Chapter 9

Human capital, learning technology, and the Mincer equation

We start with an overview of different approaches to the modeling of human capital formation in macroeconomics. Next we go into detail with one of these approaches, the life-cycle approach. In Section 9.3 a simple model of the choice of schooling length is considered. Finally, Section 9.4 presents the theory behind the empirical relationship named the Mincer equation.[1] In this connection it is emphasized that the Mincer equation should be seen as an equilibrium relationship for relative wages at a given point in time rather than as a production function for human capital.

9.1 Macroeconomic approaches to human capital

We define *human capital* as the stock of productive skills embodied in an individual. Human capital is thus a *production factor*, while by *human wealth* is meant the *present value* of expected future labor income (usually after tax).

Increases in the stock of human capital occurs through formal education and on-the-job-training. By contributing to the maintenance of life and well-being, also health care is of importance for the stock of human capital and the incentive to invest in human capital.

Since human capital is embodied in individuals and can only be used one place at a time, it is a *rival* and *excludable* good. Human capital is thus

[1] After Mincer (1958, 1974).

very different from technical knowledge. We think of *technical knowledge* as a list of instructions about how different inputs can be combined to produce a certain output. A principle of chemical engineering is an example of a piece of technical knowledge. In contrast to human capital, technical knowledge is a *non-rival* and only partially excludable good. Competence in *applying* technical knowledge is one of the skills that to a larger or smaller extent is part of human capital.

9.1.1 Modelling human capital

In the macroeconomic literature there are different theoretical approaches to the modelling of human capital. Broadly speaking we may distinguish these approaches along two "dimensions": 1) What characteristics of human capital are emphasized? 2) What characteristics of the decision maker investing in human capital are emphasized? Combining these two "dimensions", we get Table 1.

Table 1. Macroeconomic approaches to the modelling of human capital.

The character of the decision maker	*The character of human capital (hc):* Is hc treated as essentially different from physical capital? No	Yes
Solow-type rule-of-thumb households	Mankiw et al. (1992)	
Infinitely-lived family "dynasties" (the representative agent approach)	Barro&Sala-i-Martin (2004) Dalgaard&Kreiner (2001)	Lucas (1988)
Finitely-lived individuals going through a life cycle (the life cycle approach)		Ben-Porath (1967) Heijdra&Romp (2009)

My personal opinion is that for most issues the approach in the lower-right corner of Table 1 is preferable, that is, the approach treating human capital as a distinct capital good in a life cycle perspective. The viewpoint is:

First, by being *embodied* in a person and being lost upon death of this person, human capital is very different from physical capital. In addition, investment in human capital is *irreversible* (can not be recovered). Human capital is also distinct in view of the limited extend to which it can be used as a collateral, at least in non-slave societies. Financing an investment in physical capital, a house for example, by credit is comparatively easy because the house can serve as a collateral. A creditor can not gain title to a person, however. At most a creditor can gain title to a part of that person's future earnings in excess of a certain level required for a "normal" or "minimum" standard of living.

Second, educational investment is closely related to life expectancy and the life cycle of human beings: school - work - retirement. So a life cycle perspective seems the natural approach. Fortunately, convenient macroeconomic frameworks incorporating life cycle aspects exist in the form of overlapping generations models (for example Diamond's OLG model or Blanchard's continuous time OLG model).

9.1.2 Human capital and the efficiency of labor

Generally we tend to think of human capital as a combination of different skills. Macroeconomics, however, often tries (justified or not) to boil down the notion of human capital to a one-dimensional entity. So let us imagine that the current stock of human capital in society is measured by the one-dimensional index H. With L denoting the size of the labor force, we define $h \equiv H/L$. So, h is the average stock of human capital in the labor force. Further, let the "quality" (or "efficiency") of this stock in production be denoted q (under certain conditions this quality might be proxied by the average real wage per man-hour). Then it is reasonable to link q and h by some increasing *quality function*

$$q = q(h), \qquad \text{where } q(0) \geq 0, q' > 0. \tag{9.1}$$

Consider an aggregate production function, $\tilde{F}$, giving output per time unit at time t as

$$Y = \tilde{F}(K, q(h)L, t), \qquad \frac{\partial \tilde{F}}{\partial t} > 0, \tag{9.2}$$

where K is input of physical capital. The third argument of $\tilde{F}$ is time, t, indicating that the production function is time-dependent due to technical progress.

Generally the macroeconomic analyst would prefer a measure of human capital such that the quality of human capital is proportional to the stock of human capital, allowing us to write $q(h) = h$ by normalizing the factor of proportionality to be 1. The main reason is that an expedient variable representing human capital in a model requires that the analyst can decompose the real wage per working hour multiplicatively into two factors, the real wage per unit of human capital per working hour and the stock of human capital, h. That is, an expedient human capital concept requires that we can write

$$w = \hat{w} \cdot h, \tag{9.3}$$

where $\hat{w}$ is the real wage per unit of human capital per working hour. Indeed, if we have

$$Y = \tilde{F}(K, hL, t), \tag{9.4}$$

then, under perfect competition, we can write

$$w = \frac{\partial Y}{\partial L} = \tilde{F}_2(K, hL, t)h = \hat{w} \cdot h.$$

Under disembodied Harrod-neutral technical progress, (9.4) would take the form

$$Y = \tilde{F}(K, hL, t) = F(K, AhL) \equiv F(K, EL), \tag{9.5}$$

where $E \equiv A \cdot h$ is the "effective" labor input. The proportionality between E and h will under perfect competition allow us to write

$$w = \frac{\partial Y}{\partial L} = \tilde{F}_2(K, EL, t)E = w_E \cdot E = w_E \cdot A \cdot h = \hat{w} \cdot h.$$

So with the introduction of the technology level, A, an additional decomposition, $\hat{w} = w_E \cdot A$ comes in, while the original decomposition in (9.3) remains valid.

Whether or not the desired proportionality $q(h) = h$ can be obtained depends on how we model the formation of the "stuff" h. Empirically it turns out that treating the formation of human capital as similar to that of physical capital does *not* lead to the desired proportionality.

Treating the formation of human capital as similar to formation of physical capital

Consider a model where human capital is formed in a way similar to physical capital. The Mankiw-Romer-Weil (1992) extension of the Solow growth model with human capital is a case in point. Non-consumed aggregate output is split into one part generating additional physical capital one-to-one, while the other part is assumed to generate additional human capital one-to-one. Then for a closed economy in continuous time we can write:

$$\begin{aligned} Y &= C + I_K + I_H, \\ \dot{K} &= I_K - \delta_K K, \qquad \delta_K > 0, \\ \dot{H} &= I_H - \delta_H H, \qquad \delta_H > 0, \end{aligned} \tag{9.6}$$

where I_K and I_H denote gross investment in physical and human capital, respectively. This approach essentially assumes that human capital is produced by the same technology as consumption and investment goods.

Suppose the huge practical measurement problems concerning I_H have been somehow overcome. Then from long time series for I_H an index for H_t can be constructed by the *perpetual inventory method* in a way similar to the way an index for K_t is constructed from long time series for I_K. Indeed, in discrete time, with $0 < \delta_H < 1$, we get, by backward substitution,

$$\begin{aligned} H_{t+1} &= I_{H,t} + (1-\delta_H)H_t = I_{H,t} + (1-\delta_H)\left[I_{H,t-1} + (1-\delta_H)H_{t-1}\right] \\ &= \sum_{i=0}^{T}(1-\delta_H)^i I_{H,t-i} + (1-\delta_H)^{T+1}H_{t-T}. \end{aligned} \tag{9.7}$$

From the time series for I_H, an estimate of δ_H, and a rough conjecture about the initial value, H_{t-T}, we can calculate H_{t+1}. The result will not be very sensitive to the conjectured value of H_{t-T} since for large T the last term in (9.7) becomes very small.

In principle there need not be anything wrong with this approach. A snag arises, however, if, without further notice, the approach is combined with an explicit or implicit postulate that $q(h)$ is proportional to the "stuff", h, brought into being in the way described by (9.6). The snag is that the empirical evidence does *not* support this when the formation of human capital is modelled as in (9.6). This is an unintended by-product of the cross-country regression analysis by Mankiw, Romer, and Weil (1992), based on the approach in equation (9.6). One of their conclusions is that the following production function for a country's GDP is an acceptable approximation:

$$Y = BK^{1/3}H^{1/3}L^{1/3}, \tag{9.8}$$

where B stands for the total factor productivity of the country and is generally growing over time.[2] Applying that $H = hL$, we can write (9.8) this way:

$$Y = BK^{1/3}(hL)^{1/3}L^{1/3} = K^{1/3}(Ah^{1/2}L)^{2/3},$$

where $A = B^{3/2}$. That is, we end up with the form $Y = F(K, Aq(h)L)$ where $q(h) = h^{1/2}$, not $q(h) = h$. We should thus not expect the real wage to rise in proportion to h, when h is considered as some "stuff" formed in a way similar to the way physical capital is formed. (A further point is that writing a production function as in (9.8), i.e., with H and L as two separate inputs, may lead to confusion. The tangible input is L, and in this L, a certain

[2] The way Mankiw-Romer-Weil measure I_H is indirect and questionable. In addition, the way they let their measure enter the regression equation has been criticized for confounding the effects of the human capital *stock* and human capital *investment*, cf. Gemmel (1996) and Sianesi and Van Reenen (2003). It will take us too far to go into detail with these problems here.

"normal" or average h is *embodied.* In effect, varying L should immediately also imply variation of $H \equiv hL$.)

Before proceeding, a terminological point is in place. Why do we call $q(h)$ in (9.2) a "quality" function rather than simply a "productivity" function? The reason is the following. With perfect competition and *CRS*, in equilibrium the real wage per man-hour would be $w = \partial Y/\partial L = F_2'(K, Aq(h)L)Aq(h)$ $= \left[f(\tilde{k}) - \tilde{k}f'(\tilde{k})\right] Aq(h)$, where $\tilde{k} \equiv K/(Aq(h)L)$. So, with a converging $\tilde{k}$, the long-run growth rate of the real wage would in continuous time tend to be

$$g_w = g_A + g_q.$$

In this context we are inclined to identify "labor productivity" with $Aq(h)$ rather than just $q(h)$ and "growth in labor productivity" with $g_A + g_q$ rather than just g_q. So a distinct name for q seems appropriate and an often used name is "quality".

The conclusion so far is that specifying human capital formation as in (9.6) does not generally lead to a linear quality function. To obtain the desired linearity we have to specify the formation of human capital in a way different from the equation (9.6). This dissociation with the approach (9.6) applies, of course, also to its equivalent form on a per capita basis,

$$\dot{h} = (\frac{\dot{H}}{H} - n)h = \frac{I_H}{L} - (\delta_H + n)h. \tag{9.9}$$

(In the derivation of (9.9) we have first calculated the growth rate of $h \equiv H/L$, then inserted (9.6), and finally multiplied through by h.)

9.2 A life-cycle perspective on human capital

In the life-cycle approach to human capital formation we perceive h as the human capital embodied in a single individual and lost upon death of this individual. We study how h evolves over the lifetime of the individual as a result of both educational investment (say time spent in school) and work experience. In this way the life-cycle approach recognizes that human capital is different from physical capital. By seeing human capital formation as the result of individual learning, the life-cycle approach opens up for distinguishing between the production technologies for human and physical capital. Thereby the life-cycle approach offers a better chance for obtaining the linear relationship, $q(h) = h$.

Let the human capital of an individual of "age" τ (beyond childhood) be denoted h_τ. Let the total time available per time unit for study, work, and leisure be normalized to 1. Let s_τ denote the fraction of time the individual spends in school at age τ. This allows the individual to go to school only part-time and spend the remainder of non-leisure time working. If ℓ_τ denotes the fraction of time spent at work, we have

$$0 \leq s_\tau + \ell_\tau \leq 1.$$

The fraction of time used as leisure (or child rearing, say) at age τ is $1 - s_\tau - \ell_\tau$. If full retirement occurs at age $\bar{\tau}$, we have $s_\tau = \ell_\tau = 0$ for $\tau \geq \bar{\tau}$.

We measure age in the same time units as calendar time. It seems natural to assume that the increase in h_τ per unit of time (age) generally depends on four variables: current time in school, current time at work (resulting in work experience), human capital already obtained, and current calendar time itself, that is,

$$\dot{h}_\tau \equiv \frac{dh_\tau}{d\tau} = G(s_\tau, \ell_\tau, h_\tau, t), \qquad h_0 \ \geq 0 \text{ given.} \tag{9.10}$$

The function G can be seen as a production function for human capital − in brief a *learning technology.* The first argument of G reflects the role of formal education. Empirically, the primary input in formal education is the time spent by the students studying; this time is not used in work or leisure and it thereby gives rise to an opportunity cost of studying.[3] The second argument of G takes learning through work experience into account and the third argument allows for the already obtained level of human capital to affect the strength of the influence from given s_τ and ℓ_τ (the sign of this effect is theoretically ambiguous). Finally, the fourth argument, current calendar time allows for changes over time in the learning technology (organization of the learning process).

Consider an individual "born" (as a youngster) at date $v \leq t$ (v for vintage). If still alive at time t, the age of this individual is $\tau \equiv t - v$. The obtained stock of human capital at age τ will be

$$h_\tau = h_0 + \int_0^\tau G(s_x, \ell_x, h_x, v + x)dx.$$

A basic supposition in the life-cycle approach is that it is possible to specify the function G such that a person's time-t human capital embodies a time-t labor productivity proportional to this amount of human capital and thereby, under perfect competition, a real wage proportional to this human capital.

[3] We may perceive the costs associated with teachers' time and educational buildings and equipment as being either quantitatively negligible or implicit in the function symbol G.

Below we consider four specifications of the learning technology that one may encounter in the literature.

EXAMPLE 1 In a path-breaking model by the Israeli economist Ben-Porath (1967) the learning technology is specified this way:

$$\dot{h}_\tau = g(s_\tau h_\tau) - \delta h_\tau, \qquad g' > 0, g'' < 0,\ \delta > 0, \qquad h_0 > 0. \tag{9.11}$$

Here time spent in school is more efficient in building human capital the more human capital the individual has already. Work experience does not add to human capital formation. The parameter δ enters to reflect obsolescence (due to technical change) of skills learnt in school. $\square$

EXAMPLE 2 Growiec (2010) and Growiec and Groth (2015) consider the aggregate implications of a learning technology specified this way:

$$\dot{h}_\tau = (\lambda s_\tau + \xi \ell_\tau) h_\tau, \qquad \lambda > 0, \xi \geq 0, \qquad h_0 > 0. \tag{9.12}$$

Here λ measures the efficiency of schooling and ξ the efficiency of work experience. The effects of schooling and (if $\xi > 0$) work experience are here assumed proportional to the level of human capital already obtained by the individual (a strong assumption which may be questioned).[4] The linear differential equation (9.12) allows an explicit solution,

$$h_\tau = h_0 e^{\int_0^\tau (\lambda s_x + \xi \ell_x) dx}, \tag{9.13}$$

a formula valid as long as the person is alive. This result has some affinity with the familiar "Mincer equation", to be considered below.[5] $\square$

EXAMPLE 3 Here we consider an individual with exogenous and constant leisure. Hence time available for study and work is constant and conveniently normalized to 1 (as if there were no leisure at all). Moreover, in the beginning of life beyond childhood the individual goes to school full-time in S time units (years) and thereafter works full-time until death (no retirement). Thus

$$s_\tau = \begin{cases} 1 & \text{for } 0 \leq \tau < S, \\ 0 & \text{for } \tau \geq S. \end{cases} \tag{9.14}$$

We further simplify by ignoring the effect of work experience (or we may say that work experience just offsets obsolescence of skills learnt in school). The learning technology is specified as

$$\dot{h}_\tau = \eta \tau^{\eta - 1} s_\tau, \qquad \eta > 0, \quad h_0 \geq 0, \tag{9.15}$$

[4] Lucas (1988) builds on the case $\xi = 0$.

[5] In case $\xi = 0$ and $s_x = \text{constant} = 1$, while $\tau = S =$ schooling length, (9.13) reduces to $h = h_0 e^{\lambda S}$. This looks like a simple version of the "Mincer equation".

If $\eta < 1$, it becomes more difficult to learn more the longer you have already been to school. If $\eta > 1$, it becomes easier to learn more the longer you have already been under education.

The specification (9.14) implies that throughout working life the individual has constant human capital equal to $h_0 + S^\eta$. Indeed, integrating (9.15), we have for $t \geq S$ and until time of death,

$$h_\tau = h_0 + \int_0^\tau \dot{h}_x dx = h_0 + \int_0^S \eta x^{\eta-1} dx = h_0 + x^\eta|_0^S = h_0 + S^\eta. \qquad (9.16)$$

So the parameter η measures the elasticity of human capital w.r.t. the number of years in school. As briefly commented on in the concluding section, there is some empirical support for the power function specification in (9.16) and even the hypothesis $\eta = 1$ may not be rejected. □

In Example 1 there is no explicit solution for the level of human capital. Then the solution can be characterized by phase diagram analysis (as in Acemoglu, §10.3). In the examples 2 and 3 we can find an explicit solution for the level of human capital. In this case the term "learning technology" is used not only in connection with the original differential form as in (9.10), but also for the integrated form, as in (9.13) and (9.16), respectively. Sometimes the integrated form, like (9.16), is called a *schooling technology*.

EXAMPLE 4 Here we still assume the setup in (9.14) of Example 3, including the absence of both after-school learning and gradual depreciation. But the right-hand side of (9.15) is generalized to $\varphi(\tau)s_\tau$, where $\varphi(\tau)$ is some positively valued function of age. Then we end up with human capital after leaving school equal to some increasing function of S :

$$h = h(S), \qquad \text{where } h(0) \geq 0,\ h' > 0. \qquad (9.17)$$

In cross-section or time series analysis it may be relevant to extend this by writing $h = ah(S)$, $a > 0$. The parameter a could then reflect quality of schooling. In the next section we shall focus on the form (9.17) where the quality-of-schooling parameter a can be seen as implicit in the function h. □

Before proceeding, let us briefly comment on the problem of aggregation over the different members of the labor force at a given point in time. In the aggregate framework of Section 9.1 multiplicity of skill types and job types is ignored. Human capital is treated as a one-dimensional and additive production factor. In production functions like (9.4) only aggregate human capital, H, matters. So output is thought to be the same whether the input is 2 million workers, each with one unit of human capital, or 1 million workers, each with 2 units of human capital. In human capital theory this questionable assumption is called the *perfect substitutability assumption* or the *efficiency unit assumption* (Sattinger, 1980). If we are willing to impose this assumption, going from micro to macro is conceptually simple. With h denoting individual human capital and $f(h)$ being the density function at a given point in time (so that $\int_0^\infty f(h)dh = 1$), we find average human capital in the labor force at that point in time to be $\bar{h} = \int_0^\infty hf(h)dh$ and aggregate human capital as $H = \bar{h}L$, where L is the size of the labor force. To build a theory of the evolution over time of the density function, $f(h)$, is, however, a complicated matter. Within as well as across the different cohorts there is heterogeneity regarding both schooling and retirement. And the fertility and mortality patterns are changing over time.

If we want to open up for a distinction between different types of jobs and different types of labor, say, skilled and unskilled labor, we may replace the production function (9.4) with

$$Y = \tilde{F}(K, h_1 L_1, h_2 L_2, t), \tag{9.18}$$

where L_1 and L_2 indicate man-hours delivered by the two types of workers, respectively, and h_1 and h_2 are the given embodied human capital levels (measured in efficiency units for each of the two kinds of jobs), respectively. This could be the basis for studying skill-biased technical change.

Whether or not the aggregate human capital, H, is a useful concept or not in connection with production can be seen as a question about whether or not we can rewrite a production function like (9.18) as $Y = F(K, H, t)$, where $H = h_1 L_1 + h_2 L_2$. We *can* if the two types of labor are *perfectly substitutable*, otherwise not. *Perfect substitutability* in this context means that the marginal rate of substitution between the two kinds of labor in (9.18) is a constant, i.e.,

$$MRS \equiv -\frac{dL_1}{dL_2}_{\big| Y=\bar{Y}, K=\bar{K}} = \frac{\partial Y / \partial L_2}{\partial Y / \partial L_1} = \text{ a constant.} \tag{9.19}$$

This is satisfied if we can rewrite the production function such that $Y = F(K, H, t)$, where $H = h_1 L_1 + h_2 L_2$. Indeed, in this case we get $MRS = F_H h_2 / (F_H h_1) = h_2 / h_1$, a constant.

9.3 Choosing length of education

First some simplifying demographic assumptions. We assume, realistically, that expected lifetime of an individual is finite while the age at death is stochastic (uncertain) ex ante. We further assume, unrealistically, that *independently* of the already obtained age, the probability of surviving x more time units (years) is

$$P(X > x) = e^{-mx},$$

where X is remaining lifetime, a stochastic variable, while $m > 0$ is the *mortality rate* which is thus taken to be independent of age (and also independent of calendar time). Under this assumption, the "crude death rate", that is, the number of deaths per year divided by the size of population at the beginning of the year, will be approximately equal to m. Moreover, the mortality rate, m, will for an arbitrary person indicate the approximate probability of dying within one year "from now".[6]

Consider an individual's educational planning as seen from time of "birth" (entering life beyond childhood). Let the time of birth be denoted v. Suppose schooling is a full-time activity and that the individual plans to attend school in the first S years of life and after that work "full time" until death ("no retirement"). Let $\ell_{t-v}(S)$ denote the planned supply of labor (hours per year) to the labor market at age $t - v$ in the future. As $\ell_{t-v}(S)$ depends on the *stochastic* age, T, at death, $\ell_{t-v}(S)$ is itself a stochastic variable with two possible outcomes:

$$\ell_{t-v}(S) = \begin{cases} 0 & \text{when } t \leq v + S \text{ or } t > v + T, \\ \ell & \text{when } v + S < t \leq v + T, \end{cases}$$

where $\ell > 0$ is an exogenous constant ("full-time" working).

The combination of age-independent mortality rate and no retirement is sometimes called the "perpetual youth" assumption.

9.3.1 Human wealth

Let $w_t(S)$ denote the real wage received *per working hour* delivered at time t by a person who after S years in school works ℓ hours per year until death. This allows us to write the present value as seen from time v of expected lifetime earnings, i.e., the human wealth, for a person "born" at time v as

[6] If T denotes the uncertain age at death (a stochastic variable), the mortality rate (or "hazard rate" of death) at the age τ, denoted $m(\tau)$, is defined as $m(\tau) = \lim_{\Delta\tau \to 0} \frac{1}{\Delta\tau} P\left(T \leq \tau + \Delta\tau | \, T > \tau\right)$.

In the present model this is assumed equal to a constant, m. The unconditional probability of not reaching age τ is $P(T \leq \tau) = 1 - e^{-m\tau} \equiv F(\tau)$. Hence the density function is $f(\tau) = F'(\tau) = me^{-m\tau}$ and $P(\tau < T \leq \tau + \Delta\tau) \approx me^{-m\tau}\Delta\tau$. So, for $\tau = 0$, $P(0 < T \leq \Delta\tau) \approx m\Delta\tau = m$ if $\Delta\tau = 1$. Life expectancy is $E(T) = \int_0^\infty \tau me^{-m\tau} d\tau = 1/m$. All this is like in the "perpetual-youth" overlapping generations model by Blanchard (1985).

$$
\begin{aligned}
HW(v,S) &= \\
&\quad 0 + E_v\left(\int_{v+S}^{v+T} w_t(S)\ell e^{-r(t-v)}dt\right) \\
&= E_v\left(\int_{v+S}^{\infty} w_t(S)\ell_{t-v}(S)e^{-r(t-v)}dt\right) \\
&= \int_{v+S}^{\infty} E_v(w_t(S)\ell_{t-v}(S)e^{-r(t-v)})dt = \int_{v+S}^{\infty} w_t(S)e^{-r(t-v)}E_v(\ell_{t-v}(S))dt, \qquad (9.20)
\end{aligned}
$$

as in this context the integration operator $\int_{v+S}^{\infty}(\cdot)dt$ acts like a discrete-time summation operator, $\sum_{t=v}^{\infty}$. The rate of discount for potential future labor income conditional on being alive at the moment concerned is denoted r.[7] We get

$$
\begin{aligned}
HW(v,S) &= \int_{v+S}^{\infty} w_t(S)e^{-r(t-v)}\left(\ell \cdot P(T > t-v) + 0 \cdot P(T \leq t-v)\right)dt \\
&= \int_{v+S}^{\infty} w_t(S)e^{-r(t-v)}\ell e^{-m(t-v)}dt \\
&= \int_{v+S}^{\infty} w_t(S)\ell e^{-(r+m)(t-v)}dt. \qquad (9.21)
\end{aligned}
$$

In writing the present value of the expected stream of labor income this way, we have assumed that:

A1 The discount rate, r, is constant over time.

A2 There is no educational fee.

We now introduce two additional assumptions:

A3 Labor efficiency (human capital) of a person with S years of schooling is $h(S)$, so that

$$w_t(S) = \hat{w}_t h(S), \qquad h' > 0,$$

where $\hat{w}_t$ is the real wage *per unit of human capital* per working hour at time t.[8]

[7] This rate is related to the opportunity cost of going to school instead of working and depends on conditions in the credit market. Under the idealized assumption A5 below, $r =$ the risk-free interest rate.

[8] Cf. Example 4 of Section 9.2.

A4 Owing to Harrod-neutral technical progress at a constant rate $g \in [0, r+m) \geq 0$, the evolution of $\hat{w}_t$ is given by $\hat{w}_t = \hat{w}_0 e^{gt}$. So technical progress makes a given h more and more productive (there is complementarity between the technology level and human capital as in (9.5) above).

Given A3 and A4, we get from (9.21) the expected "lifetime earnings" conditional on a schooling level S :

$$\begin{aligned} HW(v,S) &= h(S)\ell \int_{v+S}^{\infty} \hat{w}_0 e^{gt} e^{-(r+m)(t-v)} dt \qquad (9.22) \\ &= \hat{w}_0 e^{gv} h(S)\ell \int_{v+S}^{\infty} e^{[g-(r+m)](t-v)} dt \quad (\text{since } e^{gt} = e^{gv} e^{g(t-v)}) \\ &= \hat{w}_0 e^{gv} h(S)\ell \left(\left. \frac{e^{[g-(r+m)](t-\nu)}}{g-(r+m)} \right|_{\nu+S}^{\infty} \right) = \hat{w}_0 e^{g\nu} h(S)\ell \frac{e^{[g-(r+m)]S}}{r+m-g}. \end{aligned}$$

Below we chose measurement units such that the "normal" working time per year is 1 rather than ℓ.

The result in (9.22) provides a convenient formula for human wealth as seen from time of "birth", v. To say something reasonable about the *choice* of S, we need to specify the set of possibilities for the individual. These possibilities depend on the market environment. In particular, we need to specify how students make a living while studying.

9.3.2 Financing education

Assuming the students are born with no financial wealth and themselves have to finance their costs of living, they have to borrow while studying. Later in life, when they receive an income, they repay the loans with interest.

In this context we shall introduce the simplifying assumption:

A5 There is a *perfect credit and life annuity market.*

Financial intermediaries will be unwilling to offer the students loans at the going risk-free interest rate. Indeed, a creditor faces the risk that the student fails in the studies, never achieves the hoped job, or dies before having paid off the debt including the compound interest. The financial intermediaries may, however, be willing to offer student loans in the form of contracts stipulating later repayment with an interest rate above the risk-free rate and with the agreement that if the debtor dies before the principal has been paid back with interest, the debtor's estate is held free of any obligations.

Given the described constant mortality rate and given existence of a perfect credit and life insurance market, it can be shown that the equilibrium interest rate on this kind of student loans is what is known as the "actuarial rate". This rate equals the risk-free interest rate plus the mortality rate, m.[9] The relevant discount rate, r, in (9.20) will under these circumstances coincide with the risk-free interest rate. So we let this rate be denoted r and write the actuarial rate as $r + m$.

If the individual later in life, after having paid off the debt and obtained a positive net financial position, places the savings on life annuity accounts in life insurance companies, the *actuarial rate*, $r + m$, will also be the equilibrium rate of return received (until death) on these deposits. At death the liability of the insurance company is cancelled which means that the deposit is transferred to the insurance company in return for the high annuity payouts while the depositor was alive. The advantage of saving in life annuities (at least for people without a bequest motive) is that life annuities imply a transfer of income from after time of death to before time of death by offering a higher rate of return than risk-free bonds, but only until the depositor dies.

9.3.3 Maximizing human wealth

Suppose that neither the educational process itself nor the resulting stock of human capital enter the utility function. That is, assume

A6 There is no "joy of going to school" and no "joy of being a learned person".

In the perspective of this assumption, human capital is only an investment good, not also a durable consumption good.[10] If moreover there is no utility from leisure, the educational decision can be separated from whatever plan for the time path of consumption and saving through life the individual may decide; this is known as the *Separation Theorem*.[11] Under the described circumstances, the only incentive for acquiring human capital is to increase the human wealth $HW(\nu, S)$ given in (9.22).

[9] See Yaari (1965). This result presupposes that the insurance companies have negligible administration costs.

Owing to asymmetric information and related credit market imperfections, in real world situations such loan contracts are rare; this is one of the reasons for public sector intervention in the provision of loans to students. These credit market imperfections are ignored by the present model, but are briefly dealt with in for instance Acemoglu (2009), pp. 761-764.

[10] For a broader conception of human capital, see for instance Sen (1997).

[11] See, e.g., Acemoglu (2009), Ch. 10.1.

By the assumptions A1, A2, . . . , A6, we have hereby reduced the problem of choosing schooling length to the unconstrained static problem of maximizing $HW(\nu, S)$ with respect to S. An interior solution to this problem satisfies the first-order condition:

$$\begin{aligned}\frac{\partial HW}{\partial S}(v,S) &= \frac{\hat{w}_0}{r+m-g}\left[h'(S)e^{[g-(r+m)]S} - h(S)e^{[g-(r+m)]S}(r+m-g)\right] \\ &= HW(v,S)\left[\frac{h'(S)}{h(S)} - (r+m-g)\right] = 0, \qquad (9.23)\end{aligned}$$

from which follows

$$\frac{h'(S)}{h(S)} = r+m-g \equiv \tilde{r}. \qquad (9.24)$$

This may be called the *schooling first-order condition*, and $\tilde{r}$ can be seen as the "required rate of return" in units of human capital. In the optimal plan the actual rate of return in units of human capital equals $\tilde{r}$, which in turn equals the risk-free interest rate adjusted for (a) the approximate probability of dying within a year from "now", $1 - e^{-m} \approx m$; and (b) wage growth due to technical progress. The trade-off faced by the individual is the following: increasing S by one year results in a higher level of human capital (higher future earning power) but postpones by one year the time when earning an income begins. The effective interest cost (opportunity cost) is diminished by g, reflecting the fact that next year the real wage per unit of human capital is $100 \cdot g$ percent higher than in the current year.

The intuition behind the first-order condition (9.24) may be easier to grasp if we put g on the left-hand-side and multiply by $\hat{w}_t$ in the numerator as well as the denominator. Then the condition looks like a standard no-arbitrage condition:

$$\frac{\hat{w}_t h'(S) + \hat{w}_t g h(S)}{\hat{w}_t h(S)} = r+m. \qquad (9.25)$$

On the left-hand side we have the rate of return (in units of consumption) obtained by "investing" one more year in education. In the numerator we have the direct increase in wage income by increasing S by one unit plus the gain arising from the fact that human capital, $h(S)$, has higher earnings capacity one year later due to technical progress. In the denominator we have the educational investment made by letting the obtained human capital, $h(S)$, "stay" one more year in school instead of at the labor market. Indeed, $\hat{w}_t h(S)$ is the size of that investment in the sense of the opportunity cost of staying in school one more year.

On the right-hand side of (9.25) appears the rate of return, $r + m$, that could be obtained by the alternative strategy, which is to leave school already after S years and then use next years's labor income to pay off study loans. This alternative would give the rate of return $r + m$.

The first-order condition (9.24) has thus similarity with a no-arbitrage equation in financial markets. (As is usual, our interpretation treats marginal changes as if they were discrete.)

Now, suppose $S = S^* > 0$ is a unique value of S satisfying (9.24). Then a sufficient (but not necessary) condition for S^* to be the unique optimal length of education for the individual is that $h'' \leq 0$ at $S = S^*$ (see Appendix A). If individuals are alike in the sense of having the same innate abilities and facing the same schooling technology $h(\cdot)$, they will all choose S^*.

EXAMPLE 5 Suppose $h(S) = S^\eta$, $\eta > 0$, as in Example 3, but with $h_0 = 0$. Then the first-order condition (9.24) gives a unique solution $S^* = \eta/(r+m-g)$; and the second-order condition (9.32) holds for all $\eta > 0$. More sharply decreasing returns to schooling (smaller η) shortens the optimal time spent in school as does of course a higher effective discount rate, $r + m - g$.

Consider two countries, one rich (industrialized) and one poor (agricultural). With one year as the time unit, let the parameter values be as in the first four columns in the table below. The resulting optimal S for each of the countries is given in the last column.

	η	r	m	g	S^*
rich country	0.6	0.06	0.01	0.02	12.0
poor country	0.6	0.12	0.02	0.00	4.3

The difference in S^* is due to r and m being higher and g lower in the poor country. □

9.4 What the Mincer equation is and is not

In this section we consider the issue whether the *exponential* form,

$$h(S) = h(0)e^{\lambda S}, \qquad \lambda > 0, \tag{9.26}$$

is a plausible specification of the production function for human capital. This specification is quite popular in the literature, and in Acemoglu (2009) it is used in connection with "levels accounting" in his Chapter 3, pp. 96-99 and is treated also theoretically in his Chapter 10.2. The form (9.26) can be seen as a special case of equation (9.13) from Example 2 above, namely the case $\xi = 0$ in combination with equation (9.14) from Example 3.

There exists a presumption in the macroeconomic literature that the famous *Mincer equation* provides an empirical foundation for the exponential form (9.26). The Mincer equation is the following semi-loglinear *cross-sectional* relationship at a given point in time, t:

$$\log w_t(S) = \log w_t(0) + \lambda S, \qquad \lambda > 0, \tag{9.27}$$

where, as in Section 9.3, $w_t(S)$ is the real wage *per working hour* delivered at time t by a person with S years' schooling level, cf. Figure 9.1. Such a semi-loglinear relationship is well documented in the empirical literature and was first discovered by the American economist Jacob Mincer (Mincer 1958, 1974).

But does it provide evidence for any particular form for the *production function* for human capital? No! *First*, as briefly commented in the concluding section, there seems to be little empirical support for an exponential production function. *Second*, as we shall now see, the microeconomic theory, proposed by Mincer (1958) as an explanation of the observed semi-loglinear relationship (9.27), has nothing to do with a specific production function for human capital.[12]

Explaining the Mincer equation

In Mincer's theory behind the observed exponential relationship called the Mincer equation, there is no role at all for any specific schooling technology, $h(\cdot)$, leading to a unique solution, S^*. The point of departure is that there is *heterogeneity in the jobs* offered to people (different educational levels not being perfectly substitutable). Assuming people are ex ante alike, they end up ex post choosing different educational levels. This outcome arises through the competitive equilibrium forces of *supply and demand in the job markets*.

Imagine, first, a case where all individuals have in fact chosen the same educational level, S^*, because they are ex ante alike and all face the same arbitrary human capital production function, $h(S)$, satisfying (9.32). Then jobs that require other educational levels will go unfilled and so the job markets will not clear. The forces of excess demand and excess supply will then tend to generate an educational wage profile different from the one presumed in (9.22), that is, different from $\hat{w}_t h(S)$. Sooner or later an *equilibrium* educational wage profile tends to arise such that people are *indifferent* as to how much schooling they choose, thereby allowing market clearing. This requires

[12] The above Example 5 follows a short note by Jones (2007) entitled "A simple Mincerian approach to endogenizing schooling". The term "Mincerian approach" should here be interpreted in a very broad sense as more or less synonymous with "life-cycle approach" rather than be associated with a particular choice regarding the form of $h(S)$.

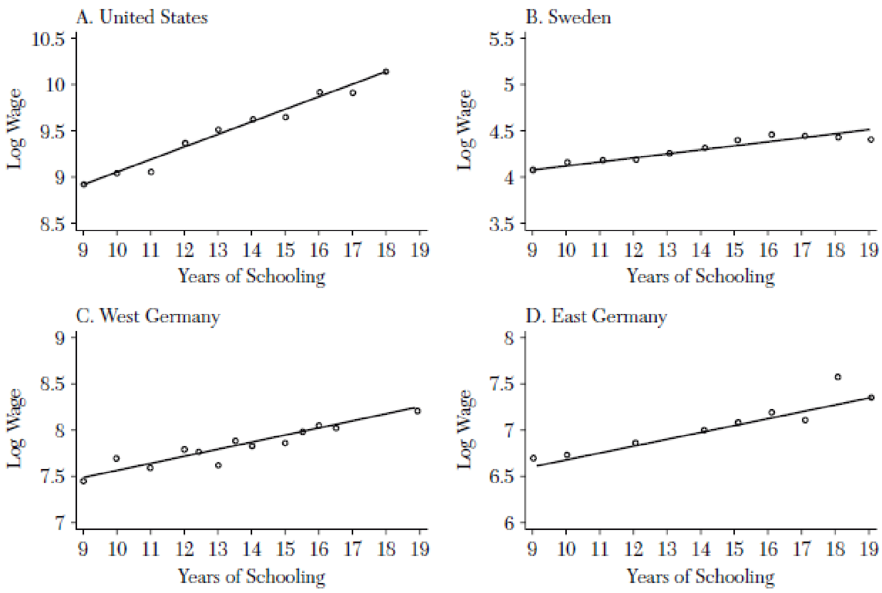

Figure 9.1: The semi-log schooling-wage relationship for fixed t. Different countries. Source: Krueger and Lindahl (2001).

a wage profile, $w_t(S)$, such that a marginal condition analogue to (9.24) holds for *all* S for which there is a positive amount of labor traded in equilibrium, say all $S \in [0, \bar{S}]$:

$$\frac{dw_t(S)/dS}{w_t(S)} = r + m - g \equiv \tilde{r} \qquad \text{for all } S \in [0, \bar{S}]\,. \tag{9.28}$$

It is here assumed, in the spirit of assumption A4 above, that technical progress implies that $w_t(S)$ *for fixed* S grows at the rate g, i.e., $w_t(S) = w_0(S)e^{gt}$, for all $S \in [0, \bar{S}]$. The equation (9.28) is a linear *differential equation* for w_t w.r.t. S, defined in the interval $0 \leq S \leq \bar{S}$, while t is fixed. And the function $w_t(S)$ is the so far unknown solution to this differential equation. That is, we have a differential equation of the form $dx(S)/dS = \tilde{r}x(S)$, where the unknown function, $x(S)$, is a function of schooling length rather than calendar time. The solution is $x(S) = x(0)e^{\tilde{r}S}$. Replacing the function $x(\cdot)$ with the function $w_t(\cdot)$, we thus have the solution

$$w_t(S) = w_t(0)e^{\tilde{r}S}. \tag{9.29}$$

Note that in the previous section, in the context of (9.24), we required the proportionate marginal return to schooling to equal $\tilde{r}$ only for a specific S, i.e.,

$$\frac{d(\hat{w}_t h(S))/dS}{\hat{w}_t h(S)} = \frac{h'(S)}{h(S)} = r + m - g \equiv \tilde{r} \qquad \text{for} \quad S = S^*. \tag{9.30}$$

This is no more than a first-order condition assumed to hold at some point, S^*. It will generally not be a differential equation the solution of which gives a Mincerian exponential relationship. A differential equation requires a derivative relationship to hold not only at one point, but in an interval for the independent variable (S in (9.28)). Indeed, in (9.28) we require the proportionate marginal return to schooling to equal $\tilde{r}$ in a whole interval of schooling levels. Otherwise, with heterogeneity in the jobs offered, there could not be equilibrium.[13]

Returning to (9.29), by taking logs on both sides and substituting $\tilde{r}$ by λ, we get (9.27), which is the *Mincer equation* in semi-loglinear form.

As mentioned, empirically, the Mincer equation does surprisingly well in cross-section regression analysis, cf. Figure 9.1.[14] Note that (9.29) also yields a theory of how the "Mincerian slope", λ, in (9.26) is *determined*, namely as

[13] It seems that Acemoglu (2009, p. 362) makes the logical error of identifying a first-order condition, (9.30), with a differential equation, (9.28).

[14] The slopes are in the interval $(0.05, 0.15)$.

the mortality- and growth-corrected real interest rate, $\tilde{r}$. The evidence for this part of the theory is more scarce.

Given the equilibrium educational wage profile, $w_t(S)$, the human wealth of an individual "born" at time 0 can be written

$$\begin{aligned} HW_0 &= \int_S^\infty w_t(0)e^{\tilde{r}S}e^{-(r+m)t}dt = e^{\tilde{r}S}\int_S^\infty w_0(0)e^{\bar{g}t}e^{-(r+m)t}dt \\ &= w_0(0)e^{\tilde{r}S}\int_S^\infty e^{[g-(r+m)]t}dt = w_0(0)e^{\tilde{r}S}\left[\frac{e^{[g-(r+m)]t}}{g-(r+m)}\right]_S^\infty \\ &= \frac{w_0(0)}{r+m-g}, \end{aligned} \tag{9.31}$$

since $\tilde{r} \equiv r+m-g$. In equilibrium the human wealth of the individual is thus *independent of* S (within an interval) according to the Mincerian theory. This is due to "compensating wage differentials", that is, the adjustment of the S-dependent wage levels so as to compensate for the S-dependent differences in length of work life after schooling. Indeed, the essence of Mincer's theory is that if one level of schooling implies a higher human wealth than the other levels of schooling, the number of individuals choosing that level of schooling will rise until the associated wage has been brought down so as to be in line with the human wealth associated with the other levels of schooling. Of course, such adjustment processes must in practice be quite time consuming and can only be approximative. Moreover, who among the ex ante similar individuals ends up with what schooling level is indeterminate in this setup.

In this context, the original schooling technology, $h(\cdot)$, for human capital formation has lost any importance. It does not enter human wealth in a long-run equilibrium in this disaggregate model where human wealth is simply given by (9.31). In this equilibrium people have different S's and the received wage of an individual per unit of work has no relationship with the human capital production function, $h(\cdot)$, by which we started in this section.

Although there thus exists a microeconomic theory behind a Mincerian relationship, this theory gives us a relationship for relative wages in a cross-section at a given point in time. It leaves open what an intertemporal production function for human capital, relating educational investment, S, to a resulting level, h, of labor efficiency, looks like. Besides, the Mincerian slope, $\tilde{r}$, is a market price, not an aspect of schooling technology.

We have up to now been silent about the fact that our simple framework in Section 9.3 does not fully embrace the case of strong convexity implied by an exponential specification of $h(S)$. Appendix B briefly comments.

9.5 Empirics relating to $h(S)$

The empirical macroeconomic literature typically measures S as the average number of years of schooling in the working-age population, taken for instance from the Barro and Lee (2001) data set.[15]

In their cross-country regression analysis de la Fuente and Domenech (2006) find a relationship essentially like that in Example 3 with $\eta = 1$. The authors find that the elasticity of GDP w.r.t. average years in school in the labor force is at least 0.60.

Similarly, the cross-country study, based on calibration, by Bills and Klenow (2000) as well as the time series study by Cervelatti and Sunde (2010) favor the hypothesis of diminishing returns to schooling. According to this, the linear term, $\tilde{r}S$, in the exponent in (9.26) should be replaced by a strictly concave function of S. These findings are in accordance with the results by Psacharopoulus (1994). They give empirical reasons for scepticism towards the linearity in h assumed in Example 2 of Section 9.2.

For $S > 0$, the power function in Example 5 can be written $h = S^{\eta} = e^{\eta \ln S}$ and is thus in better harmony with the data than the exponential function (9.26). A parameter indicating the quality of schooling may be added: $h = ae^{\eta \ln S}$, where $a > 0$ may be a function of the teacher-pupil ratio, teaching materials per student etc. See Caselli (2005).

9.6 Concluding remarks

Our formulation of the schooling length decision problem in Section 9.3 contained several simplifications such that we ended up with a static maximization problem in Section 9.3.3. More general setups lead to truly dynamic human capital accumulation problems.

This chapter considered human capital as a productivity-enhancing factor. There is a partly complementary perspective on human capital, often named the Nelson-Phelps hypothesis about the key role of human capital for technology *adoption* and technological *catching up.* An increase in human capital leads to an increase in the technology absorption capability of a nation.

A simple way of formalizing this idea is obtained by recognizing that it is not obvious that technical knowledge and human capital should enter the production function in the simple *multiplicative* way, $Y = F(K, AhL)$, as

[15] This means that complicated aggregation issues, arising from cohort heterogeneity and from the fact that individual human capital is lost upon death, are bypassed. For discussion, see Growiec and Groth (2015).

assumed in (9.5) above. The complementarity between A and h may take another form, perhaps better reflecting that workers with high skill level can use more advanced technology than workers with low skill level:

$$Y = \tilde{F}(K, hL, t) = F(K, \min(h/\eta(A), 1)AL), \qquad \eta'(A) > 0,$$

where $\eta(A)$ is the level of human capital *required* to fully exploit the current technology level, A. If actual $h < \eta(A)$, only the fraction $h/\eta(A)$ of A is utilized. Similar ideas are sketched in Jones and Vollrath (2013, Ch. 6.1) and Acemoglu (2009, Ch. 10.8). See also Exercise V.3.

Models based on the life-cycle approach to human capital typically conclude that education is productivity enhancing, i.e., more education has a positive *level* effect on income per capita but can only temporarily raise the per capita growth rate. Education is not a factor which in itself can explain sustained per capita growth. A more plausible main driving factor behind growth rather seems to be technological innovations. A higher level of per capita human capital may raise the speed of innovations, however. These themes are taken up in the next chapter (and in Exercise V.7 and V.8).

9.7 Appendix

Appendix A

Suppose $S = S^* > 0$ satisfies the first-order condition (9.24). To check the second-order condition, we consider

$$\begin{aligned}
& \frac{\partial^2 HW}{\partial S^2}(v, S^*) \\
= \; & \frac{\partial HW}{\partial S}(v, S^*)\left[\frac{h'(S^*)}{h(S^*)} - (r + m - g)\right] + HW(v, S^*)\frac{h(S^*)h''(S^*) - h'(S^*)^2}{h(S^*)^2} \\
= \; & HW(v, S^*)\frac{\frac{S^*}{h'(S^*)}h''(S^*) - \frac{S^*}{h(S^*)}h'(S^*)}{S^*h(S^*)}h'(S^*), \qquad (9.32)
\end{aligned}$$

since the first term on the right-hand side in the second row vanishes due to (9.24) being satisfied at $S = S^*$. The second-order condition, $\partial^2 HW/\partial S^2 < 0$ at $S = S^*$ holds if and only if the elasticity of h w.r.t. S exceeds that of h' w.r.t. S at $S = S^*$. A sufficient but not necessary condition for this is that $h'' \leq 0$. Anyway, since $HW(v, S)$ is a continuous function of S, if there is a unique $S^* > 0$ satisfying (9.24), and if $\partial^2 HW/\partial S^2 < 0$ holds for this S^*, then this S^* is the unique optimal length of education for the individual.

Appendix B

As alluded to at the end of Section 9.4, the strong convexity implied by the exponential specification $h(S) = h(0)e^{\lambda S}$ does not fit entirely well with the model in Section 9.3 based on the "perpetual youth" assumption of age-independent mortality and no retirement. The problem is that when $h(S) = h(0)e^{\lambda S}$, the "perpetual youth" setup implies that the first-order condition (9.24) holds for all S; moreover, we get $\partial^2 HW/\partial S^2 = 0$ for all S.

This problem reflects a limitation of the "perpetual youth" setup, where there is no conclusive upper bound for anyone's lifetime. It is not an argument for *apriori* rejection of the exponential specification.

9.8 Literature

Barro, R.J., and J. Lee, 2001, International data on educational attainment, *Oxford Economic Papers*, vol. 53 (3), 541-563.

Barro, R.J., and X. Sala-i-Martin, 2004, *Economic Growth*, 2nd edition, MIT Press.

Ben-Porath, Y., 1967, The production of human capital and the life cycle of earnings, *Journal of Political Economy* 75 (4), 352-365.

Bils, M., and P. J. Klenow, 2000. Does schooling cause growth? *American Economic Review*, 90 (5), 1160-1183.

Blanchard, O., 1985, *J. Political Economy*,

Browning, M., L.P. Hansen, and J.J. Heckman, 199, Micro data and general equilibrium models. In: *Handbook of macroecomics*, vol. , ed. by ...., 543-633.

Caselli, F., 2005, Accounting for cross-country income differences. In: *Handbook of Economic Growth*, vol. IA.

Cervellati, M., and U. Sunde, 2010, Longevity and lifetime labor supply: Evidence and implications revisited, WP.

Cohen, D., and M. Soto, 2007, Growth and human capital: good data, good results, *Journal of Economic Growth* 12, 51-76.

Cunha, F., J.J. Heckman, L. Lochner, and D.V. Masterov, 2006, Interpreting the evidence on life cycle skill formation, *Handbook of the Economics of Education*, vol. 1, Amsterdam: Elsevier.

Dalgaard, C.-J., and C.-T. Kreiner, 2001, Is declining productivity inevitable? *J. Econ. Growth,* vol. 6 (3), 187-203.

de la Fuente, A., and R. Domenech, 2006, Human capital in growth regressions: How much difference does quality data make, *Journal of the European Economic Association* 4 (1), 1-36.

Gemmell, N., 1996, , *Oxford Bulletin of Economics and Statistics* 58, 9-28.

Growiec, J., 2010, Human capital, aggregation, and growth, *Macroeconomic Dynamics* 14, 189-211.

Growiec, J., and C. Groth, 2015, On aggregating human capital across heterogeneous cohorts, *Mathematical Social Sciences*, vol. 78, 21-38.

Hall, R., and C.I. Jones, 1999, , *Quarterly Journal of Economics.*

Hanushek, E.A., and L. Woessmann, 2012, Do better schools lead to more gorwth? Cognitive skills, economic outcomes, and causation, *J. Econ. Growth*, vol. 17, 267-321.

Hazan, M., 2009, Longevity and lifetime labor supply: Evidence and implications, *Econometrica* 77 (6), 1829-1863.

Heckman, J.J., L.J. Lochner, and P.E. Todd, 2003, Fifty years of Mincer earnings regressions, NBER WP 9732.

Heijdra, B.J., and W.E. Romp, 2009, *Journal of Economic Dynamics and Control 33,* 725-744.

Hendry, D., and H. Krolzig, 2004, We ran one regression, *Oxford Bulletin of Economics and Statistics* 66 (5), 799-810.

Jones, B., 2015, The human capital stock: A generalized approach, *American Economic Review.*

Jones, C. I., 2005, *Handbook of Economic Growth,*

Jones, C. I., 2007, A simple Mincerian approach to endogenizing schooling, WP.

Krueger, A. B., and M. Lindahl, 2001. Education for growth: Why and for whom? *Journal of Economic Literature*, 39, 1101-1136.

Lucas, R.E., 1988, On the mechanics of economic development, *Journal of Monetary Economics* 22, 3-42.

Lucas, R.E., 1993, Making a miracle, *Econometrica* 61, 251-272.

Mankiw, G., 1995,

Mankiw, G., D. Romer, and D. Weil, 1992, *QJE*.

Miles, D., 1999, Modelling the impact of demographic change upon the economy, *Economic Journal* 109, 1-36.

Mincer, J., 1958,

Mincer, J., 1974, *Schooling, Experience, and Earnings,* New York: NBER Press,

Ortigueira, S., 2003, Equipment prices, human capital and economic growth, *Journal of Economic Dynamics and Control* 28, 307-329.

Pritchett, L., 2006,

Psacharopoulus, 1994,

Rosen, S., 1976. A Theory of Life Earnings, JPE 84 (4), S45-S67.

Rosen, S., 2008. Human capital. In: *The New Palgrave Dictionary of Economics*, 2nd ed., ed. by S. N. Durlauf and L. E. Blume, available at

http://www.econ.ku.dk/english/libraries/links/

Rosenzweig, M.R., 2010, Microeconomic approaches to development: Schooling, learning, and growth, Economic Growth Center Discussion Paper No. 985, Yale University.

Sattinger, M., 1980, *Capital and the Distribution of Labor Earnings*, North-Holland: Amsterdam.

Sen, A., 1997,

Sheshinski, E., 2007, *The Economic Theory of Annuities*, Princeton: Princeton University Press.

Sheshinski, E., 2009, Uncertain longevity and investment in education, WP, The Hebrew University of Jerusalem, August.

Sianesi, B., and J. Van Reenen, 2003, The returns to education: Macroeconomics, *Journal of Economic Surveys* 17 (2), 157-200.

—