing high school and college graduation, respectively. These simulations show substantially larger returns to the lower school transitions as a result of the sizeable sheepskin effects in later years. Option values are particularly large in early school years. In general, the greater the nonlinearity, the greater the option value. Estimates from a Mincer regression suggest a rate of return of only 0.060, substantially less than the true average returns or treatment on the treated estimates, which range from 0.21 to 0.46. While true returns have increased relative to Table 7a, the Mincer estimate actually declines slightly. Because most individuals are choosing to continue to higher schooling levels in this simulation, there is little difference between average returns and estimated treatment on the treated parameters.

The simulations presented in Tables 7a and 7b point to the potentially important role of both sources of option values in determining total returns to schooling. Turning to real data, we use the nonparametrically estimated earnings profiles for white males in the 1990 Census to compute the option value of high school completion and college attendance for a range of reasonable schooling transition probabilities, p, and interest rates, r. These estimates would provide unbiased measures of the option value within the framework of Comay, Melnik, and Pollatschek (1973) where $p = \pi_{s+1,s}$ for the schooling levels we examine. Within the model of sequential resolution of uncertainty, where p represents $Pr\left(\epsilon_s \leq \frac{E_s(V_{s+1})}{(1+r)\bar{W}_s}\right)$, they under-estimate the option value and return to school, since observed earnings represent $\bar{W}_s E_{s-1}\left(\epsilon_s | \epsilon_s > \frac{E_s(V_{s+1})}{(1+r)\bar{W}_s}\right)$ rather than $\bar{W}_s$ (i.e. observed earnings are based on a sample selecting not to continue).[38] Table 8 reports the average discounted lifetime earnings for individuals making different schooling choices, denoted by $\hat{W}_s$. It also reports the total expected value of a schooling choice, $E_{s-1}(V_s)$, the implied option value, $\hat{O}_{s,s-1}$, and return to school, $R_{s,s-1}$. The table reports results for interest rates of 7% and 10% and transition probabilities ranging from 0.1 to 0.5 (empirically, about half of all 1990 high school graduates attended college and about half of those went on to graduate). As expected, both the present value of earnings for each schooling choice and the option value of continuing are declining in the interest rate. Option values rise with increases in the transition probability. The option value for high school completion ranges from a low of

---

[38]We note, however, that estimates of the value function are unbiased in both models.

only \$370 when interest rates are 10% and $p = 0.1$ to a high of \$22,000 when interest rates are 7% and $p = 0.5$. Much of the option value comes from the probability of completing college and not just attending college, because the difference in earnings between high school graduates and those with some college is quite small. Accordingly, option values are noticeably higher for college attendance, reaching a high of \$35,000 when interest rates are 7% and $p = 0.5$. Simply comparing the earnings streams for two schooling levels fails to recognize a potentially important component of the returns to education. Rates of return, shown in the final two columns, increase by about 50% for college attendance when the transition probability is raised from 0.1 to 0.5. Returns to high school completion are less sensitive to assumptions about p and option values. Failing to consider option values leads to biased estimates of the true return to schooling. A basic question is whether the traditional internal rate of return has any value as an investment evaluation criterion.

In a model with sequential resolution of uncertainty, single crossings of earnings profiles, a near universal feature of schooling-earnings data, do not guarantee unique internal rates of return when option values are taken into account. Hirshleifer (1970) shows that there is always a unique positive internal rate of return when comparing two expected earnings streams which cross at only one age. This is typically the case when comparing the earnings profiles for any two schooling levels. Accounting for options to continue in school, it is possible for multiple roots to arise in the computation of more sophisticated internal rates of return that account for the option value of schooling even if earnings are monotonically increasing in schooling for workers conditional on age.

To see why, consider a model of exogenous transition probabilities (like that of Comay, Melnick, and Polatschek, 1973) for the case where earnings are zero until the end of school, age s, at which time they jump up to $\alpha_s + \beta s$ and linearly increase thereafter at rate $\beta > 0$.[39] As long as $\alpha_s > \alpha_{s'}$ for all $s > s'$, any two earnings streams will only cross once. Letting $w_s(a)$ denote the earnings for someone with s years of school at age a, we have

$$w_s(a) = \begin{cases} 0 & \text{if } a < s \\ \alpha_s + \beta a & \text{if } a \geq s. \end{cases}$$

Consider three schooling choices, $s \in \{0, s_1, s_2\}$. Suppose p is the probability that someone with $s_1 < s_2$ years of school continues on to s_2 years. The expected earnings at

---

[39]The example can easily be extended to account for tuition costs and more general lifecycle earnings profiles.

age a of someone choosing to attend s_1 years of school with the option of continuing will be $\bar{w} = (1-p)w_{s_1} + pw_{s_2}$.

For $\alpha_0 < \alpha_{s_1} < \alpha_{s_2}$, $\bar{w}$ will cross $w_0(a)$ three times whenever

$$\frac{\alpha_0 + \beta s_1}{\alpha_{s_1} + \beta s_1} < 1 - p < \frac{\alpha_0 + \beta s}{\alpha_{s_1} + \beta s}$$

for any s, where $s_1 < s < s_2$.[40] This is illustrated in Figure 8. Because there are three crossings between w_0 and $\bar{w}$, internal rate of return equations can generate up to three possible positive roots. Even if pairwise earnings streams cross only once, there may be multiple internal rates of return, invalidating their use as a guide to selecting human capital investment projects.

In general, the schooling transition probability is not exogenous. Multiple roots are even more likely in our model with sequential resolution of uncertainty, since the transition probability changes with the discount rate. Writing equations out explicitly in terms of r, we obtain

$$
\begin{aligned}
E_{s-1}\left(V_s\left(r\right)\right) &= \mathrm{Pr}_{s-1}\left(\epsilon_s \geq \frac{E_s\left[V_{s+1}\left(r\right)\right]}{(1+r)\bar{W}_s\left(r\right)}\right) \bar{W}_s\left(r\right) E_{s-1}\left(\epsilon_s | \epsilon_s \geq \frac{E_s\left[V_{s+1}\left(r\right)\right]}{(1+r)\bar{W}_s\left(r\right)}\right) \\
&\quad + \mathrm{Pr}_{s-1}\left(\epsilon_s < \frac{E_s\left[V_{s+1}\left(r\right)\right]}{(1+r)\bar{W}_s\left(r\right)}\right) \frac{E_{s-1}\left[V_{s+1}\left(r\right)\right]}{(1+r)}
\end{aligned}
$$

The IRR is the value (or values) of r that solve

$$W_s\left(r\right) = \frac{E_{s-1}\left(V_{s+1}\left(r\right)\right)}{1+r}.$$

Take a three period example. In this case, the IRR for the second level of schooling solves

$$
\begin{aligned}
\bar{W}_1\left(r\right) &= \mathrm{Pr}_1\left(\epsilon_2 \geq \frac{\bar{W}_3\left(r\right)}{(1+r)\bar{W}_2\left(r\right)}\right) \frac{\bar{W}_2\left(r\right)}{1+r} E_1\left(\epsilon_2 | \epsilon_2 \geq \frac{\bar{W}_3\left(r\right)}{(1+r)\bar{W}_2\left(r\right)}\right) \\
&\quad + \mathrm{Pr}_1\left(\epsilon_2 < \frac{\bar{W}_3\left(r\right)}{(1+r)\bar{W}_2\left(r\right)}\right) \frac{\bar{W}_3\left(r\right)}{(1+r)^2}
\end{aligned}
$$

The fact that the continuation probabilities also depend on r makes multiple roots more likely. To gain some intuition in this case, take a limiting case where the variance of ϵ_2 goes to zero. This implies that the probability of continuing to level three will be either zero or

---

[40]The left hand side of this condition ensures that $\bar{w}$ jumps from zero to some point above w_0 at age s_1. Then, $\bar{w}$ increases with age at a slower rate than does y_0. The right hand condition guarantees that at some later age s, $\bar{w}$ will be below w_0. Finally, we know at age s_2, $\bar{w}$ will jump above w_0, since both α_{s_1} and α_{s_2} are both greater than α_0.

one, depending on whether or not $\bar{W}_2$ is greater or less than $\frac{\bar{W}_3}{(1+r)}$. We may, therefore, get two valid solutions to the above IRR equation:

Case 1 (individual always continues): r_1^* satisfies

$$\bar{W}_1\left(r_1^*\right) = \frac{\bar{W}_3\left(r_1^*\right)}{\left(1+r_1^*\right)^2} > \frac{\bar{W}_2\left(r_1^*\right)}{1+r_1^*}.$$

The latter inequality guarantees that the person always wants to continue to schooling level three upon reaching level two.

Case 2 (individual never continues): r_2^* satisfies

$$\bar{W}_1\left(r_2^*\right) = \frac{\bar{W}_2\left(r_2^*\right)}{\left(1+r_2^*\right)} > \frac{\bar{W}_3\left(r_2^*\right)}{\left(1+r_2^*\right)^2}.$$

The latter inequality guarantees that the person always stops his schooling at level two.

Both of these cases may arise if log earnings are not parallel in experience. Consider the case where wage gaps are small initially and large later in the lifecycle. In this case, r_1^* would be less than r_2^*. In Case 1, the high wage differential later on is not discounted very much, so the individual always wants to attend schooling level three. A low IRR must, therefore, equate level one earnings with discounted level three earnings. On the other hand, the high late wage differential may be discounted so much with a high discount rate that the individual never chooses to go on to college at that rate. In this case, a high IRR, r_2^*, must equate level one earnings with discounted level two earnings. These examples are extreme, but multiple roots can arise more generally as long as the variance of ϵ_s is not too large. This type of multiplicity could also come more directly out of the Comay, *et al.* (1973) type of model, where the probability of continuing to level three would be either zero (if individuals do not want to continue) or p (if individuals wish to continue), depending on the discount rate. Such multiplicity is likely to be empirically relevant in recent years given the lack of parallelism in log earnings profiles.

These issues call into serious question the usefulness of internal rates of return as a measure of the return to education in an environment when the schooling decision is dynamic and sequential. A central tool of policy evaluation from classical human capital theory becomes meaningless in the presence of option values. Criterion (9) does not suffer from this criticism and is the appropriate measure of rate of return to use. In the absence of sequential resolution of uncertainty and option values, $R_{s,s-1}$ is equal to the internal rate of return.

Empirical work on the option value of schooling is in its infancy. If option values are empirically unimportant, conventional investment evaluation methods based on the IRR may well be informative. However, the analysis presented throughout this paper suggests that the Mincer model will not estimate theoretically appropriate rates of return to schooling. Even in the absence of option values, other key assumptions required to equate Mincer coefficients with internal rates of return are violated.

7   Conclusions

The earnings function is at the heart of labor economics, and Mincer's work has had a major influence on how labor economists specify earnings relationships. The Mincer earnings specification has been widely used over the last five decades and continues to be applied in recent work. It captures many important empirical regularities, such as concavity of log earnings age and experience profiles, steeper profiles for persons with more years of education, and a U-shaped interpersonal variance of earnings over the life-cycle. Mincer (1958, 1974) provided two theoretical motivations for his specification, one based on a compensating differentials principle and a second based on an accounting identity model of human capital formation. The two models are economically distinct, but both lead to very similar empirical specifications of the wage equation.

Using 1940-1990 Census data, this paper examines the empirical support for Mincer's earnings specification and the key implications of his accounting identity model. Data from 1940-1950 provide support for the model, and, in particular, for parallelism in experience-based log earnings profiles across different schooling levels. Eyeball evidence supports the hypothesis of parallelism in 1960 and 1970, even though parallelism is rejected by formal econometric tests. However, data from 1980 and 1990 Census years are not supportive and show convergence with experience in the log earnings profiles for different schooling groups linked to vintage effects. The data on U-shaped residual variances is supportive of Mincer's model in all census years. Verification of this key prediction confirms the importance of human capital investment in accounting for lifecycle earnings.

Under the simplifying assumptions used by Mincer and others, the coefficient on schooling in a log earnings regression should equal the internal rate of return to schooling. However, many of these assumptions are no longer appropriate. We show in this paper that

log earnings do not increase linearly with schooling, and experience profiles for log earnings are not parallel across schooling types in recent decades. Moreover, tuition costs and income taxes are non-negligible and should be taken into account in calculating returns to higher schooling levels. Allowing for a more flexible earnings specification and adjusting estimates of the internal rate of return to schooling for tuition and taxes leads to substantially different conclusions about the value of schooling as summarized in Tables 6a and 6b.

The Mincer specification dramatically understates the return to finishing high school relative to a more flexible model. The Mincer specification would also suggest that the IRR to high school completion rose slightly from 1940-1990, but our flexible model indicates that the IRR nearly doubled from 1970-1990. Estimates based on CPS data (both cross-section and cohort estimates) also indicate a much higher IRR to completing high school than the IRR implied by the Mincer model.

The differences in IRR estimates obtained from the different models and datasets are less pronounced for college completion. The Mincer equation leads to an overstatement of the IRR to college completion (12-16) of about 2-3% for census years 1950-1990 in comparison to a flexible model estimated on the same data. The cross-section and the cohort-based estimates derived using CPS data are fairly close to those from the Census data. Both the Census and CPS estimates for white men indicate a fall in the return to college in the 1970's and a rise in the 1980's.[41]

We also establish that estimates based on cross-section data can be misleading in times of economic transition, when estimated returns to schooling may be only loosely related to the true returns facing any cohort. In general, cross-sectional estimates will under-estimate (over-estimate) the returns to schooling for individuals making their schooling choices just prior to increases (decreases) in the price of skill. While Mincer recognized this problem in his seminal work, it was not empirically important in the data he analyzed. Now it is an important feature of the data.

The original Mincer model did not explicitly account for uncertainty. We incorporate uncertainty in both static and dynamic forms. Adjusting rates of returns for uncertainty

---

[41]These patterns are also documented in Katz and Murphy (1992) and Gottschalk (1997), which both use cross-sectional CPS data. Katz and Murphy (1992) attribute the fall in the 1970's primarily to an increase in the supply of college graduates. However, we do not observe a similar decline in returns to college for the black sample.

about future earnings in a static setting substantially reduces estimated rates of return but does not change the qualitative conclusions obtained from a model that assumes perfect certainty.

Accounting for non-geometric growth in earnings with years of schooling and sequential resolution of uncertainty gives rise to option values which can be substantial. Existence of option values calls into question the usefulness of a standard tool of human capital analysis - the internal rate of return to schooling. We produce general examples in which there are multiple internal rates of return to schooling even though earnings streams for each schooling level only cross once. This analysis calls into question the validity of the entire enterprise aimed at estimating "the" internal rate of return. We demonstrate that a more standard measure of the rate of return to schooling (grounded in modern capital theory) is widely applicable in the presence of uncertainty and option values. Under the sequential resolution of uncertainty, the Mincer model does not provide a valid estimate of this return or the internal rate of return. However, the empirical importance of uncertainty and option values in determining schooling choices remains to be demonstrated.

Although fairly accurate for earlier Census data, the Mincer model no longer produces even roughly valid estimates of rates of return to education and is not a valid guide to the evaluation of educational policy. However convenient it is, it is no longer an accurate guide to identifying pricing relationships or rates of return to schooling. A more general dynamic analysis of the earnings function that accounts for tuition, taxes, nonlinearity in schooling, non-separability between experience and schooling, and uncertainty is required.

References

Ashenfelter, O. and A. B. Krueger (1994): "Estimates of the Economic Return to Schooling from a New Sample of Twins," *Quarterly Journal of Economics*, 84(5):1157-73.

Ashenfelter, O. and C. Rouse (1998): "Income, Schooling and Ability: Evidence from a New Sample of Identical Twins," *Quarterly Journal of Economics*, 113(1):253-84.

Barro, R. and C. Sahasakul (1983): "Measuring the Average Marginal Tax Rates from the Individual Income Tax," *Journal of Business*, 56(4), 419-452.

Becker, G. S. (1964): *Human Capital.* New York: Columbia University Press for the National Bureau of Economic Research.

Becker, G. S. and Chiswick, B. R. (1966): "Education and the Distribution of Earnings," *American Economic Review, Proceedings*, 56:358-369.

Behrman, J. and N. Birdsall (1983): "The Quality of Schooling: Quantity Alone is Misleading" *American Economic Review.* Vol. 73 (5). p 928-46. December 1983.

Ben-Porath, Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 75(4), part 1, 352-65.

Bils, M. and P. Klenow (2000): "Does Schooling Cause Growth?" *American Economic Review*, 90(5), 1160-1183.

Bound, J., D. Jaeger and R. Baker (1995): "Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable Is Weak," *Journal of the American Statistical Association.* Vol. 90 (430). p 443-50.

Cameron, S. and J. Heckman (1998): "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy.* Vol. 106 (2). p 262-333.

Card, D. and A. Krueger (1992): "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States" in *Journal of Political Economy.* Vol. 100 (1). p 1-40.

Card, D. (1995): "Earnings, Schooling, and Ability Revisited," in *Research in Labor Economics,* Vol. 14, edited by S. Polachek, Greenwich, Conn: JAI Press, 23-48.

Card, D. (1999): "The Causal Effect of Education on Earnings." In *Handbook of Labor Economics, Vol. 3A*, edited by Orley Ashenfelter and David Card. (Amsterdam: North

Holland).

Card, D. and T. Lemieux (2000): "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," NBER Working Paper No. 7655.

Carneiro, P. (2002): "Heterogeneity and Self-Selection in the Returns to Schooling: Implications for Policy Evaluation," Ph.D. Dissertation, University of Chicago.

Carneiro, P., K. Hansen and J. Heckman (2001): "Removing the Veil of Igorance in Assessing the Distributional Impacts of Social Policies," *Swedish Economic Policy Review,* Vol. 8(2): 273-301.

Carneiro, P., K. Hansen and J. Heckman (2003): "Estimating Distributions of Counter-factuals with an Applicaton to the Returns to Schooling and Measurement of the Effects of Uncertainty on Schooling Choice," *International Economic Review*, forthcoming.

Carneiro, P. and J. Heckman (2002): "The Evidence on Credit Constraints in Post-Secondary Schooling," *The Economic Journal* Vol. 112(482): 705-734.

Carneiro, P., J. Heckman and E. Vytlacil (2001): "Estimating The Return to Education When It Varies Among Individuals." Paper presented as the *Economic Journal* Lecture at the Royal Economic Society meeting, Durham, England, April 2001. Also presented as the *Review of Economics and Statistics Lecture*, Harvard University., April, 2001

Cawley, J., J. Heckman, L. Lochner, and E. Vytlacil (2000): "Understanding the Role of Cognitive Ability in Accounting for the Recent Rise in the Economic Return to Education," in *Meritocracy and Economic Inequality*, ed. by K. Arrow, S. Bowles, and S. Durlauf, Princeton: Princeton University Press.

Comay, Y., A. Melnik, M.A. Pollatschek (1973): "The Option Value of Education and the Optimal Path for Investment in Human Capital" in *International Economic Review.* Vol. 14 (2). p 421-35.

Dooley, M. D. and P. Gottschalk (1984): "Earnings Inequality and Males in the U.S.: Trends and the Effect of Labor Force Growth," *Journal of Political Economy*, 92:59-89.

Eckstein, Z. and K. Wolpin (1999): "Why Youths Drop out of High School: The Impact of Preferences, Opportunities, and Abilities" in *Econometrica*, Vol. 67 (6). p 1295-1339.

Fan, J. and Gijbels, I. (1996): *Local Polynomial Modelling and Its Applications.* New York: Chapman and Hall.

Glewwe (2002): "Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes," *Journal of Economic Literature.* Vol. 40 (2). p 436-82. June

2002.

Gottschalk, Peter (1997): "Inequality, Income Growth, and Mobility: The Basic Facts," *Journal of Economic Perspectives*, Vol. 11, p. 21-40.

Griliches, Z. (1977): "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica*, 45:1-22.

Hansen, K., J. Heckman and Mullen (2003): "A Generalized Ordered Probit Model with An Application to Schooling Choices," unpublished manuscript, University of Chicago.

Hanoch, G. (1967): "An Economic Analysis of Earnings and Schooling" in *The Journal of Human Resources*, Vol. 2, No. 3. (Summer, 1967), pp. 310-329.

Hause, J. (1980): "The Fine Structure of Earnings and the On-the-Job Training Hypothesis," *Econometrica*, 48: 1013-29.

Heckman, J. J., H. Ichimura, J. Smith and P. Todd (1998): "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 66, 1017-1098.

Heckman, J., A. Layne-Farrar and P. Todd (1996): "Human Capital Pricing Equations with an Application to Estimating the Effect of Schooling Quality on Earnings" in *Review of Economics & Statistics*. Vol. 78 (4). p 562-610.

Heckman, J. and E. Vytlacil (2001): "Identifying the Role of Cognitive Ability in Explaining the Rising Return to Education," *Review of Economics and Statistics,* Vol. 83(1): 1-12.

Heckman, J. and E. Vytlacil (2003): "Econometric Evaluation of Social Programs," in *Handbook of Econometrics*, Vol. 6, ed. J. J. Heckman and E. Leamer. Amsterdam: North-Holland. (in press).

Hirshleifer, J. (1970): *Investment, Interest, and Capital*, Englewood Cliffs: Prentice-Hall, Inc.

Jaeger, D. and M. Page (1996): "Degrees Matter: New Evidence on Sheepskin Effects in Returns to Education" in Review of Economics and Statistics, Vol. 78.

Katz, L. F; Autor, D. H. "Changes in the Wage Structure and Earnings Inequality" in *Handbook of labor economics*. Volume 3A. Ashenfelter, Orley Card, David, eds., Handbooks in Economics, vol. 5. Amsterdam; New York and Oxford: Elsevier Science, North-Holland. p 1463-1555. 1999.

Katz, L., and K. Murphy (1992): "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 107(1):35-78.

Keane, M. and K. Wolpin (1997): "The Career Decisions of Young Men" in *Journal of Political Economy*, Vol. 105 (3), p 473-522.

Krueger, A. B. (1993): "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989" in Quarterly Journal of Economics, Vol. 108, No. 1., pp. 33-60.

MaCurdy, T., and T. Mroz (1995): "Measuring Macroeconomic Shifts in Wages from Cohort Specifications," Working Paper, Stanford University.

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

Mincer, J. (1974): *Schooling, Experience, and Earnings*, New York: NBER Press.

Mincer, J. (1997):"Changes in Wage Inequality, 1970-1990" in *Research in labor economics*. Volume 16. Polachek, Solomon W., ed., Associate Editor: John Robst. Greenwich, Conn. and London: JAI Press. p 1-18. 1997.

Mincer, J. and S. Polachek (1974): "Family Investment in Human Capital: Earnings of Women," *Journal of Political Economy,* Vol. 82(2, part II): S76-S108.

Mulligan, C. and J. Marion (2000): "Average Marginal Tax Rates Revisited: A Comment," Working Paper.

Murphy, K. and F. Welch (1990): "Empirical Age-Earnings Profiles" in *Journal of Labor Economics.* Vol. 8 (2). p 202-29.

Murphy, K. and F. Welch (1992): The Structure of Wages.. [Journal Article] The Quarterly Journal of Economics. Vol. 107 (1). p 285-326. February 1992.

Psacharopoulos, G. (1981): "Returns to Education: An Updated International Comparison," in *The Economic Value of Education: Studies in the Economics of Education.* ed Mark Blaug, International Library of Critical Writings in Economics, Vol. 17, Aldershot, U. K.

Rosen, S. (1977): "Human Capital: A Survey of Empirical Research," *Research in Labor Economics, Vol. 1*, ed. R. Ehrenberg, Greenwich: JAI Press, 3-40.

Schultz, T. P. (1975): "Long-Term Change in Personal Income Distribution: Theoretical Approaches, Evidence, and Explanations," in *The "Inequality" Controversy: Schooling and Distributive Justice*, ed. D. Levine and M. J. Bane. New York: Basic Books.

Smith, J. and F. Welch (1979): "Inequality: Race Differences in the Distribution of Earnings," *International Economic Review*, 20: 515-26.

Smith, James P. and Finis R. Welch (1989): "Black Economic Progress after Myrdal" in *Journal of Economic Literature,* Vol. 27, No. 2., pp. 519-564

Solon, G. and T. Hungerford (1987): "Sheepskin Effects in the Returns to Education," in Review of Economics and Statistics, Vol. 69, 175-177.

U.S. Department of Education (1993): *120 Years of Higher Education*, National Center for Education Statistics. (Available through nces.ed.gov/pubs93/93442.pdf)

U.S. Department of Education (2000): *Digest of Education Statistics, 1999*, NCES 2000-031, Washington, DC.

Weisbrod, B. (1962): "Education and Investment in Human Capital" in *The Journal of Political Economy*, Vol. 70, No. 5, Part 2: Investment in Human Beings, pp. 106-123.

Weisbrod, B. (1972): "Earnings Profile: Ability and Schooling: Comment" in *Journal of Political Economy.* Vol. 80 (3). p S139-S41. Part II, ay/June 1972.

Willis, R. J. (1986): "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in *Handbook of Labor Economics*, ed. Orley Ashenfelter and Richard Layard.

Willis, R. J. and S. Rosen (1979): "Education and Self-Selection," *Journal of Political Economy*, 87(5):part 2, S7-S36.

Appendix A: Derivation of the Overtaking Age

Based on the text,

$$
\begin{aligned}
\ln w(s,x) &= \ln E_{s+x} + \ln(1 - k_{s+x}) \\
&\approx \ln E_s + \rho_0 \sum_{j=0}^{x-1} k_{s+j} - k_{s+x}
\end{aligned}
$$

Further using the assumption of linearly declining investment yields

$$
\ln w(s,x) \approx \ln E_s + \kappa \left(\rho_0 \sum_{j=0}^{x-1} (1 - j/T) - (1 - x/T) \right).
$$

Assuming only initial earnings potential (E_s) and investment levels (κ) vary in the population, the variance of log earnings is given by

$$
\begin{aligned}
Var(\ln w(s,x)) &= Var(\ln E_s) + \left(\rho_0 \sum_{j=0}^{x-1} (1 - j/T) - (1 - x/T) \right)^2 Var(\kappa) \\
&\quad + 2 \left(\rho_0 \sum_{j=0}^{x-1} (1 - j/T) - (1 - x/T) \right) Cov(\ln E_s, \kappa)
\end{aligned}
$$

If κ and $\ln E_s$ are uncorrelated, then earnings are minimized (and equal to $Var(\ln E_s)$) when

$$
\begin{aligned}
\rho_0 \sum_{j=0}^{x-1} (1 - j/T) &= 1 - x/T, \qquad \text{or} \\
\rho_0 \left(x - \frac{x(x-1)}{2T} \right) &= 1 - x/T
\end{aligned}
$$

Clearly, $\lim_{t \to \infty} x^* = \frac{1}{\rho_0}$, so the variance minimizing age is $\frac{1}{\rho_0}$ when the work-life is long. More generally, re-arranging terms and solving for the root of this equation[42] yields the variance minimizing experience level of

$$
\begin{aligned}
x^* &= T + \frac{1}{2} + \frac{1}{\rho_0} - \sqrt{ \left(T + \frac{1}{2} + \frac{1}{\rho_0} \right)^2 - \frac{2T}{\rho_0} } \\
&\approx \left(\rho_0 + \frac{\rho_0}{2T} + \frac{1}{T} \right)^{-1},
\end{aligned}
$$

where the final approximation comes from a first order Taylor approximation of the square root term around the squared term inside. The approximation suggests that the variance minimizing age will generally be less than or equal to $\frac{1}{\rho_0}$, with the difference disappearing as T grows large.

---

[42] There is a second root which is greater than T (the maximum working age), so it is ignored.

Appendix B: Data description

Census Data

The Census samples used in this paper are taken from the 1940, 1950, 1960, 1970, 1980 and 1990 Public-Use Census Samples. The 1940 sample consists of the self-weighting subsample which represents 1% of the population. The 1950 sample consists of sample-line persons (for whom questions regarding earnings were asked) which represent about 0.303% of the population. The 1960 sample is a self-weighting 1% sample. The 1970 sample is taken from two Public-Use A samples: the 1% State sample (5% form) and the 1% State sample (15% form). It is a self-weighting sample of 2% of the population. The 1980 and 1990 Census samples are both 5% Public Use A samples. The 1980 sample is self-weighting but the 1990 sample is not. For 1990, we use person weights to re-weight the sample back to random proportions.

The following sample restrictions are imposed for each Census year:

age: Sample includes individuals age 16-64. For Census years when a *quarter-of-birth* variable is available, we take into account the quarter of birth in calculating the age of each individual from the *year of birth* variable provided in the data set.

race: Only individuals reported as being black or white are included in the analysis.

earnings: The earnings measure used is annual earnings, which includes both wage and salary and business income for the Census years when business income is available. For Census years when earnings are reported in intervals, we use the midpoint of the interval as the individuals earnings.

imputations: Individuals with imputed information on age, race, sex, education, weeks worked or income are excluded. For years when all the imputation flags are not provided, we omit individuals on the basis of the available imputation flags.

The following variables are constructed:

experience: Potential experience is measured by Age - Years of Education - 6.

years of education: For the 1940-1980 Censuses, years of education is reported as the highest grade completed. For the 1990 Census, years of education is reported differently: by categories for first through fourth grade and for fifth through eighth grade, by year for ninth through 12th grade, and then by degree attained. To maintain comparability with the other Census samples, we impute the number of years of school associated with each category or degree. For those with some college but no degree or for those with an associate degree, we assign 14 years of school. For those with a

bachelor's degree, we assign 16 years of school. For professional degrees we assign 17 years and for masters degrees and beyond, including doctoral degrees, we assign 18 years of school.

Current Population Survey (CPS) Data

The CPS samples used in this paper are taken from the 1964-2000 CPS March Supplements.

The following sample restrictions are used for each year:

age: Sample includes individuals age 18-65.

race: Sample separated into whites and all non-whites.

earnings: Annual wage and salary income (deflated using the CPI-U) is used as the earnings measure in each year.

The following variables are constructed for our analysis:

experience: Potential experience is measured by Age - Years of Education - 6.

years of education: For 1964-1991, years of education is reported as the highest grade completed. Categories of schooling include 9-11 years, 12 years, and 16 years. From 1992-2000, years of education is reported differently. Those completing 12 years of schooling but who do not receive a high school diploma are assigned 11 years. Only those with 12 years of schooling and a diploma are assigned 12 years of schooling. For those with a bachelor's degree, we assign 16 years of school.

Tuition Time Series

To estimate the private cost of college, we use the time series Total Revenue from Student Fees and Tuition obtained from the publication *120 Years of American Higher Education*(Table 33). Tables 24 and 33 of this publication provide, for all institutions, statistics on total educational revenue, total tuition revenue, and total enrollment. We divide total revenue for all institutions by total enrollment. Supplementing this data with data from the 1999 *Digest of Educational Statistics* (Tables 175 and 331), we create a consistent time series of total educational revenue, total tuition revenue, and total enrollment for 1940-1995.

Tax Rate Time Series

We obtain the average marginal tax rate time series from Barro and Sahasakul (1983) and Mulligan and Marion (2000, Table 1, column 1). The tax rates used in our progressive tax analysis are obtained from the federal schedule for a single adult with no dependents. All income is assumed to be earned income and standard deductions are assumed. To obtain after-tax income for 1960-90, we use the TAXSIM version 4.0 program available at http://www.nber.org/ taxsim/taxsim-calc4/index.html. For 1940 and 1950, we use the actual federal tax schedules (Form 1040) as reported in the Statistics of Income.

Appendix C: Local Linear Regression

In estimating the nonparametric matching regressions, we use local linear regression methods. As discussed in Fan and Gijbels (1996), the local linear estimator for the conditional expectation $E[y_i| z_i = z_0]$ can be computed from the minimization problem

$$\min_{a,b} \sum_{i=1}^{n} (y_i - a - b_1 (z_i - z_0))^2 K\left(\frac{z_i - z_0}{h_n}\right),$$

where $K(\cdot)$ is a kernel function and $h_n > 0$ is a bandwidth which converges to zero as $n \to \infty$.[43] The estimator of the conditional mean $E[y_i| z_i = z_0]$ is $\widehat{a}$. The local linear estimator can be expressed as a weighted average of the y_i observations, $\sum_{i=1}^{n} y_i W_i(z_0)$, where the weights are

$$W_i(z_0) = \frac{K_i \sum_{j=1}^{n} K_j^2 - K_i \sum_{k=1}^{n} K_k}{\sum_{k=1}^{n} K_k \sum_{j=1}^{n} K_j^2 - (\sum_{k=1}^{n} K_k)^2}.$$

Taking advantage of the fact that we have many observations with repeated z_i values, our local regression estimator is given by

$$\hat{m}(z_o) = \frac{\sum_{i=1}^{N_z} n_{z_i} y(z_i) W_i(z_0)}{\sum_{i}^{N_Z} n_{z_i} W_i(z_0)},$$

where $y(z_i)$ represents average earnings at experience level z_i, n_{z_i} represents the number of observations at experience level z_i, and N_Z represents the number of distinct values of potential experience.[44]

The asymptotic distribution of the estimator $\hat{m}(z_0)$ for $m(z_0) = E(y_i|z_i = z_0)$ is given by

$$(nh_n)^{-1/2}(\hat{m}(z_0) - m(z_0)) \tilde{} N(B_n, V_n) + o_p(1)$$

where the bias and variance expressions are given by

$$
\begin{aligned}
B_n &= h_n^2 \cdot (0.5 m''(z_0)) \cdot \int_{-\infty}^{\infty} u^2 K(u) du \\
V_n &= \sigma^2(z_0) \int_{-\infty}^{\infty} K^2(u) du,
\end{aligned}
$$

---

[43]The kernel function we use in the empirical work is the quartic kernel, given by

$$K(s) = \begin{cases} (15/16)(s^2 - 1)^2 & \text{if } |s| < 1 \\ 0 & \text{otherwise.} \end{cases}$$

The bandwidth used is equal to 5.

[44]For some of the Census years, there is a problem of non-random sampling with sampling weights provided in the data. The sampling weights are taken into account when calculating the mean log earnings at each experience level.

and where $\sigma^2(z_0) = E(\{y_i - E(y_i|z_i = z_0)\}^2|z_i = z_0)$.[45]

Tests of Parallelism

In section III of this paper, we perform nonparametric tests of whether the log-earnings-experience profiles are parallel across schooling levels. Let s_1 and s_2 denote two different schooling levels (16 years and 12 years, for example). We test whether

$$E(y_i|z_i, s = s_1) - E(y_i|z_i, s = s_2) = \text{constant across } z_i \in \{10, 20, 30, 40 \text{ years}\}$$

We select the experience values at which the hypothesis is tested to be at least 2 bandwidths apart from the other experience levels, so that the nonparametric estimates are independent from one another. Let $\hat{m}(z_i, s_1)$ denote the estimator for $E(y_i|z_i, s = s_1)$ for experience level z_i and schooling level $s = s_1$. The test statistic for testing parallelism for two different schooling levels s_1 and s_2 and two experience levels z_i and z_k is given by

$$(\hat{m}(z_i, s_1) - \hat{m}(z_i, s_2) - (\hat{m}(z_k, s_1) - \hat{m}(z_k, s_2)) \cdot$$
$$(\hat{V}_1 + \hat{V}_2 + \hat{V}_3 + \hat{V}_4)^{-1} \cdot$$
$$(\hat{m}(z_i, s_1) - \hat{m}(z_i, s_2) - (\hat{m}(z_k, s_1) - \hat{m}(z_k, s_2)),$$

where $\hat{V}_1$, $\hat{V}_2$, $\hat{V}_3$, and $\hat{V}_4$ are estimators for $V_1 = Var(\hat{m}(z_i, s_1))$, $V_1 = Var(\hat{m}(z_i, s_2))$, $V_3 = Var(\hat{m}(z_k, s_1))$, $V_3 = Var(\hat{m}(z_k, s_2))$.

Under the null hypothesis of parallelism, the bias terms cancel out, so that it is not necessary to estimate the bias expressions in performing the test.[46] To estimate the variances, we use

$$Var(\hat{m}(z_i, s_1)) = \frac{\sum_{i=1}^{N_Z} n_{z_i}\hat{\varepsilon}(z_i, s_1)^2 W_i(z_i)}{\sum_i^{N_Z} n_{z_i} W_i(z_i)},$$

where $\hat{\varepsilon}(z_i, s_1) = y(z_i, s_1) - \hat{m}(z_i, s_1)$ is the fitted residual from the nonparametric regression evaluated at experience level z_i.[47] In Table 1, we report test results based on the test statistic that straightforwardly generalizes the test statistic given above to four experience levels.

---

[45]See, e.g. Fan and Gijbels (1996), for derivation of these formulae.

[46]This cancelling only occurs with the local linear estimator and would not occur if the standard kernel estimator were used instead to generate the nonparametric estimates.

[47]Heckman, Ichimura, Smith, and Todd (1998) show that this estimator has better finite sample performance than a "plug-in" estimator based on the asymptotic variance formulae.

Table 1: Tests of Parallelism in Log Earnings Experience Profiles for Men

Sample	Experience Level	Estimated Difference Between College and High School Earnings at Different Experience Levels					
		1940	1950	1960	1970	1980	1990
Whites	10	878	1,567	2,889	4,497	6,247	16,246
	20	877	1,627	4,026	6,753	11,284	19,315
	30	1,282	1,873	4,369	7,359	11,427	23,739
	40	776	566	4,913	6,771	9,233	21,332
	p-value	0.18	0.44	<0.001	<0.001	<0.001	<0.001
Blacks	10	272	7,661	1,528	2,581	6,003	13,108
	20	108	256	1,954	2,490	5,221	15,048
	30	-77	158	338	1,160	6,297	13,960
	40	-70	633	3,446	1,538	2,923	5,162
	p-value	0.15	0.72	0.48	0.53	0.71	0.002

Notes: Data taken from 1940-90 Decennial Censuses without adjustment for inflation. Because there are very few blacks in the 1940 and 1950 samples with college degrees, especially at higher experience levels, the test results for blacks in those years refer to a test of the difference between earnings for high school graduates and persons with 8 years of education. See Appendix C for the formulae used for the test statistics.

Table 2: Estimated Coefficients from Mincer Log Earnings Regression for Men

		Whites		Blacks	
		Coefficient	Std. Error	Coefficient	Std. Error
1940	Intercept	4.4771	0.0096	4.6711	0.0298
	Education	0.1250	0.0007	0.0871	0.0022
	Experience	0.0904	0.0005	0.0646	0.0018
	Experience-Squared	-0.0013	0.0000	-0.0009	0.0000
1950	Intercept	5.3120	0.0132	5.0716	0.0409
	Education	0.1058	0.0009	0.0998	0.0030
	Experience	0.1074	0.0006	0.0933	0.0023
	Experience-Squared	-0.0017	0.0000	-0.0014	0.0000
1960	Intercept	5.6478	0.0066	5.4107	0.0220
	Education	0.1152	0.0005	0.1034	0.0016
	Experience	0.1156	0.0003	0.1035	0.0011
	Experience-Squared	-0.0018	0.0000	-0.0016	0.0000
1970	Intercept	5.9113	0.0045	5.8938	0.0155
	Education	0.1179	0.0003	0.1100	0.0012
	Experience	0.1323	0.0002	0.1074	0.0007
	Experience-Squared	-0.0022	0.0000	-0.0016	0.0000
1980	Intercept	6.8913	0.0030	6.4448	0.0120
	Education	0.1023	0.0002	0.1176	0.0009
	Experience	0.1255	0.0001	0.1075	0.0005
	Experience-Squared	-0.0022	0.0000	-0.0016	0.0000
1990	Intercept	6.8912	0.0034	6.3474	0.0144
	Education	0.1292	0.0002	0.1524	0.0011
	Experience	0.1301	0.0001	0.1109	0.0006
	Experience-Squared	-0.0023	0.0000	-0.0017	0.0000

Notes: Data taken from 1940-90 Decennial Censuses.

Table 3a: Internal Rates of Return for White Men: Earnings Function Assumptions
(Specifications Assume Work Lives of 47 Years)

	Schooling Comparisons					
	6-8	8-10	10-12	12-14	12-16	14-16
1940						
Mincer Specification	13	13	13	13	13	13
Relax Linearity in S	16	14	15	10	15	21
Relax Linearity in S & Quad. in Exp.	16	14	17	10	15	20
Relax Lin. in S & Parallelism	12	14	24	11	18	26
1950						
Mincer Specification	11	11	11	11	11	11
Relax Linearity in S	13	13	18	0	8	16
Relax Linearity in S & Quad. in Exp.	14	12	16	3	8	14
Relax Linearity in S & Parallelism	26	28	28	3	8	19
1960						
Mincer Specification	12	12	12	12	12	12
Relax Linearity in S	9	7	22	6	13	21
Relax Linearity in S & Quad. in Exp.	10	9	17	8	12	17
Relax Linearity in S & Parallelism	23	29	33	7	13	25
1970						
Mincer Specification	13	13	13	13	13	13
Relax Linearity in S	2	3	30	6	13	20
Relax Linearity in S & Quad. in Exp.	5	7	20	10	13	17
Relax Linearity in S & Parallelism	17	29	33	7	13	24
1980						
Mincer Specification	11	11	11	11	11	11
Relax Linearity in S	3	-11	36	5	11	18
Relax Linearity in S & Quad. in Exp.	4	-4	28	6	11	16
Relax Linearity in S & Parallelism	16	66	45	5	11	21
1990						
Mincer Specification	14	14	14	14	14	14
Relax Linearity in S	-7	-7	39	7	15	24
Relax Linearity in S & Quad. in Exp.	-3	-3	30	10	15	20
Relax Linearity in S & Parallelism	20	20	50	10	16	26

Notes: Data taken from 1940-90 Decennial Censuses. In 1990, comparisons of 6 vs. 8 and 8 vs. 10 cannot be made given data restrictions. Therefore, those columns report calculations based on a comparison of 6 and 10 years of schooling.

Table 3b: Internal Rates of Return for Black Men: Earnings Function Assumptions
(Specifications Assume Work Lives of 47 Years)

| | Schooling Comparisons | | | | | |
	6-8	8-10	10-12	12-14	12-16	14-16
1940						
Mincer Specification	9	9	9	9	9	9
Relax Linearity in S	18	7	5	3	11	18
Relax Linearity in S & Quad. in Exp.	18	8	6	2	10	19
Relax Linearity in S & Parallelism	11	0	10	5	12	20
1950						
Mincer Specification	10	10	10	10	10	10
Relax Linearity in S	16	14	18	-2	4	9
Relax Linearity in S & Quad. in Exp.	16	14	18	0	3	6
Relax Linearity in S & Parallelism	35	15	48	-3	6	34
1960						
Mincer Specification	11	11	11	11	11	11
Relax Linearity in S	13	12	18	5	8	11
Relax Linearity in S & Quad. in Exp.	13	11	18	5	7	10
Relax Linearity in S & Parallelism	22	15	38	5	11	25
1970						
Mincer Specification	12	12	12	12	12	12
Relax Linearity in S	5	11	30	7	10	14
Relax Linearity in S & Quad. in Exp.	6	11	24	10	11	12
Relax Linearity in S & Parallelism	15	27	44	9	14	23
1980						
Mincer Specification	12	12	12	12	12	12
Relax Linearity in S	-4	1	35	10	15	19
Relax Linearity in S & Quad. in Exp.	-4	6	29	11	14	17
Relax Linearity in S & Parallelism	10	44	48	8	16	31
1990						
Mincer Specification	16	16	16	16	16	16
Relax Linearity in S	-5	-5	41	15	20	25
Relax Linearity in S & Quad. in Exp.	-3	-3	35	17	19	22
Relax Linearity in S & Parallelism	16	16	58	18	25	35

Notes: Data taken from 1940-90 Decennial Censuses. In 1990, comparisons of 6 vs. 8 and 8 vs. 10 cannot be made given data restrictions. Therefore, those columns report calculations based on a comparison of 6 and 10 years of schooling.

Table 4: Internal Rates of Return for White & Black Men: Accounting for Taxes and Tuition
(General Non-Parametric Specification Assuming Work Lives of 47 Years)

| | | Schooling Comparisons | | | | | |
| | | Whites | | | Blacks | | |
		12-14	12-16	14-16	12-14	12-16	14-16
1940	No Taxes or Tuition	11	18	26	5	12	20
	Including Tuition Costs	9	15	21	4	10	16
	Including Tuition & Flat Taxes	8	15	21	4	9	16
	Including Tuition & Prog. Taxes	8	15	21	4	10	16
1950	No Taxes or Tuition	3	8	19	-3	6	34
	Including Tuition Costs	3	8	16	-3	5	25
	Including Tuition & Flat Taxes	3	8	16	-3	5	24
	Including Tuition & Prog. Taxes	3	7	15	-3	5	21
1960	No Taxes or Tuition	7	13	25	5	11	25
	Including Tuition Costs	6	11	21	5	9	18
	Including Tuition & Flat Taxes	6	11	20	4	8	17
	Including Tuition & Prog. Taxes	6	10	19	4	8	15
1970	No Taxes or Tuition	7	13	24	9	14	23
	Including Tuition Costs	6	12	20	7	12	18
	Including Tuition & Flat Taxes	6	11	20	7	11	17
	Including Tuition & Prog. Taxes	5	10	18	7	10	16
1980	No Taxes or Tuition	5	11	21	8	16	31
	Including Tuition Costs	4	10	18	7	13	24
	Including Tuition & Flat Taxes	4	9	17	6	12	21
	Including Tuition & Prog. Taxes	4	8	15	6	11	20
1990	No Taxes or Tuition	10	16	26	18	25	35
	Including Tuition Costs	9	14	20	14	18	25
	Including Tuition & Flat Taxes	8	13	19	13	17	22
	Including Tuition & Prog. Taxes	8	12	18	13	17	22

Notes: Data taken from 1940-90 Decennial Censuses. See discussion in text and Appendix A for a description of tuition and tax amounts.

Table 5: Internal Rates of Return for White & Black Men: Residual Adjustment
(General Non-Parametric Specification Accounting for Tuition and Progressive Taxes)

| | | | Schooling Comparisons | | | | | |
			6-8	8-10	10-12	12-14	12-16	14-16
a. Whites								
	1940	Unadjusted	12	14	24	8	15	21
		Adjusted	2	2	8	9	13	16
	1950	Unadjusted	25	26	26	3	7	15
		Adjusted	17	19	14	5	8	14
	1960	Unadjusted	21	27	29	6	10	19
		Adjusted	13	19	16	7	11	16
	1970	Unadjusted	16	27	29	5	10	18
		Adjusted	11	18	16	6	10	16
	1980	Unadjusted	14	64	41	4	8	15
		Adjusted	9	28	24	5	8	13
	1990	Unadjusted	19	19	47	8	12	18
		Adjusted	11	11	31	8	12	17
b. Blacks								
	1940	Unadjusted	11	0	10	4	10	16
		Adjusted	3	0	-8	4	6	7
	1950	Unadjusted	33	14	44	-3	5	21
		Adjusted	53	8	21	1	9	15
	1960	Unadjusted	20	14	34	4	8	15
		Adjusted	14	12	16	6	6	8
	1970	Unadjusted	14	25	39	7	10	16
		Adjusted	12	16	22	7	10	12
	1980	Unadjusted	9	43	46	6	11	20
		Adjusted	7	21	29	6	9	15
	1990	Unadjusted	16	16	57	13	17	22
		Adjusted	8	8	42	11	15	20

Notes: Data taken from 1940-90 Decennial Censuses. In 1990, comparisons of 6 vs. 8 and 8 vs. 10 cannot be made given data restrictions. Therefore, those columns report calculations based on a comparison of 6 and 10 years of schooling. See discussion in text and Appendix A for a description of tuition and tax amounts.

Table 6a: Internal Rates of Return for White Men: Best Census and CPS Estimates

Schooling Comparison	Year	Mincer	Census Data: General Spec. (No Residual Adjustment)	General Spec. (Residual Adjustment)	CPS Data: Cross Section	Cohort
10 vs. 12	1940	13	24	8	-	-
	1950	11	26	14	-	3
	1960	12	29	16	-	7
	1970	13	29	16	29	34
	1980	11	41	24	38	38
	1990	14	47	31	50	-
12 vs. 16	1940	13	15	13	-	-
	1950	11	7	8	-	14
	1960	12	10	11	-	8
	1970	13	10	10	12	10
	1980	11	8	8	8	14
	1990	14	12	12	14	-

Notes: Mincer estimates make no adjustment for taxes or tuition. Census General Specification estimates account for tuition and progressive taxes with fully non-parametric wage specification. CPS Cross Section Estimates use cross sectional data and a general wage specification accounting for tuition and flat taxes. CPS Cohort estimates follow a cohort turning age 18 in the reported year, using a general wage specification accounting for tuition and flat taxes.

Table 6b: Internal Rates of Return for Black Men: Best Census and CPS Estimates

Schooling Comparison	Year	Mincer	Census Data: General Spec. (No Residual Adjustment)	General Spec. (Residual Adjustment)	CPS Data: Cross Section	Cohort
10 vs. 12	1940	9	10	-8	-	-
	1950	10	44	21	-	4
	1960	11	34	16	-	18
	1970	12	39	22	32	49
	1980	12	46	29	55	70
	1990	16	57	42	64	-
12 vs. 16	1940	9	10	6	-	-
	1950	10	5	9	-	15
	1960	11	8	6	-	6
	1970	12	10	10	12	14
	1980	12	11	9	14	17
	1990	16	17	15	16	-

Notes: Mincer estimates make no adjustment for taxes or tuition. Census General Specification estimates account for tuition and progressive taxes with fully non-parametric wage specification. CPS Cross Section Estimates use cross sectional data and a general wage specification accounting for tuition and flat taxes. CPS Cohort estimates follow a cohort turning age 18 in the reported year, using a general wage specification accounting for tuition and flat taxes. Each CPS estimate is based on three adjoining years/cohorts worth of data.

Table 7a: Simulated Returns under Uncertainty with Option Values
(Log Wages Linear in Schooling: $\bar{W}_{s+1} = (1+r)\bar{W}_s$)

Education Level (s)	Transition Prob. ($p_{s,s-1}$)	Prop. Increase in $\bar{W}$	Prop. Increase Observed Earnings	Option/Total Value ($O_{s,s-1}/E_{s-1}(V_s)$)	Average Return ($E_{s-1}[R_{s,s-1}]$)	Treatment on Treated	Treatment on Untreated
2	0.796	0.100	0.086	0.075	0.201	0.242	0.041
3	0.746	0.100	0.082	0.060	0.182	0.231	0.037
4	0.669	0.100	0.072	0.038	0.155	0.216	0.032
5	0.520	0.100	0.016	0.000	0.111	0.196	0.019

OLS estimate of the rate of return is 0.063.

Notes: The simulated model assumes lifetime earnings for someone with s years of school equal $\bar{W}_s \epsilon_s$ where ϵ_s are independent and identically distributed $\log(\epsilon_s) \sim N(0, 0.1)$. An interest rate of $r = 0.10$ is assumed. The transition probability from $s-1$ to s is given by $p_{s,s-1} = Pr\left(\epsilon_{s-1} \leq \frac{E_{s-1}(V_s)}{(1+r)\bar{W}_{s-1}}\right)$. Observed earnings for someone with s years of school are $\bar{W}_s E_{s-1}\left(\epsilon_s | \epsilon_s > \frac{E_s(V_{s+1})}{(1+r)\bar{W}_s}\right)$, and option values are $E_{s-1}(V_s - W_s)$. The return to school year s for someone with earnings W_{s-1} is $R_{s,s-1} = \frac{E_{s-1}(V_s) - W_{s-1}}{W_{s-1}}$. Average returns reflect the expected return over the full distribution of W_{s-1}, or $E_{s-1}[R_{s,s-1}]$. "Treatment on Treated" reflects returns for those who continue to grade s, or $E_{s-1}\left[R_{s,s-1} | \epsilon_{s-1} \leq \frac{E_{s-1}(V_s)}{(1+r)\bar{W}_{s-1}}\right]$. "Treatment on Untreated" reflects returns for those who do not continue to grade s, or $E_{s-1}\left[R_{s,s-1} | \epsilon_{s-1} > \frac{E_{s-1}(V_s)}{(1+r)\bar{W}_{s-1}}\right]$. The marginal treatment effect equals $r = 0.10$. OLS estimate is the coefficient on schooling in a log earnings regression (the Mincer return).

Table 7b: Simulated Returns under Uncertainty with Option Values

(Sheepskin Effects: $\bar{W}_{s+1} = (1 + \rho_{s+1})\bar{W}_s$ with $\rho_2 = 0.1$, $\rho_3 = 0.3$, $\rho_4 = 0.1$, $\rho_5 = 0.2$)

Education Level (s)	Transition Prob. ($p_{s,s-1}$)	Prop. Increase in $\bar{W}$	Prop. Increase Observed Earnings	Option/Total Value ($O_{s,s-1}/E_{s-1}(V_s)$)	Average Return ($E_{s-1}[R_{s,s-1}]$)	Treatment on Treated	Treatment on Untreated
2	0.997	0.100	0.101	0.239	0.459	0.460	0.068
3	0.997	0.300	0.116	0.100	0.459	0.460	0.068
4	0.846	0.100	0.092	0.093	0.224	0.257	0.045
5	0.822	0.200	0.041	0.000	0.212	0.249	0.043

OLS estimate of the rate of return is 0.060.

Notes: The simulated model assumes lifetime earnings for someone with s years of school equal $\bar{W}_s\epsilon_s$ where ϵ_s are independent and identically distributed $\log(\epsilon_s) \sim N(0, 0.1)$. An interest rate of $r = 0.10$ is assumed. The transition probability from $s-1$ to s is given by $p_{s,s-1} = Pr\left(\epsilon_{s-1} \leq \frac{E_{s-1}(V_s)}{(1+r)\bar{W}_{s-1}}\right)$. Observed earnings for someone with s years of school are $\bar{W}_s E_{s-1}\left(\epsilon_s | \epsilon_s > \frac{E_s(V_{s+1})}{(1+r)\bar{W}_s}\right)$, and option values are $E_{s-1}(V_s - W_s)$. The return to school year s for someone with earnings W_{s-1} is $R_{s,s-1} = \frac{E_{s-1}(V_s) - W_{s-1}}{W_{s-1}}$. Average returns reflect the expected return over the full distribution of W_{s-1}, or $E_{s-1}[R_{s,s-1}]$. "Treatment on Treated" reflects returns for those who continue to grade s, or $E_{s-1}\left[R_{s,s-1}|\epsilon_{s-1} \leq \frac{E_{s-1}(V_s)}{(1+r)\bar{W}_{s-1}}\right]$. "Treatment on Untreated" reflects returns for those who do not continue to grade s, or $E_{s-1}\left[R_{s,s-1}|\epsilon_{s-1} > \frac{E_{s-1}(V_s)}{(1+r)\bar{W}_{s-1}}\right]$. The marginal treatment effect equals $r = 0.10$. OLS estimate is the coefficient on schooling in a log earnings regression (the Mincer coefficient).

Table 8: Present Value of Earnings, Option Values, and Returns to Schooling
(White Men, 1990 Census)

Interest Rate	Trans. Prob.	PV Lifetime Earnings (in $1000's)			Option Value (in $1000's)		Total Value (in $1000's)		Return to Schooling	
r	p	$\hat{W}_{12}$	$\hat{W}_{14}$	$\hat{W}_{16}$	$\hat{O}_{12,10}$	$\hat{O}_{14,12}$	$E(V_{12})$	$E(V_{14})$	$\hat{R}_{12,10}$	$\hat{R}_{14,12}$
0.07	0.1	226.46	274.15	394.97	1.92	7.08	228.38	281.23	0.24	0.11
0.07	0.3	226.46	274.15	394.97	9.47	21.25	235.92	295.40	0.26	0.14
0.07	0.5	226.46	274.15	394.97	21.96	35.41	248.42	309.56	0.30	0.17
0.1	0.1	149.26	181.17	266.12	0.37	3.88	149.63	185.05	0.27	0.11
0.1	0.3	149.26	181.17	266.12	3.02	11.63	152.29	192.80	0.28	0.14
0.1	0.5	149.26	181.17	266.12	8.24	19.38	157.51	200.56	0.31	0.16

Notes: Transition probability, p, represents the probability of continuing in school conditional on current education. "PV of lifetime earnings" is $\hat{W}_s = \sum_{x=0}^{65} (1+r)^{-x} \hat{w}(s,x)$ where $\hat{w}(s,x)$ is the nonparametrically estimated earnings for a white man with s years of school and x years of experience (based on the 1990 Census). "Total value", $E(V_s) = (1-p)\hat{W}_s + p(1+r)^{-1} E(V_{s+1})$, is recursively solved backward from $E(V_{16}) = \hat{W}_{16}$. "Option value" is $\hat{O}_{s,s-1} = E(V_s) - \hat{W}_s$. "Return to school" $\hat{R}_{s,s-1} = \frac{E(V_s) - \hat{W}_{s-1}}{\hat{W}_{s-1}}$ is annualized.

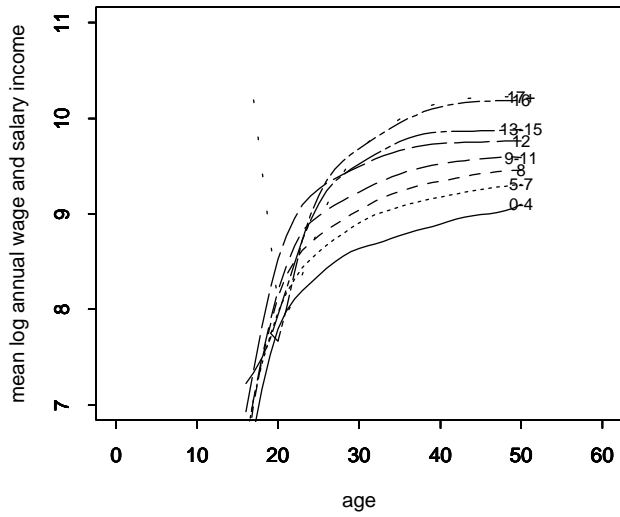# Figure 1a: Experience-Earnings Profiles, 1940-1960

1940 Census, White Males

1940 Census, Black Males

1950 Census, White Males

1950 Census, Black Males
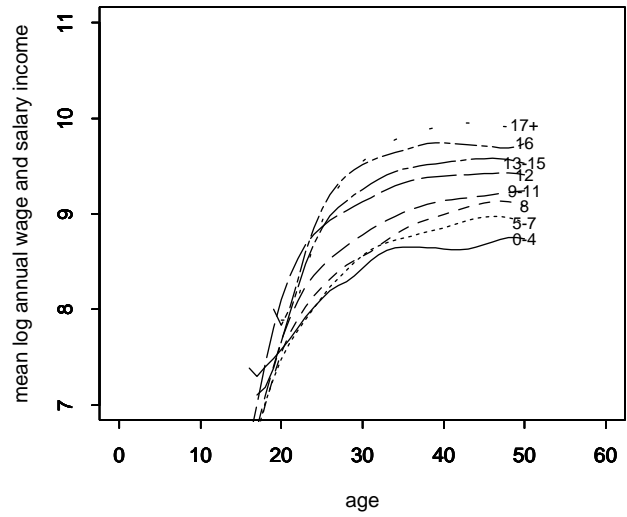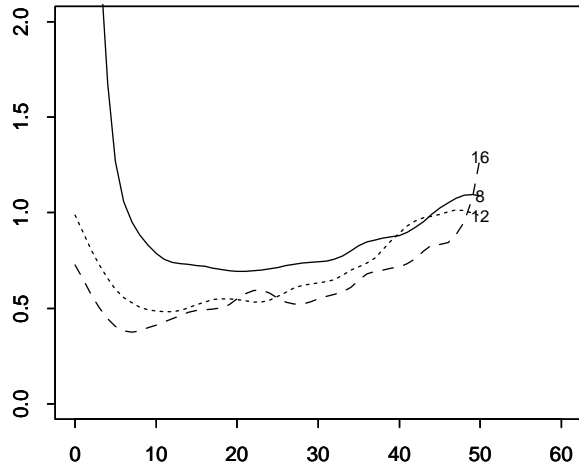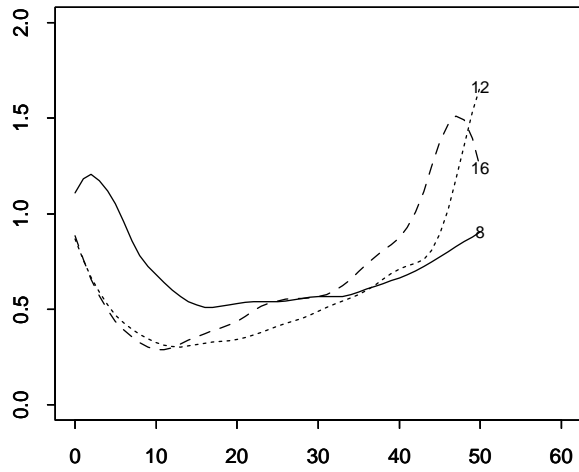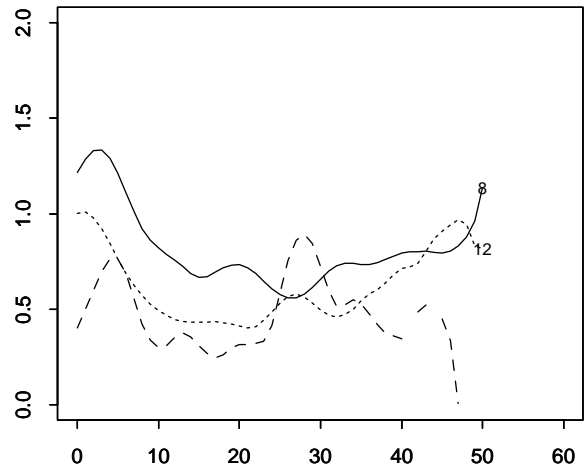
1960 Census, White Males

1960 Census, Black Males
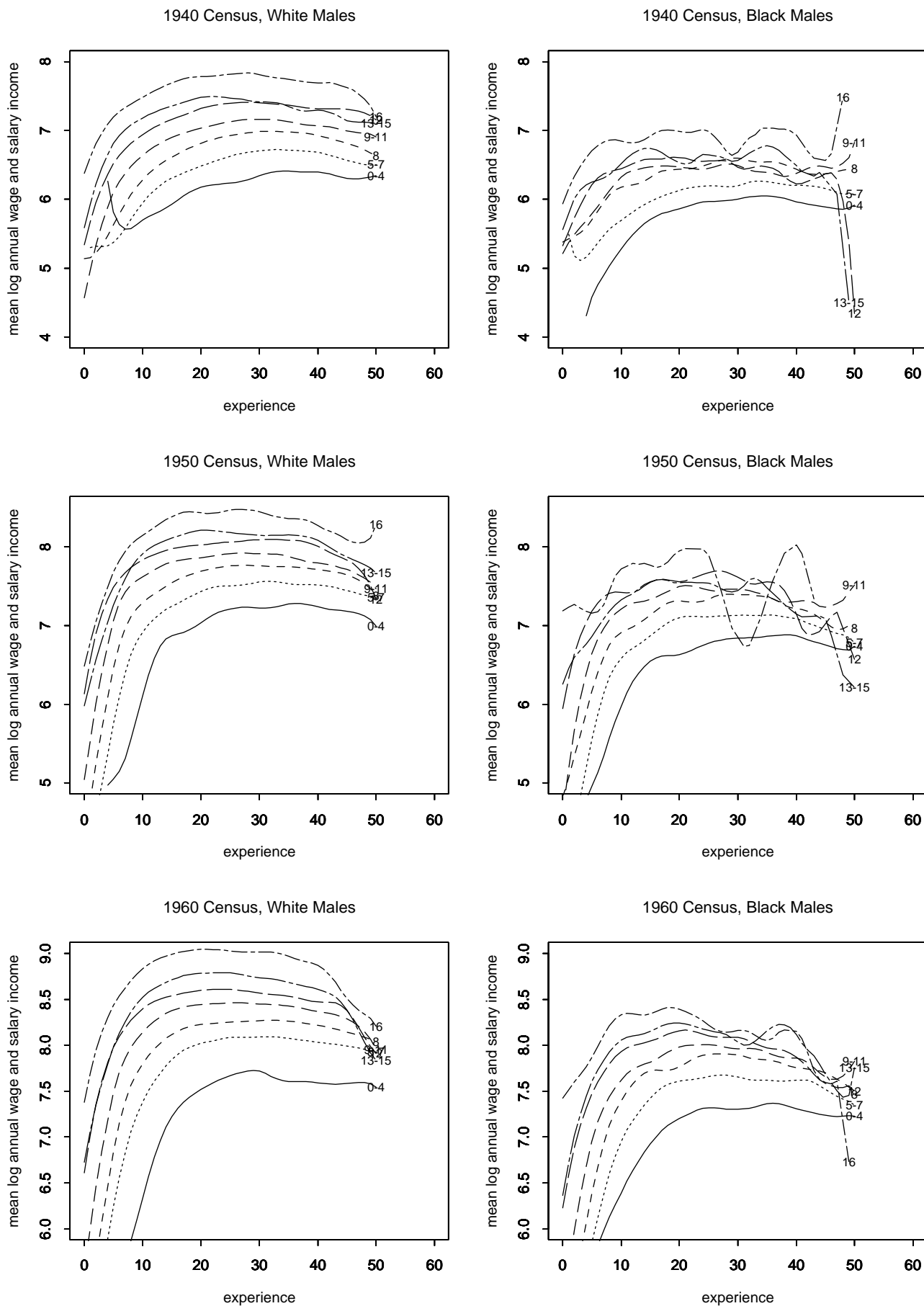
Figure 1b: Experience-Earnings Profiles, 1970-1990

1970 Census, White Males

1970 Census, Black Males

1980 Census, White Males

1980 Census, Black Males

1990 Census, White Males

1990 Census, Black Males
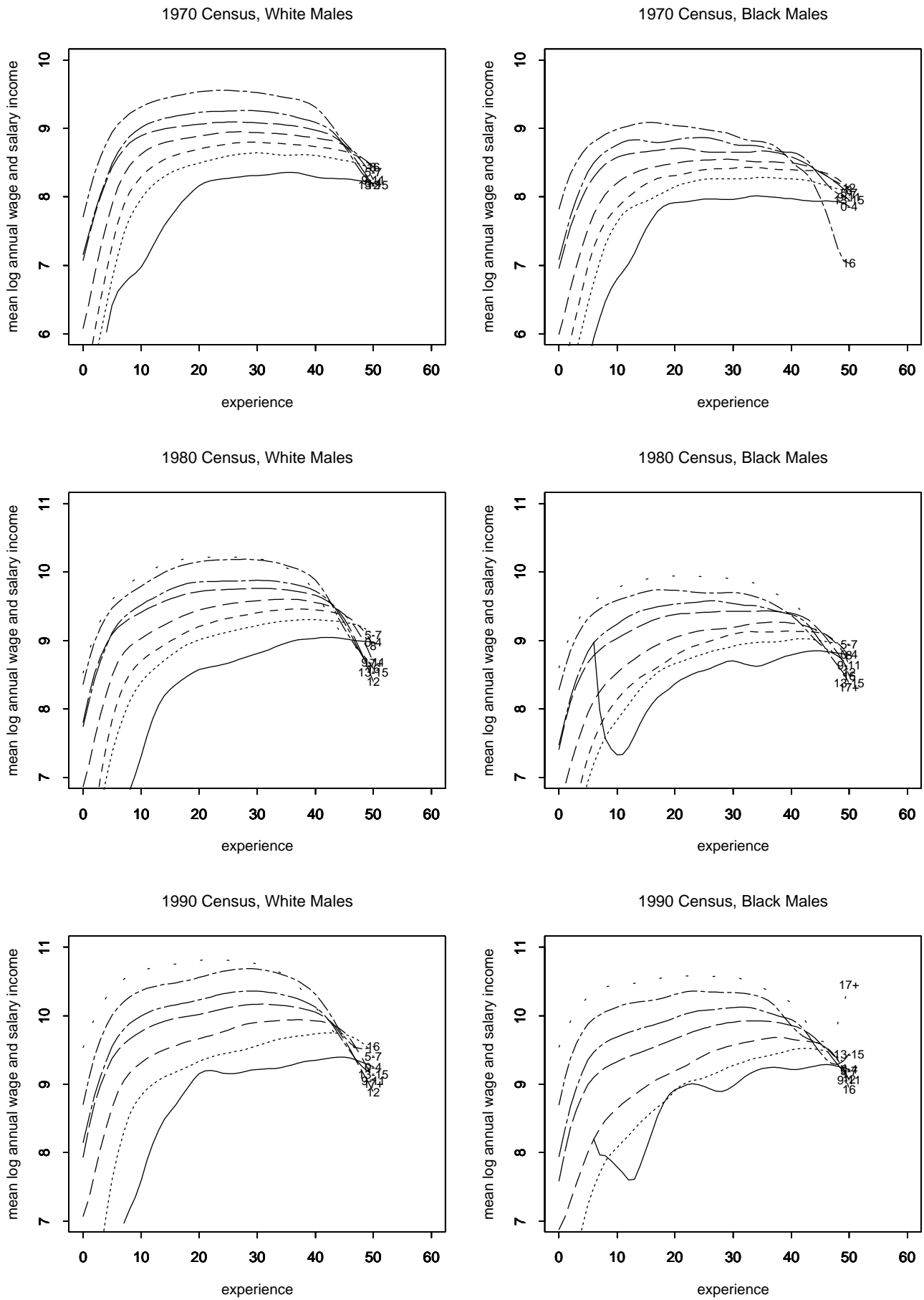
Figure 2: Age-Earnings Profiles, 1940,1960,1980

1940 Census, White Males
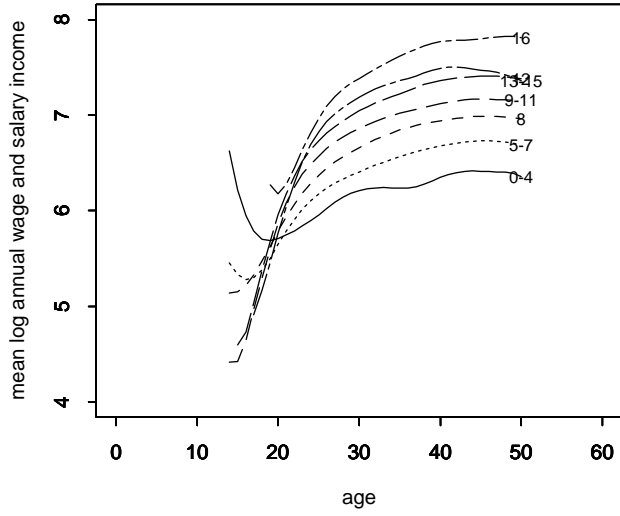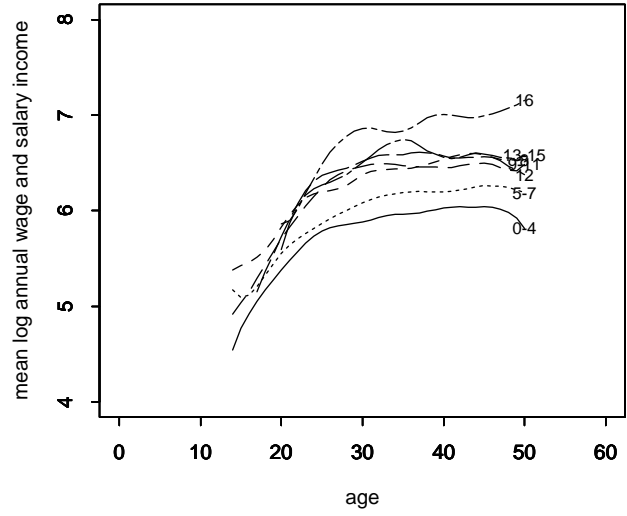
1940 Census, Black Males

1960 Census, White Males
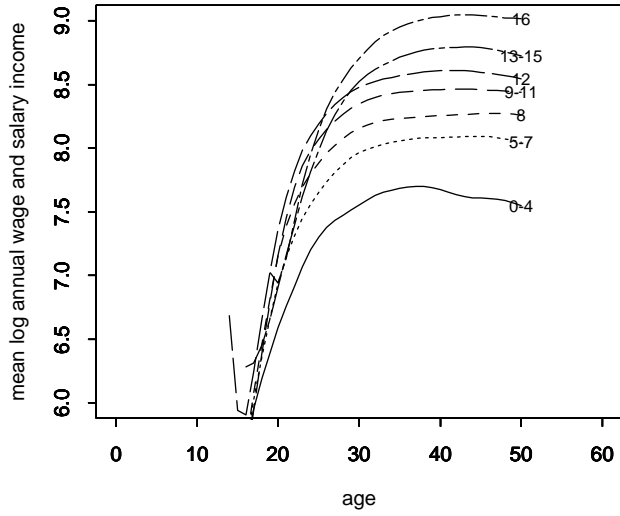
1960 Census, Black Males
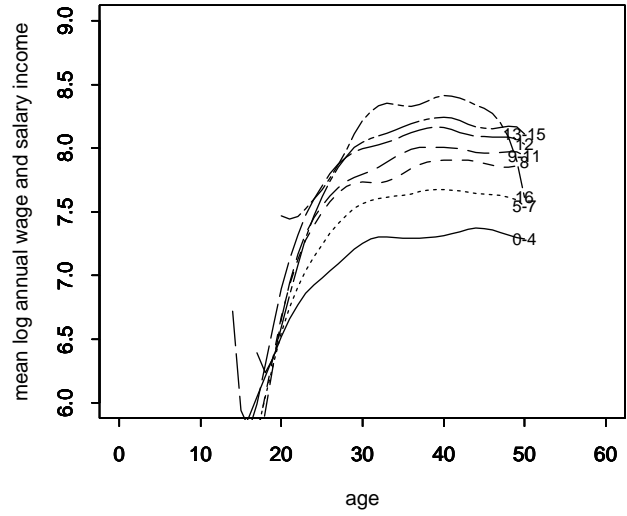
1980 Census, White Males

1980 Census, Black Males

Figure 3: Experience-Variance Log Earnings

1940 Census, White Males
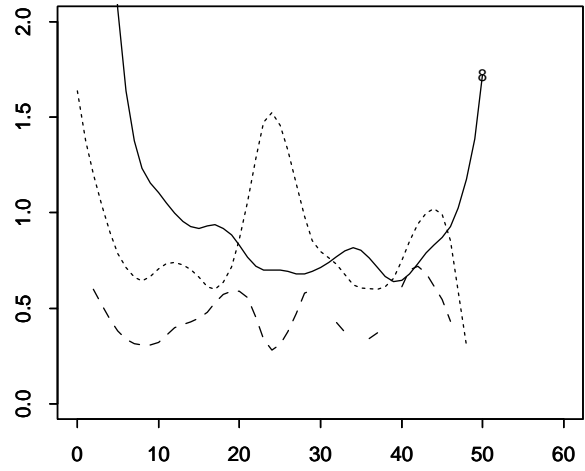
1940 Census, Black Males

1960 Census, White Males

1960 Census, Black Males

1980 Census, White Males
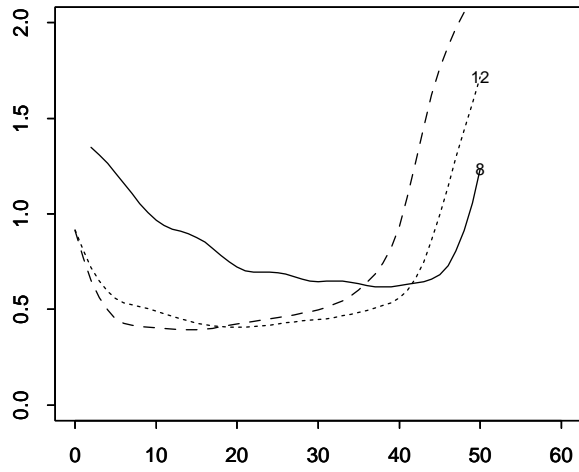
1980 Census, Black Males
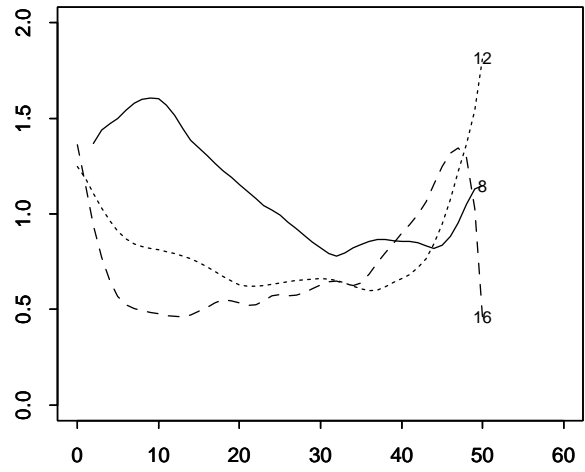
Figure 4a: Average College Tuition Paid (in 2000 dollars)
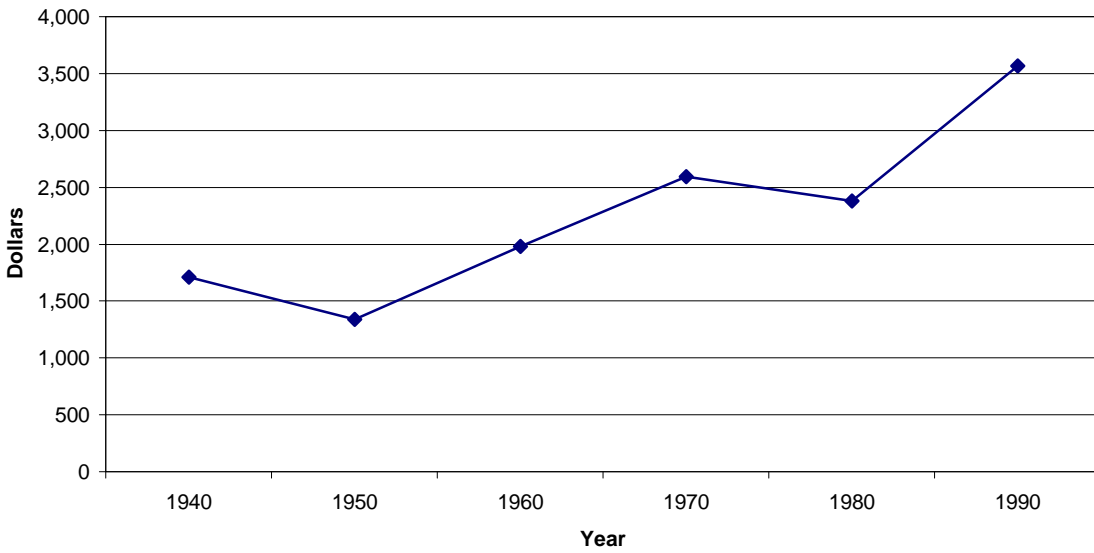
**Figure 4b: Marginal Tax Rates
(from Barro & Sahasakul, 1983, Mulligan & Marion, 2000)**
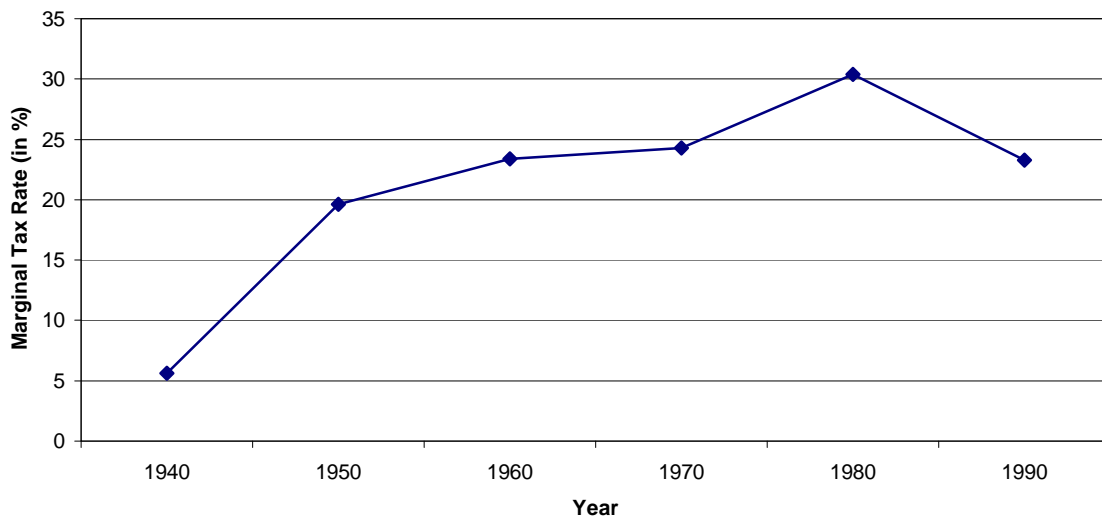
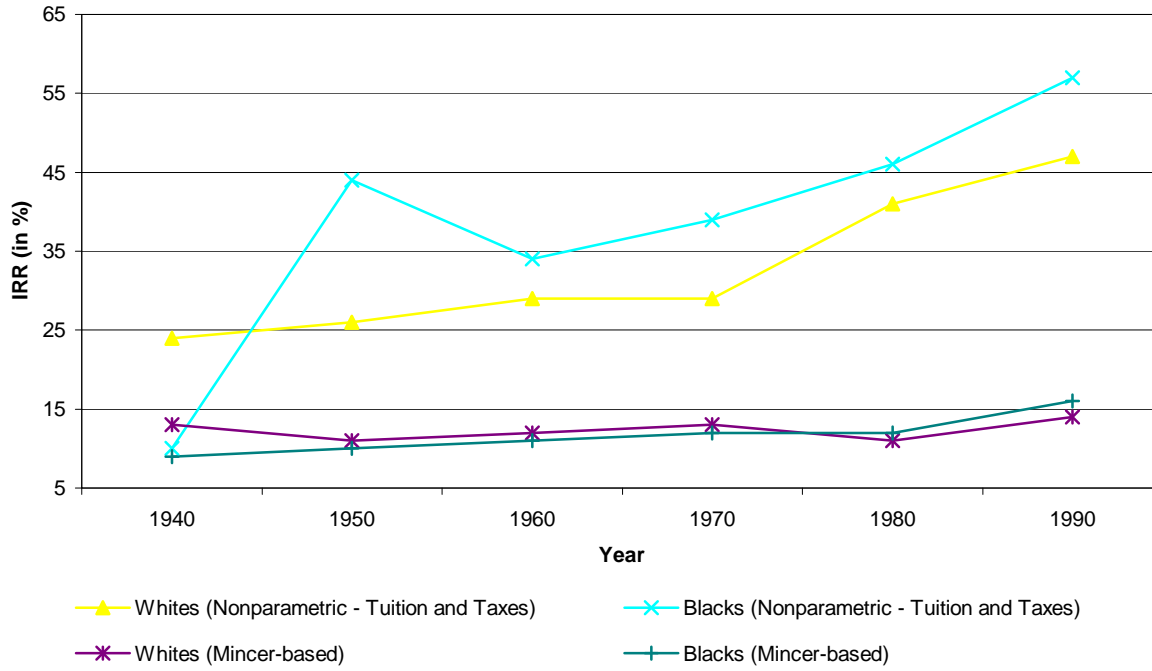**Figure 5: IRR for High School Completion (White and Black Men)**

Legend:
- ▲ Whites (Nonparametric - Tuition and Taxes)
- ✕ Blacks (Nonparametric - Tuition and Taxes)
- ✳ Whites (Mincer-based)
- ＋ Blacks (Mincer-based)

Figure 6: IRR for College Completion (White and Black Men)

Legend:
- ▲ Whites (Nonparametric - Tuition and Taxes)
- ✕ Blacks (Nonparametric - Tuition and Taxes)
- ◆ Whites (Mincer-based)
- ■ Blacks (Mincer-Based)
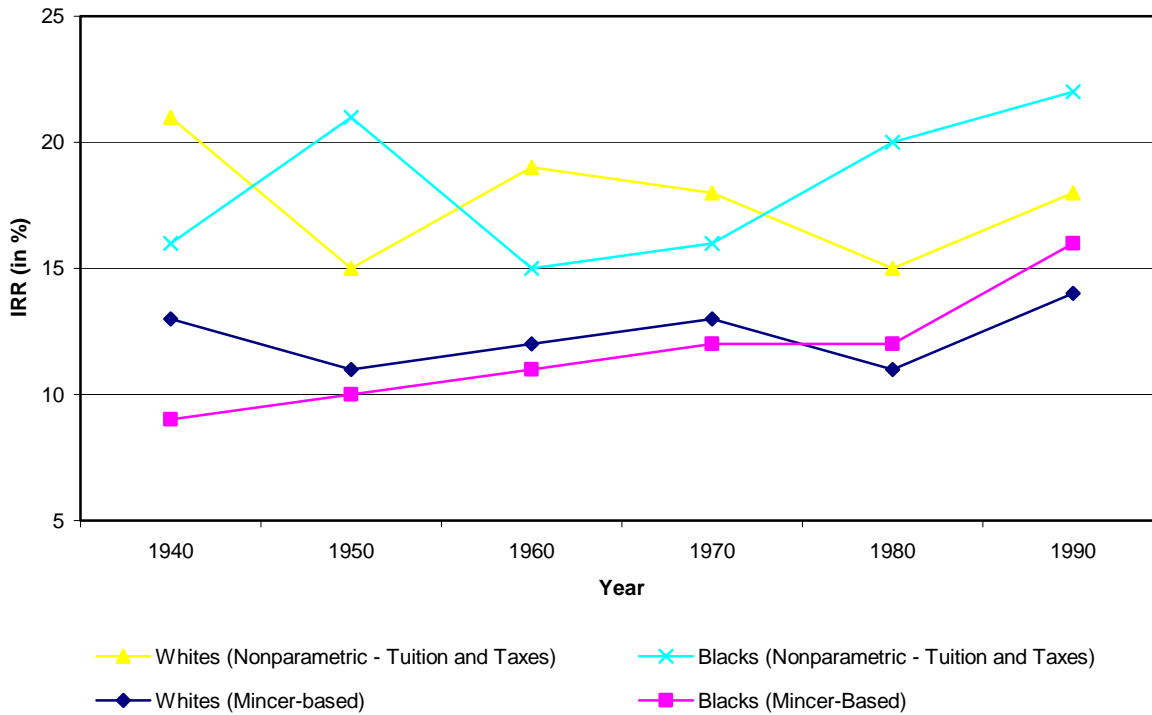
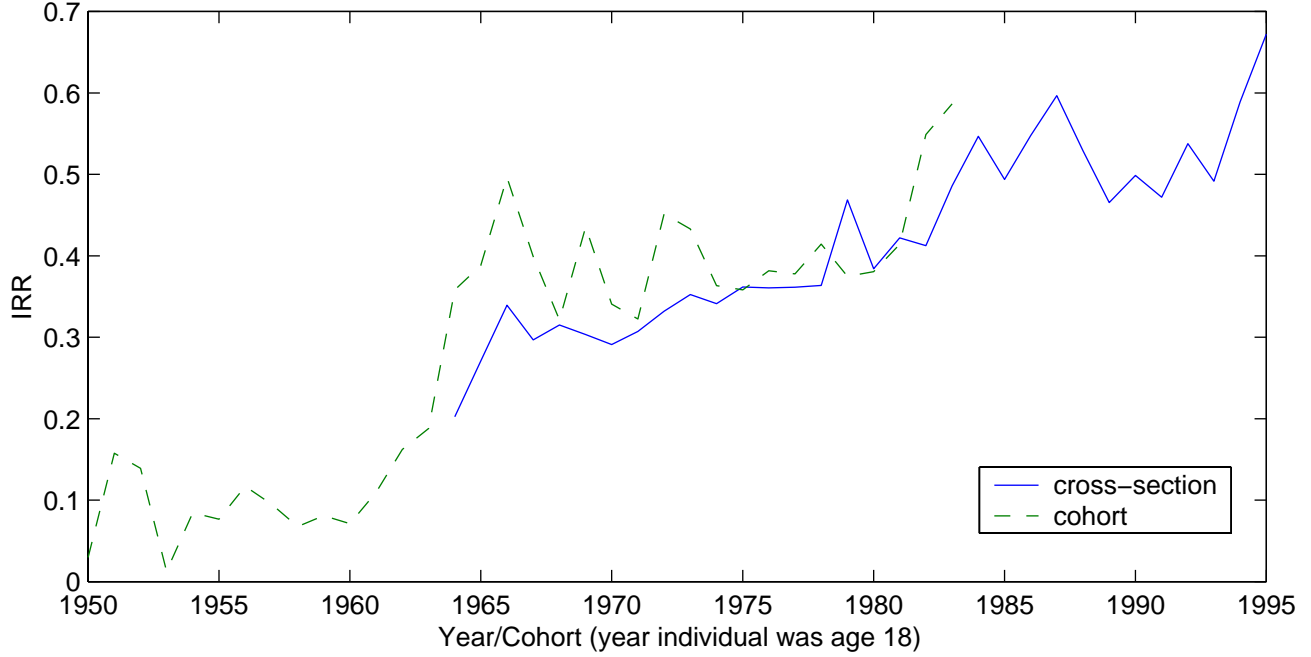Figure 7a: IRR for 10 vs. 12 Years of Education for White Men (1964–2000 CPS)

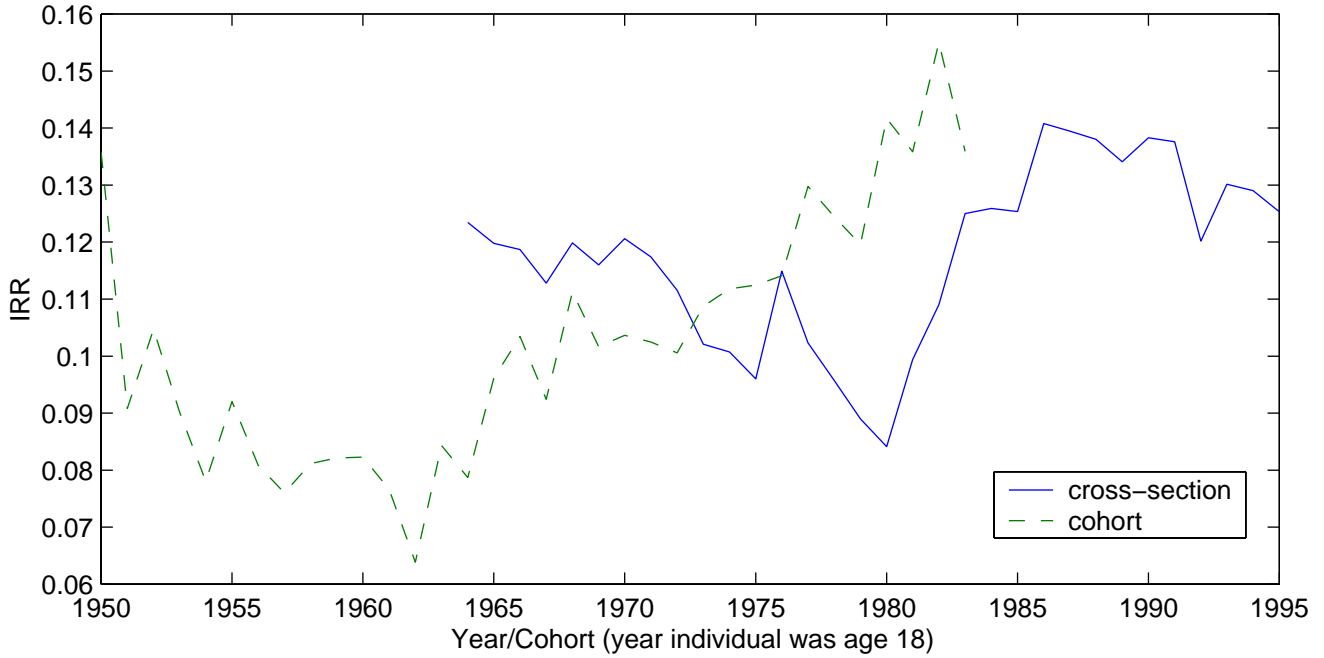Figure 7b: IRR for 12 vs. 16 Years of Education for White Men (1964–2000 CPS)

Figure 8: Three Crossings with Option Values