I Z A

Earnings Functions and Rates of Return

James J. Heckman
Lance J. Lochner
Petra E. Todd

Earnings Functions and Rates of Return

James J. Heckman
*University of Chicago, University College Dublin,
American Bar Foundation and IZA*

Lance J. Lochner
University of Western Ontario

Petra E. Todd
*University of Pennsylvania
and IZA*

Discussion Paper No. 3310
January 2008

ABSTRACT

Earnings Functions and Rates of Return[*]

The internal rate of return to schooling is a fundamental economic parameter that is often used to assess whether expenditure on education should be increased or decreased. This paper considers alternative approaches to estimating marginal internal rates of return for different schooling levels. We implement a general nonparametric approach to estimate marginal internal rates of return that take into account tuition costs, income taxes and nonlinearities in the earnings-schooling-experience relationship. The returns obtained by the more general method differ substantially from Mincer returns in levels and in their evolution over time. They indicate relatively larger returns to graduating from high school than from graduating from college, although both have been increasing over time.

Corresponding author:

James J. Heckman
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
USA
E-mail: jjh@uchicago.edu

I Introduction

The internal rate of return — the discount rate that equates the present value of two potential income streams — plays a central role in the economics of human capital (Becker, 1964). Under standard conditions, it is privately profitable to take additional schooling if the marginal internal rate of return exceeds the opportunity cost of funds.[1] It is also socially profitable if social returns and opportunity costs are captured in this calculation. Given the centrality of the marginal internal rate of return, it is surprising that so few estimates of this parameter are reported in the recent literature in labor economics which focuses almost exclusively on the college – high school wage differential. Most of the literature ignores taxes and tuition costs — two vital ingredients required to compute true economic rates of return.

The coefficient on schooling in a Mincer regression of log earnings on years of schooling and years of post-school work experience is often interpreted as an estimate of the internal rate of return. Psacharopoulos (1981), and Psacharopoulos and Patrinos (2004, 2007) report extensive tabulations of Mincer rates of return across many countries and time periods. Recent evidence (e.g., Murphy and Welch, 1990, and Katz and Autor, 1999) shows that the functional form of the Mincer model no longer adequately describes labor earnings for U.S. workers.[2] The received literature does not offer an alternative to the Mincer framework for measuring the returns to schooling.

In response to this evidence, we estimate more general models of earnings where the coefficient on schooling in a log earnings equation is not, in general, interpretable as an internal rate of return. From estimated cross section earnings functions, we compute marginal internal rates of return to education separately for black and white men across different schooling levels for different decades using data from the U.S. decennial Censuses and the Current Population Survey (CPS). Our estimates account for nonlinearities and non-separabilities

---

[1]This assumes life cycle income-education profiles cross only once for each educational level. See Hirshleifer (1970) for conditions under which the internal rate of return is a valid investment evaluation criterion. Heckman, Lochner, and Todd (2006) show how option values invalidate the internal rate of return as an investment choice criterion. Heckman and Urzua (2008) present evidence on the empirical importance of option values in schooling choices. The Mincer coefficient measures the marginal internal rate of return only under special conditions which we discuss in this paper.

[2]Heckman, Lochner, and Todd (2003, 2006) test and reject the assumptions required to use the Mincer model to estimate the internal rate of return.

in earnings functions, income taxes, and tuition. When contrasting these estimated returns with the estimated schooling coefficient from a Mincer earnings function, we find that both levels and trends in rates of return generated from the Mincer model are misleading for many schooling levels.

Most troubling, we find high estimated internal rates of return to high school completion in recent decades when we use more general non-Mincer models to compute rates of return. We also estimate returns to college completion that are roughly double typical interest rates assumed to govern schooling decisions. Our estimates call into question the validity of empirical conventions invoked in the recent literature on schooling and earnings functions. Conventional earnings functions estimated from Census or CPS data appear to be severely biased due to factors such as endogeneity of schooling, censoring of wages or both. Alternatively, the conventional income maximizing model due to Becker (1964) that motivates the literature on returns to schooling appears to be at odds with recent data.[3]

Using CPS data, we also follow cohorts of individuals over time to estimate cohort internal rates of return to high school and college in order to compare them with cross sectional estimates. The levels and patterns for the returns to college do not always agree. Most notably, cohort-based returns to college increase continuously over time for white men entering the labor market between 1960 and 1985; however, cross section estimates suggest decreasing returns throughout most of this period. Both cross section and cohort returns for blacks suggest generally rising returns over this period, although cohort-based estimates are typically higher by 2 to 3 percentage points. We estimate substantial returns to high school completion for cohorts entering the labor market in the 1970's and early 1980's, regardless of whether we use cohort or cross sectional methods. In light of the recent evidence on declining U.S. high school graduation rates (Heckman and LaFontaine, 2007), this is paradoxical. Even though returns are increasing, the relative supply is decreasing.

This paper proceeds as follows. Section II develops a simple schooling choice model extending the Mincer (1958) framework along the lines suggested by Rosen (1977) and Willis (1986). We discuss the relationship between the Mincerian rate of return and true

---

[3]A large literature, using better microdata, considers the endogeneity of schooling. See the surveys in Card, 1999, and Heckman, Lochner, and Todd, 2006.

internal rates of return to schooling for general earnings functions, accounting for income taxes, college tuition, and length of working life that may depend on schooling attainment. Section III uses seven decades of Census data (1940-2000) to calculate cross sectional internal rates of return from nonparametric estimates of earnings functions under various assumptions about income taxes, college tuition, and length of working life. We also explore differences between standard cross section estimates and cohort estimates using CPS data. Section IV concludes with a discussion of the significance of these findings.

II Internal Rates of Return

Using an income maximizing framework under perfect certainty as developed in Mincer (1958) and extended in Rosen (1977) and Willis (1986), this section explores the relationship between internal rates of return to schooling, interest rates, and the earnings function in an environment with income taxes, direct schooling costs, and length of working life that depends on schooling levels. We assume that individuals choose education levels to maximize the present value of their lifetime earnings, taking as given post-school earnings profiles.

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.[4] Individuals are assumed to choose s to maximize the present discounted value of lifetime earnings:[5]

$$V(s) = \int_{0}^{T(s)-s} (1 - \tau)e^{-(1-\tau)r(x+s)}w(s, x)dx - \int_{0}^{s} ve^{-(1-\tau)rz}dz. \qquad (1)$$

---

[4]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. Under perfect credit markets, a difference between time preferences and interest rates only affects the relative importance of non-pecuniary costs of school; otherwise, individuals maximize the present value of lifetime income (using market interest rates to discount future earnings) and allocate consumption over time according to their own rate of time preference.

[5]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.

Define $\tilde{r} = (1-\tau)r$ (the after-tax interest rate). After re-arranging terms, the first order condition for a maximum yields

$$
\tilde{r} = \frac{[T'(s)-1]e^{-\tilde{r}(T(s)-s)}w(s,T(s)-s)}{\int\limits_0^{T(s)-s} e^{-\tilde{r}x}w(s,x)dx}
$$
(Term 1)

$$
+ \frac{\int\limits_0^{T(s)-s} e^{-\tilde{r}x}w(s,x)\left[\frac{\partial log\, w(s,x)}{\partial s}\right]dx}{\int\limits_0^{T(s)-s} e^{-\tilde{r}x}w(s,x)dx} - \frac{v/(1-\tau)}{\int\limits_0^{T(s)-s} e^{-\tilde{r}x}w(s,x)dx}. \tag{2}
$$
(Term 2)   (Term 3)

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

A number of commonly invoked assumptions greatly simplify the analysis of the returns to schooling. Mincer (1958) and many other economists assume $v = 0$ (i.e., no private tuition or "psychic costs").[6] It is also commonly assumed that there is no loss of work life associated with schooling, or $T'(s) = 1$.[7] This simplifies the first order condition to

$$
\tilde{r}\int\limits_0^{T(s)-s} e^{-\tilde{r}x}w(s,x)dx = \int\limits_0^{T(s)-s} e^{-\tilde{r}x}\frac{\partial w(s,x)}{\partial s}dx.
$$

Multiplicative separability between schooling and experience components of earnings (i.e. $w(s,x) = \mu(s)\varphi(x)$) implies that log earnings profiles are parallel in experience across schooling levels. In this special case, the first order condition further reduces to $\tilde{r} = \mu'(s)/\mu(s)$. If this condition holds for all s, then wage growth must be log linear in schooling,

---

[6]In light of the substantial estimates of psychic costs presented in Carneiro, Hansen, and Heckman (2003) and Cunha, Heckman, and Navarro (2005), the assumption that $v = 0$ is very strong even if tuition costs are a small component of the present value of income. This assumption is sometimes justified by claiming that earnings in school offset tuition costs (see Becker, 1964). See the survey in Heckman, Lochner, and Todd (2006) on the evidence against this assumption.

[7]One could also assume $T(s) \to \infty$ for all s.

$\mu(s) = \mu(0)e^{\rho s}$ so $\ln w(s,x) = \ln \mu_0 + \rho s + \ln \varphi(x)$, where ρ must equal $\tilde{r}$.[8],[9]

The original Mincer model (1958) provides an equilibrium-based model of earnings determination and an easy way to estimate the internal rate of return to schooling. Mincer assumes that individuals have identical abilities and opportunities, credit markets are perfect, there are no taxes, and that the environment is perfectly certain. Occupations differ in the amount of training required. Individuals forego earnings while in school but incur no direct tuition costs. Firms offer wages that depend on the level of schooling and individuals choose their schooling levels to maximize lifetime earnings. Because individuals are assumed to be identical, a compensating wage differential is needed to induce work in occupations that require a longer training period. The compensating differential is determined by equating the present value of earnings streams net of costs associated with different levels of investment. Assuming $T(s)$ is large, and no effect of post-school work experience on wages (so $w(s,x) = w(s)$), the present value of earnings for a person with s years of schooling is

$$V(s) = \frac{e^{-rs}w(s)}{r}.$$

Equating $V(s)$ across all schooling levels, assuming homogeneity across people and using wages at $s = 0$ as a benchmark, we obtain

$$\ln w(s) = \ln w(0) + rs \tag{3}$$

where the coefficient on schooling is the marginal rate of return. Equation (3) is an arbitrage condition. The first order conditions (2) do not determine a unique schooling level under the assumptions of Mincer (1958). Because individuals are assumed to be identical and indifferent across schooling levels, schooling is indeterminate at the individual level.

In later work, Mincer (1974) draws on the analysis of Becker and Chiswick (1966) to incorporate post-school work experience as a determinant of earnings and to approximate an optimizing model of schooling choice and work experience assuming that the proportion

---

[8]Log linearity and separability are distinct. For example, the earnings function $w(s,x) = e^{\rho(x)s}w(0,x)$ is log linear in s given x, but earnings are not necessarily multiplicatively separable in schooling and experience.

[9]Optimality only requires that this condition hold at the optimal schooling level. If, as Mincer (1958) assumes, all individuals are identical and face the same schooling decision, then the condition must hold over the range of all observed schooling choices. We discuss this point further below.

of earnings foregone in investment declines linearly with experience.[10] This framework allows for heterogeneity among agents. Assuming that the proportion of income foregone in post-school investments is the same at all schooling levels, the Mincer (1974) earnings function is

$$\ln w(s, x) = \alpha_0 + \rho s + \beta_0 x + \beta_1 x^2 + \varepsilon, \tag{4}$$

where ρ is an average return to schooling across people. While ρ may vary among people, it is assumed to be uncorrelated with s.

Mincer's (1958) model and its subsequent extension to accommodate heterogeneity and post-school work experience (Mincer, 1974) provides a basis for using estimates of ρ in equation (4) as an estimate of the internal rate of return to schooling. In the Mincer (1958) model, the coefficient on schooling reflects the (after-tax) market interest rate that governs schooling choices for all individuals. It is both the marginal and average return to schooling. In Mincer (1974), the connection between the opportunity cost of funds and the coefficient on schooling is more obscure. Being an average return, ρ is not the marginal return appropriate to evaluating the optimality of schooling policy. It is the internal rate of return only if the marginal return is the average return.[11]

The elegance of the (1958) model and simplicity of the (1974) estimating equation are, no doubt, responsible for the popularity of (4) as a framework for estimating rates of return to schooling.[12] Yet, many strong assumptions are required to claim that estimates of ρ in (4) accurately measure the internal rate of return or the interest rate faced by individuals. Notably, the model assumes: (1) no direct or psychic costs of schooling, (2) no income taxes, (3) no loss of working life with additional years of schooling, (4) earnings functions are multiplicatively separable in experience and schooling, and (5) marginal returns equal average returns, which, together with multiplicative separability, implies linearity of log earnings in schooling over the observed range of schooling choices.

We find that these assumptions do not represent the U.S. economy very well in recent

---

[10]Heckman, Lochner, and Todd (2006) present a derivation of the Mincer model from first principles.

[11]Carneiro, Heckman, and Vytlacil (2006) present evidence that the marginal return does not equal the average return. See the evidence on this issue in Heckman, Lochner, and Todd (2006).

[12]See, e.g., Psacharopoulos (1981), Psacharopoulos and Patrinos (2004, 2007) and Willis (1986) for extensive surveys of the literature on Mincer returns.

decades and, when imposed, substantially affect estimated internal rates of return. Figures 1a and 1b present nonparametric estimates of the experience – log earnings profiles using 1940-2000 U.S. Census data for white males.[13] Log earnings – experience profiles for recent Censuses show convergence and not parallelism for both white and black males. Murphy and Welch (1990, 1992) and Katz and Autor (1999) document similar patterns for wages using CPS data. Formal non-parametric statistical tests, reported in Table 1 of Heckman, Lochner, and Todd (2003), reject the hypothesis of parallel experience - log earnings profiles (i.e., separability of $\log w(s, x)$ in terms of x and s) for whites for all Census years except 1940 and 1950.[14] For black males, parallelism is only rejected in later years, most likely because the samples are much smaller in the earlier years (particularly in 1940 and 1950). Formal tests of the hypothesis that log earnings are linear in education and quadratic in experience, against an alternative that allows the coefficient on education to differ across schooling levels, rejects linearity for all Census years and for both blacks and whites (p-values $< .001$).[15] Heckman, Layne-Farrar, and Todd (1996), Jaeger and Page (1996) and Hungerford and Solon (1987) provide additional evidence against linearity in log wages. Murphy and Welch (1990) argue that earnings are better represented as a quartic function of experience rather than a quadratic function.[16] In addition, college tuition costs are nontrivial and are not offset by work in school for most college students. See Table 1 of this paper for evidence on tuition costs. Carneiro, Hansen, and Heckman (2003) and Cunha, Heckman, and Navarro (2005) report estimates of substantial non-pecuniary costs of schooling, far greater than tuition costs. See the survey in Heckman, Lochner, and Todd (2006).

Rejection of these assumptions implies that the Mincer coefficient (ρ) will not, in general, equal the internal rate of return to schooling or real after-tax interest rates ($\tilde{r}$) as shown in

---

[13] Earnings correspond to annual earnings, which includes both wage and salary income and business income. Business income is not available in the 1940 Census. Appendix A provides detailed information on the construction of our data subsamples and variables. Nonparametric local linear regression is used to generate the estimates, with a bandwidth of 5 years and quartic kernel. Estimates are not very sensitive to changes in the bandwidth parameter in the range of 3-10 years.

[14]See also the estimates in Heckman, Lochner, and Todd (2006).

[15]It is also rejected for nonparametric specifications of the experience term. These results are available on request from the authors and are reported in Heckman, Lochner, and Todd (2003, 2006).

[16]However, Heckman, Lochner, and Todd (2003, 2006) show that allowing for quartic experience vs. quadratic experience has negligible effects on estimated rates of return. We present additional evidence on this question below.

equation (2). If working lives are decreasing in schooling and if tuition and psychic costs are non-negligible, the Mincer coefficient will tend to be larger than the internal rate of return or the real interest rate (see Terms 1 and 3 in equation 2). Term 2 shows how deviations from log-linearity in schooling and separability between schooling and experience may also create a wedge between estimates of ρ and $\tilde{r}$.

In the general case, $\tilde{r}$ is the marginal internal rate of return to schooling. It is the discount rate that, at an optimum, equates the net lifetime earnings for persons at the margin at all schooling levels. For people at various margins, this internal rate of return equals the after-tax interest rate in a world with perfect credit markets, once all costs and benefits from schooling are considered. Assuming income maximizing behavior, estimates of the true internal rate of return (not necessarily ρ in equation 4) can be used to determine effective interest rates faced by students. Alternatively, if interest rates are known, they may be used to judge the optimality of investment decisions.

III   Estimating Internal Rates of Return for the U.S.

We next explore how allowing for taxes, tuition, variable length of working life, and a flexible relationship between earnings, schooling and experience affects estimated internal rates of return. It is possible to calculate the internal rate of return more generally by solving for the discount rate that equates lifetime earnings streams for two different schooling levels where the earnings-schooling-experience relationships are estimated flexibly using nonparametric methods.[17] Incorporating tuition and taxes, the internal rate of return for schooling level s_1 versus s_2, $r_I(s_1, s_2)$, solves

$$
\int_0^{T(s_1)-s_1} (1-\tau)e^{-r_I(x+s_1)}w(s_1, x)dx - \int_0^{s_1} ve^{-r_I z}dz
$$
$$
= \int_0^{T(s_2)-s_2} (1-\tau)e^{-r_I(x+s_2)}w(s_2, x)dx - \int_0^{s_2} ve^{-r_I z}dz, \tag{5}
$$

---

[17]Becker (1964) states this logic, and Hanoch (1967) systematically applies it.

where we keep the (s_1, s_2) implicit in the r_I used in this expression and in our discussion.[18] As with $\tilde{r}$ above, r_I will equal the Mincer coefficient on schooling under the five assumptions described earlier. Notice that when tuition costs are negligible, proportional taxes on labor earnings have no effect on estimated internal rates of return, since they reduce earnings by the same proportion regardless of educational choices.

Using Census Data to Estimate Cross Section Internal Rates of Return

Using data for U.S. born men from 1940-2000 decennial Censuses, we examine how estimates of the internal rate of return change when different assumptions are relaxed. Tables 2a and 2b report internal rates of return to schooling for each Census year and for a variety of pairwise schooling level comparisons separately for white and black men.[19] (We only report estimates for 1960-2000 for black men due to the smaller available sample sizes for earnings in the public use Census data for 1940 and 1950.)[20] The estimates in this table 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). We initially relax only the assumptions about the functional form of the earnings equation and 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. We then predict earnings under this specification for the first 47 years of experience. The estimated IRR is the root of equation (5).[21] As a benchmark, the first row for each year reports the IRR estimate obtained from the Mincer specification for log wages (equation (4)). The IRR could equivalently be obtained from a

---

[18]We assume the absence of measured non-pecuniary costs and benefits associated with schooling. See Carneiro, Hansen, and Heckman (2003), Cunha, Heckman, and Navarro (2005) and Heckman, Lochner, and Todd (2006) for estimates of non-pecuniary costs.

[19]As lower schooling levels are reported only in broader intervals in the 1990 and 2000 Censuses, 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.

[20]Estimates for black men in these years are reported in Heckman, Lochner, and Todd (2003, 2006), but they are quite noisy for most schooling comparisons. Standard errors computed by bootstrapping are often above 4 percentage points for any IRR estimates that relax linearity in schooling prior to 1960 for blacks. In contrast, standard errors for estimates in other years are generally less than 2 percentage points for blacks and less than 1 percentage point for whites in all years.

[21]Strictly speaking, we solve for the root of the discrete time analog of equation (5). Appendix B provides a more detailed explanation of our IRR estimation procedure.

Mincer regression coefficient.[22]

Relative to the Mincer specification, row 2 relaxes the assumption of linearity in schooling by including dummy variables for the categories of schooling shown in the table. 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 for other schooling transitions. 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, which are jumps in earnings observed upon completion of a degree, are an important feature of the data.[23] Row 3 relaxes both linearity in schooling and the quadratic specification for experience (i.e., we include dummy variables for schooling categories and assume a common nonparametric experience profile), which produces similar estimates. The assumption that earnings are quadratic in experience is empirically innocuous for estimating returns to schooling once linearity in years of schooling is relaxed.

Row 4 fully relaxes all three Mincer assumptions, including parallelism in experience of log wage equations for different educational levels. Earnings functions are nonparametrically estimated as a function of experience, separately within each schooling class.[24] This procedure does not impose any assumption other than continuity on the functional form of the 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 for the estimates for recent decades, there are

---

[22]They would be identical if our internal rate of return calculations were computed in continuous time. Because we use discrete time to calculate internal rates of return, $r_I = e^\rho - 1$, which is approximately equal to ρ when it is small. Empirically, the differences between the Mincer coefficient and the Mincer return we compute are slight.

[23]We use the term "sheepskin effects" to refer to exceptionally large rates of return at degree certifying 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 except in a few years where 12th grade with and without a diploma are reported as separate categories. See Heckman, Layne-Farrar, and Todd (1996), Jaeger and Page (1996) and Hungerford and Solon (1987) for additional evidence on sheepskin effects in log wage equations.

[24]See footnote 13 for a description of the log earnings estimates and Figures 1a-1b for a graphical representation.

large differences. The functional form assumptions implied by Mincer-based estimation of the internal rate of return lead to quantitatively important biases — biases much larger than those commonly discussed as attributable to selection or ability bias (e.g., Card, 1999).

For most grade comparisons and years, the Mincer coefficient implies a lower return to schooling than do the fully nonparametric estimates, with an especially large disparity for the return to high school completion. The nonparametric estimates in Table 2a show a large increase in the return to completing high school for whites, which grew from 24% in 1940 to 52% in 2000. Much of this increase occurred between 1970 and 1990 when the IRR for high school completion rose 17 percentage points. Blacks experienced a comparable increase in returns to high school (14 percentage points) over the 1970-90 period, an increase that was already underway in the 1960s (see Table 2b). For both blacks and whites, there has been little change in the return to high school since 1990.

The nonparametrically estimated returns to college completion (14 vs. 16 years of school) are also substantially higher than the returns estimated by the Mincer model. Returns to college rose considerably in the 1980s for blacks from 23% to 31%. Returns to college completion for white men fell slightly in the 1940s, rose in the 1950s, fell again in the 1970s, and then rose in the 1980s and 1990s. During the period of rising wage returns to college from 1980-2000 (see, e.g., Katz and Autor, 1999), white men experienced an 8 percentage point rise in the internal rate of return to college completion while there was no noticeable change in this return for black men. The Mincer coefficient more closely reflects the return to completing college relative to high school (comparing 12 vs. 16 years of school). In most decades, the Mincer and nonparametric estimates are within 5 percentage points of each other. However, differences of this magnitude are considered quite large in recent debates about "the return" to schooling. For both white and black men, the time trends for this measure of college returns are similar to the trends in the return to college completion. It is interesting that the recent literature has focused so heavily on documenting and explaining the rising returns to college, when the return to high school completion appears to have risen even more substantially.[25]

---

[25]Of course, the full return to high school includes the option value of being able to go to college. Option values are ignored in computing the internal rate of return. Accounting for this option only raises the return to graduating high school. See Heckman and Urzua (2008) for estimates of the option value of schooling.

Comparing the nonparametric IRR estimates for black and white males since 1960, the return to high school completion has been 3-11 percentage points higher for black males. The internal rate of return to completing college (14 vs. 16 years) was quite similar for black and white men in 1960 and 1970, but then it jumped noticeably for blacks with little change for whites. In 1980 and 1990, the return to college completion was roughly 10 percentage points higher for blacks than whites. This gap disappeared in the 1990s. The return was only 2 percentage points higher for blacks than whites in 2000.

Estimated internal rates of return are sensitive to the assumptions imposed on the earnings model. While the assumption that log earnings profiles are quadratic in experience 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 estimates in row 1 reveals substantial differences for nearly all grade progressions across most years.

Table 3 examines how the IRR estimates for post-secondary education change when we account for income taxes (both flat and progressive) and college tuition.[26] 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 2a and 2b.) 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.[27] The average real college tuition paid by students increased steadily since 1950 as shown in Table 1. In 1990, it stood at roughly $3,500 (in 2000 dollars). Row three accounts for flat wage taxes using the estimated average marginal tax rates (τ) reported in Table 1. 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

---

[26]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. (Estimates are available upon request.) For these reasons, we only report in Table 3 the IRR estimates for comparisons of school completion levels 12 and 14, 12 and 16, and 14 and 16.

[27]We use contemporaneous college tuition costs, since cross section estimates are typically assumed to reflect the rate of return governing the behavior of cohorts making their schooling decisions that year. We discuss the comparability of cross section and cohort estimates below using CPS data.

no unearned income. See Appendix A for details on the sources and time series we use for computing tuition and taxes.

When tuition costs of schooling 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 to be the same, accounting for tuition costs has a bigger effect on the IRR estimates for blacks. 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 effects of taxes on earnings (rows 3 and 4) has little additional impact on the estimates.

As seen in equation (2), flat taxes simply multiply the effects of tuition on the IRR by the factor $\frac{1}{1-\tau}$. Progressive taxes reduce IRR estimates, because they disproportionally offset some of the increase in earnings associated with higher levels of schooling. Interestingly, the progressive nature of the tax system typically reduces rates of return by less than a percentage point compared to estimates accounting for average marginal tax rates. The U.S. tax system is not that progressive over the relevant range of earnings determining schooling decisions. Overall, failure to account for tuition and taxes leads to an overstatement of the return to college, without affecting the estimated trends over time. However, even after accounting for taxes and tuition, nonparametric estimates of the return to college completion are generally 5-10% higher than the corresponding 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. The estimates reported here 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 - work-life relationship.[28] 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 estimated internal rates of return.

---

[28]Results are available from the authors upon request.

Comparing Cross Section and Cohort IRR Estimates Using CPS Data

Previous studies 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, 2001). As a result, the rates of return to schooling estimated from cross sections of workers are likely to differ from the rates of return faced by cohorts making their schooling decisions, potentially rendering invalid the synthetic cohort approach that uses a cross section to approximate the life cycle of a representative individual.[29] We examine these differences and determine whether the differences help explain why returns to high school completion estimated for recent years appear to be out of line with plausible values of real interest rates.

We use repeated cross section data from the 1964-2006 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 life cycles. Based on our Census results, we estimate a flexible earnings specification that does not impose assumptions of log-linearity or parallelism and compute internal rates of return to high school completion (10 vs. 12 years of schooling) and four years of college (12 vs. 16 years of schooling). Our estimates take into account average marginal tax rates and tuition costs using the time series data discussed in Appendix A. Table 1 reports tuition and average marginal tax rates every ten years from 1940-2000. Because earnings are not observed at every experience level for any cohort in the sample, we require a method for extrapolating the earnings function to work experience levels not observed in the data.[30] 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. We estimate log earnings for each year or for each cohort using regressions of the form

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

---

[29]Mincer (1974) addressed this issue for earlier, more stable periods (1950 and 1960) and found that the cross section and cohort rates agreed.

[30]The earliest cohort we analyze turned age 18 in 1950 — we observe their earnings from age 32 through retirement. The latest cohort we study turned age 18 in 1985 — we follow them from ages 18-39.

where the regression coefficients are allowed to vary by schooling group.[31] 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 life cycles to produce a set of cohort-based estimates. Both sets of estimates are used to generate predicted life cycle 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.[32]

Figures 2 and 3 show cohort and cross section estimates of IRR for high school and college completion for white men. Cross section estimates are shown for each year of the sample from 1964-2000, and cohort estimates are shown for cohorts turning age 18 in 1950 through 1985.[33] Cohort and cross section estimated rates of return for high school completion (Figure 2) are fairly similar over most years of overlap, although cohort estimates are typically a few percentage points higher. Cohort bias does not explain the high estimated cross sectional internal rate of return to schooling. Cohort-based returns rose dramatically from around 10% for 1950-60 cohorts to around 40% for post-1965 cohorts. A second apparent jump in returns occurs near the end of our sample for cohorts entering the labor market in the early 1980's. Cross-section based estimates increase more slowly and consistently over most of the 1964-1995 period, although there also appears to be a noticeable jump in the early 1980s.

More noticeable differences between cohort and cross section estimates are seen in the college-going decisions of white men (Figure 3). While cross section estimates show declining returns to college over the 1970's (from 12% down to 8%), cohort-based estimates show continually increasing returns from the early 1960's to the early 1980's. The rate of return estimated from cross sections of individuals does not begin to increase until 1980. There is

---

[31]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.

[32]We also tried using a cubic and quartic function 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. Recall that for the cross section estimates, we find that both quadratics and quartics produce the same rates of return.

[33]We do not estimate returns for cohorts beyond 1985, since there are too few years of earnings observations for those cohorts to produce stable and reliable estimates.

severe cross section bias in estimating the internal rate of return to college for white men.[34]

Figures 4 and 5 report cohort and cross section returns to high school and college for black men. To increase sample sizes and reduce variability of our estimates, we group cohorts and cross sections into three-year periods. However, the estimates are still fairly noisy as can be seen in the Figures. The estimates suggest generally rising returns to high school and college throughout most of our sample period. While cross section estimates are typically lower than cohort estimates, the time trends are fairly similar.

In addition to the CPS cohort and cross section estimates, Figures 2-5 plot Mincer and nonparametric cross section estimates (accounting for tuition and average marginal tax rates) using Census data (reported in Tables 2 and 3). The nonparametric Census and the CPS cross section estimates are in remarkable agreement. The Mincer estimates are less successful overall. While they do a reasonably good job of reproducing the more general cross section estimates of the return to college compared to high school, they substantially underestimate the cross section and cohort returns to high school completion. Furthermore, the cross section returns to college do not accurately reflect the cohort estimates for white men.

IV   Summary and Conclusions

The internal rate of return to schooling is an economic parameter introduced into the human capital literature by Becker (1964). It is used to assess whether educational expenditure should be expanded or contracted. This paper considers alternative approaches for estimating the marginal internal rate of return to schooling across different schooling levels, including the widely used Mincer model of earnings. We present conditions under which the Mincer model produces true internal rates of return, but report substantial empirical evidence against the validity of those conditions. Our evidence on the empirical relevance

---

[34]The pattern for college vs. high school rates of return is relatively easy to rationalize given the well-documented fall and subsequent rise in relative wages for college workers relative to high school workers (see, e.g., Katz and Autor, 1999). The pattern for high school graduates vs. dropouts is less easily understood given the magnitude of the recent IRR estimates. Heckman (2005) shows that this pattern does not result from a steady state growth phenomenon. A secular growth in wages for all workers will imply slightly higher cohort IRR estimates relative to cross section estimates (though similar time trends), since acquiring more schooling also implies later entrance into the labor market. Cross section estimates do not capture the growth in earnings associated with later market entry.

of sheepskin effects and on nonlinearities and nonseparabilities in the earnings-education-experience relationship is generally consistent with findings reported in the empirical literature.

We use nonparametrically estimated earnings functions to generate marginal internal rates of return and examine the importance of accounting for taxes and tuition. The levels and time series patterns of marginal internal rates of return differ dramatically from those produced by a Mincer model. Deviations from parallelism and linearity in log earnings — key assumptions of the Mincer approach — are quantitatively important for estimating internal rates of return, as are the effects of taxes and tuition. Economists cannot continue to pretend that violations of the required assumptions are innocuous when using Mincer regressions to estimate 'returns to schooling'.

The standard representative agent income maximizing model that serves as the foundation for most analyses of returns to schooling suggests that marginal internal rates of return should be the same across observed schooling choices and should equal the common real interest rate faced by students. Yet, our estimates of the return to high school and college completion for recent years are substantially larger than real interest rates typically observed. One possible explanation of this evidence is the failure of the income maximizing concept as it applies to schooling decisions. Psychic costs or distaste for schooling may explain why more than fifteen percent of new cohorts of American youth do not receive a high school degree despite its high estimated monetary return. Estimates reported in Carneiro, Hansen, and Heckman (2003), Cunha, Heckman, and Navarro (2005) and Heckman, Lochner, and Todd (2006) show high psychic cost components. Some have argued that credit constraints may be important for college-going decisions. However, recent research suggests that this is not the case for most students.[35]

In a dynamic setting, uncertainty about future earnings and schooling outcomes may create a wedge between *ex post* average rates of return and real interest rates due to the option value of continuing on in school and updating of information.[36] For example,

---

[35]See Carneiro and Heckman (2002), and the evidence in Heckman, Lochner, and Todd (2006). Belley and Lochner (2007) suggest that credit constraints may have become more important recently.

[36]See the examples in Heckman, Lochner, and Todd (2006) and the empirical estimates in Heckman and Urzua (2008).

some individuals may attend college, knowing that the expected return to attending only a few years are low but that the expected return from finishing is quite high. Even if college graduation is not certain, many individuals may be willing to take the gamble of attending with the hope that they will finish successfully. Our estimated low returns to college attendance and high returns to completion are consistent with this story.[37] The distinction between *ex ante* and *ex post* returns to schooling is ignored in most of the empirical literature on returns to schooling. *Ex ante* marginal returns govern schooling choices but *ex post* returns are typically estimated. Cunha, Heckman, and Navarro (2005) and Cunha and Heckman (2007) show the empirical importance of distinguishing *ex ante* from *ex post* returns to schooling.[38] If agents are risk averse, this helps to explain high estimated returns to schooling. In addition, time inconsistent preferences and/or heavy discounting of future utility in combination with credit market constraints at the time of making schooling decisions offer possible alternative explanations for why individuals may drop out before completing a high school degree despite a high internal rate of return.[39]

A common criticism directed against the Mincer approach that also applies to the analysis in this paper is that schooling is endogenous.[40] Unfortunately, the current empirical debate on the importance of accounting for the endogeneity of schooling is far from settled. The validity of the instruments used in this literature have been challenged (see Carneiro and Heckman, 2002, and the review in Heckman, Lochner, and Todd, 2006), and many estimates make assumptions about the earnings function (e.g., log-linearity in schooling and parallelism in experience) that are invalid. It remains to be established whether the substantial increase in the returns to high school in recent decades estimated using the standard approach can be explained by changes in ability differences between high school dropouts and graduates. The evidence on the increasing schooling attainment levels of high school dropouts presented in Heckman and LaFontaine (2006) suggests otherwise.[41]

---

[37] Heckman and Urzua (2008) report substantial estimates of option values.

[38] See surveys of the evidence in Heckman, Lochner, and Todd (2006) and Abbring and Heckman (2007).

[39] See, e.g., Oreopoulos (2007).

[40] See, e.g., Card (1995, 1999); Carneiro, Hansen, and Heckman (2001); Carneiro, Heckman, and Vytlacil (2006); Griliches (1977); Heckman and Vytlacil (1998); Willis (1986); Willis and Rosen (1979); and Heckman, Lochner, and Todd (2006).

[41] This assumes that schooling attainment at the time of dropout is a measure of ability which may not be true if social promotion has been increasing.

In addition, the growth in high school dropout rates since 1970, documented in Heckman and LaFontaine (2007), and the continued increase in rates of return to high school (as measured by cross section or cohort estimates) since that time poses a serious challenge to explanations based on this premise.

The Census and CPS data are the foundation for the widely publicized evidence on the changing structure of wages and returns to schooling (see e.g., Katz and Murphy, 1992; Murphy and Welch, 1990, 1992; and Katz and Autor, 1999). Our estimated *ex post* internal rates of return are implicit in the earnings regression coefficient estimates reported in an entire recent literature. We take the earnings patterns analyzed in this literature and show the implied internal rates of return after accounting for taxes and tuition. To the extent that some of our estimates raise suspicion, they should also raise questions about a large literature on the changing wage structure. While widely used Census and CPS data have large samples, they lack good instruments. Following the tradition of the entire influential CPS-Census based literature on returns to schooling that defines the "facts" about rising returns to schooling, we have ignored censoring, ability bias and endogeneity of schooling. Determining the contribution of these factors to estimated rates of return is an important task requiring richer data than are available in the Census or CPS.

Most of the recent literature has focused on the rising returns to college. Our estimates suggest an even greater increase in the returns to high school, raising the obvious question: why do so many individuals continue to drop out of high school? The answer may rely on high 'psychic' costs of school, credit constraints, or unobserved differences in ability between dropouts and graduates, but the literature has yet to convincingly sort out these competing explanations. This is an important task which can only be done using richer microdata.[42]

---

[42]See Oreopoulos (2007) and Heckman and Urzua (2008) for analyses of this question.

Appendix A: Data description

Census Data

The Census samples used in this paper are taken from the 1940, 1950, 1960, 1970, 1980, 1990, and 2000 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 and 2000 samples are not. For 1990 and 2000, we use person weights to re-weight the samples back to random proportions.

The following sample restrictions are imposed for each Census year:

age: Sample includes individuals age 16-75. 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 black or white are included in the analysis.

country of birth/residence: Only individuals born and residing in the U.S. 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. Because we estimate models of log earnings, all men with zero earnings are dropped.

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 and 2000 Censuses, years of education is reported 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. 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 1964-2006 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 1992-2006, 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

For all higher education institutions in the U.S., we divide total revenue from tuition and fees by total enrollment to obtain estimates of the private tuition cost of college from 1940-2000. Total enrollment is taken from Snyder (1993, Table 24) and Table 171 of the *Digest of Education Statistics 2005*. Tuition and fees up to 1995 are taken from Snyder (1993, Table 33) and Table 328 in the *Digest of Education Statistics 2005*. Tuition and fees from 1996-2000 sum tuition and fees across institution types (public, private non-profit, private for-profit) from various tables and years of the *Digest of Education Statistics*. (Tuition and fees for private for-profit institutions in 1996-97 is not reported, so we use the average of 1995-96 and 1996-97 values.)

Tax Rate Time Series

We obtain the average marginal federal tax rate time series for 1940-1995 from Barro and Sahasakul (1983) and Mulligan and Marion (2001, Table 1, column 1). Average marginal federal tax rates for 1996-2005 are from calculations by Feenberg using TAXSIM and are reported at `http://www.nber.org/taxsim`; tax rates beyond 2005 are assumed to equal that of 2005.

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 1960-2000, we use the TAXSIM program available at `http://www.nber.org/taxsim`. (See Feenberg and Coutts, 1993, for details of the TAXSIM model.) For 1940 and 1950, we use the actual federal tax schedules (Form 1040) as reported in the *Statistics of Income*.

Appendix B: Information on Method for Calculating IRR

This appendix outlines the procedure used to estimate IRRs in this paper.

Step 1: Estimate log earnings functions using local linear regression

In estimating the nonparametric matching regressions, we use local linear regression

methods. Let $y_i = \ln w(s_i, x_i)$ denote log earnings, which depends on schooling level s_i and experience x_i. As discussed in Fan and Gijbels (1996), the local linear estimator for the conditional expectation $E[y_i | s_i = s_0, x_i = x_0]$ can be computed from the minimization problem

$$\min_{a,b} \sum_{i=1}^{n} (y_i - a - b_1(x_i - x_0))^2 K\left(\frac{x_i - x_0}{h_n}\right) \mathbf{1}(s_i = s_0)$$

where $K(\cdot)$ is a kernel function and $h_n > 0$ is a bandwidth which converges to zero as $n \to \infty$.[43] The indicator $\mathbf{1}(s_i = s_0)$ denotes that only observations with schooling level equal to s_0 are used in the computation. The estimator of the conditional mean $E[y_i | s_i = s_0, x_i = x_0]$ is $\widehat{a}$. The local linear estimator can be expressed as a weighted average of the y_i observations with schooling level s_0, $\sum_{i=1}^{n_{s_0}} y_i W_i(x_0)$, where the weights are

$$W_i(x_0) = \frac{K_i \sum_{j=1}^{n_{s_0}} K_j^2 - K_i \sum_{k=1}^{n_{s_0}} K_k}{\sum_{k=1}^{n_{s_0}} K_k \sum_{j=1}^{n_{s_0}} K_j^2 - (\sum_{k=1}^{n_{s_0}} K_k)^2}$$

and n_{s_0} are the number of observations with schooling level s_0.

Nonparametric regression can be computationally burdensome when there are large numbers of observations. To facilitate the computation, we take advantage of the fact that we have many earnings observations at each value of experience, x_i. Our local regression estimator for $E[y_i | s_i = s_0, x_i = x_0]$ is given by

$$\frac{\sum_{i=1}^{N_x} n_{x_i} \bar{y}_i W_i(x_0) \mathbf{1}(s_i = s_0)}{\sum_{i}^{N_x} n_{x_i} W_i(x_0) \mathbf{1}(s_i = s_0)},$$

where $\bar{y}$ represents average earnings at experience level x_i, n_{x_i} represents the number of observations at that experience level, and N_X represents the number of distinct values of potential experience.[44] Note that this expression is equivalent to estimating the nonparam-

---

[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 an additional 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 by introducing additional weights. That is, the effective number of observations at each experience level depend on the sampling weights, normalized so that the weights sum to the total number of observations in the sample.

eteric earnings-experience relationship separately for each education level. The bandwidth choice was set equal to 5 years of experience for all census years.

Step 2: Generate predicted earnings profiles

In this step, we compute $\hat{w}(s,x) = exp\left(ln[\widehat{w(s,x)}]\right)$ for all desired schooling and experience levels, where $ln[\widehat{w(s,x)}]$ is predicted log earnings from Step 1 for the fully nonparametric model.[45] For the semiparametric models shown in Tables 2 and 3, the predicted earnings are obtained directly from the earnings regression coefficients.

Step 3: Solve for IRR accounting for taxes and tuition

We use the predicted earnings profiles for two potential schooling levels s_1 and $s_2 > s_1$ computed in Step 2 along with estimated college tuition, v, and taxes, $\tau(w(s,x))$, to solve for the internal rate of return, $r_I(s_1, s_2)$, in the following equation:

$$
\sum_{x=1}^{47}(1+r_I)^{-(x+s_1)}\left[w(s_1,x) - \tau(w(s_1,x))\right] - \mathbf{1}(s_1 > 12)\sum_{z=1}^{s_1-12}(1+r_I)^{-(z+12)}v
$$
$$
= \sum_{x=1}^{47}(1+r_I)^{-(x+s_2)}\left[w(s_2,x) - \tau(w(s_2,x))\right] - \mathbf{1}(s_2 > 12)\sum_{z=1}^{s_2-12}(1+r_I)^{-(z+12)}v.
$$

The indicator function, $\mathbf{1}(\cdot)$, equals one if the argument is true and zero otherwise. This equation assumes earnings (and taxes) are zero while in school and discounts earnings net of taxes and college tuition back to the age of school entry. Taxes either equal the average marginal tax rate times earnings that period or are derived from actual tax schedules based on earnings that period. Schooling costs are assumed to equal zero for grades 1-12 and to equal the average college tuition level that year for all grades higher than 12. See Appendix A for further details on the data we use on tuition costs and taxes.

---

[45]See Heckman, Lochner, and Todd (2003, 2006) for additional adjustments used to account for differences in residual variance across schooling groups.

References

Abbring, J. H. and J. J. Heckman (2007). Econometric evaluation of social programs, part III: Distributional treatment effects, dynamic treatment effects, dynamic discrete choice, and general equilibrium policy evaluation. In J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Volume 6B, pp. 5145–5303. Amsterdam: Elsevier.

Barro, R. J. and C. Sahasakul (1983, October). Measuring the average marginal tax rate from the individual income tax. *Journal of Business 56*(4), 419–452.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* New York: National Bureau of Economic Research, distributed by Columbia University Press.

Becker, G. S. and B. R. Chiswick (1966, March). Education and the distribution of earnings. *American Economic Review 56*(1/2), 358–369.

Belley, P. and L. Lochner (2007, December). The changing role of family income and ability in determining educational achievement. *Journal of Human Capital 1*(1), 37–89.

Card, D. (1995). Earnings, schooling, and ability revisited. In S. Polachek (Ed.), *Research in Labor Economics*, pp. 23–48. Greenwich, CT: JAI Press.

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 5, pp. 1801–1863. New York: North-Holland.

Card, D. and T. Lemieux (2001, May). Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *Quarterly Journal of Economics 116*(2), 705–746.

Carneiro, P., K. Hansen, and J. J. Heckman (2001, Fall). Removing the veil of ignorance in assessing the distributional impacts of social policies. *Swedish Economic Policy Review 8*(2), 273–301.

Carneiro, P., K. Hansen, and J. J. Heckman (2003, May). Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review 44*(2), 361–422.

Carneiro, P. and J. J. Heckman (2002, October). The evidence on credit constraints in post-secondary schooling. *Economic Journal 112*(482), 705–734.

Carneiro, P., J. J. Heckman, and E. J. Vytlacil (2006). Estimating marginal and average returns to education. Unpublished manuscript, University of Chicago. Under revision.

Cunha, F. and J. J. Heckman (2007). The evolution of uncertainty in labor earnings in the U.S. economy. Unpublished manuscript, University of Chicago. Under revision.

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

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

Feenberg, D. and E. Coutts (1993, Winter). An introduction to the TAXSIM model. *Journal of Policy Analysis and Management 12*(1), 189–194.

Griliches, Z. (1977, January). Estimating the returns to schooling: Some econometric problems. *Econometrica 45*(1), 1–22.

Hanoch, G. (1967, Summer). An economic analysis of earnings and schooling. *Journal of Human Resources 2*(3), 310–329.

Heckman, J. J. (2005). Notes on cross-section and cohort internal rates of return. Unpublished manuscript, Economics 345, University of Chicago.

Heckman, J. J. and P. A. LaFontaine (2006, July). Bias corrected estimates of GED returns. *Journal of Labor Economics 24*(3), 661–700.

Heckman, J. J. and P. A. LaFontaine (2007). The American high school graduation rate: Trends and levels. Working Paper 13670, National Bureau of Economic Research.

Heckman, J. J., A. Layne-Farrar, and P. E. Todd (1996, November). Human capital pricing equations with an application to estimating the effect of schooling quality on earnings. *Review of Economics and Statistics 78*(4), 562–610.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2003). Fifty years of Mincer earnings regressions. Technical Report 9732, National Bureau of Economic Research.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006). Earnings equations and rates of return: The Mincer equation and beyond. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Chapter 7, pp. 307–458. Amsterdam: Elsevier.

Heckman, J. J. and S. Urzua (2008). Option values and rates of return to schooling. Unpublished manuscript, University of Chicago, Department of Economics.

Heckman, J. J. and E. J. Vytlacil (1998, Fall). Instrumental variables methods for the correlated random coefficient model: Estimating the average rate of return to schooling when the return is correlated with schooling. *Journal of Human Resources 33*(4), 974–987.

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

Hungerford, T. and G. Solon (1987, February). Sheepskin effects in the returns to education. *Review of Economics and Statistics 69*(1), 175–177.

Jaeger, D. A. and M. E. Page (1996, November). Degrees matter: New evidence on sheepskin effects in the returns to education. *Review of Economics and Statistics 78*(4), 733–740.

Katz, L. F. and D. H. Autor (1999). Changes in the wage structure and earnings inequality. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3, Chapter 25, pp. 1463–1555. New York: North-Holland.

Katz, L. F. and K. M. Murphy (1992, February). Changes in relative wages, 1963-1987: Supply and demand factors. *Quarterly Journal of Economics 107*(1), 35–78.

MaCurdy, T. E. and T. Mroz (1995). Measuring macroeconomic shifts in wages from cohort specifications. Unpublished manuscript, Stanford University, Department of Economics.

Mincer, J. (1958, August). 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: Columbia University Press for National Bureau of Economic Research.

Mulligan, C. B. and J. G. Marion (2001, February). Average marginal tax rates revisited: Comment. Working Paper 01.4, Harris School of Public Policy.

Murphy, K. M. and F. Welch (1990, April). Empirical age-earnings profiles. *Journal of Labor Economics 8*(2), 202–229.

Murphy, K. M. and F. Welch (1992, February). The structure of wages. *Quarterly Journal of Economics 107*(1), 285–326.

Oreopoulos, P. (2007, December). Do dropouts drop out too soon? Wealth, health, and happiness from compulsory schooling. *Journal of Public Economics 91*(11-12), 2213–2229.

Psacharopoulos, G. (1981, October). Returns to education: An updated international comparison. *Comparative Education 17*(3), 321–341.

Psacharopoulos, G. and H. A. Patrinos (2004, August). Returns to investment in education: A further update. *Education Economics 12*(2), 111–134.

Psacharopoulos, G. and H. A. Patrinos (2007, June). Returns to education: An international update. World Bank Policy Research working paper, Washington, DC.

Rosen, S. (1977). Human capital: A survey of empirical research. In R. Ehrenberg (Ed.), *Research in Labor Economics*, Volume 1, pp. 3–40. Greenwich, CT: JAI Press.

Snyder, T. D. (1993). *120 Years of American Education: A Statistical Portrait.* Washington, D.C.: U.S. Dept. of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Snyder, T. D. and C. M. Hoffman (2005). *Digest of education statistics 2005*. Washington, DC: National Center for Education Statistics, Institute of Education Science.

Willis, R. J. (1986). Wage determinants: A survey and reinterpretation of human capital earnings functions. In O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics, Volume*, pp. 525–602. New York: North-Holland.

Willis, R. J. and S. Rosen (1979, October). Education and self-selection. *Journal of Political Economy 87*(5, Part 2), S7–S36.

Table 1: Average College Tuition Paid (in 2000 Dollars) and Average Marginal Tax Rates

Year	Avg. Tuition	Avg. Marginal Tax Rate
1940	$1,710	5.6%
1950	$1,336	19.6%
1960	$1,978	23.4%
1970	$2,596	24.3%
1980	$2,380	30.4%
1990	$3,569	23.3%
2000	$4,413	25.3%

Notes: Average college tuition paid calculated from *120 Years of American Higher Education* and the 1999 *Digest of Educational Statistics*. Average marginal tax rates are taken from Barro and Sahasakul (1983) and Mulligan and Marion (2001). See Appendix A for a detailed description of these series.

Table 2a: Internal Rates of Return for White Men: Earnings Function Assumptions

| | 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
Relax Linearity in S	*	-7	39	7	15	24
Relax Linearity in S & Quad. in Exp.	*	-3	30	10	15	20
Relax Linearity in S & Parallelism	*	20	50	10	16	26
2000						
Mincer Specification	*	14	14	14	14	14
Relax Linearity in S	*	-12	44	8	15	23
Relax Linearity in S & Quad. in Exp.	*	-5	33	10	15	20
Relax Linearity in S & Parallelism	*	9	52	11	18	29

Notes: Data taken from 1940-2000 Decennial Censuses. In 1990 and 2000, comparisons of 6 vs. 8 and 8 vs. 10 cannot be made given data restrictions. For these years, the 8 vs. 10 column reports calculations based on a comparison of 6 and 10 years of schooling. All IRR calculations assume work lives of 47 years. "*" denotes cell sizes too small to obtain reliable estimates.

Table 2b: Internal Rates of Return for Black Men: Earnings Function Assumptions

	Schooling Comparisons					
	6-8	8-10	10-12	12-14	12-16	14-16
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
Relax Linearity in S	*	-5	41	15	20	25
Relax Linearity in S & Quad. in Exp.	*	-3	35	17	19	22
Relax Linearity in S & Parallelism	*	16	58	18	25	35
2000						
Mincer Specification	*	18	18	18	18	18
Relax Linearity in S	*	-8	44	15	20	24
Relax Linearity in S & Quad. in Exp.	*	-2	38	16	19	21
Relax Linearity in S & Parallelism	*	31	56	19	24	31

Notes: Data taken from 1960-2000 Decennial Censuses. In 1990 and 2000, comparisons of 6 vs. 8 and 8 vs. 10 cannot be made given data restrictions. For these years, the 8 vs. 10 column reports calculations based on a comparison of 6 and 10 years of schooling. All IRR calculations assume work lives of 47 years. "*" denotes cell sizes too small to obtain reliable estimates.

Table 3: Internal Rates of Return for White & Black Men: Accounting for Taxes and Tuition (General Non-Parametric Specification)

| | | Schooling Comparisons | | | | | |
| | | Whites | | | Blacks | | |
		12-14	12-16	14-16	12-14	12-16	14-16
1940	No Taxes or Tuition	11	18	26			
	Including Tuition Costs	9	15	21			
	Including Tuition & Flat Taxes	8	15	21			
	Including Tuition & Prog. Taxes	8	15	21			
1950	No Taxes or Tuition	3	8	19			
	Including Tuition Costs	3	8	16			
	Including Tuition & Flat Taxes	3	8	16			
	Including Tuition & Prog. Taxes	3	7	15			
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
2000	No Taxes or Tuition	11	18	29	19	24	31
	Including Tuition Costs	9	14	22	14	18	22
	Including Tuition & Flat Taxes	8	14	20	13	16	21
	Including Tuition & Prog. Taxes	8	13	19	12	16	20

Notes: Data taken from 1940-2000 Decennial Censuses. See discussion in text and Appendix A for a description of tuition and tax amounts. All IRR calculations assume work lives of 47 years. Missing rows correspond to cell sizes too small to obtain reliable estimates.

Figure 1a: Experience–Earnings Profiles, 1940–1970

1940 Census, White Males

1940 Census, Black Males

1950 Census, White Males

1950 Census, Black Males

1960 Census, White Males

1960 Census, Black Males

1970 Census, White Males

1970 Census, Black Males
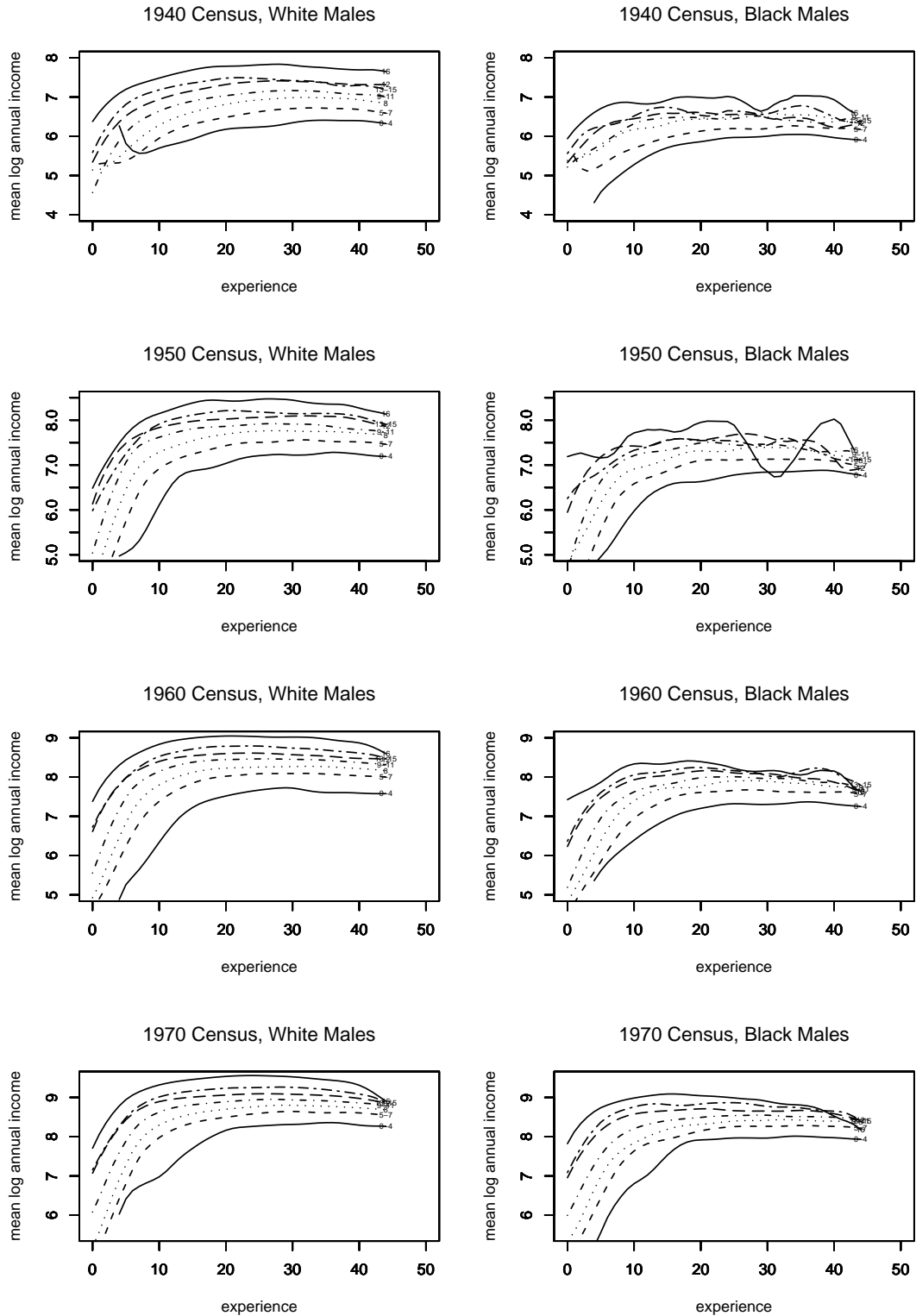
Figure 1b: Experience-Earnings Profiles 1980-2000

Figure 1b: Experience–Earnings Profiles, 1980–2000

1980 Census, White Males

1980 Census, Black Males

1990 Census, White Males

1990 Census, Black Males

2000 Census, White Males

2000 Census, Black Males
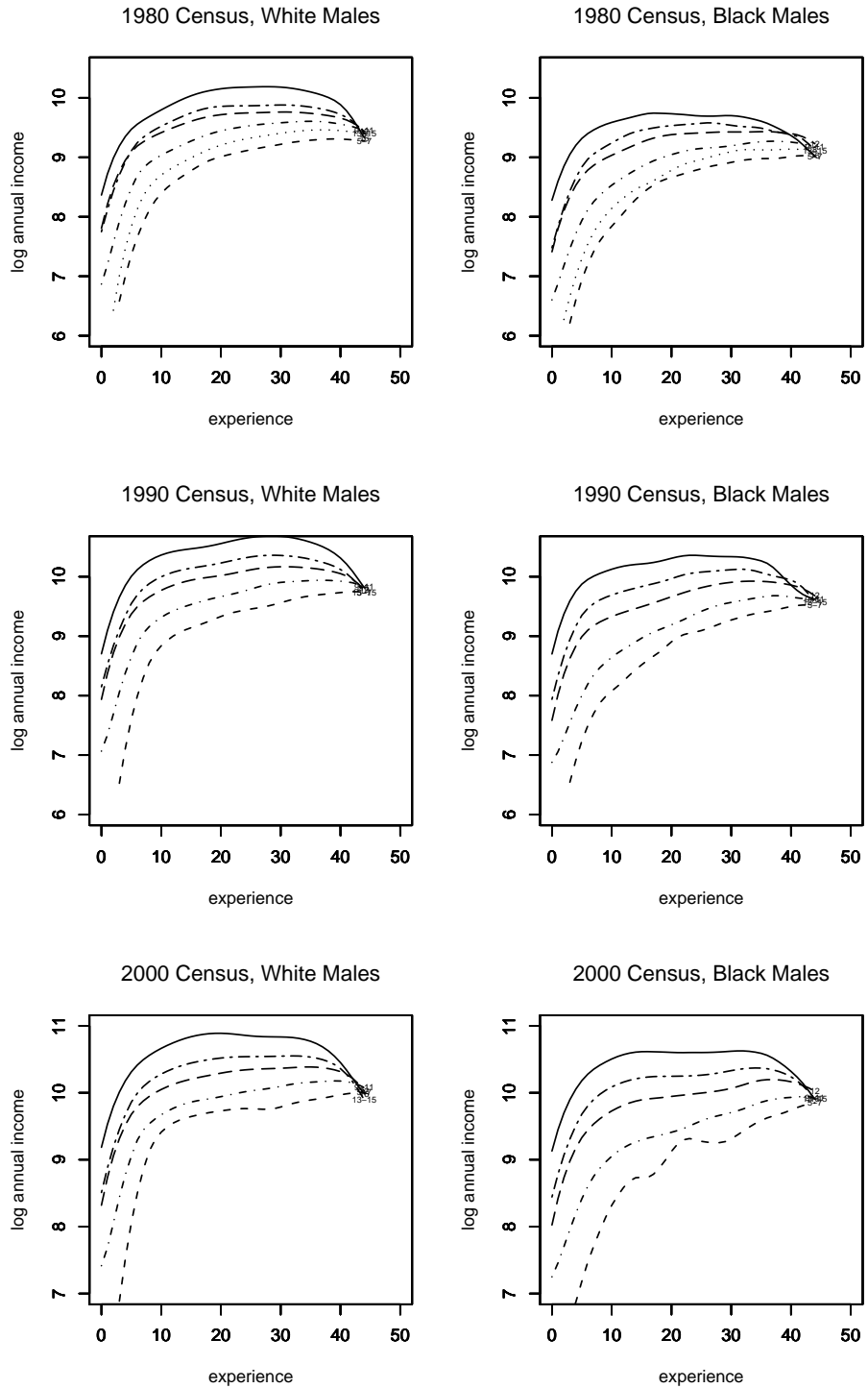
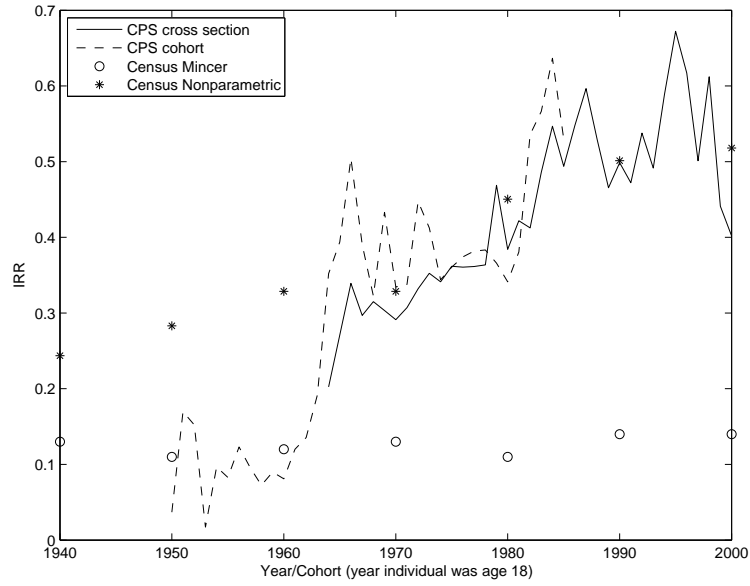Figure 2: IRR for 10 vs. 12 Years of Education, White Men (CPS and Census)

Figure 3: IRR for 12 vs. 16 Years of Education, White Men (CPS and Census)
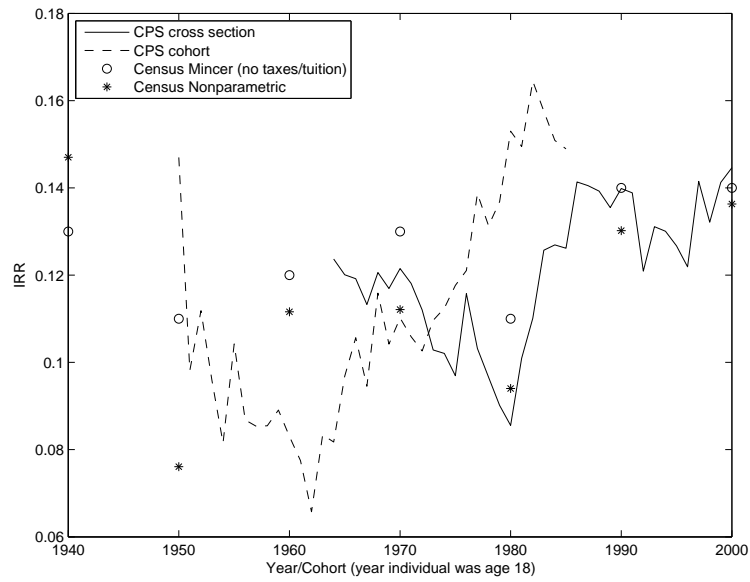
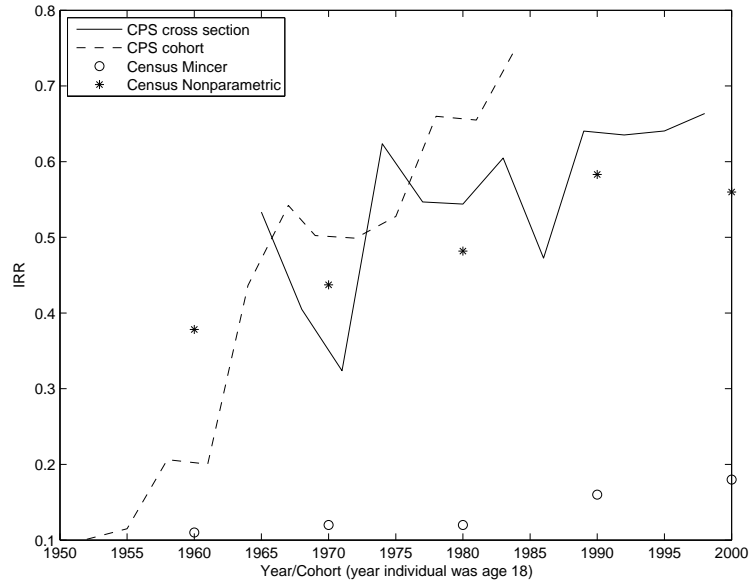Figure 4: IRR for 10 vs. 12 Years of Education, Black Men (CPS and Census)

Figure 5: IRR for 12 vs. 16 Years of Education, Black Men (CPS and Census)