The Economic Returns to Schooling

Josh Angrist

MIT 14.661 (FALL 2017)

Equalizing Differences: A Theoretical Benchmark (originates with Mincer, 1974)

- Though published after Becker's *Human Capital*, Mincer's *Schooling, Experience, and Earnings* is more squarely focused on empirical wage determination. The goal is a regression version of the human capital story that describes the causal effects of schooling and experience on earnings in a simple equilibrium model
- The theory is remarkably spare: no individual variation in ability or opportunity; perfect capital markets
  1. The cost of schooling is foregone earnings
  2. We discount the future at a common interest rate, r, at which all are free to borrow and lend
  3. Labor markets equalize the value of alternative investment plans
  4. Experience is "part-time schooling." Initially, we'll ignore this.
- $y(s)$ is how much someone (anyone) with s years of schooling will earn
- Schooling plans are valued at

$$V(s) = \int_s^\infty y(s)e^{\ rt}dt = \frac{y(s)e^{\ rs}}{r}$$

- The Equalizing Diffs equilibrium: people choose as much schooling as they like; they do so until the labor market equalizes plan values. In other words, for all s:

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

- This implies that

$$y(s) = e^{rs}y(0)$$

  or

$$\ln y(s) = \ln y(0) + rs$$

- We conclude:
  1. Schooling raises earnings; in a world of equalizing differences, the return on a year invested schooling is the same as the return on financial assets
  2. Individuals are indifferent between, say $s = 12$ and $s = 16$ because all plans yield the same PDV: quit school and be happy!
- Exercise: show that finite lifetimes raise the cross-sectional returns to schooling under equalizing diffs
- This is a highly unrealistic world - but it tells us that if labor and capital markets are functioning reasonably well, we might expect the returns to human capital to approach the cost of funds
- Since we're indifferent between, say, $s = 12$ and $s = 16$, there isn't much call to worry about selection bias. Likewise, in this simple homogenous world, the returns to schooling are reliably linear and constant.
  - Mincer's description of the CEF linking average log wages with schooling and experience has proved relevant and resilient. This simple regression is a workhorse of empirical labor economics.

Optimal HK in a Heterogeneous World (Becker, 1964)

- Abilities and opportunities vary - much more realistic - but the 'metrics is messier
- Everyone does the best they can, and the value of alternative plans is not necessarily equalized across plans
- Assume:
  1. The only cost of schooling is foregone earnings (again)
  2. Everyone has the same "opportunities," meaning we all borrow, lend, and discount the future at continuously compounded rate r
  3. Ability to convert HK into earnings differs
  4. Individuals choose schooling to max the PDV of their earnings profiles
- What's new? People are different! We'll start with a parameterization:

$$y_i(s) = g(s, a_i)$$

  where a_i is "ability" and $g(s, a_i)$ is the amount someone with ability a_i earns when they get schooling s
- Workers are paid their marginal products; more educated and more able workers are more productive

- PDV plan values are given by

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

for schooling plan s and individual i

- Pick your optimal schooling levels by setting $V_i'(s) = 0$:

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

or

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

- This implies

$$\ln y_i(s) = \ln y_i(0) + rs,$$

just like equalizing diffs

- But now we all want different schooling levels - because our returns differ (or so we must hope!)

Ability bias

- Are more educated people more productive *because* of their schooling – or are these people who would have earned more regardless? To find out, we might estimate a regression like this one:

$$\ln y_i = \alpha_1 + \rho_1 s_i + \quad a_i + \eta_{1i},$$

in other words, a regression of log wages on schooling controlling for ability. We can think of this as defining

$$a_i = \ln y_i(0)$$

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 in a wage equation is often called "ability bias" - here's the formula:

$$\rho_0 = \rho_1 + \quad _{as}$$

where $\quad _{as}$ is the regression of omitted on included

- Ability bias is surely positive: the more able get the most schooling, right?

Griliches (1977)

- Try this

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

  Whoops! This generates either ∞ or no schooling [show this]

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

$$V(s) = \int_s^{\infty} g(s, a_i) e^{\ rt} dt + \int_0^s TR e^{\ rt} dt$$

$$= \frac{g(s, a_i) e^{\ rs}}{r} + \frac{TR}{r}(1 \quad e^{\ rs})$$

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

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

  again, an MR=MC type relation.

- Now we've got sufficient curvature:

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

  Curvature is on the *cost* side: the importance of TR declines with s.

  - Define $\ _i(s) = \frac{TR}{y_i(s)}$ for $\ _i(s) \in [0, 1]$. Then the FOC becomes

$$= r[1 \qquad _i(s)] \tag{1}$$

    and the graduate stipend increases our schooling because MC is below r (intuitively we have to borrow less). Because TR is fixed, $\ _i(s)$ is declining in s. We'll also assume $\quad < r$, so that the Griliches model can get an interior solution for s by picking the implied value of that produces (1)

- The solution ain't pretty, but it's optimal:

$$s_i^* = \frac{1}{-}\{\ln TR \quad \ln \frac{r}{r} \qquad a_i\}$$

  where again we require $\quad < r$. Ability what?

- Now try this:

$$g(s, a_i) = e^{\ a_i + \ _1 s \quad \ _2 s^2 + \ _3 a_i s} \ ; \quad _3 > 0$$

  We can also vary costs:

$$TR_i = TR_0 + \ a_i$$

$$r_i = r_0 \quad r_1 a_i$$

- What's the moral here?

Estimates of the economic returns to schooling (attached)

- Controls: the good, the bad, and the ugly
- IV estimates

Discount rate bias (following Lang 1993, Card 1995, and AK-99)

- Potential earnings at s years of schooling are $g_i(s)$
- Potential schooling indexed against a Bernoulli instrument, Z_i, determines actual schooling

$$S_i = S_{0i} + (S_{1i} \quad S_{0i})Z_i$$

- The Wald estimand can be shown to be

$$E\left\{\frac{S_{1i} \quad S_{0i}}{E\left[S_{1i} \quad S_{0i}\right]}\left[\frac{g_i(S_{1i}) \quad g_i(S_{0i})}{S_{1i} \quad S_{0i}}\right.\right. \quad = E\left\{\omega_i g_i'(S_i^*)\right\}$$

  for some $S_i^* \in [S_{0i}, S_{1i}]$, where $\omega_i = \frac{S_{1i} \quad S_{0i}}{E[S_{1i} \quad S_{0i}]}$ (this uses the fact that $g_i(S_{1i}) = g_i(S_{0i}) + g_i'(S_i^*)(S_{1i} \quad S_{0i})$; for details, see Angrist, Imbens, Graddy, 2000)

  - 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, $E[\rho_i]$. [when will it?]

- Lang (1993) and Card (1995) postulate quadratic (concave) HK production:

$$g_i(s) = \alpha + \rho_1 s \quad \rho_2 s^2$$

  implying

$$S_i^* = \frac{S_{0i} + S_{1i}}{2}$$

  - These authors further sketch a scenario where the first stage is proportional to discount rates. Since people with higher discount rates get less schooling, and the schooling-earnings relation 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

Signaling

- Perhaps schooling does not boost productivity; rather, it reveals *who* the productive people are
- In this case, we say: "schooling is a signal," an idea that originates with Spence (1973) and Stiglitz (1975), first tested empirically by Lang and Kropp (1986)
- The idea in a nutshell:
    1. Schooling is easier (cheaper) to acquire for more productive workers but does not in and of itself make them so
    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 they are more productive
    3. For less productive workers, the signal is too expensive relative to the pay gap so they do without
    4. Employers know this, so in equilibrium the more educated get paid more, because their willingness to pay for schooling signals that they are indeed more productive
- If things work out this way, it's called a "separating equilibrium"
- The signaling value of a HS diploma nicely done in Martorell and Clark (2014), summarized in MM, Chpt. 6

14.661 Labor Economics I

Fall 2017