Labor Economics, 14.661. Lectures 1 and 2: Investment in Education and Skills

Daron Acemoglu

MIT

October 27 and 29, 2009.

Introduction

- Key idea in labor economics: the set of marketable skills of workers $\approx$ form of capital
- **Human capital theory**: workers make a variety of investments in their human capital (set of marketable skills). This perspective is important in understanding both investment incentives, and the structure of wages and earnings.
- How do we measure human capital?
- Two approaches:
  1. Objective measures... often imperfect
  2. Earnings potential... approximated by *actual* earnings.
- Both approaches popular in practice.
- Both have the same problem: *unobserved heterogeneity.*

Introduction (continued)

- Unobserved heterogeneity will make "objective measures" imperfect.
- But it will also mean that there are non-human capital related sources of earnings differentials.
- For example

1. Compensating differentials
2. Labor market imperfections
3. Taste-based discrimination

- What can we do?
- Use both approaches, with caution.

Uses of Human Capital

- Simplest (standard view): human capital increases a worker's earnings.
- But how?

1. The Becker view: human capital directly useful in the production process and increases productivity in a broad range of tasks.
2. The Gardener view: multi-dimensional skills; ranking impossible.
3. The Schultz/Nelson-Phelps view: human capital as capacity to adapt.
4. The Bowles-Gintis view: "human capital" as the capacity to work in organizations and obey orders.
5. The Spence view: observable measures of human capital are more a signal of ability than characteristics independently useful in the production process.

- Despite their differences, the first three views quite similar.

Sources of Human Capital Differences

- Why will human capital the for across workers?

1. Innate ability
2. Schooling
3. School quality and non-schooling investments
4. Training
5. Pre-labor market influences

- First part of these lectures about theories and empirical implications of these different channels.

The Separation Theorem

- Consider an individual with an instantaneous utility function $u(c)$ with planning horizon of T (here $T = \infty$ is allowed).
- Continuous time for simplicity.
- Discount rate $\rho > 0$ and constant flow rate of death equal to $\nu \geq 0$.
- Perfect capital markets.
- Objective of the individual:

$$\max \int_0^T \exp\left(-\left(\rho + \nu\right) t\right) u\left(c\left(t\right)\right) dt. \tag{1}$$

- Suppose that this individual is born with some human capital $h(0) \geq 0$.

The Separation Theorem (continued)

- Evolution of human capital:

$$\dot{h}(t) = G(t, h(t), s(t)), \tag{2}$$

- $s(t) \in [0,1]$ fraction of time that the individual spends for investments in schooling
- Suppose also that

$$s(t) \in \mathcal{S}(t) \subset [0,1], \tag{3}$$

  (for example, to allow for $s(t) \in \{0,1\}$ so that only full-time schooling would be possible).
- Exogenous sequence of wage per unit of human capital given by $[w(t)]_{t=0}^{T}$, so that his labor earnings at time t are

$$W(t) = w(t)[1 - s(t)][h(t) + \omega(t)],$$

- Here $1 - s(t)$ is the fraction of work time and $\omega(t)$ is non-human capital labor, with $[\omega(t)]_{t=0}^{T}$ exogenous.
- **Note:** this formulation assumes no leisure.

The Separation Theorem (continued)

- Perfect capital markets: borrowing and lending at constant interest rate equal to r.
- Therefore, sufficient to express the lifetime budget constraint

$$\int_0^T \exp(-rt)\, c(t)\, dt \leq \int_0^T \exp(-rt)\, w(t)\, [1 - s(t)]\, [h(t) + \omega(t)]\, dt. \tag{4}$$

The Separation Theorem (continued)

Theorem

Suppose $u(\cdot)$ is strictly increasing. Then the sequence $\left[\hat{c}(t), \hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ is a solution to the maximization of (1) subject to (2), (3) and (4) if and only if $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ maximizes

$$\int_0^T \exp(-rt)\, w(t)\,[1 - s(t)]\,[h(t) + \omega(t)]\,dt \tag{5}$$

subject to (2) and (3), and $[\hat{c}(t)]_{t=0}^{T}$ maximizes (1) subject to (4) given $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$. That is, human capital accumulation and supply decisions can be separated from consumption decisions.

The Separation Theorem (continued)

- Intuition: in the presence of perfect capital markets, the best human capital accumulation decisions are those that maximize the lifetime budget set of the individual.
- What are the limitations of this result?
- Is it useful for empirical applications?

Returns to Education

- Mincer's (1974) model.
- Assume that $T = \infty$.
- Suppose that $s(t) \in \{0, 1\}$ and that schooling until time S.
- At the end of the schooling interval, the individual will have a schooling level of

$$h(S) = \eta(S),$$

where $\eta(\cdot)$ increasing, continuously differentiable and concave.
- For $t \in [S, \infty)$, human capital accumulates over time (as the individual works) according to the differential equation

$$\dot{h}(t) = g_h h(t), \tag{6}$$

for some $g_h \geq 0$.

Returns to Education (continued)

- Suppose also that the wage rate wage per unit of human capital grows exponentially,
$$\dot{w}(t) = g_w w(t), \tag{7}$$
with boundary condition $w(0) > 0$.
- Let us also assume that
$$g_w + g_h < r + \nu,$$
so that the net present discounted value of the individual is finite (why is this necessary?).

Returns to Education (continued)

- Now using Theorem 1, the optimal schooling decision must be a solution to the following maximization problem

$$\max_S \int_S^\infty \exp\left(-\left(r+\nu\right)t\right) w\left(t\right) h\left(t\right) dt. \tag{8}$$

- Now using (6) and (7), this is equivalent to:

$$\max_S \frac{\eta\left(S\right) w\left(0\right) \exp\left(-\left(r+\nu-g_w\right) S\right)}{r+\nu-g_h-g_w}. \tag{9}$$

- Unique solution (why?) This:

$$\frac{\eta'\left(S^*\right)}{\eta\left(S^*\right)} = r+\nu-g_w. \tag{10}$$

- Higher interest rates and higher values of ν reduce human capital investments; higher values of g_w increase the value of human capital and thus encourage further investments.

Returns to Education (continued)

- Integrating both sides of the above equation with respect to S:

$$\ln \eta \left(S^*\right) = \text{constant} + \left(r + \nu - g_w\right) S^*. \quad (11)$$

- Therefore, wage earnings at age $\tau \geq S^*$ and at time t:

$$W\left(S, t\right) = \exp\left(g_w t\right) \exp\left(g_h \left(t - S\right)\right) \eta\left(S\right).$$

- Taking logs and using equation (11):

$$\ln W\left(S^*, t\right) = \text{constant} + \left(r + \nu - g_w\right) S^* + g_w t + g_h \left(t - S^*\right).$$

- Here $t - S^*$ is "experience" (time after schooling).
- In cross-sectional comparisons, time trend $g_w t$ will also go into the constant, so that we obtain the *canonical Mincer equation*:

$$\ln W_j = \text{constant} + \gamma_s S_j + \gamma_e \text{experience}, \quad (12)$$

  with γ_s at the margin equal to $r + \nu - g_w$.
- **But**: this is an approximation. Why? Is it important?

Returns to Education (continued)

- Economic insight: the functional form of the Mincerian wage equation is not just a mere coincidence
  - the opportunity cost of one more year of schooling is foregone earnings.
  - hence benefits to schooling must be commensurate with these foregone earnings
  - thus at the margin, one year of schoolingus should lead to a proportional increase of approximately $(r + \nu - g_w)$ in future earnings.
  - **but**, this result does **not** imply that wage-schooling relationship should be log linear everywhere. Why not?
- Empirical work using equations of the form (12) leads to estimates for γ in the range of 0.06 to 0.10.
- Reasonable estimates; r approximately 0.10, ν approximately 0.02 (expected life of 50 years), and g_w—rate of wage growth holding the human capital level of the individual constant—between 0.01 and 0.02.

Schooling Investments Under Credit Market Imperfections

- In practice, free borrowing and lending at the same rate might be unreasonable.
- How do credit market problems influence schooling decisions?
- A two-period model to illustrate the main ideas.
- In period 1 an individual (parent) works, consumes c, saves s, decides whether to send their offspring to school, $e = 0$ or 1, and then dies at the end of the period.
- Utility of household i is given as:

$$\ln c_i + \ln \hat{c}_i \tag{13}$$

- $\hat{c}$ is the consumption of the offspring.
- Heterogeneity among children: the cost of education, θ_i varies with i.
- In the second period skilled individuals (those with education) receive a wage w_s and an unskilled worker receives w_u.

Schooling Investments Under Credit Market Imperfections (continued)

- Without credit market problems, the problem is trivial.
- The Separation Theorem, Theorem 1 applies.
- "Lifetime" budget constraint:

$$c_i + \frac{\hat{c}_i}{1+r} \leq \frac{w_u}{1+r} + e_i \cdot \frac{w_s - w_u}{1+r} + y_i - e_i \cdot \theta_i$$

- Again e_i does not appear in the objective function, so the education decision will be made simply to maximize the budget set of the consumer.

Schooling Investments Under Credit Market Imperfections (continued)

- Therefore parents will choose to educate their offspring only if

$$\theta_i \leq \frac{w_s - w_u}{1 + r} \tag{14}$$

- Three important observations:
  1. Schooling decision depends on ability inversely measured by θ_i.
  2. Greater skill premium as captured by $w_s - w_u$ encourages schooling.
  3. Lower interest rate r encourages schooling.

Schooling Investments Under Credit Market Imperfections (continued)

- Now suppose that parents cannot have negative savings (why is this reasonable?).
- Than the budget constraint is replaced by

$$\begin{aligned} c_i &\leq y_i - e_i \cdot \theta_i - s_i \\ s_i &\geq 0 \\ \hat{c}_i &\leq w_u + e_i \cdot (w_s - w_u) + (1 + r) \cdot s \end{aligned}$$

- First note that for a parent with $y_i - e_i \cdot \theta_i > w_s$, the constraint of nonnegative savings is not binding, so the same solution as before will apply $\Longrightarrow$ credit constraints will only affect parents who needed to borrow to finance their children's education.
- Now focus on parents who would not choose positive savings, that is, parents with $(1 + r) y_i \leq w_u$.

Schooling Investments Under Credit Market Imperfections (continued)

- Utilities from investing and not investing in education are

$$U(e = 1 \mid y_i, \theta_i) = \ln(y_i - \theta_i) + \ln w_s$$

and

$$U(e = 0 \mid y_i, \theta_i) = \ln y_i + \ln w_u.$$

Schooling Investments Under Credit Market Imperfections (continued)

- Comparison of these expressions implies that parents with

$$\theta_i \leq y_i \cdot \frac{w_s - w_u}{w_s}$$

will invest in education.

1. This condition is more restrictive than (14) above, since

$$(1+r)\, y_i \leq w_u < w_s.$$

2. As income increases, there will be more investment in education, which contrasts with the non-credit-constrained case.
3. The skill premium, $w_s - w_u$, still has a positive effect on human capital investments. But this is not a general result. Why not?

Evidence

- Does income have an effect on schooling?
- Typical regression

$$\text{schooling}=\text{controls} + \alpha \cdot \log \text{parental income}$$

- Result: positive estimates of α, consistent with credit constraints.
- But at least two alternative explanations for why we may be estimating a positive α:

1. Children's education may also be a consumption good, so rich parents will "consume" more of this good as well as other goods.
2. The distribution of costs and benefits of education differ across families, and are likely to be correlated with income (i.e., θ_i correlated with y_i).

Evidence (continued)

- Include other characteristics to proxy for the costs and benefits of education or for attitudes toward education.
- When parents' education is also included in the regression, the role of income is substantially reduced.
- Conclusion?
- Two considerations:

1. First, parents' income may affect more the quality of education, especially through the choice of the neighborhood in which the family lives.
2. Parental income is often measured with error, and has a significant transitory component, so parental education may be a much better proxy for permanent income than income observations in these data sets.

Evidence (continued)

- To deal with the first problem, let us look at a more "encompassing" measure of skills $\rightarrow$ *earnings*.
- $\Longrightarrow$ same issues as intergenerational mobility.
- The typical regression here is

$$\text{log child income} = \text{controls} + \alpha \cdot \text{log parental income} \tag{15}$$

- Regressions of this sort were first investigated by Becker and Tomes. They found relatively small coefficients, typically in the neighborhood of 0.3.
- This means that if your parents are twice as rich as my parents, you will typically be about 30 percent as rich as I. Your children will be only 10 percent richer than my children.
- With this degree of intergenerational dependence, differences in initial conditions will soon disappear$\rightarrow$ converges to a relatively "egalitarian" society (does this mean inequality will disappear?)

Evidence (continued)

- To elaborate on this, consider the following simple model:

$$\ln y_t = \mu + \alpha \ln y_{t-1} + \varepsilon_t$$

- y_t is the income of t-th generation, and ε_t is serially independent disturbance term with variance σ_ε^2.
- Then the long-term variance of log income is:

$$\sigma_y^2 = \frac{\sigma_\varepsilon^2}{1 - \alpha^2} \tag{16}$$

- Using the estimate of 0.3 for α, equation (16) implies that the long-term variance of log income will be approximately 10 percent higher than σ_ε^2.
- Therefore, the long-run income distribution will basically reflect transitory shocks to dynasties' incomes and skills, and not inherited differences.
- But inequality could be very large if σ_ε^2 is large.

Evidence (continued)

- What does this say about credit market problems?
- Persistence of about 0.3 is not very different from what we might expect to result simply from the inheritance of IQ between parents and children, or from the children's adoption of cultural values favoring education from their parents.
- Therefore, relatively small effect of parents income on children's human capital.

Evidence (continued)

- However, econometric problems biasing α toward zero.
- First, **measurement error**.
- Second, in typical panel data sets, we observe children at an early stage of their life cycles, where differences in earnings may be less than at later stages.
- Third, income mobility may be very nonlinear, with a lot of mobility among middle income families, but very little at the tails.
- Solon and Zimmerman: dealing with the first two problems increases α to about 0.45 or even 0.55.
- A paper by Cooper, Durlauf and Johnson, in turn, finds that there is much more persistence at the top and the bottom of income distribution than at the middle.
- If $\alpha = 0.55$, then $\sigma_y^2 \approx 1.45 \cdot \sigma_\varepsilon^2$ instead of $\sigma_y^2 \approx 1.1 \cdot \sigma_\varepsilon^2$—substantial difference.

Evidence (continued)

- Second empirical issue: endogeneity.
- Either structural or instrumental variable approaches.
- Acemoglu and Pischke (2001): exploit changes in the income distribution that have taken place over the past 30 years to get a source of exogenous variation in household income.
- Basic idea: the rank of a family in income distribution is a good proxy for parental human capital.
- Therefore, conditional on that rank, income variations can be taken to be orthogonal to other determinants of schooling.
- This income gap has widened over the past 20 years and differentially across states.

Evidence (continued)

- Therefore, one can estimate regressions of the form

$$s_{iqjt} = \delta_q + \delta_j + \delta_t + \beta_q \ln y_{iqjt} + \varepsilon_{iqjt}, \quad (17)$$

  where q denotes income quartile, j denotes region, and t denotes time.
- s_{iqjt} is education of individual i in income quartile q region j time t.
- With no effect of income on education, β_q's should be zero.
- With credit constraints, we might expect lower quartiles to have positive β's.
- Acemoglu and Pischke report versions of this equation using data aggregated to income quartile, region and time cells.
- The estimates of β are typically positive and significant.
- However, the evidence does not indicate that the β's are higher for lower income quartiles.
- Why might this be?

Returns to Experience

- Workers with greater labor market experience are paid more.
- Why is this?
  - Learning by "aging"?
  - Upward sloping incentive contracts?
  - Better matches?
  - Continued investments?
- The baseline Ben-Porath:
  - continued investments.
  - human capital investments and non-trivial labor supply decisions throughout the lifetime of the individual.

The Ben-Porath Model

- Starting point for models of investment in skills on the job.
- Let $s(t) \in [0,1]$ for all $t \geq 0$.
- Suppose

$$\dot{h}(t) = \phi(s(t)h(t)) - \delta_h h(t), \tag{18}$$

- Here $\delta_h > 0$ captures "depreciation of human capital". Why?
- The individual starts with an initial value of human capital $h(0) > 0$.
- The function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ is strictly increasing, continuously differentiable and strictly concave.
- Let us also impose the the following Inada-type conditions (why?)

$$\lim_{x \to 0} \phi'(x) = \infty \text{ and } \lim_{x \to h(0)} \phi'(x) = 0.$$

- Finally, suppose that $\omega(t) = 0$ for all t, that $T = \infty$, that there is a flow rate of death $\nu > 0$, and that $w(t) = 1$ for all t.

The Ben-Porath Model (continued)

- Again using Theorem 1, maximization problem:

$$\max \int_0^\infty \exp\left(-\left(r+\nu\right)\right)\left(1-s\left(t\right)\right)h\left(t\right)dt$$

subject to (18).

- This problem can be solved by setting up the current-value Hamiltonian, which in this case takes the form

$$\mathcal{H}\left(h,s,\mu\right)=\left(1-s\left(t\right)\right)h\left(t\right)+\mu\left(t\right)\left(\phi\left(s\left(t\right)h\left(t\right)\right)-\delta_h h\left(t\right)\right).$$

- Necessary conditions:

$$\begin{aligned}\mathcal{H}_s &= -h\left(t\right)+\mu\left(t\right)h\left(t\right)\phi'\left(s\left(t\right)h\left(t\right)\right)=0\\ \mathcal{H}_h &= \left(1-s\left(t\right)\right)+\mu\left(t\right)\left(s\left(t\right)\phi'\left(s\left(t\right)h\left(t\right)\right)-\delta_h\right)\\ &= \left(r+\nu\right)\mu\left(t\right)-\dot{\mu}\left(t\right)\\ &\lim_{t\to\infty}\exp\left(-\left(r+\nu\right)t\right)\mu\left(t\right)h\left(t\right)=0.\end{aligned}$$

The Ben-Porath Model (continued)

- To solve for the optimal path of human capital investments, let us adopt the following transformation of variables:

$$x(t) \equiv s(t)\, h(t).$$

- Instead of $s(t)$ (or $\mu(t)$) and $h(t)$, we will study the dynamics of the optimal path in $x(t)$ and $h(t)$.
- Therefore:

$$1 = \mu(t)\, \phi'(x(t)), \tag{19}$$

and

$$\frac{\dot{\mu}(t)}{\mu(t)} = r + \nu + \delta_h - s(t)\, \phi'(x(t)) - \frac{1 - s(t)}{\mu(t)}.$$

- Substituting for $\mu(t)$ from (19):

$$\frac{\dot{\mu}(t)}{\mu(t)} = r + \nu + \delta_h - \phi'(x(t)). \tag{20}$$

The Ben-Porath Model (continued)

- The steady-state (stationary) solution: $\dot{\mu}(t) = 0$ and $\dot{h}(t) = 0 \Longrightarrow$

$$x^* = \phi'^{-1}(r + \nu + \delta_h), \tag{21}$$

- Therefore $x^* \equiv s^* h^*$ will be higher when the interest rate is low, when the life expectancy of the individual is high, and when the rate of depreciation of human capital is low.
- To determine s^* and h^* separately, we set $\dot{h}(t) = 0$ in the human capital accumulation equation (18):

$$\begin{aligned} h^* &= \frac{\phi(x^*)}{\delta_h} \\ &= \frac{\phi\left(\phi'^{-1}(r + \nu + \delta_h)\right)}{\delta_h}. \end{aligned} \tag{22}$$

- Since $\phi'^{-1}(\cdot)$ is strictly decreasing and $\phi(\cdot)$ is strictly increasing, this equation implies that the steady-state solution for the human capital stock is uniquely determined and is decreasing in r, ν and δ_h.

The Ben-Porath Model: Dynamics

- More interesting than the stationary (steady-state) solution are the dynamics.
- Differentiate (19) with respect to time to obtain

$$\frac{\dot{\mu}(t)}{\mu(t)} = \varepsilon_{\phi'}(x)\frac{\dot{x}(t)}{x(t)},$$

where

$$\varepsilon_{\phi'}(x) = -\frac{x\phi''(x)}{\phi'(x)} > 0$$

- Combining this equation with (20), we obtain

$$\frac{\dot{x}(t)}{x(t)} = \frac{1}{\varepsilon_{\phi'}(x(t))}\left(r + \nu + \delta_h - \phi'(x(t))\right). \tag{23}$$

- Together with the $\dot{h}$ equation, two differential equations in two variables, x and h.

The Ben-Porath Model: Dynamics

- Solution:

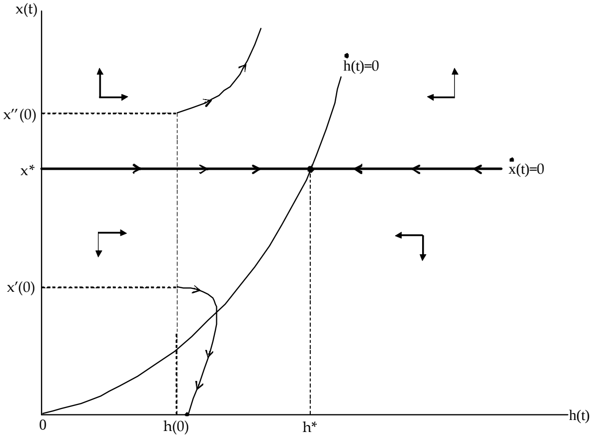
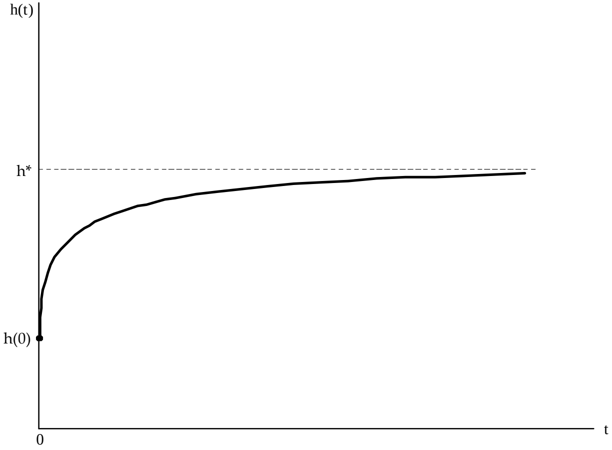


Figure:

The Ben-Porath Model: Dynamics (continued)

- Time path of human capital:



The Ben-Porath Model: Dynamics (continued)

- Smooth and monotonic.
- In the original Ben-Porath model, which involves the use of other inputs in the production of human capital and finite horizons, the constraint for $s(t) \leq 1$ typically binds early on in the life of the individual, and the interval during which $s(t) = 1$ can be interpreted as full-time schooling.
- After full-time schooling, the individual starts working (i.e., $s(t) < 1$).
- But even on-the-job, the individual continues to accumulate human capital (i.e., $s(t) > 0$), which can be interpreted as spending time in training programs or allocating some of his time on the job to learning rather than production.

The Ben-Porath Model: Empirical Implications

- This model also provides us with a useful way of thinking of the lifecycle of the individual, which starts with higher investments in schooling.
- Then there is a period of "full-time" work (where $s(t)$ is high), but this is still accompanied by investment in human capital and thus increasing earnings.
- The increase in earnings takes place at a slower rate as the individual ages.
- Earnings may also start falling at the very end of workers' careers, though this does not happen in the version presented here (how would you modify it to make sure that earnings may fall in equilibrium?).
- The available evidence is consistent with the broad patterns suggested by the model.
- But, this evidence comes from cross-sectional age-experience profiles, so caution (why?).

The Ben-Porath Model: Empirical Implications (continued)

- Perhaps more worrisome for interpretation: the increase in earnings may reflect not the accumulation of human capital due to investment, but either:

1. simple age effects; individuals become more productive as they get older. Or
2. simple experience effects: individuals become more productive as they get more experienced—this is independent of whether they choose to invest or not.

- Difficult to distinguish between the Ben-Porath model and the second explanation. But there is some evidence that could be useful to distinguish between age effects vs. experience effects (automatic or due to investment).

The Ben-Porath Model: Empirical Implications (continued)

- Josh Angrist's paper on Vietnam veterans: workers who served in the Vietnam War lost the experience premium associated with the years they served in the war.
- Presuming that serving in the war has no productivity effects, this evidence suggests that much of the age-earnings profiles are due to experience not simply due to age.
- But still consistent both with direct experience effects on worker productivity, and also a Ben Porath type explanation where workers are purposefully investing in their human capital while working, and experience is proxying for these investments.
- Potential area for future work...
- How would one distinguish between different approaches?

The Ben-Porath Model: Conclusions

- Why is the Ben-Porath model useful?
  - Useful framework for thinking about age-earnings profiles.
  - New questions on the table related to the effect of age, experience and learning on wages.
  - Emphasis on the fact that schooling is not the only way in which individuals can invest in human capital and there is a continuity between schooling investments and other investments in human capital.
  - Implication: in societies where schooling investments are high we may also expect higher levels of on-the-job investments in human capital$\longrightarrow$ possibility of systematic mismeasurement of the amount or the quality human capital across societies.