Narodowy Bank Polski

NBP Working Paper No. 279

Do Mincerian wage equations inform how schooling influences productivity?

Christian Groth, Jakub Growiec

NBP Working Paper No. 279

Do Mincerian wage equations inform how schooling influences productivity?

Christian Groth, Jakub Growiec

Economic Research Department
Warsaw, 2018

Christian Groth – University of Copenhagen; chr.groth@econ.ku.dk
Jakub Growiec – Narodowy Bank Polski and SGH Warsaw School of Economics;
jakub.growiec@nbp.pl

Acknowledgements.
For helpful comments we thank an anonymous reviewer, Michał Gradzewicz, Jakub Mućk, and participants at the international workshop "Economic Growth, Macroeconomic Dynamics and Agent Heterogeneity", European University at St. Petersburg, May 2017. The opinions expressed in the paper belong to the authors and have not been endorsed by NBP.

Published by:
Narodowy Bank Polski
Education & Publishing Department
ul. Świętokrzyska 11/21
00-919 Warszawa, Poland
www.nbp.pl

ISSN 2084-624X

Contents

Abstract

We study the links between the Mincerian wage equation (the cross-sectional relationship between wages and years of schooling) and the human capital production function (the causal effect of schooling on labor productivity). Based on a stylized Mincerian general equilibrium model with imperfect substitutability across skill types and *ex ante* identical workers, we demonstrate that the mechanism of compensating wage differentials renders the Mincerian wage equation uninformative for the human capital production function. Proper identification of the human capital production function should take into account the equilibrium allocation of individuals across skill types.

Keywords: Mincerian wage equation, human capital production function, skill distribution, compensating wage differentials, golden rule of skill formation.

JEL Classification Numbers: E24, I26, J24.

1 Introduction

Do cross-sectional wage equations provide evidence on how schooling influences productivity? Several influential authors assume that this is the case. Lucas (1988) motivates his assumption of an exponential relationship between human capital and years of schooling – an exponential *human capital production function* – with its consistence "with the evidence we have on individual earnings" (p. 19). Bils and Klenow (2000) do the same thing "precisely (...) [to] draw on the large volume of micro evidence" (p. 1162). So do Hall and Jones (1999) who construct their macro-level human capital production function by drawing from a number of country-specific cross-sectional Mincerian return estimates (cf. Psacharopoulos and Patrinos, 2004). Caselli (2005), referring to Hall and Jones (1999), explains this step even more forcefully: "Given our production function, perfect competition in factor and good markets implies that the wage of a worker with s years of education is proportional to his human capital. Since the wage-schooling relationship is widely thought to be log-linear, this calls for a log-linear relation between h and s as well, or something like $h = \exp(\phi_s s)$, with ϕ_s a constant (...) at country level" (p. 686). The chapter by Klenow and Rodriguez-Clare (1997) features an entire section titled "Using Mincer Regression Evidence to Estimate Human Capital Stocks".

The present paper elucidates one important pitfall of this approach and suggests an extension to the standard Mincerian wage regression which allows to avoid it. Namely, when individuals are allowed to endogenously choose the number of years of schooling, s, the standard Mincerian wage equation (the cross-sectional relationship between wages and years of schooling, $w(s)$) is insufficient for identifying the underlying human capital production function $h(s)$, and may even fail to convey *any* useful information in this regard. We present a full dynamic general equilibrium model which exactly exposes the reverse causal link from wages to individuals' schooling decisions, lying at the heart of the difficulty of identifying the shape of the human capital production function from Mincerian wage equations. While the literature appears to play down the role of this reverse causal link, we show that it can actually be crucial[1] and can be addressed only if the identification of the human capital production function is adequately augmented with the endogenous distribution of skills, as captured by our *extended Mincer equation*.

Our theoretical approach departs from the usual simplifying assumption that skill levels are perfectly substitutable. The reason is that under this assumption, if furthermore individuals are *ex ante* identical and the link between skills and wages is deterministic, generally a unique equilibrium skill level s^* is obtained (as in C. Jones, 2007). In consequence, there is no skill heterogeneity in the population. Hence, even though it is then true that wages are proportional to human capital levels in equilibrium, this statement of little use because the cross-sectional wage equation cannot be identified in the first place.

In contrast, if skill levels are *im*perfectly substitutable, there is demand for varying skill levels in equilibrium (as in B. Jones, 2014), leading to a non-degenerate equilibrium skill distribution. On the one hand, presence of variation in skill levels allows for identifying the cross-sectional wage equation. On the other hand, however, imperfect substitutability enables the mechanism of *compensating wage differentials* (Jovanovic, 1998) which implies that skill-specific productivity differences will be incorporated in the equilibrium skill distribution, so that wages will be no longer proportional to human capital levels in equilibrium. In fact, in the stylized overlapping generations model which we put forward in this paper, productivity differences are *fully* accounted for in the skill demand profiles, leaving the cross-sectional wage equation to be identified only by the underlying demographics and retirement pattern. It then carries *no* information on the human capital production function.

Looking from an empirical angle, the extended Mincer equation derived in this paper implies that if a large share of the population chooses high skill levels – as it is arguably the case in contemporary developed economies – then the human capital production function $h(s)$ should increase more sharply with s than the cross-sectional wage equation $w(s)$. In such a case, using the

[1] Our argumentation expands on the criticism formulated by Jones (2008). The focus of that paper is on how cross-country income differences may arise through strategic complementarities in joint decisions regarding "breadth" and duration of education.

cross-sectional wage equation to identify the human capital production function would lead to downward biased estimates of "true" returns to education (i.e., measured in the units of human capital, not its equilibrium remuneration). The lower equilibrium dispersion in $w(s)$ compared to $h(s)$ follows from compensating wage differentials.

The remainder of the paper is structured as follows. In Section 2 we present our model, treating human capital in a life-cycle perspective. Section 3 solves for the equilibrium allocation of skills; a by-product of the analysis is the *golden rule of skill formation*. Section 4 provides a discussion of our theoretical argument that the cross-sectional relationship between wages and years of schooling may not convey any information on the underlying human capital production function when education decisions are endogenous. In Section 5 a few simple analytical examples are advanced. Section 6 concludes.

2 The model

We consider a closed economy where labor is the only factor of production. Workers are allowed to differ in their skills. The only source of variation in skills is the number of years of schooling s. Labor services provided by workers with different skills are imperfectly substitutable, with a constant elasticity of substitution. Firms employ workers in order to produce the unique final consumption good. They operate in a perfectly competitive environment. *Ex ante* identical individuals maximize the expected value of their discounted lifetime utility from consumption. With this aim they choose length of their education as in Mincer (1958). Education precludes working but is otherwise cost-free. There is no on-the-job learning. Labor supply by working-age individuals is inelastic. There is exogenous, skill-neutral exponential technological progress at a constant rate g. Time is continuous and flows from $-\infty$ to $+\infty$. People have no bequest motive. Individuals face a known age-specific hazard rate of death at each instant. There is a perfect credit and life annuity market. In equilibrium, wages are going to be such that individuals are exactly indifferent across various lengths of education (compensating differentials, Jovanovic, 1998). We assume a stationary age structure of the population (Growiec, 2010) and concentrate on the steady-state equilibrium.

Demographics. Individuals are born continuously with a fixed birth rate $b > 0$. The unconditional probability of survival until age τ is independent of calendar time and given by a function $m(\tau)$ such that $m(0) = 1$, m is non-increasing with τ, and there is a maximum lifetime, T, such that $m(\tau) > 0$ for $\tau \in [0, T)$ and $m(\tau) = 0$ for $\tau \geq T$. The survival law $m(\tau)$ implies an age-specific hazard rate of death, $-m'(\tau)/m(\tau) \equiv d(\tau), 0 < \tau < T$.

We denote the population size at time t as $N(t)$. Then, by the Law of Large Numbers, there are $P(t, \tau) \equiv bN(t-\tau)m(\tau)$ people aged τ in the population at time t. We assume a stationary age structure of the population, signifying that the shares of population at a given age, $P(t, \tau)/N(t)$, are independent of calendar time t.

Stationarity of the population age structure implies a constant aggregate death rate $\bar{d}$ which is uniquely determined by the assumed survival law $m(\tau)$. The population size at time t is thus $N(t) = N(0)e^{nt}$, where $n \equiv b - \bar{d}$ is the constant population growth rate. Accordingly, $P(t,\tau)/N(t) = be^{-n\tau}m(\tau)$. It is found that $n > 0$ (respectively, $n < 0$) if the birth rate b is above (respectively, below) the reciprocal of the life expectancy at birth (see Growiec, 2010; Growiec and Groth, 2015, for derivations).

Under these assumptions the size of a population cohort aged τ at time t (and thus born in the vintage $v \equiv t - \tau$) is equal to

$$P(t,\tau) = bN(0)e^{n(t-\tau)}m(\tau) \equiv e^{n(t-\tau)}m(\tau), \tag{1}$$

normalizing initial population size at $N(0) = 1/b$ so that the cohort born at time 0 is of unit size.

A function that will repeatedly appear in the formulas to follow, with different specifications of the parameter α, is the following:

$$M_\alpha(s) \equiv \int_s^T e^{-\alpha\tau}m(\tau)d\tau, \qquad \alpha \in \mathbb{R}, s \in [0,T]. \tag{2}$$

The function $M_\alpha(s)$ may be interpreted as "expected remaining α-discounted lifetime" of an individual at age s. The function $M_\alpha(s)$ takes non-negative values and is differentiable and non-increasing in s. Moreover, if $\alpha_1 > \alpha_2$, then $M_{\alpha_1}(s) \leq M_{\alpha_2}(s)$ and the ratio $M_{\alpha_1}(s)/M_{\alpha_2}(s)$ is non-increasing in s.

Time profiles of schooling and work. We assume that the individuals spend their first s years of life at school, and then they work full time, providing one unit of labor per time unit at every age $\tau \geq s$. There is no retirement. The total expected stream of working time through life provided by an individual with exactly s years of schooling is then $\int_s^T m(\tau)d\tau \equiv M_0(s)$.

We denote as $\pi(s)$ the fraction of any population vintage v who have decided to obtain exactly s years of schooling. The maximum demanded skill level (maximum number of years of schooling) is set at $\bar{s} > 0$ where $M_0(\bar{s}) > 0$ so that even among those most educated some manage to do at least some work before they die. By definition, $\int_0^{\bar{s}} \pi(s)ds = 1$. We assume $\pi(s)$ to be independent of vintage v and calendar time t, signifying that we concentrate on a steady-state equilibrium.

Integrating across past vintages v, we find that the measure of workers with exactly s years of schooling in the population at time t equals:

$$\int_{t-T}^{t-s} \pi(s)P(t,t-v)dv = e^{nt}\pi(s)\int_s^T e^{-n\tau}m(\tau)d\tau = e^{nt}\pi(s)M_n(s), \qquad s \in [0,\bar{s}]. \tag{3}$$

Because of our normalization $bN(0) = 1$, the factor $M_n(s)$ can be interpreted as a measure of workers at least s years old at time 0. Naturally, this measure declines with s. In addition it depends negatively on n because, looking backward from time 0, the cohorts decline faster the higher is n.

The human capital production function. The level of human capital (productive skills) of an individual who has completed s years of schooling – the *human capital production function* – is denoted as $h(s)$. We assume that $h(s)$ takes positive values and is differentiable and increasing in s; it requires no other inputs beyond the individual's time.

The firm's optimization problem. We assume that firms operate in a competitive environment and face a CES production technology with respect to labor services $h(s)L_t(s)$, where $L_t(s)$ measures working hours per time unit at time t delivered by workers of skill type s. There is also constant exogenous technological progress at a rate $g \geq 0$:

$$Y_t = e^{gt} \left(\int_0^{\bar{s}} (h(s)L_t(s))^{\theta} \, ds \right)^{\frac{1}{\theta}}, \qquad \theta < 1. \tag{4}$$

The substitutability parameter θ determines the elasticity of substitution between skill types as $\sigma = (1-\theta)^{-1}$. The case $0 < \theta < 1$ captures the (empirically relevant) case where skill levels are gross substitutes ($\sigma > 1$), so that an increase in the supply of a given skill type increases its competitive income share. The opposite case $\theta < 0$ implies that skill types are gross complements.[2]

Focusing on the steady state, we shall use the notations $L(s) \equiv L_0(s) = e^{-nt}L_t(s)$ and $Y \equiv Y_0 = \left(\int_0^{\bar{s}} (h(s)L(s))^{\theta} \, ds \right)^{\frac{1}{\theta}} = e^{-(g+n)t}Y_t$ to single out the time-invariant component of skill-specific labor and aggregate output, respectively.

The representative firm chooses its demand for every skill type, $\{L(s)\}_{s=0}^{\bar{s}}$, in order to maximize its static profit given by:

$$\Pi_t(\{L(s)\}_{s=0}^{\bar{s}}) = e^{(g+n)t} \left(\int_0^{\bar{s}} (h(s)L(s))^{\theta} \, ds \right)^{\frac{1}{\theta}} - e^{nt} \int_0^{\bar{s}} w_t(s)L(s)ds.$$

[2] In the limiting case $\theta = 0$ ($\sigma = 1$), factor shares in income are constant and consequently the CES formula should be replaced with its Cobb-Douglas counterpart, $Y_t = e^{gt + \int_0^{\bar{s}} \ln(h(s)L_t(s))ds}$.

The first-order conditions for labor of different skill types are

$$e^{gt}Y^{1-\theta}h(s)^{\theta}L(s)^{\theta-1} = w_t(s) = w(s)e^{gt}, \qquad s \in [0,\bar{s}], \tag{5}$$

where $w(s) \equiv w_0(s)$ is the time-invariant component of the wage rate. Given that (5) can be written as $w(s)L(s) = Y^{1-\theta}(h(s)L(s))^{\theta}$, we have

$$\int_0^{\bar{s}} w(s)L(s)ds = Y^{1-\theta}\int_0^{\bar{s}}(h(s)L(s))^{\theta} = Y. \tag{6}$$

So, in accordance with constant returns to scale, the firm's total production cost will equal output and profits will be zero in equilibrium.

By (3), clearing in the labor markets amounts to

$$L(s) = \pi(s)M_n(s), \qquad s \in [0,\bar{s}]. \tag{7}$$

The individual's optimization problem. We assume that every individual born at time v, subject to the usual budget constraint, maximizes her expected lifetime utility from consumption and with this aim optimally chooses her number of years of schooling. As mentioned, there is a perfect credit and life annuity market. People have no bequest motive. Hence they are born with zero net financial assets and early in life they take life-insured loans to finance consumption while at school. Apart from the uncertain lifetime, there is no uncertainty. At birth individuals are alike.

The problem at hand admits the "separation theorem" (Acemoglu, 2009, Chapter 10), thanks to which we may first solve for the optimal number of years of schooling and then turn to the consumption decision.

The schooling decision. An individual born at time v chooses the length of education s in order to maximize human wealth – discounted expected lifetime earnings – as seen from time v (Mincer, 1958; Heckman, Lochner, and Todd, 2003):

$$\begin{aligned} HW(v,s) &= \int_{v+s}^{T} w_t(s)e^{-r(t-v)}m(t-v)dt = e^{gv}w(s)\int_s^T e^{-(r-g)\tau}m(\tau)d\tau \\ &= e^{gv}w(s)M_{r-g}(s), \end{aligned} \tag{8}$$

where r is the risk-free interest rate, perceived by the individual as exogenous. Subject to subsequent confirmation, we tentatively consider the rate r as constant over time. The time-invariant component, $w(s)$, of the skill-specific wage rate can be taken in front of the integral because it does not depend on the individual's age τ (no on-the-job learning). This also implies that once an individual enters the workforce, she will receive exponentially growing flows of earnings. Thus we rule out the usual hump-shaped age–earnings profiles (Ben-Porath, 1967).

Solving the individual's optimization problem regarding schooling yields the following first-order condition:

$$w'(s)M_{r-g}(s) = -w(s)M'_{r-g}(s). \tag{9}$$

This condition equates the marginal benefit of one more year of schooling to the marginal opportunity cost in terms of earnings forgone by entering the labor market one year later.

From the firm's first-order condition (5) it follows that there will be positive demand for labor of every skill level $s \in [0, \bar{s}]$. To be willing to supply any of these different skill levels, the *ex ante* identical individuals must be exactly indifferent when choosing length of education s. Hence the individuals' first-order condition (9) must hold for all $s \in [0, \bar{s}]$. So (9) makes up a linear differential equation for w as a function of s. The solution is

$$w(s) = w(0)e^{-\int_0^s \frac{M'_{r-g}(x)}{M_{r-g}(x)}dx} = w(0)\frac{M_{r-g}(0)}{M_{r-g}(s)}, \qquad s \in [0, \bar{s}]\,. \tag{10}$$

In equilibrium, human wealth, $HW(v, s)$ in (8), is therefore the same for any $s \in [0, \bar{s}]$. The intuition behind the second equality in (10) is that according to (9), the augmentation rate of the wage rate with respect to schooling is the same as the rate of decline with respect to schooling of the expected stream of discounted working time through life. Hence, the augmentation factor of the wage rate with respect to schooling when comparing 0 to s years of schooling – the *compensating wage differential* – equals the corresponding decay factor of the expected stream of discounted working time through life.

The consumption-saving decision. Having made her optimal schooling decision, the individual of vintage v with planned schooling level s maximizes her discounted expected lifetime utility from consumption (we assume CRRA utility):

$$\max_{\{c(v,t)\}_{t=v}^{v+T}} \int_{v}^{v+T} \frac{c(v,t)^{1-\eta}}{1-\eta} e^{-\rho(t-v)} m(t-v) dt, \qquad \eta > 0, \rho \geq 0,$$

subject to the dynamic budget constraint:

$$\dot{a}(v,s,t) = (r + d(t-v))a(v,s,t) + \tilde{w}_t(s) - c(v,t), \qquad a(v,s,v) = 0, \quad (11)$$

where $a(v,s,t)$ is net assets held at time t. We use the notation $\tilde{w}_t(s) = w_t(s)$ if $t - v > s$ (so that the individual is in her working age) and $\tilde{w}_t(s) = 0$ otherwise (when the individual is still at school). In (11) the term $d(t-v)a(v,s,t)$ captures the life insurance part of annuity payments (or annuity receipts if the individual has positive $a(v,s,t)$) covering the hazard rate of death. So $r + d(t-v)$ is the "actuarial interest rate" at age $t - v$. When the individual dies, the obligation or the entitlement is canceled. Upon birth, the individual holds no assets. Subject to subsequent confirmation, we tentatively consider the consumption path of any individual of vintage v to be independent of the chosen s.

The individual also faces the solvency condition implying that, in *expected value*, accumulated discounted primary saving at death is nonnegative:

$$\int_{v}^{v+T} (\tilde{w}_t(s) - c(v,t)) e^{-r(t-v)} m(t-v) dt \geq 0. \quad (12)$$

Equation (12) is required to hold only in expected value thanks to the assumption of a perfect life annuity market (Yaari, 1965, p. 147-148).

Solving for the optimal path of consumption yields the Keynes-Ramsey rule:

$$\frac{\dot{c}(v,t)}{c(v,t)} = \frac{r - \rho}{\eta} \equiv \gamma(r). \quad (13)$$

Individual consumption will be either growing, constant, or declining across the individual's lifetime, depending on the relation between r and the rate of time preference, $\rho \geq 0$. Solving for $c(v,t)$ yields:

$$c(v,t) = c(v,v)e^{\gamma(r)(t-v)} = e^{gv} c_0 e^{\gamma(r)(t-v)}, \quad (14)$$

where $c_0 \equiv c(0,0)$ and we have imposed that in steady state $c(v,v) = e^{gv} c_0$, to be confirmed subsequently.

Integrating the asset equation (11) over time t, we find[3] that net asset holdings at time t of an individual of vintage v, having decided schooling level s, are equal to

$$a(v,s,t) = \begin{cases} -\underbrace{\frac{e^{r(t-v)}}{m(t-v)} e^{gv} c_0 \int_0^{t-v} e^{-(r-\gamma(r))\tau} m(\tau) d\tau}_{\textit{accumulated consumption spending}}, & \text{if } t \in [v, v+s], \\ \underbrace{\frac{e^{r(t-v)}}{m(t-v)} e^{gv} \left(w(s) \int_s^{t-v} e^{-(r-g)\tau} m(\tau) d\tau \right.}_{\textit{accumulated income}} & \\ \qquad \left. - \underbrace{c_0 \int_0^{t-v} e^{-(r-\gamma(r))\tau} m(\tau) d\tau}_{\textit{accumulated consumption spending}} \right), & \text{if } t \in (v+s, v+T). \end{cases}$$

It remains to determine the exact value for c_0 via the necessary transversality condition that the solvency condition (12) holds with strict equality:

$$\begin{aligned} \int_v^{v+T} (\tilde{w}_t(s) - c(v,t)) e^{-r(t-v)} m(t-v) dt &= e^{gv} \left(w(s) \int_s^T e^{-(r-g)\tau} m(\tau) d\tau \right. \\ \left. -c_0 \int_0^T e^{-(r-\gamma(r))\tau} m(\tau) d\tau \right) &= e^{gv} (w(s) M_{r-g}(s) - c_0 M_{r-\gamma(r)}(0)) = 0, \end{aligned}$$

using the definition of the function $M_\alpha(s)$ in (2). Thus

$$c_0 = w(s) M_{r-g}(s) / M_{r-\gamma(r)}(0) = w(0) M_{r-g}(0) / M_{r-\gamma(r)}(0), \qquad (15)$$

where the latter equality is implied by (10). It is hereby confirmed that the consumption path of any individual of vintage v is independent of the chosen s.

[3] Recall that since $d(\tau) \equiv -m'(\tau)/m(\tau)$, and $m(0) = 1$, $m(t-v) = \exp(-\int_v^t d(u-v) du)$.

3 Intertemporal equilibrium

Clearing in the market for loanable funds. The "life-cycle" of net assets is such that early in life the individual borrows to finance her consumption while at school. She then joins the workforce, which allows her to gradually repay the initial debt and on average accumulate positive net wealth.

The only store of value is loans. Aggregate financial wealth, $A(t)$, is thus zero for all t:

$$A(t) \equiv \int_0^{\bar{s}} \left(\int_{t-T}^{t-s} a(v,s,t)\pi(s)P(t,t-v)dv \right) ds = 0.$$

Hence, also aggregate saving, $S(t)$, is nil for all t:

$$\dot{A}(t) = S(t) = \int_0^{\bar{s}} w_t(s)L_t(s)ds - C(t) = 0, \tag{16}$$

where $C(t) \equiv \int_{t-T}^{t} c(v,t)P(t,t-v)dv$ is aggregate consumption. Concentrating on $t = 0$, we obtain the following equation for the interest rate r in equilibrium:

$$\begin{aligned} C(0) &= c_0 \int_0^T e^{-(g+n-\gamma(r))\tau} m(\tau)d\tau \equiv c_0 M_{g+n-\gamma(r)}(0) = w(0)M_{r-g}(0)\frac{M_{g+n-\gamma(r)}(0)}{M_{r-\gamma(r)}(0)} \\ &= \int_0^{\bar{s}} w(s)\pi(s)M_n(s)ds = w(0)M_{r-g}(0)\int_0^{\bar{s}} \pi(s)\frac{M_n(s)}{M_{r-g}(s)}ds, \end{aligned} \tag{17}$$

where we have first applied (14) and (1), then the definition of M, then the transversality condition, then (16) combined with clearing in the labor market, i.e., (7), and finally the compensating wage differential in (10). Equating the last term in the first line to that in the second gives $r = g + n$,[4] in view of the identity $\int_0^{\bar{s}} \pi(s)ds = 1$.

In this way we have confirmed that r is constant over time. More importantly, we find that the equilibrium interest rate r equals the steady-state growth rate of final output, $g + n$. The interest rate matters, through the Keynes-Ramsey rule (13), for *growth* of individual consumption, and through the transversality condition it matters for the individual's *level* of consumption, cf. (15). The equality $r - g = n$ ensures that the initial consumption level is

[4]Generically, this is the *unique* solution.

the same for all skill types and remains consistent with clearing in the output market, cf. (17).

Note that the equilibrium interest rate is thus independent of the individuals' preference parameters ρ and η. This is because we have for simplicity ignored physical capital. In the absence of physical capital accumulation, there is no trade-off between less consumption now, i.e., more capital accumulation now, and higher consumption in the future. Instead, the interest rate is free to adjust to the golden rule level, $g + n$. This level of the interest rate takes into account that (i) each consecutive generation has higher consumption than the previous one due to exogenous technological progress (presupposing $g > 0$), and (ii) each consecutive generation is also more populous than the previous one (if $n > 0$). At the same time it turns out that this interest rate ensures the *highest* technically feasible steady-state level of trend-corrected per capita consumption, see below. Thus, also in our present context of *human* capital accumulation is the name *golden rule interest rate* justified for the interest rate $r = g + n$.

Equilibrium skill distribution and wage structure. To prepare the ground for the main result, we shall determine the distribution of skills and the resulting wage structure in a steady-state equilibrium. In the next section we then conclude that thanks to the mechanism of compensating wage differentials, the wage distribution in a steady-state equilibrium is independent of the human capital production function $h(s)$.

As noted in (6), the firm's total production cost equals output. With clearing in the labor markets this implies

$$\int_0^{\bar{s}} w(s)\pi(s)M_n(s)ds = Y.$$

Plugging (10) into this and substituting $r = g + n$ gives

$$w(s)M_n(s) = w(0)M_n(0) = Y \text{ for all } s \in [0, \bar{s}], \tag{18}$$

since $\int_0^{\bar{s}} \pi(s)ds = 1$. That is, because $r = g + n$ in equilibrium, the mechanism of compensating wage differentials equalizes the total input cost of each skill type, making it proportional to total cost (i.e., total output at time 0, Y). The factor of proportionality is one because we have normalized cohort 0 to be of size one, $bN(0) = 1$. Note that the human capital production function $h(s)$ does not enter (18).

From clearing in the market for labor of skill type s, together with the firm's first-order condition (5), we get

$$\pi(s)M_n(s) = L(s) = Yw(s)^{-\frac{1}{1-\theta}}h(s)^{\frac{\theta}{1-\theta}} = Y^{-\frac{\theta}{1-\theta}}M_n(s)^{\frac{1}{1-\theta}}h(s)^{\frac{\theta}{1-\theta}},$$

where the last equality comes from inserting (18). It follows that

$$\pi(s) = Y^{-\frac{\theta}{1-\theta}}\left(h(s)M_n(s)\right)^{\frac{\theta}{1-\theta}}. \tag{19}$$

This leads to the final solution for the time-invariant component of output, the distribution of skills, and the wage structure in a steady-state equilibrium. Integrating over s in (19) and solving for equilibrium output at time 0 yields

$$Y = \left(\int_0^{\bar{s}}\left(h(s)M_n(s)\right)^{\frac{\theta}{1-\theta}}ds\right)^{\frac{1-\theta}{\theta}}. \tag{20}$$

Plugging this into (19) gives the equilibrium distribution of skills as

$$\pi(s) = \left(\int_0^{\bar{s}}\left(h(s)M_n(s)\right)^{\frac{\theta}{1-\theta}}ds\right)^{-1}\left(h(s)M_n(s)\right)^{\frac{\theta}{1-\theta}}, \text{ for all } s \in [0,\bar{s}]. \tag{21}$$

Finally, plugging (20) into (18) gives the equilibrium wage structure:

$$w(s) = \frac{Y}{M_n(s)} = \frac{1}{M_n(s)}\left(\int_0^{\bar{s}}\left(h(s)M_n(s)\right)^{\frac{\theta}{1-\theta}}ds\right)^{\frac{1-\theta}{\theta}}, \text{ for all } s \in [0,\bar{s}]. \tag{22}$$

The golden rule of skill formation. To demonstrate that the above equilibrium allocation is consistent with the golden rule, consider a social planner solving the following problem: among all technically feasible steady-state paths, choose the one maximizing the trend-corrected level of consumption $e^{-(g+n)t}C(t) = C(0)$. The social planner will choose the same $\pi(s)$ as that given in (21), which results in $C(0) = Y$ as given in (20).[5] Hence, the skill distribution, $\pi(s)$, obtained in a steady-state equilibrium of our competitive economy without externalities and without capital accumulation complies with the *golden rule of skill formation*: the skill profile required to obtain the highest situated sustainable path of per capita consumption is obtained when the interest rate equals $g + n$, the golden rule interest rate.

[5] The proof, using $\pi(s)$ as control variable over the interval $[0, \bar{s}]$, is available from the authors upon request.

4 Can the human capital production function be identified from the Mincerian wage equation?

We are now ready to answer the main question: do cross-sectional Mincerian wage equations inform how schooling influences productivity? Our answer is no, or at least not when the distribution of skills arises endogenously, driven by compensating wage differentials.

To see this in our model, we take logs of both sides of equation (22):

$$\ln w(s) = \ln Y - \ln M_n(s). \tag{23}$$

Hence, in equilibrium, the cross-sectional wage regression equation cannot be used for inferring the human capital production function $h(s)$. If individuals are *ex ante* identical and the population age structure is stationary, then what the Mincerian wage regression actually captures is not skill-specific productivities, $h(s)$, but measures of people at least s years old, for $s = s_1, s_2, \ldots$, and still in the labor force. Due to the presence of compensating wage differentials, the skill-specific productivities are fully incorporated in the equilibrium skill allocation (21), leaving the wage equation to be identified only by the underlying demographics (and retirement pattern in an extended model[6]), but *not* the human capital production function.[7]

The extended Mincer equation. It is also instructive to isolate $h(s)$ in equation (21), take logs, and use (23) to get

$$\ln h(s) = \ln w(s) + \left(\frac{1-\theta}{\theta}\right) \ln \pi(s). \tag{24}$$

We may call this the *extended Mincer equation.* Furthermore, this equation may, in view of clearing in the labor markets, $\pi(s)M_n(s) = L(s)$, be also given an alternative formulation:

$$\ln h(s) = \left(\frac{1}{\theta}\right) \ln w(s) + \left(\frac{1-\theta}{\theta}\right) \ln \frac{L(s)}{Y}. \tag{25}$$

[6] The model is easily generalized to the case of age-dependent labor supply with exponential retirement.

[7] Note also that due to compensating wage differentials, accumulated lifetime incomes (human wealth $HW(v,s)$) are fixed within cohorts, equation (10).

These two last expressions for $\ln h(s)$ imply that if one wants to identify the shape of the human capital production function from a cross-sectional wage regression equation, then one should also account for the information conveyed by either the equilibrium skill distribution within each cohort, $\pi(s)$, or the equilibrium skill distribution in the cross-section of the labor force, represented by $L(s)/Y$.

A few observations are due here. First, the extended Mincer equation (23) is reduced to its standard form only if $h(s) \propto (M_n(s))^{-1}$. In other words, the human capital production function $h(s)$ can be identified with the cross-sectional wage equation $w(s)$ only if they both happen to be inversely proportional to the demographic profile of the population, summarized by $M_n(s)$; an unlikely coincidence.

Second, under endogeneity of the schooling decision the curvature of the human capital production function $h(s)$ is related to the skewness of the endogenous skill distribution $\pi(s)$. If relatively more individuals within a cohort choose high skill levels, so that $\pi(s)$ is increasing in s, then $h(s)$ increases more sharply than $w(s)$ – and the other way round if relatively more individuals choose low skill levels. Preliminary empirical investigation based on US Current Population Survey data for 2016 favors the former alternative and implies that in the US, human capital $h(s)$ may be increasing more sharply with s than the wage profile $w(s)$ would suggest. This would mean that using the cross-sectional wage equation to identify the human capital production function would in fact lead to downward biased estimates of "true" returns to education (in human capital units). The bias follows from endogeneity of the skill level s, its direction is determined by the mechanism of compensating wage differentials, and its absolute magnitude is inversely related to the elasticity of substitution between skill types, σ (or equivalently, θ).[8] In future research, one may want to reconcile this finding with the mounting cross-country evidence that Mincerian rates of return tend to fall with s (Hall and

[8] These preliminary empirical results are available upon request. They are however subject to many caveats. Other than in the model, in reality individuals are not born identical but inherit material wealth, social and cultural capital, and unobserved innate abilities from their parents. Some individuals may be credit constrained and thus unable to equalize human wealth with their unconstrained peers. Life annuity markets are not perfect in reality. Education is not costless and its costs may vary at different stages of education. Etc.

Jones, 1999; Bils and Klenow, 2000; Psacharopoulos and Patrinos, 2004).

Third, in the limiting case of perfect substitutability across skill levels, i.e., the case $\theta = 1$, with competitive firms maximizing their static profits and *ex ante* identical workers, one directly obtains from (5) $w(s) = h(s)$. That is, for two different educational levels, s_1 and s_2, to be supplied in equilibrium, in view of (10), $h(s_2)/h(s_1)$ must be exactly equal to the required compensating wage differential $M_{r-g}(s_1)/M_{r-g}(s_2)$. Given that we have not imposed any functional restrictions on h, apart from $h' > 0$, this would be an unlikely coincidence. Hence, the schooling first-order condition (9), with w replaced by h, generally has at most one solution, s^* (Jones, 2007). So there will generally be no skill heterogeneity in the population, necessary for identifying the Mincerian wage equation. Hence, the equality $w(s) = h(s)$ alone does not justify the use of the Mincerian cross-sectional wage equation as an indirect human capital production function representation.

Discussion. The model presented in this paper features a number of simplifying assumptions, necessary for obtaining closed-form solutions but also crucially affecting the results. Let us now briefly discuss its potential extensions.

First, the model can be easily generalized to the case of age-dependent labor supply with exponential retirement where labor supply of individuals aged τ equals $\ell(\tau) = e^{-\mu\tau}$, $\mu > 0$. Exactly the same algebra is required to incorporate exponential decline in human capital (unit labor productivity) due to depreciation, such that $h(s, \tau - s) = h(s, 0)e^{a(\tau-s)}$, with $a < 0$. By the same token, exponential increase in human capital due to on-the-job experience accumulation would imply the same formula but with $a > 0$ (Growiec and Groth, 2015). If both kinds of influences are operative, then a would reflect their net effect and could be of any sign. These generalizations do not overturn any of our results.

Non-exponential retirement or work experience accumulation patterns do change (complicate) the algebra, though. However, the key insight that the equilibrium skill distribution is driven by compensating wage differentials, necessitating modifications of the standard Mincer equation, remains intact.

Second, one could relax the assumption that the age structure of population is stationary. Then each birth cohort would face a different reward structure to s years spent on schooling, so that the equilibrium skill distributions would be cohort-specific, $\pi_v(s)$. Furthermore, within-cohort skill structures $\pi_v(s)$ would then be decoupled from cross-cohort (cross-sectional) skill structures $L_t(s)$. However, apart from making the analysis essentially intractable, this change would not overturn the key observation that the equilibrium skill distribution would still be non-degenerate only because of compensating wage differentials, and thus the endogeneity of the schooling decision would remain vital for identifying the human capital production function.

Third, one could think of relaxing the assumptions that (i) individuals are *ex ante* identical and (ii) the relationship between skills and wages is deterministic. That would constitute a major change in the workings of the model as compensating wage differentials would then cease to be the only source of equilibrium heterogeneity, yielding some space to, respectively, (i) *endowments* and (ii) *luck* (Jovanovic, 1998). Neither of these mechanisms is included in the extended Mincer equation (24), so we cannot be sure how much of the heterogeneity lending identification to the cross-sectional wage equation would be related to the human capital production function in such an extended model.

5 Analytical examples

In this section two simple analytical examples will illustrate our main point that in a model where equilibrium skill heterogeneity accrues thanks to compensating wage differentials, the cross-sectional wage equation carries no information useful for inferring the shape of the human capital production function. This is so even if the cross-sectional wage equation is well approximated by the famous log-linear Mincerian form.

Fixed lifetimes. This is the case $m(\tau) = 1$ for $\tau < T$ and $m(\tau) = 0$ for $\tau \geq T$. So here individuals' lifespans are deterministically equal to T. If $n \neq 0$, then $M_n(s) = (e^{-ns} - e^{-nT})/n$ and $M_n'(s)/M_n(s) = -n/(1 - \exp(-n(T-s)))$. Under this survival law, the cross-sectional wage equation (23) becomes:

$$\ln w(s) = const - \ln\left(e^{-ns} - e^{-nT}\right) \approx new\ const + ns, \tag{26}$$

where the last approximation assumes that T is "large" and thus the aggregate death rate $\bar{d}$ is "small" (Heckman, Lochner, and Todd, 2003).[9]

If $n = 0$, however, $M_n(s)$ becomes $M_0(s) = T - s$, with $M_0'(s)/M_0(s) = -1/(T-s)$, and (26) is replaced by

$$\ln w(s) = const - \ln(T - s).$$

The "perpetual youth" survival law, Blanchard (1985). This is the case $m(\tau) = e^{-\bar{d}\tau}$. Lifetime is uncertain but has no upper bound. Our above results are easily generalized to this case. Allowing $T = \infty$ in the definition of $M_\alpha(s)$ in (2), we get, under this survival law, in view of $n = b - \bar{d}$, that $M_n(s) = e^{-(n+\bar{d})s}/(n+\bar{d}) = e^{-bs}/b$ and $M_n'(s)/M_n(s) = b$. Note that also in this case is $M_n(s)$ finite.

The cross-sectional wage equation (23) becomes:

$$\ln w(s) = const + bs, \tag{27}$$

which is the exact Mincerian (log-linear) specification. It does not reflect the economy's human capital production function, however, which essentially

[9] This may be a rather bad approximation, however, as it requires $n \approx b$. In modern days, most advanced economies tend to have $n \approx 0$, with $b \approx \bar{d}$.

can be any increasing function $h(s)$. The relationship (27) only reflects the underlying demographics.

In both cases, the approximate Mincerian (log-linear) relationships obtained do not reflect the human capital production function but only how the demographics shape the wage profile.

6 Conclusion

Do cross-sectional wage equations inform how schooling influences productivity? Several influential authors have been assuming that it is the case. We have, however, presented a theoretical argument that such claims should be treated with caution due to the endogeneity of schooling decisions. Our simple Mincerian model highlights the role of compensating wage differentials in shaping these decisions. And when skill-specific productivity differentials are fully incorporated in the demand function for skills, then the cross-sectional Mincerian wage equation does not carry any information on the human capital production function.

Our model is highly stylized, though. One could imagine an economy where the cross-sectional wage equation conveys at least *some* information on the human capital production technology. This could obtain, for example, because of heterogeneity in workers' innate abilities, credit market frictions limiting individuals' education choices, or technological change that gives rise to continued change in the educational composition of the labor force. While working these cases out is left for further research, it can still be generally concluded that when identifying the shape of the human capital production function from cross-sectional wage regressions, it is advised not to omit the information conveyed by the equilibrium skill allocation. Typically, this allocation has been taken as exogenous in the associated literature, though.

Bibliography

ACEMOGLU, D. (2009): *Introduction to Modern Economic Growth*. Princeton University 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, 1160–1183.

BLANCHARD, O. (1985): "Debt, Deficits, and Finite Horizons," *Journal of Political Economy*, 93, 223–247.

CASELLI, F. (2005): "Accounting for Cross-Country Income Differences," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf. North-Holland.

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*, 8, 21–38.

HALL, R. E., AND C. I. JONES (1999): "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *Quarterly Journal of Economics*, 114, 83–116.

HECKMAN, J., L. LOCHNER, AND P. TODD (2003): "Fifty Years of Mincer Earnings Regressions," Working paper, NBER WP 9732.

JONES, B. F. (2008): "The Knowledge Trap: Human Capital and Development Reconsidered," NBER Working Papers 14138, National Bureau of Economic Research, Inc.

——— (2014): "The Human Capital Stock: A Generalized Approach," *American Economic Review*, 104, 3752–3777.

JONES, C. I. (2007): "A Simple Mincerian Approach to Endogenizing Schooling," Working paper, University of California, Berkeley.

JOVANOVIC, B. (1998): “Vintage Capital and Inequality,” *Review of Economic Dynamics*, 1(2), 497–530.

KLENOW, P. J., AND A. RODRIGUEZ-CLARE (1997): “The Neoclassical Revival in Growth Economics: Has It Gone Too Far?,” in *NBER Macroeconomics Annual 1997*, ed. by B. Bernanke, and J. Rotemberg. MIT Press.

LUCAS, R. E. (1988): “On the Mechanics of Economic Development,” *Journal of Monetary Economics*, 22, 3–42.

MINCER, J. (1958): “Investment in Human Capital and Personal Income Distribution,” *Journal of Political Economy*, 66, 281–302.

PSACHAROPOULOS, G., AND H. A. PATRINOS (2004): “Returns to Investment in Education: A Further Update,” *Education Economics*, 12(2), 111–134.

YAARI, M. E. (1965): “Uncertain Lifetime, Life Insurance, and the Theory of the Consumer,” *Review of Economic Studies*, 32, 137–150.

www.nbp.pl