Human Capitalism

J.Angrist
MIT 14.661 Fall 2024

The Big Picture

- Economists view time spent in school as investment in *human capital.* By sacrificing earnings, students pay for the skills that generate higher earnings later on (some also pay tuition).
- Students mostly study when young, leaving a long working life to reap the rewards. Indeed, education appears to be a major determinant of wages (few college grads are impoverished once they've worked a few years).
- HK Q's: Who chooses more education and why? Are more educated workers really more productive? Should the government subsidize schooling? We address these questions with theoretical guidance from three important models: equalizing differences, optimal HK, and signaling.
- Following this bit of theory, we consider the econometrics of Mincer-style wage equations, seeking causal effects of schooling and experience on wages.
- Time-permitting, we also consider the contribution of on-the-job training (OJT) to human capital and earnings, pondering the question of whether and how OJT differs from schooling.

A Equalizing Differences (Mincer, 1974)

- Published after Becker's *Human Capital*, Mincer's *Schooling, Experience, and Earnings* focuses on empirical models of wage determination. Mincer pursued a regression version of the human capital story that describes the causal effects of schooling and experience on earnings in a simple equilibrium model
- Mincer theory is remarkably spare, with no individual variation in ability, opportunity, or access to funds – yet, the Mincer model explains important facts with few moving parts. That's why it's worth studying!
- Key assumptions
    1. The only cost of schooling is foregone earnings (not unrealistic for most people)

2. The future is discounted at a common interest rate, i, at which all are free to borrow and lend
3. Labor markets equalize the value of alternative HK investment plans
4. Learning that takes place on the job can be seen as "part-time schooling." Initially, however, we ignore this.

- Suppose $y(s)$ is earnings with s years of schooling and $y(0)$ is earnings with none
- Assuming (as young people typically believe) we live forever, schooling plans are valued at:

$$V(s) = (s \cdot 0) + \sum_{t=s+1}^{\infty} \frac{y(s)}{(1+i)^t}$$

- The *equalizing differences* model assumes we can choose as much schooling as we like, but in equilibrium the labor market equates plan values. In other words, for all s:

$$V(s) = V(0)$$

  Why would this happen?

- Note that

$$V(s) = y(s) \sum_{t=s+1}^{\infty} \frac{1}{(1+i)^t} = \frac{y(s)}{(1+i)^s} \sum_{t=1}^{\infty} \frac{1}{(1+i)^t},$$

  so

$$y(s) = y(0)(1+i)^s,$$

  and

$$\ln y(s) \approx \ln y(0) + is$$

- This model implies:
  1. Schooling raises earnings; in a world of equalizing differences, the return to a year invested schooling is the same as the return to financial assets
  2. Individuals are indifferent between, say, $s = 12$ and $s = 16$ because all plans yield the same PDV. Quit school now and be happy! (or at least indifferent)

- Exercises: (a) repeat the analysis here using continuously compounded interest rates; (b) show that finite lifetimes raise the cross-sectional return to investment in education in the eq. diffs world
- This is a highly unrealistic model - but it suggests that if labor and capital markets work reasonably well, we can expect the returns to human capital to approach the cost of the funds needed to cover consumption when not working

- Because we're indifferent between $s = 12$ and $s = 16$, schooling might as well be randomly assigned! And the returns to schooling (defined as differences in earnings between more and less educated workers) are linear and constant.
  - Mincer's description of the CEF linking average log earnings with schooling and experience has proved resilient. This simple regression is a workhorse of empirical labor economics.

B Optimal HK in a Heterogeneous World (Becker, 1964)

- Ability and opportunity varies
- Everyone does the best they can; because we differ, the value of alternative plans is not equalized
- Assume
  1. The only cost of schooling is foregone earnings (again)
  2. Equal opportunity (initially), meaning we all borrow, lend, and discount the future at continuously compounded rate r
  3. Ability to convert schooling into earnings differs
  4. Individuals choose schooling to max PDV of earnings profiles
- What's new? *People differ*! We'll start with a parameterization of *potential* earnings:

$$y_i(s) = g(s, a_i), \tag{1}$$

  where a_i is "ability" and $g(s, a_i)$ is a potential outcome: the amount someone with ability a_i earns when they get schooling s (a_i might be <u>defined</u> as potential earnings when $s_i = 0$)
  - What to assume about derivatives of $g(s, a_i)$, w.r.t. s?
- Workers are paid their marginal products; more educated and more able workers are more productive
- With continuous compounding, PDV plan values are given by:

$$V_i(s) = \int_s^\infty g(s, a_i)e^{-rt}dt = g(s, a_i)\int_s^\infty e^{-rt}dt = \frac{g(s, a_i)e^{-rs}}{r}$$

  for schooling plan s and individual i (integration is summation with t divided increasingly finely)
- Pick your optimal schooling levels by setting $V_i'(s) = 0$:

$$\frac{\partial g(s, a_i)}{\partial s} = rg(s, a_i)$$

so

$$\frac{d\ln y_i(s)}{ds} = \frac{\partial g(s,a_i)/\partial s}{g(s,a_i)} = r, \tag{2}$$

and MR=MC type relation

- This implies optimal schooling, s_i^* satisfies

$$\ln y_i(s_i^*) = \ln y_i(0) + rs_i^*, \tag{3}$$

for individual i (reminiscent of equalizing diffs)

- But now we all want different schooling levels: optimal s_i^* depends on a_i by virtue of (1) (pretty subtle!)

Ability bias

- Are more educated people more productive *because* of their schooling – or would those fortunate enough to be more educated have earned more regardless? To find out, we might estimate a regression like this:

$$\ln y_i = \alpha_1 + \rho_1 s_i + \gamma a_i + \eta_{1i}.$$

We can think of this as defining:

$$\ln y_i(0) = \alpha_1 + \gamma a_i + \eta_{1i}$$

But ability (a potential outcome) is hard to measure. We must therefore settle for the short regression

$$\ln y_i = \alpha_0 + \rho_0 s_i + \eta_{0i}$$

(This typically includes imperfect ability controls, but not $y_i(0)$)

- OVB in a wage equation with schooling on the RHS is called "ability bias", expressed as

$$\rho_0 = \rho_1 + \lambda_{as}\gamma,$$

where λ_{as} is the regression of omitted ability on schooling

- Ability bias must be positive ...

Griliches (1977)

- Try this parameterization:

$$g(s,a_i) = e^{\beta s + \gamma a_i}$$

Whoops! This generates either ∞ or zero schooling [show this]

- How 'bout a graduate stipend? (TR for "transfer" in Griliches 1977)

$$V(s) = \int_s^\infty g(s,a_i)e^{-rt}dt + \int_0^s TRe^{-rt}dt$$
$$= \frac{g(s,a_i)e^{-rs}}{r} + \frac{TR}{r}(1-e^{-rs})$$

- Setting $V_i'(s) = 0$:

$$\frac{\partial g(s,a_i)}{\partial s} = r[g(s,a_i) - TR]$$

again, an MR=MC type relation.

- Now we've got sufficient curvature to get an interior optimum:

$$\frac{d\ln y_i(s)}{ds} = r\left[1 - \frac{TR}{y_i(s)}\right]$$

  - Define $\phi_i(s) = \frac{TR}{y_i(s)}$ for $\phi_i(s) \in [0,1]$, so the stipend equals at most foregone earnings. Then, the FOC becomes:

$$\beta = r[1-\phi_i(s)]. \tag{4}$$

    Because TR is fixed, $\phi_i(s)$ is declining in s.
  - Assume $\beta < r$ (why is this necessary?), giving:

$$s_i^* = \frac{1}{\beta}\{\ln TR - \ln\frac{r-\beta}{r} - \gamma a_i\}$$

    Ability bias does what to OVB?

- Card's (2001) version of this story can be told using:

$$g(s,a_i) = e^{\gamma a_i + \beta_1 s - \beta_2 s^2 + \beta_3 a_i s}\ ;\ \beta_2, \beta_3 > 0 \tag{5}$$
$$r_i = r_0 - r_1 a_i. \tag{6}$$

  - The optimal schooling FOC becomes:

$$\beta_1 - 2\beta_2 s_i + \beta_3 a_i = r_0 - r_1 a_i.$$

  - Here, plausible parameter values yield $Cov(s_i^*, a_i) > 0$

Estimates of the economic returns to schooling (attached)

- Ability controls: nice if you can get 'em! But beware the bad (see MM Chapter 6), and note also that measurement error in schooling makes ability bias look worse
- IV estimates – less biased ... perhaps. Certainly noisier!

AKR twins in MM chpt 6

TABLE 6.2
Returns to schooling for Twinsburg twins

	Dependent variable			
	Log wage (1)	Difference in log wage (2)	Log wage (3)	Difference in log wage (4)
Years of education	.110 (.010)		.116 (.011)	
Difference in years of education		.062 (.020)		.108 (.034)
Age	.104 (.012)		.104 (.012)	
Age squared/100	−.106 (.015)		−.106 (.015)	
Dummy for female	−.318 (.040)		−.316 (.040)	
Dummy for white	−.100 (.068)		−.098 (.068)	
Instrument education with twin report	No	No	Yes	Yes
Sample size	680	340	680	340

Notes: This table reports estimates of the returns to schooling for Twinsburg twins. Column (1) shows OLS estimates from models estimated in levels. OLS estimates of models for cross-twin differences appear in column (2). Column (3) reports 2SLS estimates of a levels regression using sibling reports as instruments for schooling. Column (4) reports 2SLS estimates using the difference in sibling reports to instrument the cross-twin difference in schooling. Standard errors appear in parentheses.

from AK-91

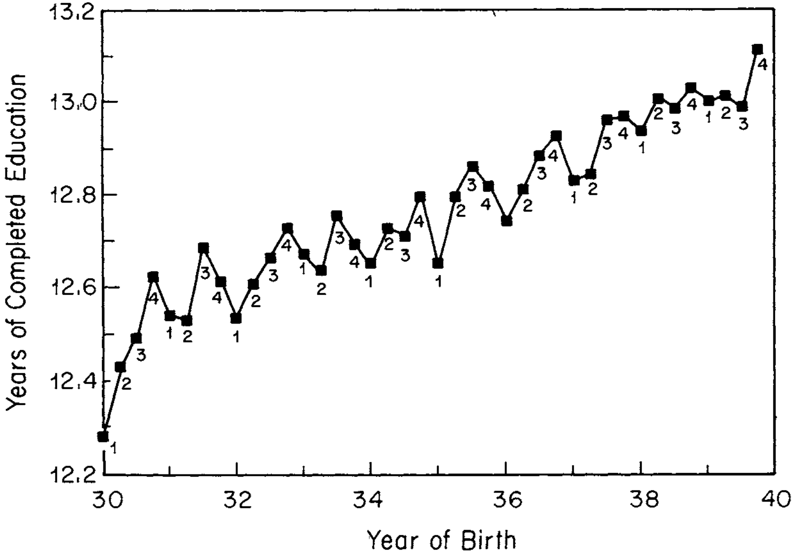

FIGURE I
Years of Education and Season of Birth
1980 Census
Note. Quarter of birth is listed below each observation.

TABLE I
THE EFFECT OF QUARTER OF BIRTH ON VARIOUS EDUCATIONAL OUTCOME VARIABLES

Outcome variable	Birth cohort	Mean	Quarter-of-birth effect[a] I	II	III	F-test[b] [P-value]
Total years of education	1930–1939	12.79	−0.124	−0.086	−0.015	24.9
			(0.017)	(0.017)	(0.016)	[0.0001]
	1940–1949	13.56	−0.085	−0.035	−0.017	18.6
			(0.012)	(0.012)	(0.011)	[0.0001]
High school graduate	1930–1939	0.77	−0.019	−0.020	−0.004	46.4
			(0.002)	(0.002)	(0.002)	[0.0001]
	1940–1949	0.86	−0.015	−0.012	−0.002	54.4
			(0.001)	(0.001)	(0.001)	[0.0001]
Years of educ. for high school graduates	1930–1939	13.99	−0.004	0.051	0.012	5.9
			(0.014)	(0.014)	(0.014)	[0.0006]
	1940–1949	14.28	0.005	0.043	−0.003	7.8
			(0.011)	(0.011)	(0.010)	[0.0017]
College graduate	1930–1939	0.24	−0.005	0.003	0.002	5.0
			(0.002)	(0.002)	(0.002)	[0.0021]
	1940–1949	0.30	−0.003	0.004	0.000	5.0
			(0.002)	(0.002)	(0.002)	[0.0018]
Completed master's degree	1930–1939	0.09	−0.001	0.002	−0.001	1.7
			(0.001)	(0.001)	(0.001)	[0.1599]
	1940–1949	0.11	0.000	0.004	0.001	3.9
			(0.001)	(0.001)	(0.001)	[0.0091]
Completed doctoral degree	1930–1939	0.03	0.002	0.003	0.000	2.9
			(0.001)	(0.001)	(0.001)	[0.0332]
	1940–1949	0.04	−0.002	0.001	−0.001	4.3
			(0.001)	(0.001)	(0.001)	[0.0050]

a. Standard errors are in parentheses. An $MA(+2, -2)$ trend term was subtracted from each dependent variable. The data set contains men from the 1980 Census, 5 percent Public Use Sample. Sample size is 312,718 for 1930–1939 cohort and is 457,181 for 1940–1949 cohort.

b. F-statistic is for a test of the hypothesis that the quarter-of-birth dummies jointly have no effect.

TABLE V

OLS AND TSLS ESTIMATES OF THE RETURN TO EDUCATION FOR MEN BORN 1930–1939: 1980 CENSUS[a]

Independent variable	(1) OLS	(2) TSLS	(3) OLS	(4) TSLS	(5) OLS	(6) TSLS	(7) OLS	(8) TSLS
Years of education	0.0711	0.0891	0.0711	0.0760	0.0632	0.0806	0.0632	0.0600
	(0.0003)	(0.0161)	(0.0003)	(0.0290)	(0.0003)	(0.0164)	(0.0003)	(0.0299)
Race (1 = black)	—	—	—	—	−0.2575	−0.2302	−0.2575	−0.2626
					(0.0040)	(0.0261)	(0.0040)	(0.0458)
SMSA (1 = center city)	—	—	—	—	0.1763	0.1581	0.1763	0.1797
					(0.0029)	(0.0174)	(0.0029)	(0.0305)
Married (1 = married)	—	—	—	—	0.2479	0.2440	0.2479	0.2486
					(0.0032)	(0.0049)	(0.0032)	(0.0073)
9 Year-of-birth dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
8 Region-of-residence dummies	No	No	No	No	Yes	Yes	Yes	Yes
Age	—	—	−0.0772	−0.0801	—	—	−0.0760	−0.0741
			(0.0621)	(0.0645)			(0.0604)	(0.0626)
Age-squared	—	—	0.0008	0.0008	—	—	0.0008	0.0007
			(0.0007)	(0.0007)			(0.0007)	(0.0007)
χ^2 [dof]	—	25.4 [29]	—	23.1 [27]	—	22.5 [29]	—	19.6 [27]

a. Standard errors are in parentheses. Sample size is 329,509. Instruments are a full set of quarter-of-birth times year-of-birth interactions. The sample consists of males born in the United States. The sample is drawn from the 5 percent sample of the 1980 Census. The dependent variable is the log of weekly earnings. Age and age-squared are measured in quarters of years. Each equation also includes an intercept.

Discount rate bias (HK theory meets IV 'metrics)

- Potential earnings at s years of schooling are $g_i(s)$
- Potential schooling indexed against a Bernoulli instrument, Z_i, determines actual schooling
$$S_i = S_{0i} + (S_{1i} - S_{0i})Z_i$$
- The Wald estimand using Z_i to instrument S_i can be shown to be
$$E\left\{\frac{S_{1i} - S_{0i}}{E\left[S_{1i} - S_{0i}\right]}\left[\frac{g_i(S_{1i}) - g_i(S_{0i})}{S_{1i} - S_{0i}}\right]\right\} = E\left\{\omega_i g_i'(S_i^*)\right\}$$
for some $S_i^* \in [S_{0i}, S_{1i}]$, where $\omega_i = \frac{S_{1i} - S_{0i}}{E[S_{1i} - S_{0i}]}$ (this uses the fact that $g_i(S_{1i}) = g_i(S_{0i}) + g_i'(S_i^*)(S_{1i} - S_{0i})$; for details, see AK99 HOLE Sec. 2.3.4)
  - IV captures a weighted average return to schooling over a range and for a set of individuals determined by the normalized first stage, ω_i
  - If $g_i(s) = \alpha + \rho_i$, then the IV estimand is $E[\omega_i \rho_i]$. Could be, but not likely, that this equals the pop avg return, $E[\rho_i]$
- Lang (1993) and Card (1995) postulate quadratic (concave) HK production:
$$g_i(s) = \alpha + \rho_1 s - \rho_2 s^2$$
implying
$$S_i^* = \frac{S_{0i} + S_{1i}}{2}$$
  - These authors consider a scenario where the first stage is proportional to discount rates. Since people with higher discount rates get less schooling, and the schooling-earnings relationship is concave, the Wald estimand in this case tends to exceed the pop average derivative, $E[g_i'(S_i)]$
  - Lang (1993) called this "discount rate bias", an idea that spawned a literature interpreting IV estimates of the returns to schooling

C Signaling

- Schooling need not boost productivity; perhaps schooling merely reveals *who* is smart and hence productive (sort of like Real Analysis!)
- In this case, we say "schooling is a signal," an idea that originates with Spence (1973) and Stiglitz (1975), first tested empirically for schooling by Lang and Kropp (1986)
  - See Borjas' labor text Chapter 6
- The idea in a nutshell:
  1. Schooling is easier (cheaper) to acquire for more productive (smarter) workers but does not in an of itself make them so. Schooling is harder (more costly) for less productive workers.
  2. More productive workers will find it worthwhile to pay for schooling if the premium paid to more productive workers is high enough and the fact that they are more educated convinces employers that they are indeed more productive
  3. For less productive workers, the signal is too expensive relative to the pay gap, so they do without
  4. Employers observe that the more educated are more productive without knowing why, so in a *signaling equilibrium*, the more educated indeed get paid more
- This is a special case of a *separating equilibrium* often seen in markets with imperfect information: in such equilibria, people's choices identify their type (in insurance markets, only the sick buy health insurance when insurance is very expensive)
- Separating equilibria can bring a market crashing down (hence the ACA coverage mandate)
- If schooling is not actually productivity-enhancing, the case for subsidizing it through policy is weaker
  - while workers gain by signaling their productivity, schooling is an expensive way to do this, and socially wasteful since valuable resources are used merely to send signals
  - if schooling spreads more widely, signaling value is diluted

Modern tests of signaling

- The signaling value of a HS diploma is explored in Martorell and Clark (2014), summarized in MM, Chpt. 6

- Aryal, Bhuller, and Lange (2022) deploy a new test, distinguishing IV estimates according to whether or not employers observe the instruments at hand. This paper argues:
  - An instrument unseen by employers identifies private returns to schooling (these returns include signaling value)
  - Schooling obtained in response to an instrument observed by employers has no signaling value, and so the returns this instrument identifies can be said to be social in the sense that they reflect enhanced productivity only
  - Signaling returns should fall with experience (they claim): employers learn workers' true productivity and can then ignore signals
  - Differences in IV estimates constructed using more- and less-salient compulsory attendance reforms in Norway, and variation in estimated returns with experience, suggest private returns are due mostly to higher productivity

Table 1—First-Stage Estimates on Years of Schooling

	Full sample (1)	Hidden IV sample (2)	Transparent IV sample (3)
Instrument			
Exposure to compulsory schooling reform	0.237	0.228	0.240
	(0.025)	(0.034)	(0.032)
Controls			
Municipality fixed effects	✓	✓	✓
Cohort fixed effects	✓	✓	✓
F-statistic (instrument)	87.7	45.7	55.5
Sample mean years of schooling	12.36	12.27	12.50
Standard deviation years of schooling	2.50	2.46	2.56
Number of observations	14,746,755	8,697,979	6,048,776

Notes: The full sample (column 1) consists of Norwegian males born 1950–1980, observed any time in earnings data over 1967–2014 with years of potential experience between 0 and 30 years and annual earnings above 1 SGA threshold ($N = 14{,}746{,}755$). The hidden IV sample (column 2) further drops individuals who grew up in the municipality with the largest population size in each of the 160 labor market regions in Norway ($N = 8{,}697{,}979$), while the transparent IV sample (column 3) retains only individuals who grew up in the municipality with the largest population size in each labor market ($N = 6{,}048{,}776$). All estimations include fixed effects for birth cohort and childhood municipality. We cluster the standard errors at the local labor market region.

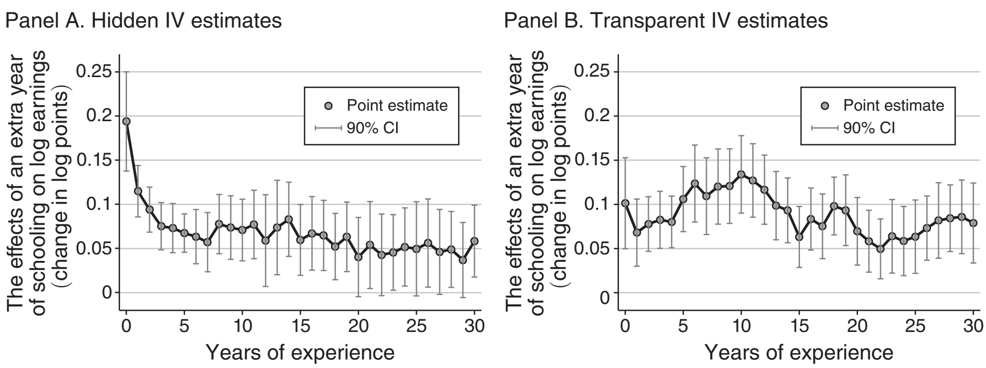

Figure 3. Hidden and Transparent IV Estimates of the Returns to Schooling

Notes: The estimation sample consists of Norwegian males born 1950–1980, observed in earnings data over years 1967–2014 with years of potential experience between 0 and 30 years and annual earnings above 1 SGA threshold ($N = 14{,}746{,}755$). The hidden IV sample (panel A) further drops individuals who grew up in the municipality with the largest population size in each of the 160 labor market regions in Norway ($N = 8{,}697{,}979$), while the transparent IV sample (panel B) retains only individuals who grew up in the municipality with the largest population size in each labor market ($N = 6{,}048{,}776$). Panel A and B display IV estimates from separate estimations of (15) for each year of experience using the hidden and the transparent IV samples. All estimations include fixed effects for birth cohort and childhood municipality. Standard errors are clustered at the local labor market region. Vertical bars denote the 90 percent confidence intervals.

approach allows differences in the experience-invariant part of the social returns across workers from central and noncentral locations.[26]

[26] Notably, Assumption 4 requires that the *experience-varying* component of returns to skill is identical across samples. However, this assumption does not restrict the *experience-invariant* component of social returns. Such heterogeneity in social returns could, for instance, exist due to differences in inputs factor in the production of human