The Rate of Return to Schooling: The Mincer Equation and Beyond

James J. Heckman
University of Chicago

Econ 350, Winter 2023

Partly based on

Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond

James J. Heckman
University of Chicago

Lance Lochner
University of Western Ontario

Petra Todd
University of Pennsylvania

- Schooling is a form of human capital
- Schooling is an important determinant of earnings
- We will examine
  a. Who gets various levels of schooling and why?
  b. What is the "return" to schooling? How much does it depend on ability and family background? What are the proper measures of ability?
  c. What are trends in returns to schooling?
  d. Are there effective policy initiatives?

- “Mincer returns” are the coin of the realm
- But what are they? Are they a guide to policy? Is schooling policy an effective strategy for promoting social mobility?

- **Challenges faced in estimating rates of return**
  a. Lifetime earnings profiles required
  b. Observed earnings profiles subject to selection bias
  c. Identifying and quantifying nonmarket benefits and nonpecuniary costs is a difficult task
  d. Identifying agent expectations about future returns
  e. Cohort effects (school quality; labor market entry effects)
- **We address each point**

1. Mincer's Earnings Equation

$$\ln Y(s,x) = \alpha + \rho_s s + \beta_0 x + \beta_1 x^2 + \varepsilon \qquad (1)$$

- $Y(s,x)$: wage or earnings at schooling level s.
- Work experience x.
- ρ_s: *"rate of return to schooling"* (same for all schooling levels) $E(\varepsilon|s,x) = 0$.

Two Models

- Two conceptually different frameworks.
- Algebraically similar.
- Different economic content.

Mincer 1

- Agents *ex ante* identical
- Schooling costs: foregone earnings
- $Y(s)$: annual earnings of an individual with s years of education.
- r: interest rate
- T: the length of working life, assumed not to depend on s.
- The present value of earnings of schooling s:

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

- Indifference $\Rightarrow$ equalizing differentials:

$$\begin{aligned}\ln Y(s) &= \ln Y(0) + rs + \ln \quad ((1 - e^{-rT})/(1 - e^{-r(T-s)})) \\ &= \ln Y(0) + rs \qquad\qquad T \to \infty.\end{aligned}$$

- **Hedonic function**
- Coefficient on s:
  1. (a) "Return" to schooling
  2. (b) Interest rate
- The model tautologically assumes equilibrium
- Ignores life cycle evolution of inequality

Mincer Equation 2: The Accounting Identity Model

THE UNIVERSITY OF CHICAGO

- Adds post-school investment (x).
- Persons *ex ante* heterogeneous.
- The compensating differences motivation of the first Mincer model is absent.
- ρ_s varies in the population to reflect heterogeneity in "returns."
- Key issue in recent literature.

$$\ln Y(s, x) = \alpha + \rho_s s + \beta_0 x + \beta_1 x^2 + \varepsilon \tag{1}$$

- A hedonic (pricing) equation.
- When is a price a rate of return?

THE UNIVERSITY OF CHICAGO

- P_t: potential earnings at age t.
- Costs of investments in training C_t: $C_t := k_t P_t$.
- k_t: proportion of earnings potential spent investing (foregone earnings).
- ρ_t: *average* return to training investments made at age t.
- Potential earnings at t:

$$P_t := P_{t-1}(1 + k_{t-1}\rho_{t-1}) := \prod_{j=0}^{t-1}(1 + \rho_j k_j)P_0, t \geq 1$$

THE UNIVERSITY OF CHICAGO

- Formal schooling: years spent in full-time investment ($k_t = 1$)
- Assumed to take place at the beginning of life and to yield a rate of return ρ_s that is constant across all years of schooling

- Assume that the "rate of return" to post-school investment is (a) constant over ages and (b) equals ρ_0:

$$\begin{aligned}
\ln P_t &:= \ln P_0 + s \ln(1 + \rho_s) + \sum_{j=s}^{t-1} \ln(1 + \rho_0 k_j) \\
&\cong \ln P_0 + s\rho_s + \rho_0 \sum_{j=s}^{t-1} k_j
\end{aligned}$$

- Last step assumes $\rho_s \cong 0$ and $\rho_0 \cong 0$.
- $\ln(1 + x) \cong x$.

- Mincer approximates the Ben-Porath (1967) model by assuming a linearly declining rate of post-school investment
- $k_{s+x} = \kappa \left(1 - \frac{x}{T}\right)$ 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

$$\ln P_{x+s} \approx \ln P_0 + \rho_s s + \left(\rho_0 \kappa + \frac{\rho_0 \kappa}{2T}\right) x - \frac{\rho_0 \kappa}{2T} x^2.$$

Observed Earnings

- Observed earnings: potential earnings less investment costs (which are foregone earnings)
- $\therefore$ Mincer equation:

$$
\begin{aligned}
& \ln Y(s, x) \\
& \cong \ln P_{x+s} - \kappa \left(1 - \frac{x}{T}\right) \\
& = [\ln P_0 - \kappa] + \rho_s s + \left(\rho_0 \kappa + \frac{\rho_0 \kappa}{2T} + \frac{\kappa}{T}\right) x - \frac{\rho_0 \kappa}{2T} x^2 .
\end{aligned} \tag{2}
$$

- Equation (1) without an error term.

Random Coefficients

- Mincer formulates a more general model.
- κ and ρ_s may differ across persons.
- **Random coefficient model:**

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

- Assume means exist.
- Let $\bar{\alpha} = E(\alpha_i)$, $\bar{\rho}_s = E(\rho_{si})$, $\bar{\beta}_0 = E(\beta_{0i})$, $\bar{\beta}_1 = E(\beta_{1i})$.
- Write the above equation as

$$\begin{aligned} &\ln Y(s,x) \\ &\quad = \bar{\alpha} + \bar{\rho}_s s + \bar{\beta}_0 x + \bar{\beta}_1 x^2 \\ &\qquad + \underbrace{[(\alpha - \bar{\alpha}) + (\rho_s - \bar{\rho}_s)s + (\beta_0 - \bar{\beta}_0)x + (\beta_1 - \bar{\beta}_1)x^2]}_{\text{error term (for whole expression)}} \end{aligned}$$

- Hedonic wage equation with individual-specific prices.

Correlated Random Coefficient Model

- ρ_s maybe correlated with s.
- Topic of huge recent literature on "heterogenous treatment effects."

THE UNIVERSITY OF CHICAGO

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

Estimates Based on Cross-Section

Invokes synthetic cohort assumption: cross-section profiles are valid predictors of life cycle outcomes (sometimes called "rational expectations" assumption)

- Will examine validity of this assumption
- Whole literature has emerged about forecasting and measuring expectations

(i)

Log-earnings experience profiles parallel across schooling levels:

$$\left(\frac{\partial \ln Y(s,x)}{\partial s \partial x} = 0\right)$$

(ii)

Log-earnings age profiles diverge with age across schooling levels:

$$\left(\frac{\partial \ln Y(s,x)}{\partial s \partial t} = \frac{\rho_0 \kappa}{T} > 0 \right).$$

Exercise: Show this. Differentiate Equation (1) with respect to t.

Overtaking Age

(iii)

The variance of earnings over the life cycle has a U-shaped pattern.

Intuition

- There is an age in the life cycle at which the interpersonal variance in earnings is minimized.
- Consider the accounting identity for observed earnings in levels at experience x and schooling s:

$$Y(s,x) := P_s + \underbrace{\rho_0 \sum_{j=s}^{s+x-1} C_j}_{\text{returns}} - \underbrace{C_{s+x}}_{\substack{\text{current cost}\\\text{(earnings}\\\text{foregone)}}} .$$

- "Overtaking age" when returns $\backsimeq$ costs.

- In logs,

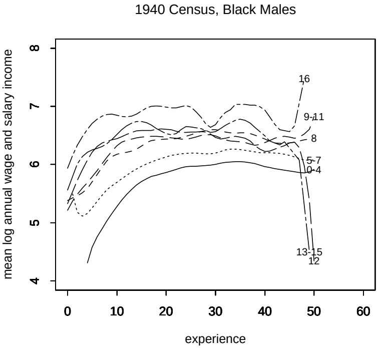
$$\ln Y(s,x) \approx \ln P_s + \rho_0 \sum_{j=0}^{x-1} k_{s+j} - k_{s+x}.$$

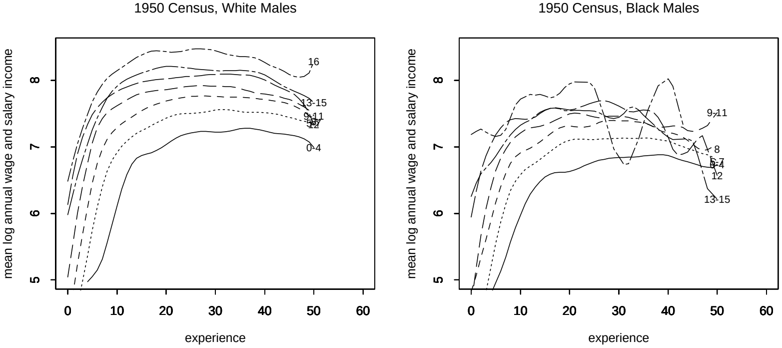
- Interpersonal differences in observed log earnings of individuals with the same P_0 and ρ_0 arise because of differences in $\ln P_s$ and in post-school investment patterns as determined by k_j.
- When $\ln P_s$ and κ or the k_{s+j} are uncorrelated, the variance of log earnings reaches a minimum when experience is approximately equal to $1/\rho_0$.
- Intuitively, the last two terms cancel.
- See Heckman et al. (2006) for proof.
  Exercise: Work through the derivation of the overtaking age.
- Relevant for measuring life cycle return to schooling.

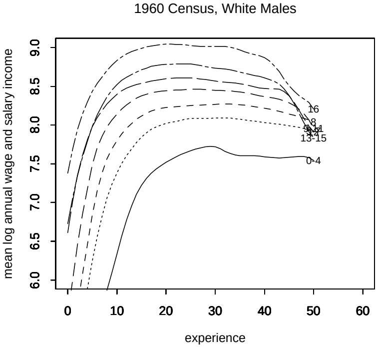
3. Empirical Evidence on the Mincer Model

Earnings By Experience

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




1950 Census, White Males
mean log annual wage and salary income
experience
16
13-15
9-11
12
0-4
1950 Census, Black Males
mean log annual wage and salary income
experience
9-11
8
12
13-15


1960 Census, White Males
mean log annual wage and salary income
experience
16
8
13-15
0-4

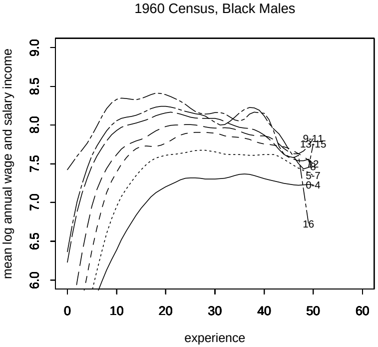
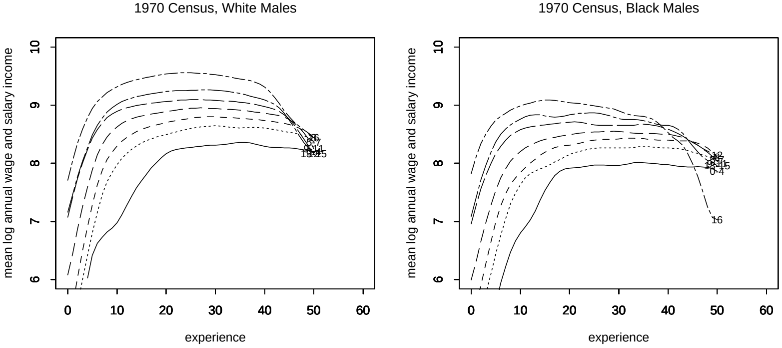
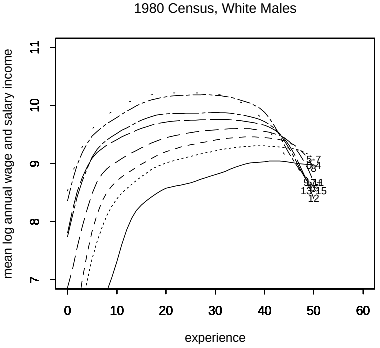
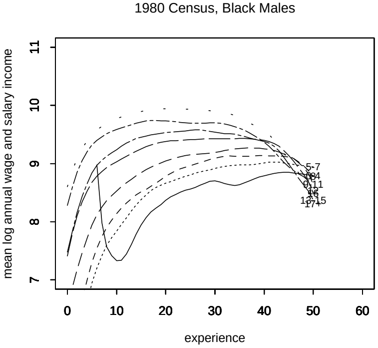
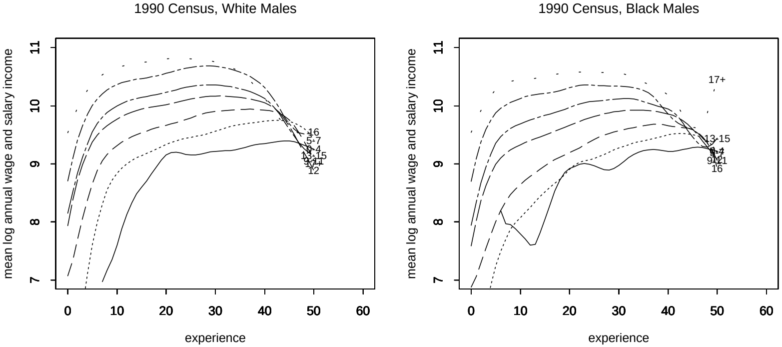

1960 Census, Black Males
mean log annual wage and salary income
experience
9-11
13-15
12
8
5-7
0-4
16

Figure 1b: Experience-Earnings Profiles, 1970-1990



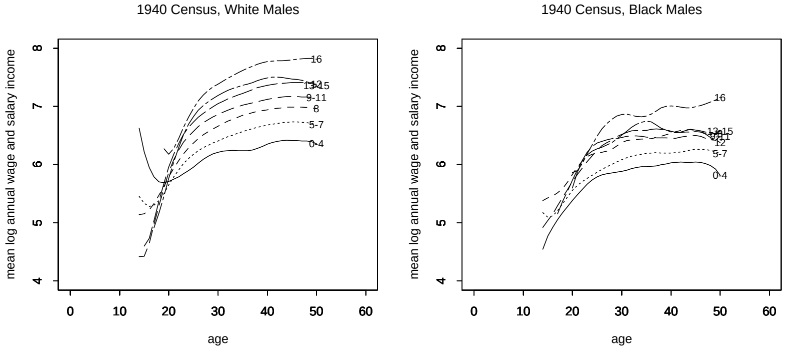

1980 Census, White Males
mean log annual wage and salary income
experience
7
8
9
10
11
0
10
20
30
40
50
60
5-7
8
13-15
12

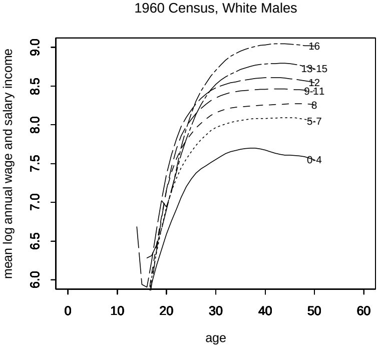

1980 Census, Black Males
mean log annual wage and salary income
experience
7
8
9
10
11
0
10
20
30
40
50
60
5-7
9-11
12
16
13-15

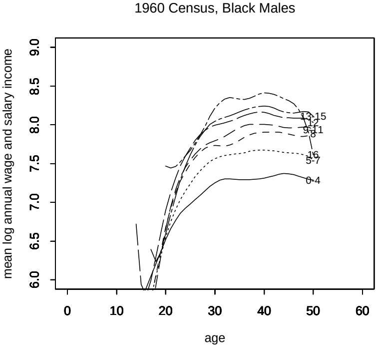

1990 Census, White Males
mean log annual wage and salary income
experience
16
5-7
0-4
13-15
12
1990 Census, Black Males
mean log annual wage and salary income
experience
17+
13-15
16
7
8
9
10
11
0
10
20
30
40
50
60

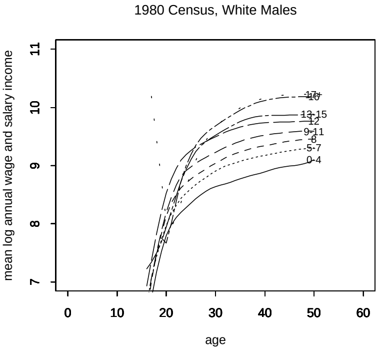
By Age

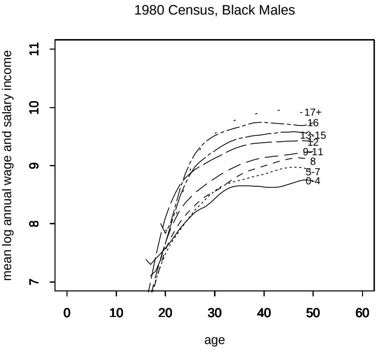
Figure 2: Age-Earnings Profiles, 1940,1960,1980




1960 Census, White Males
mean log annual wage and salary income
age
16
13-15
12
9-11
8
5-7
0-4

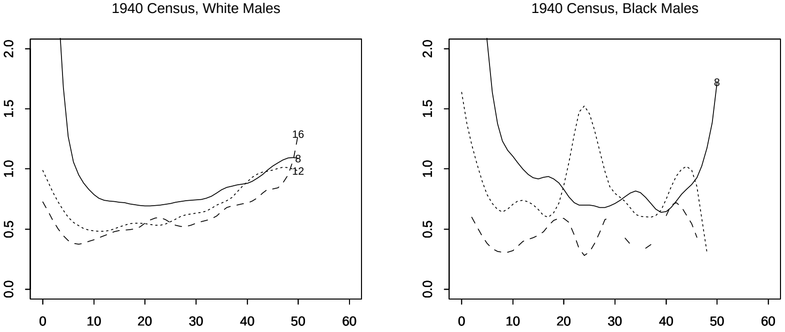

1960 Census, Black Males
mean log annual wage and salary income
age
13-15
12
9-11
8
16
5-7
0-4

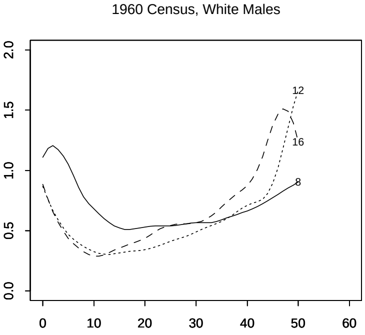

1980 Census, White Males
mean log annual wage and salary income
age
17+
16
13-15
12
9-11
8
5-7
0-4

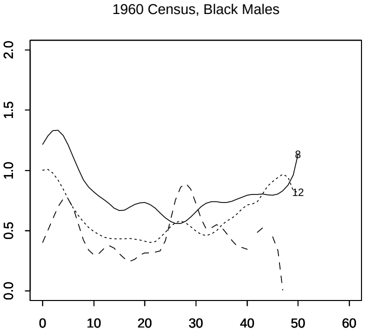

1980 Census, Black Males
mean log annual wage and salary income
age
17+
16
13-15
12
9-11
8
5-7
0-4

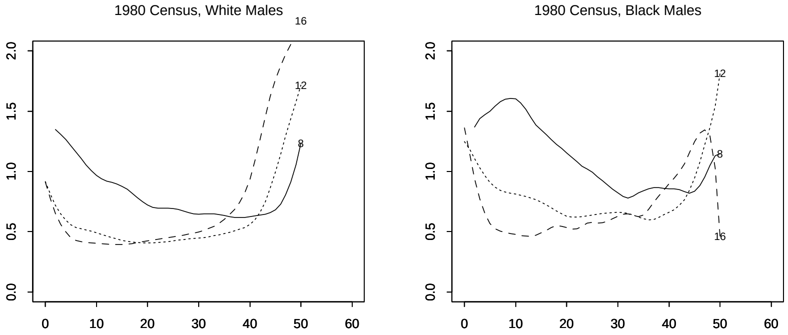
Variances

Figure 3: Experience-Variance Log Earnings




1960 Census, White Males
0.0
0.5
1.0
1.5
2.0
0
10
20
30
40
50
60
12
16
8


1960 Census, Black Males
0.0
0.5
1.0
1.5
2.0
0
10
20
30
40
50
60
8
12



1980 Census, White Males
16
12
8
2.0
1.5
1.0
0.5
0.0
0
10
20
30
40
50
60
1980 Census, Black Males
12
8
16
2.0
1.5
1.0
0.5
0.0
0
10
20
30
40
50
60

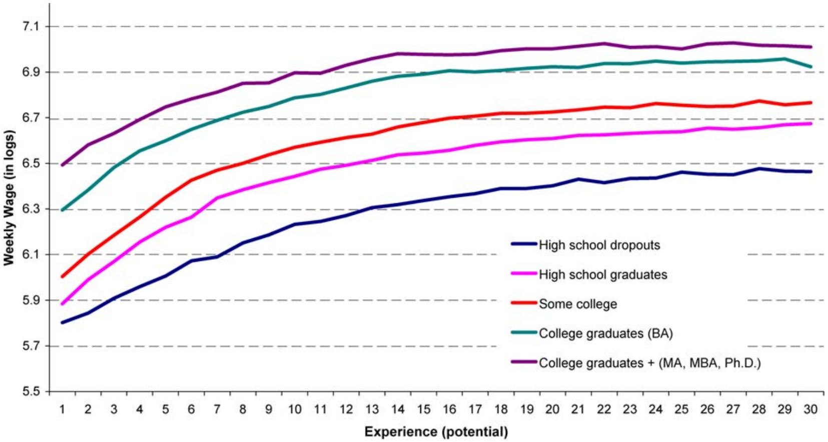
Table 2: Estimated Coefficients from Mincer Log Earnings Regression for Men

		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. See Appendix B for data description.

- Much of the growth of earnings occurs in the early years of work experience.

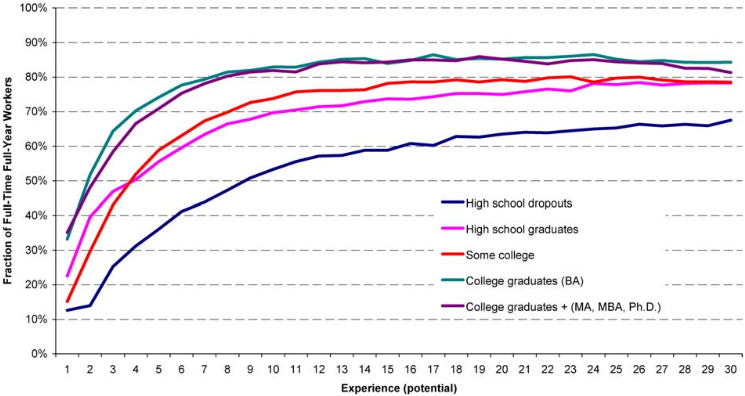
Figure 1: Mean weekly wages (in logs) by education and (potential) experience, white males, full-time full-year workers (52 weeks), CPS, March supplements, 1964–2002



Source: Rubinstein and Weiss (2006).

Figure 2: Fraction of full-time full-year workers and average weekly hours of employed workers by education and experience, CPS, March supplements, 1964–2002

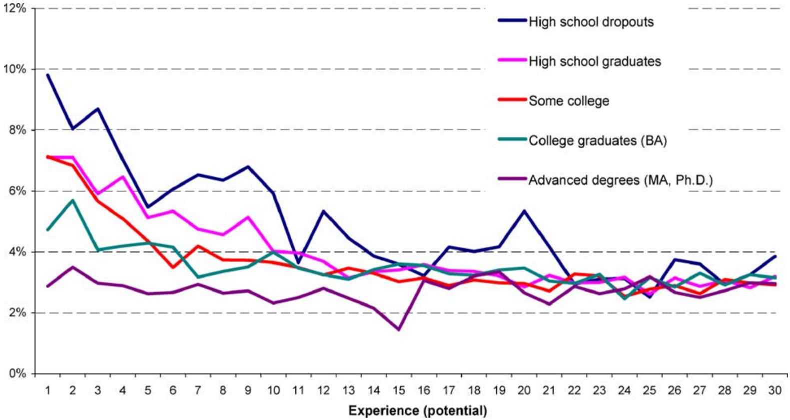
(a) Fraction of full-time full-year workers



Source: Rubinstein and Weiss (2006).

Figure 3: Proportion of workers who changed occupation, industry or employers by education and experience, full-time workers, CPS-ORG, 1998–2002, and NLSY, 1979–2000

(b) Proportion of workers who changed industry (within one month), CPS-ORG, 1998–2002



Source: Rubinstein and Weiss (2006).

4. Becker (1964): Over-Investment or Under-Investment in Education?

Internal Rates of Return: Conventional Measure of Return

- Based on ex post earnings (decisions based on ex ante earnings)

- $Y(s,x)$: 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: proportional income tax rate
- r: the before-tax interest rate.
- Agents maximize the present discounted value of lifetime earnings:

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

- FOC:

$$[T'(s)-1]e^{-(1-\tau)r(T(s)-s)}Y(s,T(s)-s) \qquad (4)$$
$$-(1-\tau)r\int_0^{T(s)-s}e^{-(1-\tau)rx}Y(s,x)dx$$
$$+\int_0^{T(s)-s}e^{-(1-\tau)rx}\frac{\partial Y(s,x)}{\partial s}dx - v/(1-\tau)=0.$$

- **Exercise:** Verify that SOC holds.

- Define $\tilde{r} = (1-\tau)r$ (the after-tax interest rate).
- Rearranging terms yields an optimality condition:

$$
\begin{aligned}
\tilde{r} \;=\; & \underset{\text{(Life Extension Effect)}}{\frac{[T'(s)-1]e^{-\tilde{r}(T(s)-s)}Y(s,T(s)-s)}{\int_0^{T(s)-s} e^{-\tilde{r}x}Y(s,x)dx}} \\
& + \underset{\substack{\text{(Effect of a change in } s\\ \text{on per period earnings)}}}{\frac{\int_0^{T(s)-s} e^{-\tilde{r}x}\left[\frac{\partial \log\ Y(s,x)}{\partial s}\right] Y(s,x)dx}{\int_0^{T(s)-s} e^{-\tilde{r}x}Y(s,x)dx}} \\
& - \underset{\text{(Tuition and Other Costs of Schooling)}}{\frac{v/(1-\tau)}{\int_0^{T(s)-s} e^{-\tilde{r}x}Y(s,x)dx}}. \qquad (5)
\end{aligned}
$$

Commonly Invoked Assumptions:

- $T'(s) = 1$ (i.e., no loss of work life from schooling)
- No tuition or psychic costs
- These assumptions simplify (3):

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

- Mincer's model implies multiplicative separability between the schooling and experience components of earnings

$$Y(s,x) = \mu(s)\varphi(x).$$
$$\therefore \tilde{r} = \mu'(s)/\mu(s).$$

- The "Mincer return"

THE UNIVERSITY OF CHICAGO

Adding Costs

- Incorporating tuition (and psychic costs) and taxes
- The internal rate of return for schooling level s_1 versus s_2, $r_I(s_1, s_2)$, solves (keep the argument of $r_I(s_1, s_2)$ implicit)

$$\int_0^{T(s_1)-s_1} (1-\tau)e^{-r_I} Y(s_1, x)dx - \int_0^{s_1} ve^{-r_I z}dz \qquad (6)$$
$$= \int_0^{T(s_2)-s_2} (1-\tau)e^{-r_I} Y(s_2, x)dx - \int_0^{s_2} ve^{-r_I z}dz.$$

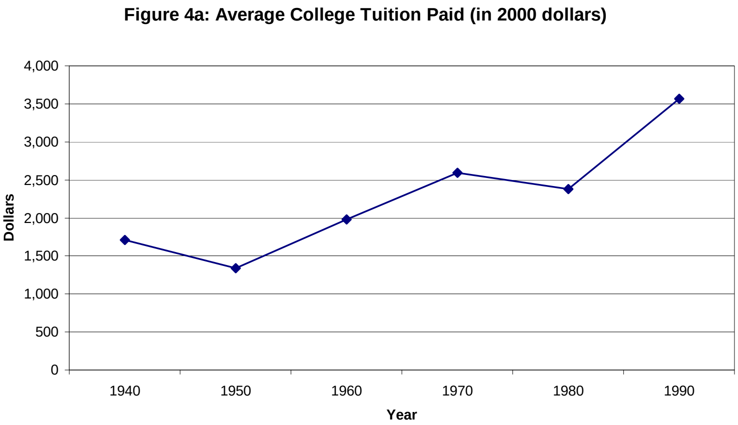


Figure 4a: Average College Tuition Paid (in 2000 dollars)
4,000
3,500
3,000
2,500
2,000
1,500
1,000
500
0
Dollars
1940
1950
1960
1970
1980
1990
Year

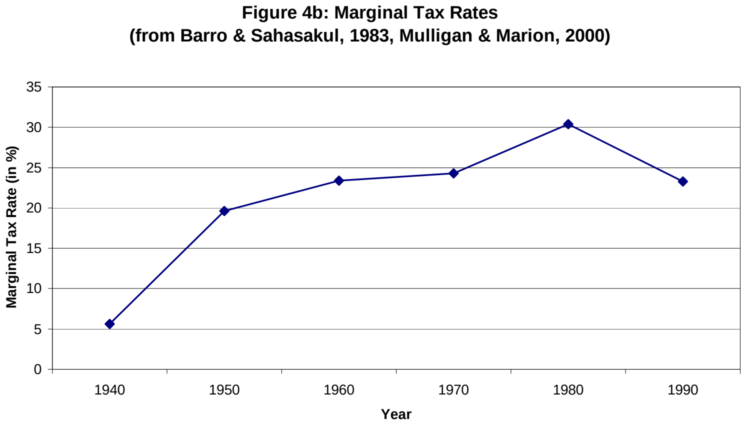


Figure 4b: Marginal Tax Rates
(from Barro & Sahasakul, 1983, Mulligan & Marion, 2000)
Marginal Tax Rate (in %)
0
5
10
15
20
25
30
35
1940
1950
1960
1970
1980
1990
Year

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

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 Appendix B for data description.

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
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. See Appendix B for data description.

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

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 Appendix B for data description.

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
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. See Appendix B for data description.

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

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

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
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 B for a description of tuition and tax amounts.

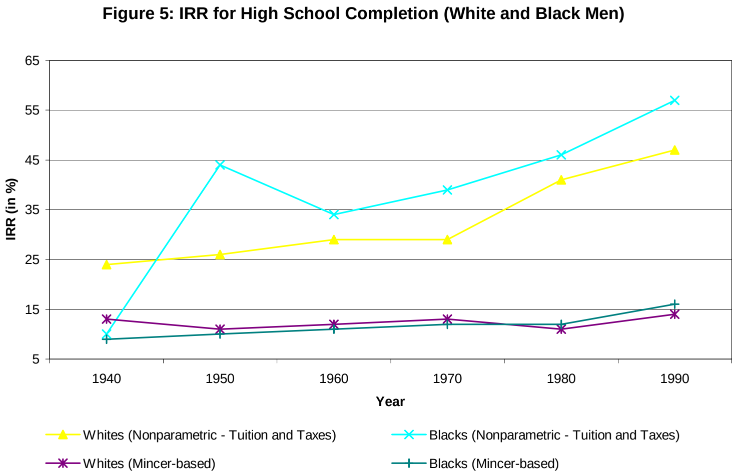


Figure 5: IRR for High School Completion (White and Black Men)
65
55
45
35
25
15
5
IRR (in %)
1940
1950
1960
1970
1980
1990
Year
Whites (Nonparametric - Tuition and Taxes)
Blacks (Nonparametric - Tuition and Taxes)
Whites (Mincer-based)
Blacks (Mincer-based)

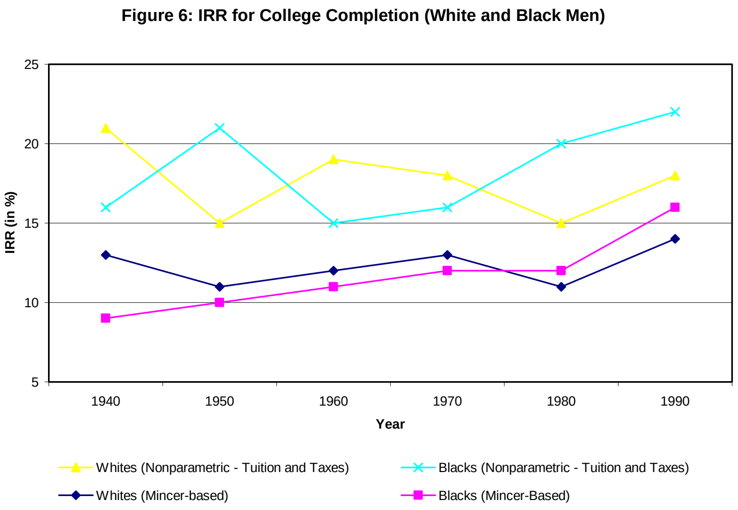


Figure 6: IRR for College Completion (White and Black Men)
25
20
15
10
5
IRR (in %)
1940
1950
1960
1970
1980
1990
Year
Whites (Nonparametric - Tuition and Taxes)
Blacks (Nonparametric - Tuition and Taxes)
Whites (Mincer-based)
Blacks (Mincer-Based)

What is “The” Rate of Return to Schooling?

- **Previous estimates traditional: based on a single cross-section**
- **Challenges faced in estimating rates of return**
  a. Lifetime earnings profiles required
  b. Observed earnings profiles subject to selection bias
  c. Identifying and quantifying nonmarket benefits and nonpecuniary costs is a difficult task
  d. Identifying agent expectations about future returns (what are their information sets?)
  e. Cohort effects (school quality; labor market entry effects)
- **Challenges in understanding schooling decisions made**
  a. What is model of schooling choices: Family? Individual?
  b. What is in agent information set?
  c. What constraints does agent confront?
  d. Role of parental transfers and altruism/paternalism

What is Proper Measure of Causal Rate of Return? Use a Slightly Less Cumbersome Notation:

$$\ln \underbrace{Y(S_i, X_i)}_{\text{earnings}} = \gamma_i + \rho_i \underbrace{S_i}_{\substack{\text{years of}\\ \text{schooling}}} + \phi(\underbrace{X_i}_{\substack{\text{work}\\ \text{experience}}}) \tag{7}$$

- Becker and Chiswick (1966); Mincer (1974)
- γ_i: "ability to earn" – source of "ability bias"
- ρ_i: "rate of return"

- Causal relationships generated by hypothetical variations of each of $\gamma_i,\ \rho_i$ and $\phi(\boldsymbol{X}_i)$.
- Correlation between γ_i and S_i is the source of "ability bias" (e.g., Griliches, 1977).
- Strictly speaking, γ_i may or may not be related to ability.
- ρ_i: "return to a unit of schooling" for person i.

THE UNIVERSITY OF CHICAGO

ρ_i: "The" Causal Effect of Education

- *Ceteris paribus* statement
- Marginal increase in earnings
- *Fixing vs. conditioning*

Fixing vs. Conditioning

- $Y = X\beta + U$
- (X, U) random variables
- Define U as a hypothetical random variable (e.g., ability)
- $E(U) = 0$

$$E(Y \mid X = x) = X\beta + E(U \mid X = x): \quad \text{conditional expectation}$$

$$E(Y \mid Fix\ X = x) = x\beta + E(U \mid Fix\ x) = x\beta \quad \text{causal effects of } x$$

Interpreting Returns to Education

- As illustrated above, ρ_i is not, in general, an *internal rate of return* for individual i.
- What is it?
- *Ex post* causal effect of increasing final schooling on annual earnings by exactly one year from any base state of schooling, holding γ_i and $\boldsymbol{X}_i$ *fixed*.
- **Slope of hedonic wage function**.
- ρ_i: ignores *continuation values* arising from the dynamic sequential nature of the schooling decision where information is updated and schooling.

What "Effect" of Education?

- Let $R_{s,i}$ be the lifetime return of s
- Mean causal effect for the whole population: $E(\rho)$? (Card, 1999)
- $E(R_s)$?
- Direct return to schooling for those who *choose* to be at a given level of schooling $E(\rho|S = s)$?
- $E(R_s|S = s)$
- People indifferent between different levels of schooling?

Example of Distinctions Among Possible Candidate "Causal Effects": The Causal Effect of a Compulsory Schooling Law

- Consider a compulsory schooling policy that forces all persons to take a minimum level of schooling ($S \geq \underline{s}$).
- What causal effect is identified by this "natural experiment?"
- Any estimated treatment effect is defined conditional on the set of people who change their schooling from below $\underline{s}$ to at or above $\underline{s}$.
- Thus, the experiment that evaluates the effects of this policy does not, in general, estimate $E(\rho)$ or even $E(\rho | S = \underline{s})$.
- Allow ρ_i to depend on the origin and destination schooling states: $(\rho_{i,s,s'})$ for $s' > s$.

Ability Bias:

$$\gamma \not\perp\!\!\!\perp S$$

$$E(\gamma|S) \neq 0$$

Approaches to Identifying Causal Effects and Causal Rates of Return

- **Two approaches to solving selection problem:**
  1. Structural models that jointly analyze choices and outcomes
  2. Treatment effect models using matching or instrumental variables (including experiments; RDD)

- The structural approach explicitly models agent decision rules that generate $P_{s,l,i}$ and the dependence between ρ_i and S_i.
- The treatment effect approach is typically agnostic about agent decision rules and relies on exclusion restrictions to identify its estimands.
- Rarely distinguishes *ex ante* from *ex post* returns (see, however, Cunha and Heckman, *JOLE*, 2016).
- LATE typically does not identify returns at the various margins of choice that generate outcomes or the sub-populations (defined in terms of observables and unobservables) affected by the instruments used.

Structural Models

- Dynamic discrete choice models facilitate the interpretation of intertemporal choices
- Assume risk-neutral agents who have a finite choice set
- Let $B_n(a) = 1$ if alternative n is chosen at age a and zero otherwise
- Let $R_n(a)$ be the current flow reward at age a from alternative n
- The current reward per period at any age a is

$$R(a) = \sum_{n=1}^{N} \underbrace{R_n(a)}_{\substack{\text{per} \\ \text{period} \\ \text{reward} \\ \text{from} \\ \text{choice } n}} \underbrace{B_n(a)}_{\substack{\text{choice} \\ \text{indicator}}} .$$

- Individual's state at age a: $\boldsymbol{H}(a)$.
- Discount factor: δ.
- The value function:

$$V\left(\boldsymbol{H}(a), a\right) = \max_{B_n(\tau)\in\mathcal{B}(a)} E\left[\sum_{\tau=a}^{\bar{a}} \delta^{\tau-a} \sum_{n=1}^{N} R_n(\tau)B_n(\tau) \mid \boldsymbol{H}(a)\right]$$

- $\mathcal{B}(a)$: the feasible set at age a.

- The alternative-specific functions, $V_n\left(\boldsymbol{H}(a), a\right)$ can be written as

$$V_n\left(\boldsymbol{H}(a), a\right) = R_n\left(\boldsymbol{H}(a), a\right) + \delta E \underbrace{\left[V\left(\boldsymbol{H}(a+1), a+1\right) \mid \boldsymbol{H}(a), B_n(a) = 1\right]}_{\text{Continuation Value}}$$

for $a < \overline{a}$, where $V_n\left(\boldsymbol{H}(\overline{a}), \overline{a}\right) = R_n\left(\boldsymbol{H}(\overline{a}), \overline{a}\right)$

- The decision rule is $B_n(a) = 1$ if $n = \underset{j \in \{1,\ldots,N\}}{\operatorname{argmax}} \left\{V_j(\boldsymbol{H}(a), a)\right\};$

$$B_n(a) = 0 \text{ otherwise}$$

Rate of Return

- Define next best set

$$\{1, \ldots, N\} \backslash \{n\}$$

- Next best alternative

$$\hat{n} = \underset{j \in \{1,\ldots,N\} \backslash \{n\}}{\operatorname{argmax}} \{V_j(H(a), a)\}$$

- "Return" is

$$\frac{V_n(H(a), a) - V_{\hat{n}}(H(a), a)}{V_{\hat{n}}(H(a), a)}$$

In a sequential model with only two choices

- Say V_n, V_{n-1}
- Separated by "one period" at interest rate r we get

$$\begin{aligned}
\frac{V_n}{1+r} - V_{n-1} &\geq 0 && \text{"go forward to } n\text{"} \\
V_n - (1+r)V_{n-1} &\geq 0 && (r > 0) \\
V_n - V_{n-1} &\geq rV_{n-1} \\
\frac{V_n - V_{n-1}}{V_{n-1}} &> r
\end{aligned}$$

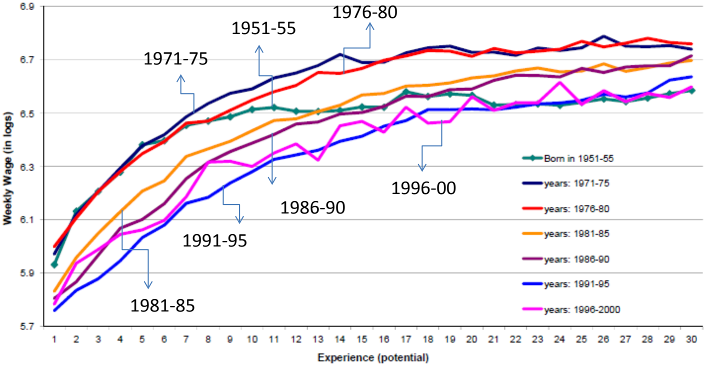
- Interest rate shows up, but not always.
- **Question:** What is the rate of return to a year of schooling in the Keane-Wolpin (1997) model?

Switch Over to Eisenhauer et al. IER Slides

How Valid is the Synthetic Cohort Assumption? What Life Cycle Parameters Do We Identify in Cross-Sections?

Figure 4: Cohort and cross-section wage profiles for high school graduates and college graduates, white males, CPS, March supplements, 1964-–2002

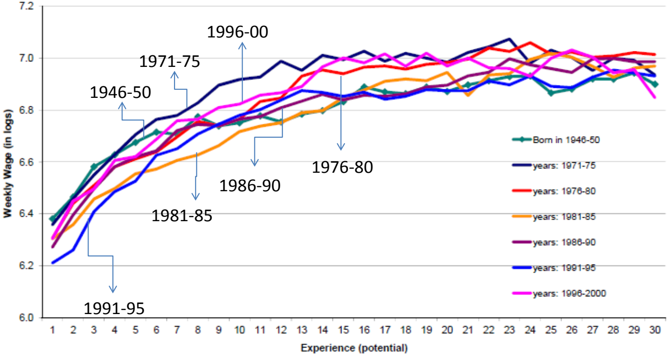
(a) High school graduates born in 1951-1955



Source: Rubinstein and Weiss (2006).
Note: The chart shows the wage-experience profiles for the cohort of high school graduates born in 1951-1955. Added to the graphs is the evolution of the cross section wage-experience profiles from 1971 to 2000 in five year intervals, where each such cross section profile shows the mean weekly wages (in log) of workers with the indicated schooling and experience in a given time interval.

Figure 4: Cohort and cross-section wage profiles for high school graduates and college graduates, white males, CPS, March supplements, 1964–2002

(b) College graduates born in 1946-1950



Source: Rubinstein and Weiss (2006).
Note: The chart shows the wage-experience profiles for the cohort of college graduates born in 1946-1950. Added to the graphs is the evolution of the cross section wage-experience profiles from 1971 to 2000 in five year intervals, where each such cross section profile shows the mean weekly wages (in log) of workers with the indicated schooling and experience in a given time interval.

Age, Period, and Cohort Effects

Add Age, Cohort, and Year Effects to Mincer Equation for Person i

$$\ln W_i = \alpha_0 + \underbrace{\alpha_1 a_i}_{\substack{\text{age} \\ \text{``maturation''}}} + \underbrace{\alpha_2 y}_{\substack{\text{year} \\ \text{aggregate} \\ \text{effects}}} + \underbrace{\alpha_3 e_i}_{\substack{\text{experience} \\ \text{(lbd or OJT)}}} + \underbrace{\alpha_4 s_i}_{\text{schooling}} + \underbrace{\alpha_5 c_i}_{\substack{\text{``vintage''} \\ \text{or cohort} \\ \text{(quality of} \\ \text{cohorts)}}} + u_i \tag{8}$$

Two Identities

$$e_i = a_i - s_i \qquad \text{"experience"} \tag{9}$$

$$y_i = a_i + c_i \qquad c_i = \text{birth year} \tag{10}$$

- Take a pedagogically convenient case where schooling and experience are initially dropped ($\alpha_3 = \alpha_4 = 0$):

$$\ln W(a, y, c) = \beta_0 + \underset{\text{(age)}}{\beta_1 a_i} + \underset{\text{(year)}}{\beta_2 y_i} + \underset{\text{(cohort)}}{\beta_3 c_i} + u_i$$

$$y_i = a_i + c_i,$$

- y_i: year wage observed.
- c_i: year of birth.
- For a given cross-section (fixing y_i) equal and exact one linear dependence: c_i and a_i move in opposite directions.
- Older people come from earlier cohorts.
- Age effects might be cohort effects, and vice versa.

- Substitute $c_i = y_i - a_i$.

$$\begin{aligned} \ln W_i &= \alpha_0 + \beta_1 a_i + \beta_2 y_i + \beta_3 (y_i - a_i) + u_i \\ &= \alpha_0 + (\beta_1 - \beta_3) a_i + (\beta_2 + \beta_3) y_i + u_i \end{aligned}$$

- Can identify only combinations of coefficients.
- Cross section: y_i is the same for everyone.
- Intercept:

$$\left[\alpha_0 + (\beta_2 + \beta_3) y_i\right].$$

- Identify $(\beta_1 - \beta_3)$: age minus cohort effect.
- If $\beta_3 > 0$, we underestimate true β_1 (e.g., schooling quality improving).
- Does longitudinal data solve the problem? Not necessarily.
- In panels, y_i generally moves over time.
- Recall that $y_i = a_i + c_i$, so we still have exact linear dependence.
- Panel data identifies model if there are no year effects.

- Supposed we substitute out for a_i?
- $y_i = a_i + c_i$.

$$\begin{aligned} \ln W_i &= \alpha_0 + \beta_1 + \beta_2(y_i - c_i) + \beta_3 c_i + u_i \\ &= \alpha_0 + (\beta_1 + \beta_2 + \beta_3)y_i + (\beta_3 - \beta_1)c_i + u_i \end{aligned}$$

- If $\beta_1 > 0$, understate cohort effects.
- **Exercise:** What are the competing stories of the labor market for each identifying assumption?

- Acquire similar problems in models with nonlinear terms:

$$y = a + c$$

- $$\begin{aligned} \ln W = \quad & \beta_0 + \beta_1 a + \beta_2 y + \beta_3 c + \beta_4 a^2 + \beta_5 ac \\ & + \beta_6 ay + \beta_7 cy + \beta_8 c^2 + \beta_9 y^2 + u, \end{aligned}$$

    - $$\left.\begin{aligned} y^2 &= a^2 + 2ac + c^2 \\ ay &= a^2 + ac \\ cy &= ca + c^2 \end{aligned}\right\} \text{3 linear dependencies}$$

- Cannot identify all of the parameters (only 3 parameters associated with second order (square and interaction terms) are identified out of 6 total.

Theorem. In a model with interactions of order k with j variables and one linear restriction among the j variables, then of the $\binom{j+k-1}{k}$ coefficients of order k, only $\binom{j+k-2}{k}$ are estimable. (Heckman and Robb, in S. Feinberg and W. Mason, *Age, Period and Cohort Effects: Beyond the Identification Problem*, Springer, 1986).

E.g. $k = 2$, $j = 3$; 6 coefficients and 3 are estimable, as in the preceding example.

Theorem. In a model with ℓ restrictions on the j variables, then $\binom{j+k-\ell-1}{k}$ kth order coefficients are estimable (Heckman and Robb, 1986).

- **Exercise:** Show the linear dependencies that arise if analysts use dummy variables in place of continuous variables.

General Case: Add Back Experience and Schooling

- Substitute out for c_i and a_i, using (9) and (11):

$$\begin{aligned}\ln W_i = \alpha_0 + (\alpha_2 + \alpha_5)y + (\alpha_1 + \alpha_3 - \alpha_5)e_i \\ + (\alpha_1 + \alpha_4 - \alpha_5)s_i + u_i.\end{aligned}$$

- In a single cross section, y is the same for everyone.
- Intercept is $\alpha_0 + (\alpha_2 + \alpha_5)y$.
- Experience "effect" $= \alpha_1 + \alpha_3 - \alpha_5 = \alpha_3 + (\alpha_1 - \alpha_5)$.
- If later vintages acquire more skills, $\alpha_5 > 0$.
- Downward bias (*e.g.* higher quality of schooling).
- If there is an aging effect ($\alpha_1 > 0$, *e.g.*, maturation) produces upward bias for $\alpha_{3.}$

THE UNIVERSITY OF CHICAGO

Schooling Coefficient

- $\alpha_1 + \alpha_4 - \alpha_5 = \alpha_4 + (\alpha_1 - \alpha_5)$
- Vintage (cohort) effects lead to downward bias in the estimated rate of return to schooling.
- Age (maturation) effects produce upward bias.
- Experience coefficient-schooling coefficient:

$$= (\alpha_1 + \alpha_3 - \alpha_5) - (\alpha_1 + \alpha_4 - \alpha_5) = \alpha_3 - \alpha_4.$$

- Can identify difference in "returns" to experience net of schooling.

- Observe that even if $\alpha_1 = 0$ (no aging effect), still can't identify true schooling or experience effects.
- Does **longitudinal data** (observations on the same people over time) or **repeated cross section data** (observations on the same population over time but sampling different persons) help? No: because year effects activated (acquire new parameters in each wave).

- If $\alpha_2 = 0$, (no year effects), can identify α_5.
- Alternatively, for each c_i, can estimate $\alpha_1 + \alpha_3$, and hence can estimate α_5.
- We also know $\alpha_1 + \alpha_4$.
- If $\alpha_1 = 0$ (no age effect), then $\alpha_{3,}$ $\alpha_{4,}$ α_5 identified.

- If year effects are present, no gain to going to longitudinal or repeated cross section data.
- We gain a new parameter when we move to each wave of the panel or repeated cross sectional data.

Solutions in Literature

1. Redefine vintage (cohort) e.g. vintage fixed over period of years (e.g. a cohort of Depression babies).
    - Then $\ln W = (\alpha_0 + \alpha_5 c) + \alpha_1 a + \alpha_2 y + \alpha_3 e + \alpha_4 s + u$.
    - In single cross section, c and y are fixed.

- Substitute for e:

$$e = a - s$$

- $$\ln W = [\alpha_0 + \alpha_5 c + \alpha_2 y] + (\alpha_1 + \alpha_3)a + (\alpha_4 - \alpha_3)s$$
- Can estimate $\alpha_1 + \alpha_3$ and $\alpha_4 - \alpha_3$, and thus $\alpha_1 + \alpha_4$.
- Successive time periods for the same vintage gives us α_2 directly [since c doesn't move].
- If no age effect, identify $\alpha_3, \alpha_4, \alpha_2$, and from successive vintages, we get α_5.

2. If we measure actual work experience, $a \neq e + s$ (agents do not fully participate in market over their life cycles), can break the linear dependence.
    - However, better proxies may be endogenous, e.g., if experience $=$ cumulated hours of work.

Nonlinear Models

- Results carry over in an obvious way to nonlinear models.

Link to Example of Interpretive Pitfall

Link to Cohort vs. Cross-Section Internal Rate of Return

Appendix

Example of Interpretive Pitfall

1. Johnson and Stafford (AER, 1974)
2. Weiss and Lillard (JPE, 1979)
   - **Fact:** Disparity in real wages between recent Ph.D. entrants and experienced workers rose in *physics* and *mathematics* in the late 60s and early 70s. Not observed in the *social sciences*.
   - **Why?** — Johnson-Stafford story.
   - Supplies of Ph.D.s enlarged by federal grants while demand for scientific personnel declined.
   - Wage rigidity at the top end motivated by specific human capital.
   - Spot market/entrant market bears the brunt of the burden (observed in many markets).

THE UNIVERSITY OF CHICAGO

Alternative Story

- Weiss & Lillard: "experience–vintage" interaction (ec).
- Ignore age effect:

$$\begin{aligned} \ln W(e,c,s,y) &= \varphi_0 + \varphi_1 e + \varphi_2 c + \varphi_3 y + \varphi_4 s \\ &\quad + \varphi_5 e^2 + \varphi_6 c^2 + \varphi_7 ec \\ &\quad + \varphi_8 ey + \varphi_9 cy + \varphi_{10} y^2 \end{aligned}$$

- Assume other powers and interactions are zero. Assume $\varphi_{10} = 0$.
- Johnson-Stafford: $\varphi_8 > 0$ or $\varphi_9 < 0$
- Weiss-Lillard: $\varphi_7 > 0$
- Remember $y = e + s + c$.

- Weiss-Lillard ignore year effects.
- We get Weiss-Lillard by substituting for y:

$$\begin{aligned}\ln W(e,c,s) &= \varphi_0 + (\varphi_1+\varphi_3)e + (\varphi_3+\varphi_4)s \\ &\quad +(\varphi_2+\varphi_3)c + (\varphi_5+\varphi_8)e^2 \\ &\quad +\varphi_8 es + (\varphi_7+\varphi_8+\varphi_9)ec \\ &\quad +(\varphi_6+\varphi_8)c^2\end{aligned}$$

- Note that if $\varphi_7 = 0$ but $\varphi_9 > 0$, we get ec interaction, but it is "really" a year effect. If entry level wages fall relative to wages of experienced workers, the wage / experience profile is steeper in more recent cross-sections.

- Looking at social scientists where no interaction appears favors Johnson-Stafford.
- Moral: auxiliary evidence and theory break the identification problem.

Return to main text

Cohort vs. Cross-Section Internal Rate of Return

Cohort Rate of Return

- Suppose we consider the return to high school (h) versus dropping out (d).
- Initially, ignore year effects even though this is the story most frequently told in recent literature.
  1. $Y^h_{a,c}$ is the earnings of a high school graduate of cohort c at age a.
  2. $Y^d_{a,c}$ is the earnings of a dropout of cohort c at age a.
  3. $\rho_c = IRR_c$ (cohort internal rate of return) solves

$$\sum_{a=0}^{A} \frac{Y^h_{a,c} - Y^d_{a,c}}{(1+\rho_c)^a} = 0.$$

  4. A: maximum age.
  5. Missing: life cycle profiles in both states for any person and generally lack full life cycle profile in any observed state.

Cross-Section Return

- Cross-sections consist of members of different cohorts.
- Recall $c + a = y$.
- Cross-section internal rate of return IRR_y in year y solves

$$\sum_{a=0}^{A} \frac{\left(Y^h_{a,y-a} - Y^d_{a,y-a}\right)}{\left(1+\rho_y\right)^a} = 0.$$

When Does $\rho_c = \rho_y$?

- Can occur if the environment is stationary.
- Justifies the synthetic cohort assumption.
- Consider a steady state geometric growth of outcomes across cohorts at rate $(1+g)$.
- Define age-specific differential

$$\Delta_{a,c}^{h,d} = Y_{a,c}^{h} - Y_{a,c}^{d} \tag{11}$$
$$\Delta_{a,c+j}^{h,d} = \left(\Delta_{a,c}^{h,d}\right)(1+g)^{j}$$

- $g = 0 \Rightarrow \rho_c = \rho_y$.

Consider a Model with Two Cohorts

- $c = 0, 1$
- $y = a + c$
- Cohorts are active for two periods $a = 0, 1$.
- Rate of return for cohort $c = 0$.
- ρ_c root of:

$$0 = Y^h_{0,0} - Y^d_{0,0} + \frac{Y^h_{1,0} - Y^d_{1,0}}{1 + \rho_c}$$

- Cross-section ρ_y at $y = 1$, when cohort $c = 0$ becomes active:

$$0 = Y^h_{0,0} - Y^d_{0,0} + \frac{Y^h_{1,-1} - Y^d_{1,-1}}{1 + \rho_y}$$

$$0 = Y^h_{0,0} - Y^d_{0,0} + \frac{Y^h_{1,0} - Y^d_{1,0}}{(1 + \rho_y)(1 + g)}$$

- $\therefore \rho_y < \rho_c$ if $g > 0$.

General Case

- For cohort c, a benchmark cohort ρ_c is the IRR that solves

$$\sum_{a=0}^{A} \frac{\left(Y_{a,c}^{h} - Y_{a,c}^{d}\right)}{\left(1+\rho_c\right)^{a}} = 0.$$

THE UNIVERSITY OF CHICAGO

- Cross section in year $y = c$ produces the equation

$$\sum_{a=0}^{A} \frac{\left(Y^{h}_{a,y-a} - Y^{d}_{a,y-a}\right)}{\left(1+\rho_y\right)^{a}} = 0,$$

- $\rho_y = \text{IRR}_y$ (goes across cohorts).
- How does ρ_y compare to ρ_c?

- In a given cross section y
- For $0 \leq a \leq A$, $y - A \leq c \leq y$.
- Can adjust cross section values in y to find corresponding cohort values.

- Suppose we want the return for cohort $\bar{c}$, $\quad y - A \leq \bar{c} \leq y$.

$$Y_{a,y-a} \overbrace{(1+g)^{\bar{c}-(y-a)}}^{\substack{\text{adjusts for growth} \\ \text{(or shrinkage)} \\ \text{between cohort} \\ (y-a) \text{ and } \bar{c}}} = Y_{a,\bar{c}}$$

$$\therefore Y_{a,y-a} = Y_{a,\bar{c}}(1+g)^{(y-a)-\bar{c}}$$

$$\therefore \sum_{a=0}^{A} \frac{Y^h_{a,y-a} - Y^d_{a,y-a}}{(1+\rho_y)^a} =$$

$$\sum_{a=0}^{A} \frac{(Y^h_{a,\bar{c}} - Y^d_{a,\bar{c}})(1+g)^{(y-a)-\bar{c}}}{(1+\rho_y)^a}.$$

$$\rho_y \text{ solves } (1+g)^{\bar{c}-y} \sum_{a=0}^{A} \left(Y^h_{a,\bar{c}} - Y^d_{a,\bar{c}}\right) \left(\frac{1+g}{1+\rho_y}\right)^{-a} = 0$$

- ρ_c: root of

$$\sum_{a=0}^{A} \frac{Y^h_{a,\bar{c}} - Y^d_{a,\bar{c}}}{(1+\rho_c)^a} = 0.$$

$$(1+g)^{\bar{c}-y} \sum_{a=0}^{A} \frac{\left(Y^h_{a,y-a} - Y^d_{a,y-a}\right)}{(1+\rho_c)^a} \left[\frac{(1+\rho_c)}{(1+\rho_y)}(1+g)\right]^{-a} = 0,$$

- If $\frac{(1+\rho_c)(1+g)}{(1+\rho_y)} = 1$, $\rho_y = \rho_c$ (independent of $\bar{c}$, so long as $g \geq -1$).
- If $g \geq 0, \rho_c \geq \rho_y$.

THE UNIVERSITY OF CHICAGO

- Another way to see what is going on: ρ_c is root of

$$\sum_{a=0}^{A} \frac{(Y^h_{a,y-a} - Y^d_{a,y-a})(1+g)^{\bar{c}-(y-a)}}{(1+\rho_c)^a} = 0$$

$$\sum_{a=0}^{A} \frac{(Y^h_{a,y-a} - Y^d_{a,y-a})}{(1+\rho_c)^a}(1+g)^{\bar{c}-(y-a)}$$

$$(1+g)^{\bar{c}-y} \sum_{a=0}^{A} \frac{(Y^h_{a,y-a} - Y^d_{a,y-a})}{(1+\rho_c)^a}(1+g)^a$$

- ρ_c solves:

$$\sum_{a=0}^{A} \frac{(Y^h_{a,y-a} - Y^d_{a,y-a})(1+g)^a}{(1+\rho_c)^a} = 0$$

$$\therefore \rho_c \geq \rho_y.$$

More General Notation to Account for Time Effects

- Again, keep schooling fixed.
- Experience: $e = a - s$.
- $h =$ high school and $d =$ dropout.
- a and e increase at the same rate unit for unit.
- Outcomes at age a for high school graduates and dropouts of cohort c, age a, and year y: $Y^h_{a,c,y}$; $Y^d_{a,c,y}$

$$\Delta^{h,d}_{a,c,y} = Y^h_{a,c,y} - Y^d_{a,c,y}.$$

- If no cohort effects: $Y^j_{a,c,y} = Y^j_{a,-,y} \ \forall c$.
- "–" sets the argument to the same constant over years and ages: $j \in \{h, d\}$.

Single Cross-Section

- Assume

$$\begin{aligned} y &= 0 \\ \text{Cohort Effects} & \qquad (1+g) \\ \text{Year Effects} & \qquad (1+\varphi): \\ \text{(year effects assumed to operate} & \\ \text{the same over all cohorts)} & \end{aligned}$$

- Estimated IRR_y:

$$\sum_{a=0}^{A+1} \frac{Y^h_{a,c,y} - Y^d_{a,c,y}}{(1+\rho_y)^a} = 0$$

$$y = 0 \Rightarrow a = -c$$

THE UNIVERSITY OF CHICAGO

- IRR_y is the root of

$$\sum_{a=0}^{A+1} \frac{Y^h_{a,-a,0}, - Y^d_{a,-a,0}}{(1+\rho_y)^a} = 0$$

for $y = 0$

- More generally, for year $y = \bar{y}$, ρ_y is a root of

$$\sum_{a=0}^{A} \frac{Y^h_{a,\bar{y}-a,\bar{y}} - Y^d_{a,\bar{y}-a,\bar{y}}}{(1+\rho_y)^a} = 0$$

- Cohort Rate of Return for Cohort $c = \bar{c}$
- IRR_c is the root of

$$\sum_{a=0}^{A} \frac{Y^h_{a,c,a+c} - Y^d_{a,c,a+c}}{(1+\rho_c)^a} = 0$$

- **Question:** Should year effects experienced by cohorts c at age $a, (y = a + c)$ be counted as part of the return?
- When will $\rho_c = \rho_{\bar{y}}$?

THE UNIVERSITY OF CHICAGO

- Sufficient Condition:

$$Y^j_{a,\bar{y}-a,\bar{y}} = Y^j_{a,c,a+c}.$$

- In the case of general growth

$$Y^j_{a,\bar{y}-a,\bar{y}}(1+g)^{c-(\bar{y}-a)}(1+\varphi)^{a+c-\bar{y}} = Y^j_{a,c,a+c}$$

$$\text{i.e., } Y^j_{a,\bar{y}-a,\bar{y}}\left[\frac{(1+g)}{(1+\varphi)}\right]^{c-(\bar{y}-a)} = Y^j_{a,c,a+c}$$

- **Exercise:** Compare ρ_c to ρ_y when $g > 0, \varphi < 0$. Show you need $g = \frac{-\varphi}{1+\varphi}$ for $\rho_c = \rho_y$. (Hint: Adjust across cross-section incomes by age to form proper cohort values.)

Return to main text