Chapter 6: Human capital investment – Part 1: Education

1. Human capital: application of Adam Smith and compensating wage diffs
   A. wage diffs for jobs with greater “difficulty and expence of learning them”
   B. human capital is...
      • “capital,” i.e., entails an up-front cost and a stream of returns into the future
      • “human,” because it’s embodied in the person
   C. examples of human capital include:
      education, on-the-job training (classroom or informal), job search costs,
      some medical care (maybe even nose jobs!)

2. Education as a form of human capital
   A. compare two earnings streams, with and without additional education
      (NB: stream w/more education could include negative earnings early on
       e.g., tuition and books – but not room and board (why?))
2. Education as a form of human capital (continued)
   B. costs of education include foregone earnings as well as tuition, books, etc.
   benefits of education include higher future earnings, better working conditions (?)
   C. most of the benefits are received well into the future,
   so must **discount** these future sums to their **present value** (i.e., their value **today**)

   e.g., if interest rate = 5%, then “present value” (as of today)
   of getting $100 a year from now (the future value) is approximately $95 –
   in other words, we would have to put only about $95 in the bank today at 5% interest
   in order to have $100 a year from now

   more generally (and more precisely), $PV = \frac{FV_1}{1 + r}$

   where $PV = \text{present value (as of today)}, \quad r = \text{interest rate}$
   $FV_1 = \text{future value of an amount received one year from now}$

   still more generally, $FV_1 = PV (1 + r)$
   and $FV_2 = FV_1 (1 + r) = PV (1 + r)(1 + r)$
   and so $FV_1 = PV (1 + r)^2$

   and so, generalizing, $FV_n = PV (1 + r)^n$ \hspace{1cm} (future value of an amount received n yrs from now)
   so that $PV = \frac{FV_n}{(1 + r)^n}$ \hspace{1cm} (present value of amount received n years from now)
D. Note: when \( r \) rises, PV of any given future amount falls (more heavily discounted) (thus, higher \( r \) reduces the PV of the benefits from educational investment)

E. Optimal schooling investment: invest in education provided
   \[ \text{PV of benefits} > \text{PV of costs} \]

3. The wage-schooling locus: a “stopping rule”
   A. wage-schooling locus shows earnings at each schooling level
      positive-sloped: more schooling \( \Rightarrow \) more pay
         (slope of the locus measures increase in earnings from one more year of school)
      concave: diminishing returns to schooling in the “production” of earnings
   B. slope of wage-schooling locus in percentage terms =
      “marginal rate of return to schooling”
C. so here, the “stopping rule” is: invest in schooling until marginal rate of return to schooling = interest rate (another kind of benefit-cost comparison) (see the Appendix, below)

D. thus, differences in rate of discount, \( r \), and in returns to schooling (due to differences in ability) will affect investment in schooling

E. ceteris paribus, higher discount rate (\( r \)) \( \Rightarrow \) less schooling (e.g., Al vs. Bo, graphs below)
F. greater ability $\Rightarrow$ higher return to schooling at all levels of schooling
   $\Rightarrow$ even more schooling (!)

NB: the earnings difference between people with different schooling
does *not* necessarily tell us what the less-schooled person would earn
from more schooling!
likewise, the earnings difference does *not* necessarily tell us
what the person with more schooling gained from the extra schooling!
4. “Ability bias” and the return to schooling

A. these results indicate that the earnings diff btw people w/different schooling may contain a component due to greater ability (provided greater ability \(\Rightarrow\) higher return to schooling at all levels of schooling)

B. then, considered as an estimate of the effect of schooling, the earnings diff contains an “ability bias” (or a “selection bias”)

C. in the graph of Ace and Bob (see previous page),
   the *estimate* of the earnings gain that Ace would get from more schooling is \(w_B - w_A\) (= the difference between Bob’s earnings and Ace’s earnings) but the *true* payoff for education for Ace is only \(w_A' - w_A\).
so the *bias* in the estimate of the payoff for Ace is \(w_B - w_A'\).
5. Estimating the rate of return to schooling with a wage equation
   
   A. typical estimating equation:
      \[ \log w = a + bS + \text{other variables} + e \]

   where
   - \( w \) = wage or earnings,
   - \( S \) = schooling,
   - \( e \) = “error term” (unmeasured/omitted variables)
   - \( b \) = coefficient on schooling (“rate of return to schooling”)

   remember that \( e \) refers, by definition, to omitted variables that increase \( \log w \)
   (e.g., “laziness” would enter with a minus sign!)

   B. omitted variables bias:
      if \( e \) is correlated with \( S \), for given values of all the other variables,
      then the OLS estimate of \( b \) will be biased

      if \( e \) is positively correlated with \( S \) for given values of all the other variables,
      then the OLS estimate of \( b \) will be upward biased (more positive/less negative)

      if \( e \) is negatively correlated with \( S \) for given values of all the other variables,
      then the OLS estimate of \( b \) will be downward biased (less pos/more neg)
\[ y_i = a + b x_i + e_i \]
\[ \overline{y} = a + b \overline{x} + \overline{e} \]

so

\[ y_i = b x_i + e_i \]
\[ y_i x_i = b x_i^2 + e_i x_i \]
\[ \sum y_i x_i = b \sum x_i^2 + \sum e_i x_i \]

so

\[ \frac{\sum y_i x_i}{\sum x_i^2} = b + \frac{\sum e_i x_i}{\sum x_i^2} \]

so

\[ \hat{b} = b + \frac{\sum e_i x_i}{\sum x_i^2} \]

where

\[ E\{\sum e_i x_i\} = 0 \]

\[ \hat{b} \] is uncorrelated

positively correlated

negatively correlated

So if \( e, x \) positively correlated, \( \hat{b} > b \) (\( \hat{b} \) upward biased)

negatively

\( \hat{b} < b \) (\( \hat{b} \) downward biased)
6. Estimating the rate of return to schooling: avoiding “ability bias”

A. “natural experiments”: compare twins’ earnings – results are quite mixed (some estimates much lower than OLS, others are much higher) also, why would twins earn different returns? Because of different discount rates? then how can we say they’re “identical”?

B. instrumental variables: Vietnam draft lottery (random selection into draft, and out of school) – estimate of pure return to schooling here is about 7%

C. instrumental variables: diff-in-diff using school construction in Indonesia
   • 0.11 extra years of education in high-construction areas (vs. to other areas)
   • 0.03 greater change in log-earnings in high-construction areas (vs. other areas)
   • So 0.03/0.11 = 0.27, or 27% higher earnings (assumes no ability differences in different construction areas or age cohorts)
   • Subsequent estimates revised down to 10%

D. school quality and earnings
   Card & Krueger: return to schooling for workers in a particular state is positively correlated with measures of school quality (e.g., teacher/student ratio)
   STAR experiments: random assignment of students and teachers to different class sizes – students in smaller classes do better on tests
   Maimonides’ rule (Israel): random variation in class size smaller class sizes ⇒ better academic achievement
7. Schooling as a signal

A. What if schooling doesn’t affect productivity at all?
What if schooling merely indicates ("signals") what productivity has been, *all along*? (e.g., does Harvard Business School really do a better of job training students, or are HBS students better to begin with?)

B. basic “signaling” model of educational investment: key ideas
- it’s costly to determine workers’ productivities
- high-productivity workers can get education at lower personal cost
- so firms offer more $ to workers with more education, to get them to self-identify as highly productive

C. simple signaling model:
q workers are low-ability, 1-q are high-ability

<table>
<thead>
<tr>
<th></th>
<th>college cost/year</th>
<th>lifetime earnings</th>
</tr>
</thead>
<tbody>
<tr>
<td>high-ability workers</td>
<td>$20,000.</td>
<td>$300,000.</td>
</tr>
<tr>
<td>low ability workers</td>
<td>30,000.</td>
<td>200,000.</td>
</tr>
</tbody>
</table>

* low-productivity worker won’t attend college for y* years if
  200,000 > 300,000 – (25,001 x y*)  or if y* > 3.999
* high-productivity worker will attend college for y* years if...
  200,000 < 300,000 – (20,000 x y*)  or if y* < 5

so competitive threshold education level will satisfy 3.999 < y* < 5
(competition will drive this down to 4)
7. Schooling as a signal (continued)

D. so (see next page) setting $y^* = 4$ separates low- from high-productivity people
none of the low-productivity workers go to college,
all of the high-productivity workers go to college
thus, people “self-identify” as to their level of productivity
by going, or not going, to college

E. Paradox: even if education has no effect on productivity,
we’ll still see higher earnings for more education!
(thus, even in a screening world, coefficient on schooling will be positive)

F. in screening model, giving everyone more education *won’t* change productivity,
or output, *at all*!
remember, in this model, education doesn’t make you more productive –
it simply reveals that you were already more productive, and always will be

G. thus, in this model, Harvard Business School graduates would have done
just as well for themselves as they actually have done,
even if they had never gone to HBS!
(although they became successful by going to HBS, thereby revealing to everyone
that they always have been, and always will be, more productive)
Earnings: \[ \frac{\text{Low - productivity worker}}{\text{High - productivity worker}} = \frac{200,000}{280,000} = \frac{10}{14} = \frac{5}{7} \]

Costs: \[ \text{Costs} = 25,000 \times 200,000 \]

Slope: \[ \text{slope} = 25,000 \]
1. Age-earnings profiles: Some stylized facts
   A. Earnings rise at a decreasing rate with age (“concave”)
   B. highly-educated workers earn more than less-educated workers
   C. age-earnings profiles of education groups “fan out” (diverge) with age
      ➔ implies education-experience complementarity

2. Post-schooling investments
   A. to understand age-earnings profiles, must understand post-schooling investments
   B. key distinction: general OJT vs. firm-specific OJT
3. General on-the-job training
   A. “general” OJT is useful “anywhere” (by definition)

   B. **key conclusions** (require a bit of thought!):
      - worker will pay all the costs of general OJT, usually via lower wage at first
      - worker will receive all the benefits of general OJT, usually via higher wage later on

   C. logic underlying the conclusions:
      - if firm paid the costs, worker could get training for free
      - relative to jobs with no OJT, jobs with OJT would be very attractive
      - so workers would apply for jobs with OJT in huge numbers
      - employers will then realize that they can offer lower wages in jobs with OJT, but workers will still be willing to seek jobs with OJT (because the training is general, it is portable – so workers can cash in on the OJT even if they go to work elsewhere)

   D. so workers getting general OJT will have lower wages at first, but higher wages later, relative to their pay if they had not gotten any general OJT
      in effect, workers pay for their general OJT by accepting lower wages at first, and collect the benefits of their general OJT by getting higher wages later on

   E. so the story for general OJT looks a lot like the story for education (which is not terribly different from general OJT)
3. General on-the-job training (continued)
   F. slope of earnings profile with general OJT is steep at first, flattens out later
      • worker is investing a lot at first (no reason to invest late in life, because not much
time to collect the benefits!)
      • as time passes, time remaining before retirement/death is shrinking
      • so payoff period to the investment is less, so worker gradually reduces
investment as time passes
      • so worker is acquiring skill rapidly at first, so earning power grows rapidly at first
      • since worker pays for the investment via a low wage, wage rises as investment
declines
      • later in life, skills and earning-power depreciate
   G. NB: turnover is NOT important to a firm with general OJT
5. Firm-specific on-the-job training
   A. firm-specific OJT is useful ONLY in the firm where it was acquired
   B. conclusions (require a bit of thought!)
      • Worker and firm will end up sharing both the costs and benefits of firm-specific OJT
      • At first, worker’s wage will be below what s/he could get elsewhere (= worker’s cost of investment); later, wage will exceed what worker could get elsewhere (= worker’s benefit (return) from investment)
      • At first, worker’s wage will be greater than productivity (= firm’s cost of investment); later, it will be smaller than productivity (= firm’s benefit from investment)
   C. logic underlying these conclusions:
      • since firm-specific OJT isn’t portable, workers would be reluctant to pay for it
      • suppose firm pays for it: it can profit on its investment (which will make workers more productive), but only if the workers stay at the firm
      • to encourage workers to stay, firm will offer higher wages after the training is completed
      • but this will make jobs with firm-specific OJT more attractive than other jobs
      • so workers will apply for jobs with firm-specific OJT in large numbers
      • so employers will realize that they can cut starting wages in jobs with firm-specific OJT, and workers will still be willing to apply (because they know they will get higher wages later on)
5. Firm-specific OJT (continued)
   D. so, workers getting firm-specific OJT will have lower wages at first, but higher wages later on, than in absence of firm-specific OJT
      thus, firm-specific OJT story is much like general OJT story
   E. caveats:
      worker can’t get the rewards of firm-specific OJT unless he stays on at the firm
      firm can’t get the rewards of firm-specific OJT unless worker stays on at the firm
   F. so firm and worker are “bonded” – each has strong incentive to stick with the other
      turnover is a crucial concern in firm with firm-specific OJT

6. Mincer “human capital earnings function”

\[
\log W = r S + b_1 X + b_2 X^2 + c_1 F + c_2 F^2 + \text{other terms} + e
\]

   e = “error term” (missing/unobserved variables)
   S = years of schooling
   X = total years of work experience (prior experience, plus years at current firm)
   F = years at current firm

semilog form implies that coefficients represent percentage effects, e.g.:
   \( \frac{\partial \log W}{\partial S} = \frac{1}{W} \frac{\partial W}{\partial S} = r \) ( = % increase in wage from
   one-unit increase in schooling
   = “return to schooling”)
6. Mincer “human capital earnings function” (continued)

quadratic ("squared") terms in X and F allow for concavity of age-earnings profile (provided \( b_2 < 0, c_2 < 0 \))

partial derivative w.r.t. X (with F constant!) gives effect of general OJT:

\[
\frac{\partial \log W}{\partial X} = \frac{1}{W} \frac{\partial W}{\partial X} = b_1 + 2 b_2 X = \text{effect of general OJT on } W
\]

with F constant, a rise in X must mean more prior work experience, which can raise W (pay at current employer) only if it provided general OJT

partial derivative w.r.t. F (with X constant!) gives effect of firm-specific OJT:

\[
\frac{\partial \log W}{\partial F} = \frac{1}{W} \frac{\partial W}{\partial F} = c_1 + 2 c_2 F = \text{effect of firm-specific OJT on } W
\]

with X constant, a rise in F must mean both (a) more current experience and (b) less prior work experience, which can raise W (pay at current employer)
only if it provided firm-specific OJT (the “general” component is removed)

Mincer earnings function used in many different contexts (e.g., studies of pay differences by race, sex, government/private sector, union/nonunion, etc.)
7. Government job-training programs

A. typical analysis: run earnings function... \[ Y = a + bT + \text{other terms} + e \]
   \[ e = \text{error term (unmeasured/omitted variables)} \]
   \[ T = \text{dummy variable for trainee (}= 1 \text{ if trainee, } = 0 \text{ if not)} \]

B. potential problem: omitted-variables bias (or, equivalently, selection bias)
   • If \( e \) is positively correlated with \( T \), OLS estimate of \( b \) will be upward-biased (e.g., program “creams” the best, most motivated people; some of their greater earnings are due to this, rather than to the program)
   • If \( e \) is negatively correlated with \( T \), OLS estimate of \( b \) will be downward-biased (e.g., program “scrapes” to get hard-core unemployed; some of their lower earnings are due to this, rather than to any low quality of the program)

C. social experiments to determine effect of training program – experimentalists get training, controls do not
diff-in-diff estimate of training program effect: earnings gain of $1407/year (implies breakeven of benefits vs. costs in > 10 years, and approx 10% return)

D. caveats: even experiments may not provide accurate estimate of effect of training
   • treatment and control groups aren’t necessarily a random sample of all the disadvantaged – rather, they’re self-selected
   • control group doesn’t necessarily get no other training, no other assistance
   • treatment group doesn’t necessarily all get training