



On aggregating human capital across heterogeneous cohorts[☆]



Jakub Growiec^{a,b,*}, Christian Groth^c

^a Warsaw School of Economics, Department of Quantitative Economics, Poland

^b Narodowy Bank Polski, Warsaw, Poland

^c University of Copenhagen, Department of Economics, Denmark

HIGHLIGHTS

- Mincerian log-linear relation between human capital, years of schooling, experience.
- We assume that the relation holds at micro level and study it upon aggregation.
- The exact relation is lost upon aggregation, except under implausible demographics.
- Numerically, the macro-Mincer equation is a good approximation of the true relation.
- We allow heterogeneity in years of schooling, retirement age and demographics.

ARTICLE INFO

Article history:

Received 2 March 2015

Received in revised form

4 August 2015

Accepted 11 September 2015

Available online 21 September 2015

ABSTRACT

This paper studies the question: Can the microeconomic Mincerian (log-linear) functional relationship between human capital, years of schooling and work experience be recovered in some similar form at the macroeconomic level? A large macroeconomic literature assumes so, warranting that the question is of interest. We first examine the question at a theoretical level and find that except under very special assumptions, the answer is in the negative. On the other hand, we also show numerically that a macro-Mincer relationship can nevertheless be perceived as a quantitatively reasonable approximation of the theoretically derived “true” relationship, at least if the observed heterogeneity comes only from differences in the number of years of schooling, retirement age, or demographic survival laws.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The Mincerian equation where the logarithm of individual wages (or human capital stocks) is explained by years of schooling and work experience (henceforth the *micro-Mincer* equation) is a cornerstone of a large body of microeconomic literature (Mincer, 1974; Heckman et al., 2003). Numerous studies have then carried it forward to country-level data on aggregate human capital stocks, average years of schooling, and average work experience in the population (e.g., Klenow and Rodríguez-Clare, 1997, Krueger and

Lindahl, 2001, Bloom et al., 2004). The aim of the present paper is to examine whether such simple aggregation is warranted. On the one hand, we formulate the necessary and sufficient conditions admitting the microeconomic Mincerian (log-linear) functional relationship to be theoretically recovered in some similar form at the macroeconomic level. On the other hand, since we find that the necessary conditions are very restrictive, we also add a quantitative edge to our analytical results. Based on a numerical study we evaluate that the macro-Mincer relationship can nevertheless be perceived as, to a large extent, a quantitatively reasonable approximation of the theoretically derived, more complex aggregate relationship.

At the theoretical level the standard micro–macro analogy has already been shown to rest on the following questionable simplifying assumptions:

- the macro-Mincer approach requires perfect substitutability between unskilled and skilled labor (Pandey, 2008; Jones, 2014a,b),

[☆] We are grateful to Holger Strulik for suggesting us to carry out the numerical study discussed in Section 3. We also thank anonymous Referees of this Journal as well as the *NBP Working Papers* series, conference participants at EEA/ESEM (Göteborg 2013), OCDGND (Vienna 2012) and EcoMod (Seville 2012) for useful comments and discussions. All errors are our responsibility.

* Correspondence to: Katedra Ekonomii Ilościowej SGH, al. Niepodległości 162, 02-554 Warszawa, Poland.

E-mail addresses: jakub.growiec@sgh.waw.pl (J. Growiec), chr.groth@econ.ku.dk (C. Groth).

- it assumes that each individual's skill level can be summarized by a single number and thus there is no heterogeneity in tasks (e.g., Jones, 2014b),
- it neglects the fact that maintaining a constant aggregate level of human capital in the society across time requires replacement investment because human capital is embodied in people whose lifetimes are finite (Growiec, 2010).

Violation of any of the above assumptions has been shown to lead to departures from the baseline macro-Mincer relationship between the aggregate human capital stock and average years of schooling and work experience, even if the micro-Mincer relationship holds perfectly at the individual level.

The current paper concentrates on the last of the above points. Thus, our analysis maintains the simplifying assumption that skill levels are perfectly substitutable, there is no intra-cohort heterogeneity of tasks or skills, and returns to schooling and experience are equal across countries. We only allow for heterogeneity in individuals' human capital levels following from the fact that people are born at different times and gradually accumulate human capital across their lives. By choosing such a restrictive framework, we isolate the effects coming from the heterogeneity of human capital due to demographics alone.

The contribution of the current paper to the literature is twofold. First, having clarified the outstanding problems related to the definitions of the aggregate human capital stock, aggregate years of schooling, and aggregate work experience under heterogeneity across population cohorts, we carry out a theoretical study leading to the conclusion that even if the micro-Mincer functional relationship holds exactly in a cross-section of individuals, the macro-Mincer (log-linear) equation generally does not. The only exceptions to this rule which we are able to identify are: (a) cases *inconsistent with heterogeneity*, insofar they require all aggregated individuals to have equal human capital stocks; (b) under the scenario where individuals first attend school full time and then work full time (and there are positive returns from work experience), the macro-Mincer equation can be recovered in the unique case where the demographic survival law has the "perpetual youth" property (Blanchard, 1985), which is empirically implausible. Under the scenario where people also retire at a certain age (and there are positive returns from work experience), the macro-Mincer equation cannot be recovered under any admissible survival law. For the cases where the macro-Mincer equation does not hold, we derive the true aggregate relationships.

Second, based on a numerical Monte Carlo study of the general theoretical model we find that, although theoretically misspecified, the macro-Mincer equation can nevertheless be perceived as an empirically reasonable approximation of the theoretically derived "true" relationship between the aggregate human capital stock, average years of schooling, and average work experience. We conclude that distortions to the log-linear shape of the macro-Mincer relationship caused by aggregating human capital across heterogeneous cohorts are quantitatively minor, at least under standard calibrations. Moreover, we also confirm numerically that average social returns to schooling are typically correctly identified with private returns (net of human capital depreciation).

The remainder of the article is structured as follows. In Section 2, we lay out the framework and discuss our theoretical results. We begin by discussing three particularly tractable special cases and then move on to our most general theoretical result. Section 3 complements this analysis with a numerical study. We begin with a presentation of the design of our Monte Carlo exercise and then present its results. Section 4 concludes. Proofs of propositions, further details and various extensions have been relegated to Appendices A–D.

2. Aggregation of human capital across population cohorts

2.1. Framework

We are concerned with human capital as a production factor. We consider human capital as a one-dimensional stock of productive skills embodied in an individual and accumulated through schooling and on-the-job learning. Our main focus is on productivity of this stock and not on its remuneration – wages – under different market forms.

We denote the current calendar time as t , and a person's age as τ . A person who is τ years old in year t must have thus been born at $t - \tau$. At time t , there is a continuum of mass $N(t)$ of individuals. We make the following assumption.

Assumption 1. Human capital of an individual τ years old, born at time j , is accumulated using a *linear* production function:

$$\frac{\partial}{\partial \tau} h(j, \tau) = [\lambda \ell_h(j, \tau) + \mu \ell_Y(j, \tau)] h(j, \tau), \quad (1)$$

where $\lambda \geq 0$ denotes the unit productivity of schooling, and $\mu \geq 0$ denotes the unit productivity of on-the-job learning (experience accumulation). $\ell_h(j, \tau) \in [0, 1]$ is the fraction of time spent by an individual born at j and aged τ on formal education, whereas $\ell_Y(j, \tau) \in [0, 1]$ is the fraction of time spent at work. We assume $\ell_h(j, \tau) + \ell_Y(j, \tau) \leq 1$ for all $j, \tau \geq 0$, and take $h(j, 0) \equiv h_0 > 0$.

Even though the current framework singles out the time spent on education and work only, it can easily accommodate other uses of time, such as leisure or childrearing. We thus also allow for retirement. We say that these alternative possibilities are exercised when $\ell_h(j, \tau) + \ell_Y(j, \tau) < 1$.¹

Integrating Eq. (1) with respect to the individual's age yields the formula for the human capital stock of an individual born at $t - \tau$, aged τ :

$$h(t - \tau, \tau) = h_0 \exp \left[\underbrace{\lambda \int_0^\tau \ell_h(t - \tau, s) ds}_{\text{years of schooling}} + \underbrace{\mu \int_0^\tau \ell_Y(t - \tau, s) ds}_{\text{work experience}} \right]. \quad (2)$$

This is directly the micro-Mincer equation, signifying the log-linear relationship between the individuals' human capital and their cumulative stocks of education and work experience. The quadratic experience term, typically also included in Mincerian equations (cf. Heckman et al., 2003), does not appear here because in Eq. (1) we have assumed human capital accumulation to be linear and not concave in work experience.²

Assumption 2. At every age $\tau \geq 0$, the individual may either survive or die. The unconditional survival probability is denoted by $m(\tau)$, with $m(0) = 1$, $\lim_{\tau \rightarrow \infty} m(\tau) = 0$ and with $m(\tau)$ weakly decreasing in its whole domain. The survival probability does not depend on calendar time t .

¹ The current framework can be also straightforwardly generalized to allow for (exponential) human capital depreciation, without altering any of the qualitative results. Please consult Appendix B.

² Although there exist models providing microfoundations for the quadratic experience term in Mincerian equations, Hamlen and Hamlen (2012) claim that it is actually inconsistent with the usual assumptions of utility maximization. These authors argue that other functional forms should be used instead.

Please note that by assuming the survival law to be independent of t , we exclude the possibility of declining mortality due to, e.g., progress in medicine. Accommodating this possibility is left for further research.

Assumption 3. The age structure of the society (the cumulative density function) is stationary. At time t , there are $P(t, \tau) = bN(t - \tau)m(\tau)$ people aged τ in the population. The total population alive at time t is $N(t)$, with

$$N(t) = \int_0^\infty P(t, \tau) d\tau = \int_0^\infty bN(t - \tau)m(\tau) d\tau. \quad (3)$$

The total labor force at time t is computed as

$$\begin{aligned} L(t) &= \int_0^\infty P(t, \tau) \ell_Y(t - \tau, \tau) d\tau \\ &= \int_0^\infty bN(t - \tau)m(\tau) \ell_Y(t - \tau, \tau) d\tau. \end{aligned} \quad (4)$$

By the Law of Large Numbers, the above assumption implies that the aggregate birth rate b and death rate d are constant. This in turn implies a constant population growth rate, and thus $N(t) = N_0 e^{(b-d)t}$. In consequence, the shares of all cohorts in the total population are indeed constant:

$$\begin{aligned} \frac{P(t, \tau)}{N(t)} &= bm(\tau) \frac{N(t - \tau)}{N(t)} \\ &= bm(\tau) e^{-(b-d)\tau}, \quad \text{independently of } t. \end{aligned} \quad (5)$$

Furthermore, the death rate d is computed uniquely from the given survival law $m(\tau)$. If the number of surviving offspring per person, i.e., the birth rate times life expectancy at birth, exceeds unity, then $b > d$ and thus the total population is growing. If it is less than unity, then $b < d$ and thus the population is declining (for the derivation, please refer to Appendix A.6 in [Growiec, 2010](#)).

The first corollary from our [Assumptions 2 and 3](#) is that, under a stationary age structure, and assuming that time profiles of education and work are independent of calendar time t , i.e., $\ell_h(t - \tau, \tau) \equiv \ell_h(\tau)$ and $\ell_Y(t - \tau, \tau) \equiv \ell_Y(\tau)$, it must be the case that the human capital stock of an individual $h(t - \tau, \tau)$ depends only on her age τ , but not on the year when she was born, $t - \tau$. Hence, without loss of generality we can write $h(t - \tau, \tau) \equiv h(\tau)$: even though each individual's human capital may grow exponentially with her age across her whole lifetime, the aggregate human capital in the population stock does not grow with calendar time because dying individuals with high human capital levels are continuously replaced by newborns with little human capital.

Under the aforementioned assumptions of a stationary age structure, it follows that the employment rate in the economy $\frac{L(t)}{N(t)}$ is independent of calendar time t , too:

$$\begin{aligned} \frac{L(t)}{N(t)} &= \frac{\int_0^\infty bN(t - \tau)m(\tau) \ell_Y(\tau) d\tau}{N(t)} \\ &= \int_0^\infty b e^{-(b-d)\tau} m(\tau) \ell_Y(\tau) d\tau. \end{aligned} \quad (6)$$

Let us now place some restrictions on the considered stationary time profiles of education and work. We shall deal with three alternative, naturally understandable scenarios which can be considered as limiting cases of more general time profiles:

- **Scenario “S + W”.** First attend school full time, until you reach S years of age, and then work full time until death:

$$\ell_h(\tau) = \begin{cases} 1, & \tau \leq S, \\ 0, & \tau > S, \end{cases} \quad \ell_Y(\tau) = \begin{cases} 0, & \tau \leq S, \\ 1, & \tau > S. \end{cases} \quad (7)$$

- **Scenario “S + W + R”.** First attend school full time, until you reach S years of age; then work full time, until you reach R years of age, then retire, and stay retired until death:

$$\begin{aligned} \ell_h(\tau) &= \begin{cases} 1, & \tau \leq S, \\ 0, & \tau > S, \end{cases} \\ \ell_Y(\tau) &= \begin{cases} 0, & \tau \in [0, S] \cup [R, +\infty), \\ 1, & \tau \in (S, R). \end{cases} \end{aligned} \quad (8)$$

- **Scenario “Fix”.** Spend fixed fractions of time on schooling and work throughout your entire life:

$$\ell_h(\tau) \equiv \bar{\ell}_h, \quad \ell_Y(\tau) \equiv \bar{\ell}_Y. \quad (9)$$

Finally, to be able to meaningfully aggregate human capital stocks, years of schooling and work experience across heterogeneous population cohorts, we need to ensure that all respective aggregative concepts are appropriately defined. To this end, we also need to take a stance on the degree of substitutability across skill levels in the aggregate production function. Keeping in mind that in the current study all skill heterogeneity comes from demographics only, the following assumption – although restrictive – appears to be a relatively good approximation of reality.

Assumption 4. Labor services provided by individuals of all ages are perfectly substitutable.

2.2. Aggregation across cohorts

The general framework for aggregating human capital across heterogeneous cohorts, building on [Assumptions 1–4](#), is consistent with the following definitions.

Definition 1. The aggregate human capital stock of the labor force at time t is given by:

$$H(t) = \int_0^\infty P(t, \tau) \ell_Y(\tau) h(\tau) d\tau. \quad (10)$$

The average human capital stock in the labor force is $h(t) = \frac{H(t)}{L(t)}$.

Definition 2. Cumulative years of schooling in the labor force at time t are given by:

$$Q(t) = \int_0^\infty P(t, \tau) \ell_Y(\tau) \left(\int_0^\tau \ell_h(s) ds \right) d\tau. \quad (11)$$

The average number of years of schooling in the labor force is $q(t) = \frac{Q(t)}{L(t)}$.

Definition 3. Cumulative work experience in the labor force at time t is given by:

$$X(t) = \int_0^\infty P(t, \tau) \ell_Y(\tau) \left(\int_0^\tau \ell_Y(s) ds \right) d\tau. \quad (12)$$

Average work experience in the labor force is $x(t) = \frac{X(t)}{L(t)}$.

We are now in a position to define the macro-Mincer equation as a relationship between the aforementioned aggregative concepts.

Definition 4. The macro-Mincer equation takes the following form:

$$h(t) = h_0 \exp(\alpha q(t) + \beta x(t)). \quad (13)$$

The parameters $\alpha \geq 0$ and $\beta \geq 0$ will be called the Mincerian schooling coefficient and the Mincerian experience coefficient, respectively.

Note that **Definitions 1–3** refer to the labor force but could be rewritten for the whole population alive at t by dropping the $\ell_Y(\tau)$ term in the integrals. One could then also define the macro-Mincer equation for population-wide averages $h_{POP}(t)$, $q_{POP}(t)$ and $x_{POP}(t)$ analogously to **Definition 4**. Such a case is, however, somewhat less relevant to the macroeconomic analysis which is preoccupied primarily with *productive* uses of human capital. As shown in a working paper version of the current article (**Growiec and Groth, 2012**), however, such an alternative case is much less likely to reproduce the macro-Mincer result than the case considered here.³

We shall now present our results under three specific survival laws $m(\tau)$, and then provide more general considerations relating to the analytical (im)possibility of obtaining the exact macro-Mincer relationship.

2.3. Results under the “perpetual youth” survival law

Apart from **Assumptions 1–4**, let us now also assume the **Blanchard’s (1985)** simple “perpetual youth” survival law $m(\tau) = e^{-d\tau}$, where d is directly the aggregate death rate. It implies that the probability of dying in the next unit of time is independent of the individual’s age. Under this condition, the stationary age structure satisfies $\frac{P(t, \tau)}{N(t)} = be^{-b\tau}$.

We observe that:

- In the scenario “S + W”, $H(t)$ is computed by aggregating the human capital embodied in individuals above the age S . In this scenario, the (constant) share of the working population is equal to $\frac{L(t)}{N(t)} = e^{-bS}$.
- In the scenario “S + W + R”, $H(t)$ is computed by aggregating the human capital embodied in individuals aged between S and R . In this scenario, the share of the working population is equal to $\frac{L(t)}{N(t)} = e^{-bS} - e^{-bR}$.
- The scenario “Fix” has already been considered by **Growiec (2010)**, who however concentrated on $H_{POP}(t)$ and did not compute $H(t)$ for the labor force. With a fixed share of time spent on work irrespective of individuals’ age, it is however clear that $H(t) = \bar{\ell}_Y H_{POP}(t)$, so that the qualitative results for both aggregates are identical up to a multiplicative constant. Also, the share of the working population is naturally $\frac{L(t)}{N(t)} = \bar{\ell}_Y$, implying $h(t) = h_{POP}(t)$.

Finally, we also note that to ensure that the aggregate human capital stock remains finite under the considered survival law, we must assume that $\mu < b$ in the scenario “S + W”, and $\lambda\bar{\ell}_h + \mu\bar{\ell}_Y < b$ in the scenario “Fix”.

The analytical results for this case are presented in **Table 1** and can be summarized in the following proposition.

Table 1
Average human capital, years of schooling and work experience under the “perpetual youth” survival law.

Scenario	S + W	S + W + R	Fix
$h(t)$	$\frac{bh_0}{b-\mu} e^{\lambda S}$	$\frac{bh_0 e^{(\lambda-\mu)S}}{b-\mu} \frac{e^{(\mu-b)S} - e^{(\mu-b)R}}{e^{-bS} - e^{-bR}}$	$\frac{bh_0}{b-\lambda\bar{\ell}_h - \mu\bar{\ell}_Y}$
$q(t)$	S	S	$\frac{\bar{\ell}_h}{b}$
$x(t)$	$\frac{1}{b}$	$\frac{1}{b} - \frac{R-S}{e^{b(R-S)} - 1}$	$\frac{\bar{\ell}_Y}{b}$

³ While most of the empirical literature is indeed preoccupied with human capital in the labor force, in some articles the macro-Mincer approach is applied to the whole population, or at least to the whole working-age population (which is somewhat closer to our definition, cf. **Caselli and Coleman, 2006**). Our analysis in **Growiec and Groth (2012)** strongly suggests that these concepts should not be used interchangeably.

Proposition 1 (Sufficient Conditions for Macro-Mincer). *Let Assumptions 1–4 hold with $\mu \in [0, b)$ and assume the “perpetual youth” survival law. Then the macro-Mincer equation holds:*

- under the “S + W” scenario,
- under the “S + W + R” scenario, but only if there is no on-the-job learning ($\mu = 0$).

In both cases the Mincerian schooling coefficient α equals the individual rate of return to education λ , whereas the Mincerian experience coefficient β is zero. Apart from these two cases, the macro-Mincer equation does not hold.

Hence, even under the “perpetual youth” survival law and even when assuming, as our model does, that all individuals of the same age have identical human capital levels, the scope for the macro-Mincer relationship to hold exactly is very limited. It is obtained only if retirement is absent or if accumulated work experience does not affect workers’ human capital stocks.⁴

It is also interesting to observe that the Mincerian experience coefficient $\beta = 0$ in the “S + W” case even if $\mu > 0$: when individuals’ death hazard is independent of their age, there is no additional function of experience accumulation at the aggregate level beyond partially alleviating the depreciation of aggregate human capital due to the continuous flow of young, unexperienced individuals, with human capital $h_0 e^{\lambda S}$, into the workforce.

2.4. Results under fixed lifetimes

Let us now substitute the **Blanchard (1985)** “perpetual youth” survival law with the assumption that individuals’ lifetimes are deterministically fixed at T , i.e., $m(\tau) = 1$ for $\tau < T$ and $m(\tau) = 0$ for $\tau \geq T$, with $T > S$ and $T \geq R$. Under this condition, the age structure satisfies $\frac{P(t, \tau)}{N(t)} = be^{-(b-d)\tau}$ for $\tau < T$ and zero otherwise. The aggregate death rate d is related to the age T via the equality $T = \frac{\ln b - \ln d}{b-d}$. It is obtained that $b > d$ if and only if $T > 1/b$, and conversely, $b < d$ if $T < 1/b$. In the case $T = 1/b$ we get $b = d$, rendering the population size constant across time. The results for the case of fixed lifetimes are presented in **Table 2**.

Under the currently considered survival law where lifetimes are bounded, aggregate human capital is always finite. From **Table 2**, it should also be clear that under fixed lifetimes, reproducing the macro-Mincer equation is possible if and only if there is no on-the-job learning ($\mu = 0$)⁵:

Proposition 2 (Sufficient Conditions for Macro-Mincer). *Let Assumptions 1–4 hold and assume that the individuals have a fixed lifetime T . Then the macro-Mincer equation holds:*

- under the “S + W” scenario with $\mu = 0$,
- under the “S + W + R” scenario with $\mu = 0$.

In both cases $h(t) = h_0 e^{\lambda S}$, that is, the Mincerian schooling coefficient α equals the individual rate of return to education λ , whereas the Mincerian experience coefficient β is zero.

Apart from these two cases, inconsistent with heterogeneity of the aggregated individuals, the macro-Mincer functional relationship does not hold.

⁴ As shown by **Growiec and Groth (2012)**, the exact macro-Mincer equation is never obtained for the whole population, even if the survival law has the “perpetual youth” property.

⁵ In the case $\lambda = b - d$, the formula $\frac{bh_0}{b-d-\lambda} (1 - e^{(\lambda-(b-d))S})$ should be replaced by $bh_0 S$. Furthermore, if $\mu = b - d$, then the formula $\frac{bh_0}{b-d-\mu} (e^{(\mu-(b-d))S} - e^{(\mu-(b-d))T})$ should be replaced by $bh_0(T - S)$, and the formula $\frac{bh_0}{b-d-\mu} (e^{(\mu-(b-d))S} - e^{(\mu-(b-d))R})$ by $bh_0(R - S)$. Wherever needed, analogous substitutions must be made for all considered survival laws.

Table 2
Average human capital, years of schooling and work experience under fixed lifetimes.

Scenario	S + W	S + W + R	Fix
Exponentially growing or declining population ($T \neq 1/b, b \neq d$)			
$h(t)$	$\frac{(b-d)h_0 e^{(\lambda-\mu)S} e^{\mu(b-d)S} - e^{(\lambda-\mu)S} e^{\mu(b-d)T}}{b-d-\mu}$	$\frac{(b-d)h_0 e^{(\lambda-\mu)S} - e^{(\lambda-\mu)S} e^{\mu(b-d)R}}{b-d-\mu}$	$\frac{bh_0}{b-d-\lambda\bar{e}_h-\mu\bar{e}_y} \left(1 - e^{(\lambda\bar{e}_h+\mu\bar{e}_y-(b-d)T)}\right)$
$q(t)$	S	S	$\frac{\bar{e}_h(1-dT)}{b-d}$
$x(t)$	$\frac{1}{b-d} - \frac{T-S}{e^{(b-d)(T-S)} - 1}$	$\frac{1}{b-d} - \frac{R-S}{e^{(b-d)(R-S)} - 1}$	$\frac{\bar{e}_y(1-dT)}{b-d}$
Constant population size ($T = 1/b, b = d$)			
$h(t)$	$\frac{h_0 e^{\lambda S}}{\mu(T-S)} (e^{\mu(T-S)} - 1)$	$\frac{h_0 e^{\lambda S}}{\mu(R-S)} (e^{\mu(R-S)} - 1)$	$\frac{h_0}{T(\lambda\bar{e}_h+\mu\bar{e}_y)} (e^{(\lambda\bar{e}_h+\mu\bar{e}_y)T} - 1)$
$q(t)$	S	S	$\frac{\bar{e}_h T}{2}$
$x(t)$	$\frac{T-S}{2}$	$\frac{R-S}{2}$	$\frac{\bar{e}_y T}{2}$

2.5. Results under a realistic survival law

Let us now move towards a much more realistic survival law than the two theoretical benchmarks discussed above. As argued by Boucekine et al. (2002) and further discussed by Azomahou et al. (2009), real-world demographics can be reasonably well approximated by a survival law $m : [0, T^*] \rightarrow [0, 1]$ of the following form:

$$m(\tau) = \frac{e^{-k\tau} - a}{1 - a}, \quad a > 1, k < 0. \tag{14}$$

The maximum lifetime of an individual is given under this survival law by $T^* = -\frac{\ln a}{k}$, whereas individuals' life expectancy is equal to $E = \frac{1}{k} + \frac{a \ln a}{(1-a)k}$. As lifetimes are always bounded here, aggregate human capital is always finite.

The results for the current case are presented in Table 3. We again observe that reproducing the macro-Mincer equation is possible if and only if there is no on-the-job learning ($\mu = 0$):

Proposition 3 (Sufficient Conditions for Macro-Mincer). *Let Assumptions 1–4 hold and assume the survival law (14). Then the macro-Mincer equation holds:*

- under the “S + W” scenario with $\mu = 0$,
- under the “S + W + R” scenario with $\mu = 0$.

In both cases $h(t) = h_0 e^{\lambda S}$, that is, the Mincerian schooling coefficient α equals the individual rate of return to education λ , whereas the Mincerian experience coefficient β is zero.

Apart from these two cases, inconsistent with heterogeneity of the aggregated individuals, the macro-Mincer functional relationship does not hold.

2.6. The case without on-the-job learning

The case without on-the-job training ($\mu = 0$) has already stood out as a very specific case in our above calculations. It is no coincidence. Actually, we can straightforwardly generalize our results, yielding the following proposition:

Proposition 4 (Sufficient Condition for Macro-Mincer). *Let Assumptions 1–4 hold and assume $\mu = 0$. Then under the “S + W” and “S + W + R” scenarios, the macro-Mincer equation holds regardless of the underlying survival law $m(\tau)$. The Mincerian schooling coefficient α is equal to the individual rate of return to education λ . The Mincerian experience coefficient β is zero.*

Proof. See Appendix A.

The above result is driven by two facts. First, the “S + W” and “S + W + R” scenarios assume that all working individuals have the same number of years of schooling. Second, the assumption $\mu = 0$ (absence of on-the-job learning) implies that all working individuals also have the same human capital level. Aggregation is thus effected across entirely homogeneous population cohorts. In such a situation, it is no surprise that the Mincerian relationship between human capital and years of schooling is directly transferred from the individual to the aggregate level.

Otherwise, when aggregation is effected across truly heterogeneous population cohorts, the fact that these cohorts also differ in size (due to natural mortality) ceases to be neutral for the aggregation procedure. The average human capital level in the labor force becomes a function of the demographical survival law $m(\tau)$ as well as the time profiles of schooling and work effort. The theoretical question if there exists a specific survival law able to reinstate the general validity of the macro-Mincer equation will be handled in the following subsection.

2.7. Necessary conditions for the macro-Mincer equation

Having obtained some positive results under very specific and empirically implausible assumptions, we shall ask the converse, much more general question: For which survival law $m(\tau)$ will the macro-Mincer functional relationship be recovered exactly from the micro-level Mincerian equation? Instead of that, however, we shall consider an equally important but analytically more tractable question: For which survival law $m(\tau)$ will the simplified macro-Mincer equation, disregarding work experience as in:

$$h(t) = h_0 \exp(\alpha q(t)), \tag{15}$$

be obtained from the micro-level Mincerian equation?

The importance of the last question stems from the fact that the related applied literature⁶ is preoccupied primarily with estimating cross-country rates of return to an additional year of schooling, while considering returns to work experience as a parallel issue, tangent but not central to the empirical arguments discussed in those articles. The decisive difference in analytical tractability, on the other hand, follows from what was already apparent in Table 2: in general, average work experience x can be influenced not only by the survival law $m(\tau)$ and the demographic parameter b , but also – nonlinearly – by years of schooling S and retirement age R .

Growiec (2010) has already addressed the aforementioned question for the scenario “Fix”, showing that recovering the macro-Mincer equation from micro-level Mincerian relationships is not

⁶ See e.g., Bils and Klenow (2000), Krueger and Lindahl (2001) and Psacharopoulos and Patrinos (2004).

Table 3
Average human capital, years of schooling and work experience under the Boucekkine et al. (2002) survival law.

Scenario	S + W	S + W + R
Exponentially growing or declining population, $k \neq b(1 + \frac{a \ln a}{1-a})$, $b \neq d$		
$h(t)$	$h_0 e^{(\lambda-\mu)S} \frac{e^{(\mu-k-b+d)T^*} - e^{(\mu-k-b+d)S} - a \frac{e^{(\mu-b+d)T^*} - e^{(\mu-b+d)S}}{\mu-b+d}}{e^{(\mu-k-b+d)T^*} - e^{(\mu-k-b+d)S} - a \frac{e^{(\mu-b+d)T^*} - e^{(\mu-b+d)S}}{\mu-b+d}}$	$h_0 e^{(\lambda-\mu)S} \frac{e^{(\mu-k-b+d)R} - e^{(\mu-k-b+d)S} - a \frac{e^{(\mu-b+d)R} - e^{(\mu-b+d)S}}{\mu-b+d}}{e^{(\mu-k-b+d)R} - e^{(\mu-k-b+d)S} - a \frac{e^{(\mu-b+d)R} - e^{(\mu-b+d)S}}{\mu-b+d}}$
$q(t)$	S	S
$x(t)$	$\frac{(T^*-S)e^{-(k-b+d)T^*} - e^{-(k-b+d)T^*} - a \left(\frac{(T^*-S)e^{-(b+d)T^*}}{-b+d} - \frac{e^{-(b+d)T^*} - e^{-(b+d)S}}{(-b+d)^2} \right)}{e^{-(k-b+d)T^*} - e^{-(k-b+d)S} - a \frac{e^{-(b+d)T^*} - e^{-(b+d)S}}{-b+d}}$	$\frac{(R-S)e^{-(k-b+d)R} - e^{-(k-b+d)R} - a \left(\frac{(R-S)e^{-(b+d)R}}{-b+d} - \frac{e^{-(b+d)R} - e^{-(b+d)S}}{(-b+d)^2} \right)}{e^{-(k-b+d)R} - e^{-(k-b+d)S} - a \frac{e^{-(b+d)R} - e^{-(b+d)S}}{-b+d}}$
Constant population size, $k = b(1 + \frac{a \ln a}{1-a})$, $b = d$		
$h(t)$	$h_0 e^{(\lambda-\mu)S} \frac{e^{(\mu-k)T^*} - e^{(\mu-k)S} - a \frac{e^{\mu T^*} - e^{\mu S}}{\mu}}{e^{-kT^*} - e^{-kS} - a(T^*-S)}$	$h_0 e^{(\lambda-\mu)S} \frac{e^{(\mu-k)R} - e^{(\mu-k)S} - a \frac{e^{\mu R} - e^{\mu S}}{\mu}}{e^{-kR} - e^{-kS} - a(R-S)}$
$q(t)$	S	S
$x(t)$	$\frac{(T^*-S)e^{-kT^*} - e^{-kT^*} - a \frac{(T^*-S)^2}{2}}{e^{-kT^*} - e^{-kS} - a(T^*-S)}$	$\frac{(R-S)e^{-kR} - e^{-kR} - a \frac{(R-S)^2}{2}}{e^{-kR} - e^{-kS} - a(R-S)}$
Scenario	Fix	
Exponentially growing or declining population, $k \neq b(1 + \frac{a \ln a}{1-a})$, $b \neq d$		
$h(t)$	$h_0 \frac{e^{(\lambda \bar{\ell}_h + \mu \bar{\ell}_Y - k - b + d)T^*} - 1 - a \frac{e^{(\lambda \bar{\ell}_h + \mu \bar{\ell}_Y - b + d)T^*} - 1}{\lambda \bar{\ell}_h + \mu \bar{\ell}_Y - k - b + d}}{e^{(\lambda \bar{\ell}_h + \mu \bar{\ell}_Y - k - b + d)T^*} - 1 - a \frac{e^{(-b+d)T^*} - 1}{-b+d}}$	
$q(t)$	$\bar{\ell}_h \frac{T^* e^{-(k-b+d)T^*} - e^{-(k-b+d)T^*} - a \left(\frac{T^* e^{-(b+d)T^*}}{-b+d} - \frac{e^{-(b+d)T^*} - 1}{(-b+d)^2} \right)}{e^{-(k-b+d)T^*} - 1 - a \frac{e^{-(b+d)T^*} - 1}{-b+d}}$	
$x(t)$	$\bar{\ell}_Y \frac{T^* e^{-(k-b+d)T^*} - e^{-(k-b+d)T^*} - a \left(\frac{T^* e^{-(b+d)T^*}}{-b+d} - \frac{e^{-(b+d)T^*} - 1}{(-b+d)^2} \right)}{e^{-(k-b+d)T^*} - 1 - a \frac{e^{-(b+d)T^*} - 1}{-b+d}}$	
Constant population size, $k = b(1 + \frac{a \ln a}{1-a})$, $b = d$		
$h(t)$	$h_0 \frac{e^{(\lambda \bar{\ell}_h + \mu \bar{\ell}_Y - k)T^*} - 1 - a \frac{e^{(\lambda \bar{\ell}_h + \mu \bar{\ell}_Y)T^*} - 1}{\lambda \bar{\ell}_h + \mu \bar{\ell}_Y - k}}{e^{-kT^*} - 1 - aT^*}$	
$q(t)$	$\bar{\ell}_h \frac{T^* e^{-kT^*} - e^{-kT^*} - a \frac{(T^*)^2}{2}}{e^{-kT^*} - 1 - aT^*}$	
$x(t)$	$\bar{\ell}_Y \frac{T^* e^{-kT^*} - e^{-kT^*} - a \frac{(T^*)^2}{2}}{e^{-kT^*} - 1 - aT^*}$	

possible unless the survival function depends on $\bar{\ell}_h$ in one crucial and arguably implausible way. For the scenario “S + W” considered in the present article, however, it is possible if the survival law satisfies the “perpetual youth” property. We have also already shown that if one disregards on-the-job learning (by assuming $\mu = 0$), then this result also follows in the “S + W” and “S + W + R” scenarios under a wide range of survival laws. In that case, however, all individuals in the labor force share exactly the same human capital level $h_0 e^{\lambda S}$, and it is precisely this homogeneity that drives the result.

It turns out that if $\mu > 0$, so there is some heterogeneity in human capital across working cohorts, then the simplified macro-Mincer equation can be reproduced under the “S + W” scenario *only* in the “perpetual youth” case, and cannot be reproduced under the “S + W + R” scenario at all. The following proposition holds.

Proposition 5 (Necessary Conditions for Macro-Mincer with S + W). *Let Assumptions 1–4 hold with $\mu \in (0, b)$ and assume that the simplified macro-Mincer equation holds. Then, under the “S + W” scenario where the individuals stay at school until the age S and then work full-time until death, the survival law must be $m(\tau) = e^{-d\tau}$, i.e., it must satisfy the “perpetual youth” property. The implied macro-Mincer equation is $h(t) = \frac{bh_0}{b-\mu} e^{\lambda S}$.*

Proof. See Appendix A.

This result, by linking the macro-Mincer functional relationship to the “perpetual youth” survival law, seriously limits the

applicability of the relationship: the “perpetual youth” survival law is highly implausible empirically because it implies that irrespective of age, individuals face the same probability of dying next year. According to empirical evidence (cf. e.g., Boucekkine et al., 2002), this is clearly not the case, not even approximately.⁷

We shall now pass to the “S + W + R” scenario. It turns out that if $\mu > 0$, so there is some heterogeneity in human capital across working cohorts, then the simplified macro-Mincer equation cannot be reproduced under the “S + W + R” scenario (with any fixed S and R) at all.

Proposition 6 (Macro-Mincer Impossible with S + W + R). *Let Assumptions 1–4 hold with $\mu \in (0, b)$. Then under the “S + W + R” scenario where the individuals stay at school until age S and then work full-time until retirement age R, there is no admissible survival law compatible with the simplified macro-Mincer equation.*

Proof. See Appendix A.

This theoretical result further restricts the practical applicability of the (simplified) macro-Mincer relationship between average human capital and average years of schooling. It turns out to be generally inconsistent with the presence of retirement.

⁷ It might be an approximate description of survival laws only in very poor, war-ridden regions, or ancient times.

3. The macro-Mincer equation as an approximation

Having obtained the above theoretical results, with rather negative conclusions for the exact validity of the macro-Mincer equation, let us now ask a closely related question which is certainly vital from the point of view of applied research: How well does the macro-Mincer equation approximate the theoretically derived “true” relationship between average human capital and average years of schooling and work experience, despite being misspecified? In the current section, this quantitative question will be answered numerically. We shall first define the setup of our study and the baseline calibration of the underlying parameters. Then we will present several stylized examples and, finally, pass to the description of our comprehensive results based on a Monte Carlo study. Details and extensions have been relegated to [Appendix C](#).

Our main finding is that, in spite of the theoretical misspecification, the approximation precision of the macro-Mincer equation is remarkably good. We also observe that macro- and micro-level returns to schooling (evaluated under the assumption that wages are proportional to human capital) tend to be typically roughly equal. On the other hand, the approximation quality becomes rather unsatisfactory in the case of the *simplified* macro-Mincer equation which omits the work experience term x , and especially so if the underlying experience coefficient μ is large.⁸ In this latter case, we also find notable discrepancies between the (inconsistently estimated) Mincerian schooling coefficient and micro-level returns to schooling.

On the other hand, it should also be noted that this main finding depends crucially on the assumption of no cross-country heterogeneity in returns to schooling and work experience (λ and μ). In the likely case that individual-level returns to education and work experience are *not* the same in all countries, the discrepancy can in fact be substantial (see [Appendix D](#)).⁹

3.1. Setup of the numerical study

The numerical calculations are based on our theoretical framework from the previous section, allowing for heterogeneity of human capital stocks across cohorts, coupled with a realistic survival law put forward by [Boucekkine et al. \(2002\)](#), cf. Eq. (14).

We concentrate on the “S + W + R” scenario—which is arguably a reasonable first approximation of the time profiles of schooling and work effort observed empirically around the world. We fix $t = 0$, so that $N = N_0$ and $P(t, \tau) \equiv P(\tau)$. All functions defined originally on the real domain, i.e., $m(\tau)$, $P(\tau)$, $h(\tau)$, $\ell_h(\tau)$, $\ell_y(\tau)$, are now discretized, i.e., evaluated on a finite grid of points in the domain. The parameters of our framework are calibrated so that they roughly match their respective estimates based on real-world data. The baseline calibration will be discussed in the following subsection.

⁸ This last result is valid also when one controls for (exponential) human capital depreciation with a rate $\delta \in (0, \mu)$. Please consult [Appendix C](#).

⁹ We relegate this extension to [Appendix D](#) because it has not been included in our analytical study which concentrates only on skill heterogeneity coming from the demographics. There is mounting evidence that returns to schooling and experience can indeed vary largely across countries, however, with quantitative consequences for the role of human capital in explaining the observed gaps in labor productivity ([Psacharopoulos and Patrinos, 2004](#); [Caselli, 2005](#); [Lagakos et al., 2012](#)). We find that the approximation quality indeed deteriorates very quickly with increasing cross-country heterogeneity in returns to schooling or work experience. However, in a sense, this negative result follows *by construction*: we are fitting a single macro-Mincer equation to a sample of countries endowed with heterogeneous schooling and work experience coefficients. This auxiliary result also indirectly implies that the empirical validity of the macro-Mincer equation should be rather limited unless the sample is conveniently restricted by the researcher.

For every parameter configuration, we are going to compute the “true” average human capital stock in the labor force h , as well as cumulative years of schooling q and cumulative work experience x , based on our analytical framework. We shall identify each parameter configuration with a “country”, assuming that the micro-Mincer equation holds exactly in every country, with equal Mincerian coefficients, and there is no cross-border migration of individuals between countries.

Obviously, if every country in the sample were endowed with exactly the same survival law $m(\tau)$, years of schooling S , retirement age R , magnitude of returns to education λ , and returns to work experience μ , they would be homogeneous in terms of their aggregate human capital stocks as well. In such a case, the macro-Mincer equation would be unidentified. Hence, to assess the approximation precision of the macro-Mincer equation, we need to consider a group of countries differing in at least one parameter. In our stylized examples, we will either vary each parameter separately or covary them jointly, in selected configurations. We shall first assume that these parameters are equidistributed along a predefined interval. Later, in our Monte Carlo study, they will be drawn from a certain pre-defined (multivariate Gaussian) joint distribution.

Having obtained the direct, precise measures of average human capital stocks, we shall approximate them with the macro-Mincer (log-linear) equation. The parameters of the approximating equation will be identified by estimating the regression:

$$\ln h_i = c + \alpha q_i + \beta x_i + \varepsilon_i \quad (16)$$

with ordinary least squares, based on artificial data computed from the “true” model ($N = 100$ observations). The “goodness-of-fit” of the macro-Mincer equation to the “true” model will be assessed by comparing the R^2 of the regression as well as within-sample mean absolute percentage error (MAPE). We shall also compare our macro estimates of returns to schooling α with the micro-level return λ (which is known *a priori*), to see if they are under- or overestimated in the macro data. The same procedure will be applied to β and μ , respectively.

Thus we will not only check if the log-linear functional specification fits the true model well or not, but also assess whether it is reasonable to carry forward the micro-level magnitudes of returns λ and μ to macro data. This is important because it has been argued in the literature that even if the macro-Mincer relationship is maintained, indirect effects appearing upon aggregation might lead to differences between micro- and macro-level Mincerian rates of return (see [Hsieh and Klenow, 2010](#), for a discussion). Moreover, microeconomic estimates of returns to schooling (e.g., [Psacharopoulos and Patrinos, 2004](#)) are frequently used in the literature when constructing aggregate human capital stocks (e.g., [Klenow and Rodríguez-Clare, 1997](#), [Hsieh and Klenow, 2010](#)). Reassuringly, at least in the case of the full macro-Mincer specification, we do not find any major discrepancies here.

Concurrently, however, we shall also report the respective numerical results for *simplified* macro-Mincer equation, obtained by omitting the experience variable, i.e., setting $\beta = 0$ in Eq. (16). Comparing the estimates for this simplified specification with their counterparts from the fully specified macro-Mincer equation, we shall assess the magnitude of omitted variable bias incurred in the estimation of the simplified equation—and we shall find it to be quantitatively important.

3.2. Calibration

The baseline calibration for parameters used in our numerical exercise is the following: (a) following [Boucekkine et al. \(2002\)](#), we assume $a = 5.44$, $k = -0.0147$, implying a life expectancy of 73 years and maximum lifespan of 115 years; (b) the population

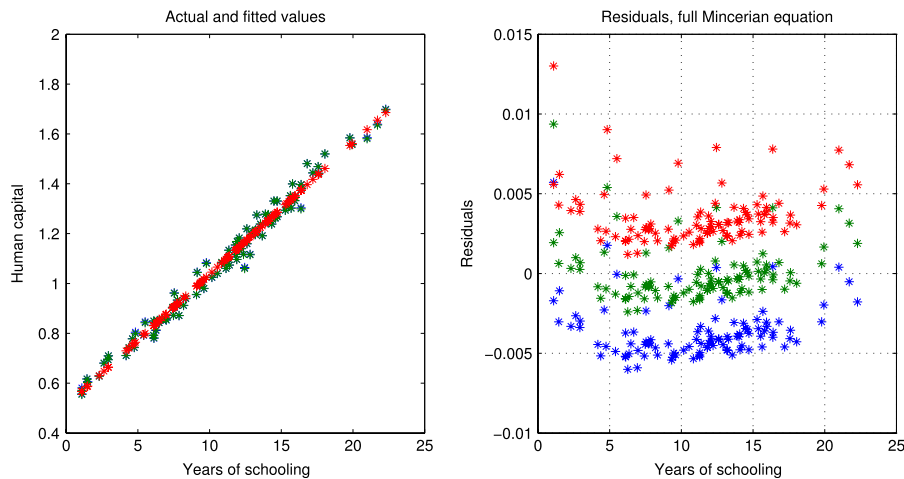


Fig. 1. Quality of approximation of average human capital with the macro-Mincer equation: the case of randomly varying S , R , a . *Notes.* Estimated equations: $\widehat{\ln h_i} = -0.0591 + 0.0600q_i + 0.0250x_i$, with $R^2 \approx 1$ and MAPE = 0.1%; $\ln h_i = 0.5107 + 0.0525q_i$ with $R^2 \approx 0.98$ and MAPE = 2.62%. Left: blue asterisks denote true data points, green asterisks denote their macro-Mincer approximations, and red asterisks denote their simplified macro-Mincer approximations. Right: blue and red asterisks delineate the boundaries of a 95% confidence interval around green asterisks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

growth rate is set at $n = 0.02$ per annum and the birth rate b is set to match this statistic given the assumed survival law; (c) initial human capital is normalized to unity, $h_0 = 1$, without loss of generality; (d) the micro-level rate of return to education is fixed at $\lambda = 0.06$ per annum¹⁰; (e) the rate of return to work experience is assumed to be $\mu = 0.02$ per annum; (f) the number of years of schooling is set to $S = 10$ (ignoring 6 preschool years); (g) the retirement age is set at $R = 59$ (again, ignoring 6 preschool years)—so that the working age is calibrated as 16–65 years.

Unless stated otherwise, the above calibration will be maintained at all times. In particular, whenever we consider the consequences of heterogeneity in a given set of parameters, all other parameters of the theoretical framework are fixed at their aforementioned baseline values.

3.3. Heterogeneity in years of schooling, retirement age, and life expectancy

As demonstrated in Appendix C, in a sample of countries differing only in the number of obligatory years of schooling S , only in the retirement age R , or only in parameters of the survival law, a and k (which map uniquely into the life expectancy E and the maximum lifespan T^*), the estimated macro-Mincer equation fits the “true” relationship between average human capital, years of schooling, and work experience almost perfectly. The simplified macro-Mincer equation, which omits the work experience variable in the regressions, is also a very tight approximation, albeit the residuals are somewhat larger in its case.

Hence, one may tentatively conclude that as long as returns to education λ and work experience μ are fixed across countries,¹¹ and the experience coefficient μ is sufficiently low, aggregation across heterogeneous cohorts does not distort the macro-Mincer relationship sufficiently strongly to call for a more general model.

3.4. Larger extent of cross-country heterogeneity

The next step of our numerical analysis consists in assessing the approximation precision of the macro-Mincer functional relationship in a case where the country-specific number of years of schooling S , retirement age R , and survival law parameter a , are drawn randomly from Gaussian distributions. The parameter R is generated independently of the two other variables, whereas S and a are assumed to be positively correlated, capturing the fact that in reality, wealthier countries tend to have both better education outcomes and a greater life expectancy.¹²

As shown in Fig. 1, the macro-Mincer equation fits the theoretically derived “true” human capital levels in the labor force remarkably well in the current case, despite substantial cross-country heterogeneity. The reason is that the two key parameters – rates of return λ and μ – are assumed to be the same across the countries, and most importantly, the experience coefficient μ is calibrated at a sufficiently low value of 0.02, making the differences in human capital levels across the aggregated cohorts manageably small.

In comparison to the fully specified macro-Mincer equation, its simplified version which omits the experience variable provides a relatively inferior fit to the data—though still probably acceptable in empirical applications.

In a sense, all these results could be anticipated from our theoretical findings because they draw from the fact that μ is sufficiently low in our numerical exercises: when $\mu \approx 0$ (or $\mu \approx \delta$ in the case allowing for human capital depreciation, discussed in Appendix B), then – extrapolating from Proposition 4 by continuity of the general model – human capital levels should be approximately equal across all working-age cohorts, and thus the macro-Mincer equation should fit the data well. It is thus of crucial importance to check the goodness of fit of the macro-Mincer equation for greater magnitudes of μ as well.

As demonstrated in Fig. 2, the conclusion remains positive in such a case as well: the macro-Mincer equation fits the true relationship between average human capital h , average years of schooling q , and average work experience x in the labor force

¹⁰ According to Psacharopoulos and Patrinos (2004) data, the mean rate of return to an additional year of education in European Union countries amounts to 6.5%, with a standard deviation of 1.9%, and goes up to 9.6% in the whole sample, displaying substantial cross-country heterogeneity (standard deviation amounts to 4.3%).

¹¹ The alternative case is considered in Appendix D.

¹² As confirmed in a series of further numerical experiments, the current results tend to be robust to arbitrary changes in the assumed multidimensional distribution of S , R , a , and k , as long as the rates of return λ and μ are kept fixed.

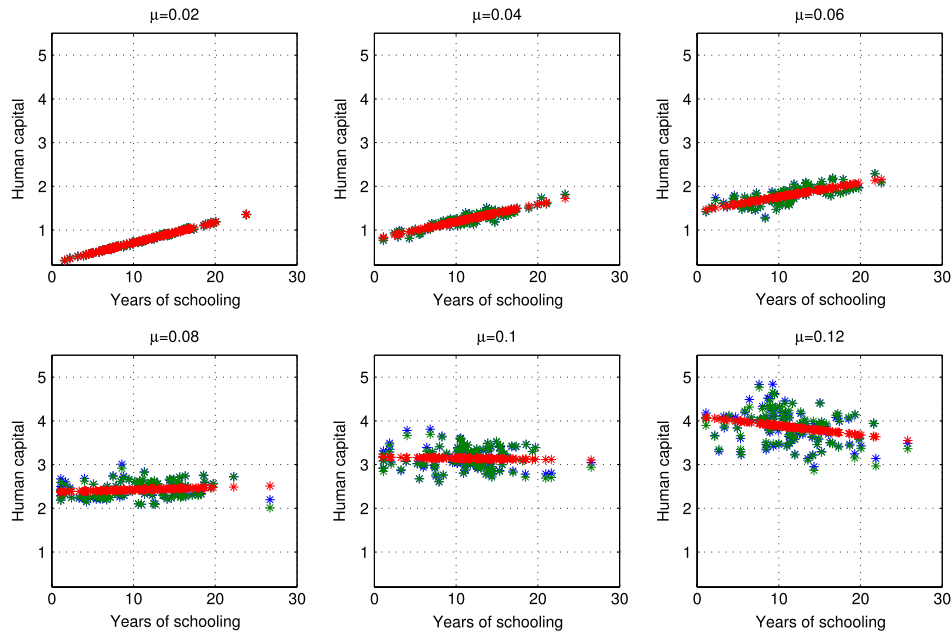


Fig. 2. Quality of approximation of average human capital h with the macro-Mincer equation: dependence on the coefficient of returns to work experience μ (increasing from 0.02 to 0.12). *Notes:* Blue asterisks denote true data points, green asterisks denote their macro-Mincer approximations, and red asterisks denote their simplified macro-Mincer approximations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

reasonably well not only for small values of μ , but even if μ is as large as 0.12 which is beyond any empirical estimates.

On the other hand, the goodness of fit of the *simplified* macro-Mincer equation to the aggregate data deteriorates very quickly with μ : at $\mu = 0.08$, the R^2 of the estimated simplified macro-Mincer equation is already close to zero, and the estimated schooling coefficient becomes statistically insignificant, and – for larger μ – even negative, reflecting both the incorrectly specified functional form of the estimated equation and the omitted experience variable. We conclude that for the macro-Mincer equation to be useful in empirical applications, it is of great importance that the Mincerian experience coefficient is not restricted to zero.

3.5. Monte Carlo study

Having illustrated the approximation precision of the true functional relationship between average human capital, years of schooling, and work experience with the macro-Mincer functional relationship on the basis of a few stylized examples, let us now address this issue more systematically. To this end, we have carried out a Monte Carlo study based on $B = 2000$ iterations of the numerical exercise described above, with the three parameters S, R, a randomly varying, and accounting for human capital depreciation with $\delta = 0.01$. In each of the iterations, the sample consists of 100 hypothetical countries, for which we compute the “true” human capital stocks.¹³ Then we estimate the macro-Mincer equation across the countries. We collect the estimates of the macro-Mincer equation from each iteration of the Monte Carlo procedure, as well as goodness of fit measures, i.e., the R^2 and MAPE, reported in Table 4.

¹³ Conceptually, this numerical exercise covers also several cases where some of these parameters – such as the length of the schooling period S or the retirement age R – are allowed to be chosen optimally by utility-maximizing agents. Such cases are included provided that the decision makers operate in a sufficiently uncertain environment. Otherwise, one could potentially obtain functional relationships between the considered parameters. This is left for further research.

Table 4

Results of the Monte Carlo study—dependence on returns to work experience (parameter μ).

	R^2	MAPE [%]	c	α	β
Micro-level benchmarks					
			0	0.05	$\mu - \delta$
$\mu = 0.02$					
Mean	1.0000	0.0427	-0.0140	0.0500	0.0112
S.D.	0.0000	0.0062	0.0009	0.0000	0.0000
$\mu = 0.04$					
Mean	0.9996	0.2261	-0.1355	0.0501	0.0414
S.D.	0.0001	0.0290	0.0090	0.0001	0.0004
$\mu = 0.06$					
Mean	0.9960	0.4455	-0.3887	0.0504	0.0825
S.D.	0.0014	0.0581	0.0262	0.0003	0.0012
$\mu = 0.08$					
Mean	0.9889	0.6348	-0.7556	0.0510	0.1332
S.D.	0.0034	0.0767	0.0510	0.0006	0.0024
$\mu = 0.10$					
Mean	0.9862	0.7774	-1.2023	0.0517	0.1912
S.D.	0.0036	0.0959	0.0771	0.0010	0.0036
$\mu = 0.12$					
Mean	0.9860	0.8789	-1.6842	0.0526	0.2536
S.D.	0.0034	0.1101	0.1091	0.0014	0.0051

Notes: The regression parameters c, α, β are defined in Eq. (16).

It is instructive to repeat our Monte Carlo study for various magnitudes of returns to work experience μ , which is the key parameter driving the extent of heterogeneity across the aggregated cohorts. For the case of aggregating human capital across the labor force, it is confirmed that if μ (or more directly, $\mu - \delta$) is low, then the macro-Mincer equation fits the data remarkably well, and its fit deteriorates somewhat with increases in μ , though it remains useful for empirical applications even if μ is as large as 0.12.

Apart from the general finding of the very good fit of the macro-Mincer equation to our data, it also stands out in Table 4 that the estimated value of the Mincerian schooling coefficient $\alpha \approx \lambda - \delta$. It must be concluded that despite human capital heterogeneity across population cohorts, average social returns to schooling in the labor force remain very close to private returns, at

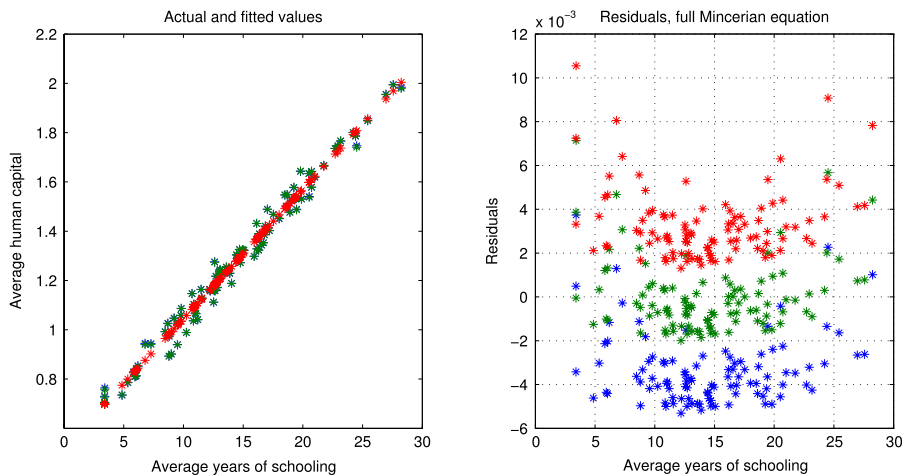


Fig. 3. Quality of approximation of average human capital with the macro-Mincer equation: the case of randomly varying S , R , a , with intra-cohort heterogeneity (fixed share of high-skill individuals $\psi = 0.5$, each having $\tilde{S} = 5$ additional years of schooling). *Notes.* Estimated equations: $\ln \hat{h}_i = -0.0400 + 0.0599q_i + 0.0246x_i$, with $R^2 \approx 1$ and MAPE = 0.09%; $\ln \hat{h}_i = 0.5257 + 0.0524q_i$ with $R^2 \approx 0.98$ and MAPE = 2.39%. Left: blue asterisks denote true data points, green asterisks denote their macro-Mincer approximations, and red asterisks denote their simplified macro-Mincer approximations. Right: blue and red asterisks delineate the boundaries of a 95% confidence interval around green asterisks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

least under the baseline calibration of other model parameters. We observe, however, that the macro-level estimate of the Mincerian experience coefficient β systematically exceeds its micro-level counterpart $\mu - \delta$.

3.6. Allowing for intra-cohort heterogeneity

Another issue which can potentially compromise the goodness of fit of the macro-Mincer equation to cross-country data on average years of schooling and average human capital, is intra-cohort heterogeneity. In reality, as opposed to the theoretical framework discussed in Section 2, every cohort consists of an entire distribution of individuals who may possess less or more years of schooling, less or more work experience, and thus less or more accumulated human capital. Although addressing this broad issue is beyond the scope of the current paper, we are nevertheless able to provide at least a partial argument that the macro-Mincer equation can in fact be a quantitatively good approximation of the theoretically derived “true” relationship even with intra-cohort heterogeneity.

In the current numerical exercise, we shall assume that the population consists of a given fraction ψ of high-skill individuals who study for $S + \tilde{S}$ years before entering the workforce (with $\tilde{S} > 0$), and a remaining fraction $1 - \psi$ of low-skill individuals who only complete S years of education. Average years of schooling q , work experience x and human capital h in the labor force is then computed as a weighted average of the respective stocks pertaining to the high-skill and low-skill individuals.

Fig. 3 repeats the numerical exercise from Fig. 1, assuming a fixed share $\psi = 0.5$ of high-skill individuals and a fixed number $\tilde{S} = 5$ of additional years of schooling completed by the high-skill population.

We find that, as long as returns to schooling and work experience are common across all countries and both skill groups, intra-cohort heterogeneity does not quantitatively affect the goodness of fit of the macro-Mincer equation to the theoretically derived “true” data: the approximation remains very close. In Appendix C.5 we confirm this result also for the cases allowing for cross-country heterogeneity in the shares of high-skill individuals ψ or the number of additional years of schooling \tilde{S} (albeit with some additional caveats).

4. Conclusion

This article demonstrates, based on a general framework for calculating the aggregate human capital stock under heterogeneity across population cohorts, that the Mincerian (log-linear) functional relationship between human capital, years of schooling and work experience is generally lost upon aggregation. It can be recovered exactly only in a few very special cases, which either require the aggregated individuals to have equal human capital stocks, or – in a scenario where individuals first attend school full time and then work full time until death – require that the demographic survival law has the empirically implausible “perpetual youth” property (Blanchard, 1985).

On the other hand, we have also shown numerically that the macro-Mincer equation can still be perceived as a quantitatively reasonable approximation of the theoretically derived “true” relationship between average human capital stocks, years of schooling, and work experience, at least if the observed heterogeneity comes only from differences in the number of years of schooling, retirement age, or demographic survival laws.

Appendix A. Proofs of propositions

Proof of Proposition 3. Using Eqs. (5)–(6), under the “S + W” scenario we have:

$$\begin{aligned} h(t) &= \int_0^\infty \frac{P(t, \tau)}{L(t)} \ell_Y(t - \tau, \tau) h(t - \tau, \tau) d\tau \\ &= \int_S^\infty h_0 e^{\lambda S} b e^{-(b-d)\tau} m(\tau) \frac{N(t)}{L(t)} d\tau \\ &= h_0 e^{\lambda S} \frac{\int_S^\infty b e^{-(b-d)\tau} m(\tau) d\tau}{\int_S^\infty b e^{-(b-d)\tau} m(\tau) d\tau} = h_0 e^{\lambda S}. \end{aligned} \quad (17)$$

Using Eqs. (5)–(6) again, under the “S + W + R” scenario we have:

$$\begin{aligned} h(t) &= \int_0^\infty \frac{P(t, \tau)}{L(t)} \ell_Y(t - \tau, \tau) h(t - \tau, \tau) d\tau \\ &= \int_S^R h_0 e^{\lambda S} b e^{-(b-d)\tau} m(\tau) \frac{N(t)}{L(t)} d\tau \\ &= h_0 e^{\lambda S} \frac{\int_S^R b e^{-(b-d)\tau} m(\tau) d\tau}{\int_S^R b e^{-(b-d)\tau} m(\tau) d\tau} = h_0 e^{\lambda S}. \quad \blacksquare \end{aligned} \quad (18)$$

Proof of Proposition 4. Under the “S + W” scenario, a person of age $\tau \geq S$ has human capital $h(t - \tau, \tau) = h_0 e^{\lambda S + \mu(\tau - S)}$. Upon aggregation, we have:

$$h(t) = \int_S^\infty h(t - \tau, \tau) \frac{P(t, \tau)}{L(t)} d\tau = h_0 e^{(\lambda - \mu)S} \frac{\int_S^\infty b e^{(\mu - (b-d)\tau)} m(\tau) d\tau}{\int_S^\infty b e^{-(b-d)\tau} m(\tau) d\tau}. \tag{19}$$

We shall use the notation:

$$\varphi(S) = \frac{\int_S^\infty b e^{(\mu - (b-d)\tau)} m(\tau) d\tau}{\int_S^\infty b e^{-(b-d)\tau} m(\tau) d\tau} \tag{20}$$

which implies $h(t) = \varphi(S) \cdot h_0 e^{(\lambda - \mu)S}$. Since $\mu > 0$ and $\tau \geq S$ in all the considered integrals, it is easily verified that for all $S \geq 0$, $\frac{\varphi(S)}{e^{\mu S}} > 1$. Furthermore, applying l'Hôpital's rule twice, we obtain:

$$\begin{aligned} \lim_{S \rightarrow +\infty} \frac{\varphi(S)}{e^{\mu S}} &= \lim_{S \rightarrow +\infty} \frac{\int_S^\infty e^{(\mu - (b-d)\tau)} m(\tau) d\tau}{e^{\mu S} \int_S^\infty e^{-(b-d)\tau} m(\tau) d\tau} \\ &= \frac{1}{1 - \lim_{S \rightarrow +\infty} \frac{\mu \int_S^\infty e^{-(b-d)\tau} m(\tau) d\tau}{m(S) e^{-(b-d)S}}} \\ &= \frac{1}{1 + \lim_{S \rightarrow +\infty} \frac{\mu}{\frac{m'(S)}{m(S)} - (b-d)}}. \end{aligned} \tag{21}$$

We are looking for functional specifications of $m(\tau)$ for which $\varphi(S) = Ge^{HS}$ for some $G > 0$ and $H \in \mathbb{R}$, so that consistently with (19), the relationship between aggregate human capital and aggregate years of schooling S is of a log-linear type. Assuming this functional relationship, it follows that

$$\lim_{S \rightarrow +\infty} \frac{\varphi(S)}{e^{\mu S}} = \lim_{S \rightarrow +\infty} Ge^{(H - \mu)S}. \tag{22}$$

Since $\frac{\varphi(S)}{e^{\mu S}} > 1$ for all $S \geq 0$, then it must be the case that $H \geq \mu$. Furthermore, one must set $G > 1$ so that $\varphi(0) > 1$. The cases $H > \mu$ and $H = \mu$ will be addressed separately.

We shall now pass to the central part of the proof. Posing $\varphi(S) = Ge^{HS}$ and rearranging in (20) yield:

$$\int_S^\infty b e^{(\mu - (b-d)\tau)} m(\tau) d\tau = Ge^{HS} \int_S^\infty b e^{-(b-d)\tau} m(\tau) d\tau. \tag{23}$$

Eq. (23) is a *functional identity* and thus it holds for all $S \geq 0$. It is also possible to differentiate both sides of (23) with respect to S . Doing this twice and rearranging terms, we obtain:

$$\frac{m'(S)}{m(S)} = \frac{(\mu - H - b + d)e^{\mu H S} - G(d - b + H)}{G - e^{\mu H S}}. \tag{24}$$

Consider first the case $H > \mu$. In such a case we obtain $\lim_{S \rightarrow +\infty} Ge^{(H - \mu)S} = +\infty$. Coupled with Eq. (21), this implies:

$$\lim_{S \rightarrow +\infty} \frac{m'(S)}{m(S)} = b - d - \mu. \tag{25}$$

Comparing (25) and (24) we obtain:

$$\lim_{S \rightarrow +\infty} \frac{m'(S)}{m(S)} = b - d - H = b - d - \mu, \tag{26}$$

and thus $H = \mu$, a contradiction. The case $H > \mu$ is thus ruled out.

Now, consider the remaining case $H = \mu$. Inserting the condition $H = \mu$ into (24) and simplifying we obtain:

$$\frac{m'(S)}{m(S)} = -\frac{(G - 1)(-b + d) + G\mu}{G - 1} = (b - d) - \frac{G\mu}{G - 1}. \tag{27}$$

Solving this differential equation for $m(S)$ and using the border condition $m(0) = 1$, we obtain the only survival law $m(\tau)$ consistent with the macro-Mincer formulation:

$$m(S) = \exp\left(\left((b - d) - \frac{G\mu}{G - 1}\right)S\right), \quad \forall (S \geq 0). \tag{28}$$

Please note that this survival law is exponential and thus has the “perpetual youth” property. Let us now make the parametrization of $m(\tau)$ in Eq. (28) consistent with its interpretation, i.e. ensure that the implied death rate is indeed equal to d . Under a stationary age structure, this is achieved by checking the following demographic identity:

$$N(t) = \int_{-\infty}^t bN(s)m(t - s)ds = N_0 e^{(b-d)t}. \tag{29}$$

From (28) and (29) it follows that

$$\int_{-\infty}^t b \exp\left(\left(\frac{G}{1 - G}\right)\mu(t - s)\right) ds = 1. \tag{30}$$

Computing the last integral reveals that $G = \frac{b}{b - \mu} > 1$. Plugging this into (28), we obtain $m(\tau) = e^{-d\tau}$. Also, $\varphi(S) = \frac{b}{b - \mu} e^{\mu S}$ and thus $h(t) = \frac{bh_0}{b - \mu} e^{\lambda S}$ so that the macro-Mincer equation holds with the Mincerian coefficient λ . ■

Proof of Proposition 5. Under the “S + W + R” scenario, a person of age $\tau \in [S, R]$ has human capital $h(t - \tau, \tau) = h_0 e^{\lambda S + \mu(\tau - S)}$. Upon aggregation, we have:

$$h(t) = \int_S^R h(t - \tau, \tau) \frac{P(t, \tau)}{L(t)} d\tau = h_0 e^{(\lambda - \mu)S} \frac{\int_S^R b e^{(\mu - (b-d)\tau)} m(\tau) d\tau}{\int_S^R b e^{-(b-d)\tau} m(\tau) d\tau}. \tag{31}$$

We shall use the notation:

$$\varphi(S) = \frac{\int_S^R b e^{(\mu - (b-d)\tau)} m(\tau) d\tau}{\int_S^R b e^{-(b-d)\tau} m(\tau) d\tau} \tag{32}$$

which implies $h(t) = \varphi(S) \cdot h_0 e^{(\lambda - \mu)S}$. Since $\mu > 0$ and $\tau \geq S$ in all the considered integrals, it is easily verified that for all $S \in [0, R)$, $\frac{\varphi(S)}{e^{\mu S}} > 1$. Furthermore, applying l'Hôpital's rule, we obtain:

$$\begin{aligned} \lim_{S \rightarrow R} \frac{\varphi(S)}{e^{\mu S}} &= \lim_{S \rightarrow R} \frac{\int_S^R e^{(\mu - (b-d)\tau)} m(\tau) d\tau}{e^{\mu S} \int_S^R e^{-(b-d)\tau} m(\tau) d\tau} \\ &= \frac{1}{1 - \lim_{S \rightarrow R} \frac{\mu \int_S^R e^{-(b-d)\tau} m(\tau) d\tau}{m(S) e^{-(b-d)S}}} = 1. \end{aligned} \tag{33}$$

The last equality follows from the fact that $m(R) > 0$ —otherwise no one would survive until retirement age and the “S + W + R” scenario would boil down to the “S + W” scenario, already considered above.

We are looking for functional specifications of $m(\tau)$ for which $\varphi(S) = Ge^{HS}$ for some $G > 0$ and $H \in \mathbb{R}$, so that consistently with (31), the relationship between aggregate human capital and

aggregate years of schooling S is of a log-linear type. Assuming this functional relationship, it follows that

$$\lim_{S \rightarrow R} \frac{\varphi(S)}{e^{\mu S}} = Ge^{(H-\mu)R}, \tag{34}$$

and thus $G = e^{(\mu-H)R}$ and consequently $\varphi(S) = e^{HS+(\mu-H)R}$. Since $\frac{\varphi(S)}{e^{\mu S}} = e^{(\mu-H)(R-S)} > 1$ for all $S \in [0, R)$, then it must be the case that $H < \mu$. It follows that $G > 1$.

We shall now pass to the central part of the proof. Positing $\varphi(S) = Ge^{HS}$ and rearranging in (32) yield:

$$\int_S^\infty be^{(\mu-(b-d))\tau} m(\tau) d\tau = Ge^{HS} \int_S^\infty be^{-(b-d)\tau} m(\tau) d\tau. \tag{35}$$

Eq. (35) is a *functional identity* and thus it holds for all $S \in [0, R)$. It is also possible to differentiate both sides of (35) with respect to S . Doing this twice and rearranging terms, we obtain:

$$\frac{m'(S)}{m(S)} = \frac{(\mu - H - b + d)e^{(\mu-H)S} - G(d - b + H)}{G - e^{(\mu-H)S}}. \tag{36}$$

Solving (36) under the assumption $H < \mu$ we obtain:

$$m(S) = e^{(b-d-H)S} \left(\frac{e^{(\mu-H)R} - 1}{e^{(\mu-H)R} - e^{(\mu-H)S}} \right)^{\frac{\mu-2H}{\mu-H}}, \tag{37}$$

$\forall (S \in [0, R)).$

Please note that the denominator necessarily tends to infinity as $S \rightarrow R_-$. The implications of this fact can be threefold, depending on the value of H relative to $\mu/2$. First, if $\mu - 2H < 0$ then $m(R) = 0$, so nobody survives until retirement, contradicting the “S + W + R” scenario. Second, if $\mu - 2H > 0$ then $\lim_{\tau \rightarrow R_-} m(\tau) = +\infty$, so m cannot be a survival law. Finally, if $\mu - 2H = 0$ so that $m(\tau) = e^{(b-d-\mu/2)\tau}$, then m takes the known exponential “perpetual youth” form. Making it consistent with interpretation requires imposing $b - d - \mu/2 = -d$, and thus $\mu = 2b$, contradicting the assumption that $\mu < b$. We conclude that the macro-Mincer equation cannot be reconciled with the “S + W + R” scenario under any admissible survival law. ■

Appendix B. Extension: allowing for human capital depreciation

In the main text, we have assumed that individuals’ human capital does not depreciate: once one has acquired a certain skill, she will be able to use it with equal efficiency ever after. In reality, however, probably a majority of people’s skills (e.g., language skills, manual skills, knowledge of facts and methods) tend to naturally deteriorate if not applied sufficiently often. Also, some skills might become obsolete due to technological progress: the recent proliferation of ICT technologies worldwide is just a demonstration that the set of skills and abilities required in any productive activity might change over time. For all these reasons, allowing for human capital depreciation might seem a natural extension of our theoretical results. As we shall see, such a modification of our framework does not lead to any qualitative changes of the results.

B.1. Modification of the framework

Let us now consider the case which allows for gradual human capital depreciation within individuals’ lifetimes. The human capital accumulation equation is modified in the following way:

Assumption 5 (*Modification of Assumption 1*). Human capital of an individual τ years old, born at time j , is accumulated using a linear production function:

$$\frac{\partial}{\partial \tau} h(j, \tau) = [\lambda \ell_h(j, \tau) + \mu \ell_Y(j, \tau) - \delta] h(j, \tau) \tag{38}$$

where $\lambda \geq 0$ denotes the unit productivity of schooling, and $\mu \geq 0$ denotes the unit productivity of on-the-job learning (experience accumulation). The parameter $\delta \geq 0$ captures the rate of human capital depreciation. $\ell_h(j, \tau) \in [0, 1]$ is the fraction of time spent by an individual born at j and aged τ on formal education, whereas $\ell_Y(j, \tau) \in [0, 1]$ is the fraction of time spent at work. We assume $\ell_h(j, \tau) + \ell_Y(j, \tau) \leq 1$ for all $j, \tau \geq 0$, and take $h(j, 0) \equiv h_0 > 0$.

Eq. (38) can be straightforwardly integrated, yielding:

$$h(t - \tau, \tau) = h_0 \exp \left[\underbrace{\lambda \int_0^\tau \ell_h(t - \tau, s) ds}_{\text{years of schooling}} + \underbrace{\mu \int_0^\tau \ell_Y(t - \tau, s) ds}_{\text{work experience}} - \delta \tau \right]. \tag{39}$$

We shall keep all other features of our framework unchanged.

B.2. Sufficient conditions for the macro-Mincer equation

The results following from the above modification of our framework are as follows. First, it is easily verified that if the survival law has the “perpetual youth” property ($m(\tau) = e^{-d\tau}$), then the macro-Mincer equation is still recovered from the micro-Mincer one in the scenario “S + W”, and not recovered in the scenario “S + W + R”. If the macro-Mincer equation holds, the Mincerian schooling coefficient amounts to $\lambda - \delta$, i.e., the individual rate of return to schooling is corrected for human capital depreciation. Furthermore, if additionally $\mu = \delta$, i.e. if the rate of on-the-job learning is exactly equal to the rate of human capital depreciation, then the macro-Mincer equation is also recovered in the “S + W + R” case. If $\mu \neq \delta$ then it is not.

Second, sufficient conditions for the macro-Mincer equation in the case of fixed lifetimes or the Boucekkine et al. (2002) survival law are fully equivalent as well, the only difference being that the condition $\mu = 0$ is replaced with $\mu = \delta$. We find that the macro-Mincer equation is obtained in the case of fixed lifetimes if and only if $\mu = \delta$. This is obtained both under the “S + W” and the “S + W + R” scenario.

This last result is an epitome of a more general phenomenon, though: if $\mu = \delta$ then the rate of human capital depreciation is exactly matched by the rate of on-the-job learning, and thus the whole labor force has exactly the same human capital level. Aggregation is then effected across entirely homogeneous population cohorts. The logic is exactly the same as in the case without human capital depreciation, as summarized by the following general proposition:

Proposition 7 (*Sufficient Condition for Macro-Mincer*). *Let Assumptions 2–5 hold and assume $\mu = \delta$. Then under the “S + W” and “S + W + R” scenarios, the macro-Mincer equation holds regardless of the underlying survival law $m(\tau)$. The Mincerian schooling coefficient is equal to the individual rate of return to education minus the rate of human capital depreciation, $\lambda - \delta$.*

The proof is a straightforward modification of the proof of Proposition 3. It is available from the authors upon request.

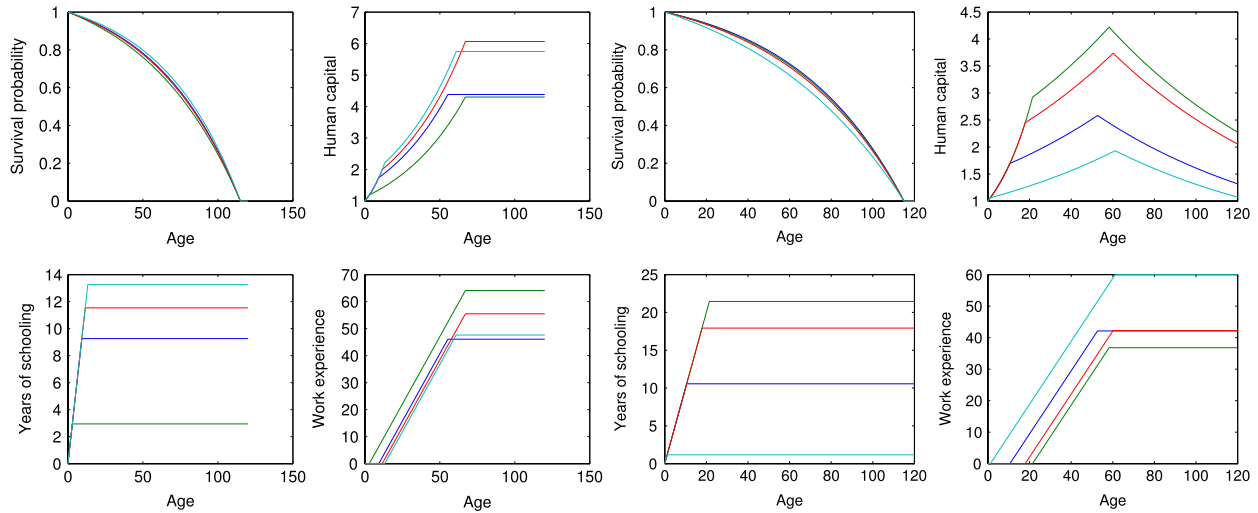


Fig. 4. Time profiles of selected variables with no human capital depreciation (left panel) and with human capital depreciation at a rate $\delta = 0.01$ (right panel).

B.3. Necessary conditions for the macro-Mincer equation

Turning to necessary conditions for the macro-Mincer equation, it turns out that – just like in the case without human capital depreciation – if there is some heterogeneity in human capital across working cohorts (which is represented now by the condition $\mu \neq \delta$), then the simplified macro-Mincer equation can be reproduced under the “S + W” scenario only in the “perpetual youth” case, and cannot be reproduced under the “S + W + R” scenario at all. The following propositions hold.

Proposition 8 (Necessary Conditions for Macro-Mincer with S + W). Let Assumptions 2–5 hold with $\mu \neq \delta$ and $b > \mu - \delta$. Assume that the simplified macro-Mincer equation holds. Then, under the “S + W” scenario where the individuals stay at school until the age S and then work full-time until death, the survival law must be $m(\tau) = e^{-d\tau}$, i.e., it must satisfy the “perpetual youth” property. The implied macro-Mincer equation is $h(t) = \frac{bh_0}{b-\mu+\delta} e^{(\lambda-\delta)S}$.

Proposition 9 (Macro-Mincer Impossible with S + W + R). Let Assumptions 2–5 hold with $\mu \neq \delta$ and $b > \mu - \delta$. Then under the “S + W + R” scenario where the individuals stay at school until age S and then work full-time until retirement age R, there is no admissible survival law compatible with the simplified macro-Mincer equation.

Proofs of the above propositions are straightforward modifications of proofs of Propositions 4–5. They are available from the authors upon request. Please note that in the case where the macro-Mincer equation holds, the implied macro-level rate of return to human capital accumulation is equal to $\lambda - \delta$, the individual rate of return to an additional year of schooling minus the rate of human capital depreciation.

Appendix C. Details of numerical experiments

The current appendix presents a detailed elaboration of a range of stylized numerical examples discussed in Section 3.

C.1. Human capital evolution with $\delta = 0$ and $\delta > 0$

Let us first illustrate the workings of our numerical analysis. The assumed survival law $m(\tau)$ as well as the implied time profiles of individuals’ human capital, cumulative years of schooling, and work experience, are illustrated in Fig. 4. We use a baseline calibration of $\delta = 0.01$ in the example which allows for $\delta > 0$,

so that the assumed human capital depreciation rate is lower than returns to schooling and work experience—and thus the net effect of both activities remains strictly positive. In such a case, human capital gradually decays for the retired population, though.

The functions illustrated in Fig. 4 are the individual-level time profiles underlying our general aggregation framework discussed in Section 2.

C.2. Heterogeneity in years of schooling

The first stylized numerical experiment (with $\delta = 0$ for simplicity) consisted in generating a sample of countries differing only in the number of obligatory years of schooling S, holding all other parameters fixed at their baseline values. Fig. 5 illustrates that in such a case, the estimated macro-Mincer equation fits the “true” relationship between average human capital and years of schooling almost perfectly, rendering a negligible mean absolute percentage error. The simplified macro-Mincer equation, which omits the work experience variable in the regressions, is also a reasonable approximation, albeit the residuals are somewhat larger in its case.

The estimated macro-level return to schooling is somewhat lower than the micro-level return λ , and so is the macro-level return to work experience as compared to the micro-level return μ . In the simplified macro-Mincer equation, returns to schooling are underestimated more strongly.

C.3. Heterogeneity in retirement age

The second experiment consisted in generating a sample of countries differing in the retirement age R only, holding other parameters fixed. The approximation of the “true” relationship between average human capital, years of schooling and work experience with the macro-Mincer equation is somewhat worse in such a case than in the previous example, but remains very good. The results can be seen in Fig. 6. The simplified macro-Mincer equation is omitted there because it is unidentified when all countries share the same S.

In the current case, the macro-level returns to schooling are underestimated as compared to λ , whereas the returns to work experience are overestimated as compared to μ . Furthermore, the estimated log-linear equation – although quantitatively close to the “true” relationship – cannot match its curvature, leading to systematic errors.

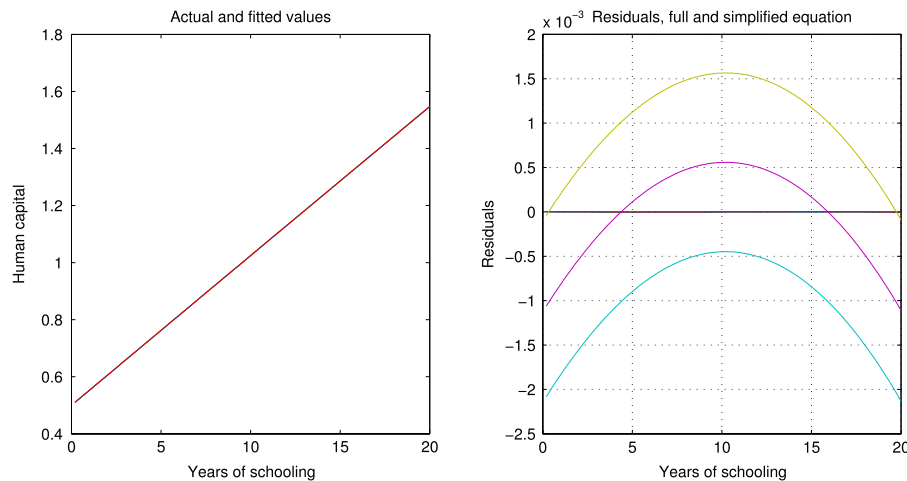


Fig. 5. Quality of approximation of average human capital with the macro-Mincer equation: the case of varying S . *Notes.* Estimated equations: $\widehat{\ln h}_i = 0.2438 + 0.0559q_i + 0.0115x_i$, with $R^2 \approx 1$ and $\text{MAPE} \approx 0.00\%$; $\widehat{\ln h}_i = 0.5005 + 0.0524q_i$, with $R^2 \approx 1$ and $\text{MAPE} = 0.05\%$. Left: blue line denotes true data points, green line denotes their macro-Mincer approximation, red line denotes their simplified macro-Mincer approximation. Right: upper and lower lines delineate the boundaries of a 95% confidence interval around the middle line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

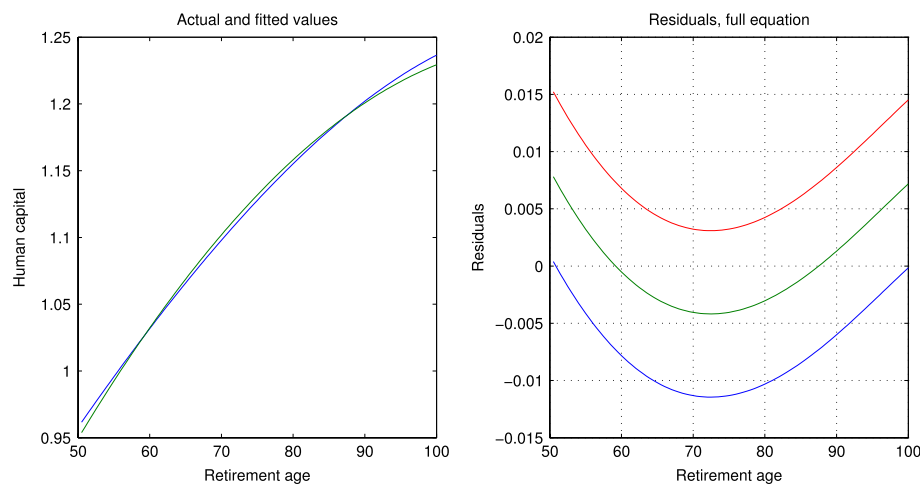


Fig. 6. Quality of approximation of average human capital with the macro-Mincer equation: the case of varying R . *Notes.* Estimated equation: $\widehat{\ln h}_i = 0.0000 + 0.0495q_i + 0.0274x_i$, with $R^2 \approx 0.998$ and $\text{MAPE} \approx 0.28\%$. Left: blue line denotes true data points, green line denotes their macro-Mincer approximation. Right: upper and lower lines delineate the boundaries of a 95% confidence interval around the middle line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

C.4. Heterogeneity in life expectancy

As it is visible in Fig. 7, if the only source of cross-country heterogeneity is located in the parameters of the survival law, a and k , mapping uniquely into the measures of life expectancy E and the maximum lifespan T^* , the macro-Mincer equation fits the data almost perfectly again.¹⁴ The bias in estimates of macro-level returns to schooling and work experience is very small in the current case.

C.5. Intra-cohort heterogeneity: additional results

Let us now present our additional results under the case allowing for intra-cohort heterogeneity.

¹⁴ Fig. 7 considers the case where the parameters a and k are varied simultaneously in a way that the maximum lifespan T^* is kept constant. The alternative where a is varied keeping k constant (and thus both E and T^* are variable) is available upon request. The results are qualitatively and quantitatively similar in both cases.

First, we see in Fig. 8 that if the only source of cross-country heterogeneity is located in the share of high-skill individuals in the population ψ , the macro-Mincer equation fits the data almost perfectly again. We observe however that the Mincerian experience coefficient is not uniquely identified in the current case: the regressor matrix is rank deficient, and both the full and the simplified macro-Mincer equation yield exactly the same fit to the data. This peculiar result disappears when one allows for additional sources of cross-country heterogeneity, though.

Second, in Fig. 9 we consider a different dimension of cross-country heterogeneity. We confirm that under the case where all heterogeneity comes from the number of additional years of schooling among the high-skill individuals \tilde{S} , the fit of the macro-Mincer equation to the theoretically derived “true” data is close to perfect as well. The fit of the simplified Mincer equation is somewhat worse but still reasonably good. What is striking here, however, is that in this case the macro-Mincer estimates of aggregate returns to schooling and work experience are negative and way off their micro-level counterparts. In the case of the simplified macro-Mincer equation, the fit to the “true” data is only slightly worse, and thus still very good. In such case, the estimated Mince-

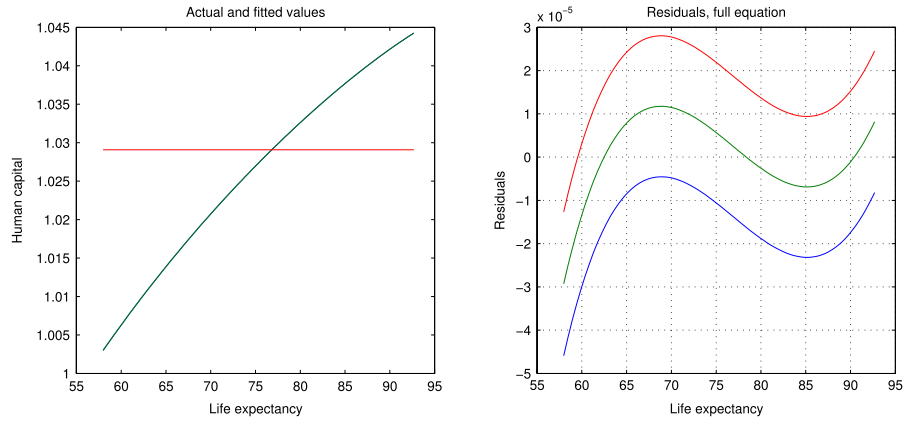


Fig. 7. Quality of approximation of average human capital with the macro-Mincer equation: the case of varying E for a fixed T^* . *Notes.* Estimated equation: $\widehat{\ln h}_i = 0.0000 + 0.0617q_i + 0.0211x_i$, with $R^2 \approx 1$ and $\text{MAPE} \approx 0.00\%$. Left: blue line denotes true data points, green line denotes their macro-Mincer approximation, red line denotes their simplified macro-Mincer approximation. Right: upper and lower lines delineate the boundaries of a 95% confidence interval around the middle line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

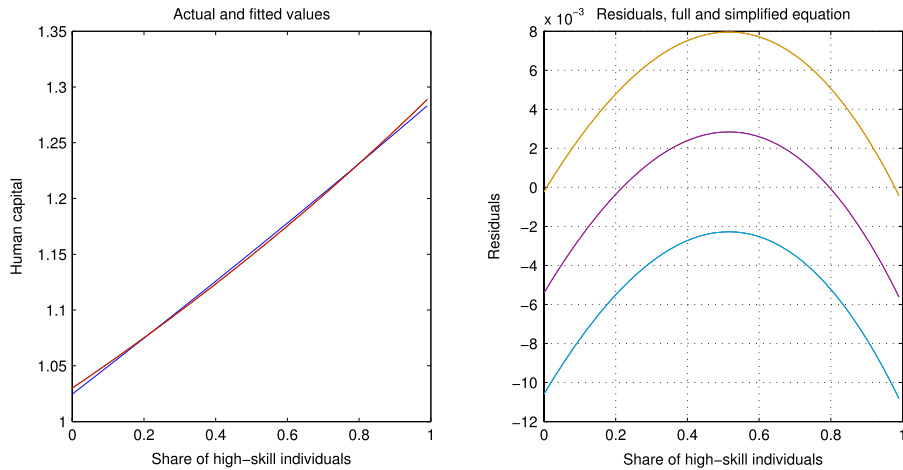


Fig. 8. Quality of approximation of average human capital with the macro-Mincer equation, with intra-cohort heterogeneity: varying shares of high-skill individuals $\psi \in (0, 1)$ (each having $\bar{S} = 5$ additional years of schooling). *Notes.* Estimated equations: $\ln h_i = 0.0000 + 0.0595q_i + 0.0225x_i$, with $R^2 \approx 0.999$ and $\text{MAPE} = 0.19\%$; $\ln h_i = 0.5060 + 0.0524q_i$, with $R^2 \approx 0.999$ and $\text{MAPE} = 0.19\%$. Both equations yield the same outcome; the Mincerian experience coefficient is unidentified. Left: blue line denotes true data points, green line denotes their macro-Mincer approximation, red line denotes their simplified macro-Mincer approximation. Right: upper and lower lines delineate the boundaries of a 95% confidence interval around the middle line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

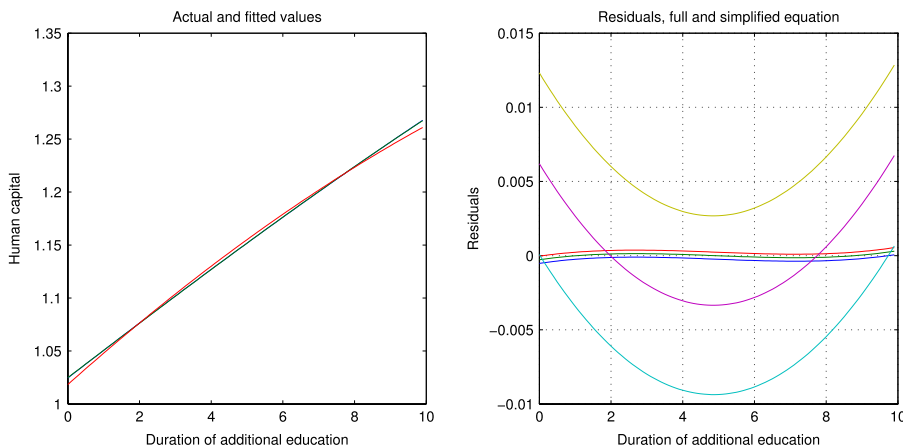


Fig. 9. Quality of approximation of average human capital with the macro-Mincer equation, with intra-cohort heterogeneity: varying additional years of schooling among the high-skill individuals \bar{S} (fixed share of high-skill individuals $\psi = 0.5$). *Notes.* Estimated equations: $\ln h_i = 13.3524 - 0.1250q_i - 0.5729x_i$, with $R^2 \approx 1$ and $\text{MAPE} = 0.01\%$; $\ln h_i = 0.4161 + 0.0602q_i$, with $R^2 \approx 0.998$ and $\text{MAPE} = 0.22\%$. Left: blue line denotes true data points, green line denotes their macro-Mincer approximation, red line denotes their simplified macro-Mincer approximation. Right: upper and lower lines delineate the boundaries of a 95% confidence interval around the middle line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

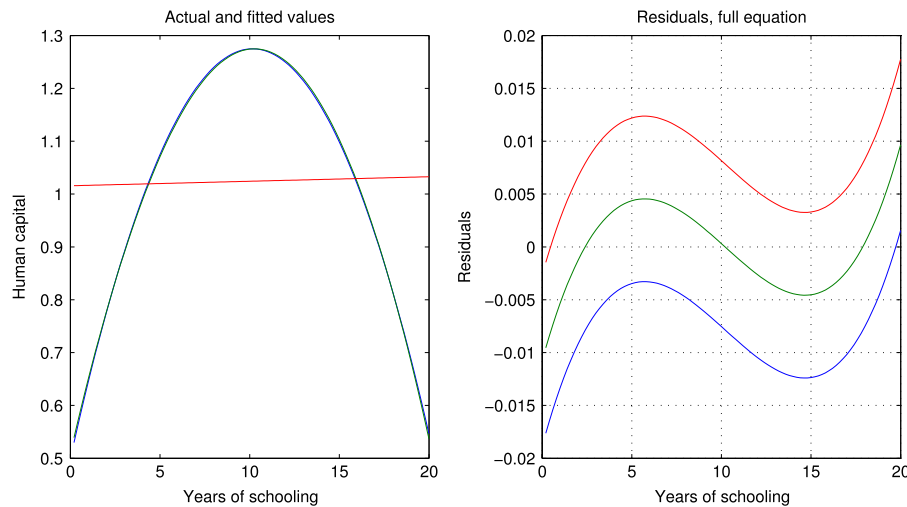


Fig. 10. Quality of approximation of average human capital with the macro-Mincer equation: the case of varying S , perfectly negatively correlated with a varying λ . *Notes.* Estimated equation: $\ln \hat{h}_i = -114.0726 + 1.5900q_i + 5.1420x_i$, with $R^2 \approx 0.999$ and $MAPE \approx 0.37\%$. Left: blue line denotes true data points, green line denotes their macro-Mincer approximation, red line denotes their simplified macro-Mincer approximation. Right: upper and lower lines delineate the boundaries of a 95% confidence interval around the middle line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

rian schooling coefficient α is close to its micro-level benchmark of 0.06.

Considering more sophisticated cases of intra-cohort heterogeneity (e.g., with cohort-specific shares of high-skill individuals) is left for further research.

Appendix D. Heterogeneity in returns to education or work experience

As hinted in the main text, the fit of the macro-Mincer equation to the data deteriorates very quickly once one allows for cross-country heterogeneity in the returns to education parameter λ or the experience coefficient μ . The reason for this result is that by construction, the macro-Mincer equation implies a unique value for the measured returns to education at the country level. Hence, if actual returns have different magnitudes across countries, the macro-Mincer equation misses all the relevant variation: the best it can do is to capture the *average* level of returns across the whole sample. This is precisely what happens when the equation is estimated with ordinary least squares.

D.1. Allowing returns to education to be correlated with years of schooling

In a related stylized example, we have considered a case which allows for simultaneous variation in returns to education λ and in years of schooling S . Indeed, as it has been discussed in the relevant

Table 5
Results of the Monte Carlo study—the case of heterogeneous returns to schooling and work experience.

	R^2	MAPE [%]	c	α	β
Micro-level benchmarks			0	0.05	0.01
$S.D.(\lambda) = S.D.(\mu) = 0.01$					
Mean	0.4725	516.4286	-0.0208	0.0499	0.0122
S.D.	0.0751	18 556	0.3915	0.0073	0.0172
$S.D.(\lambda) = S.D.(\mu) = 0.001$					
Mean	0.9887	3.0938	-0.0129	0.0500	0.0112
S.D.	0.0022	0.2936	0.0403	0.0008	0.0018
$S.D.(\lambda) = S.D.(\mu) = 0$					
Mean	1.0000	0.0427	-0.0140	0.0500	0.0112
S.D.	0.0000	0.0062	0.0009	0.0000	0.0000

Notes: The regression parameters c , α , β are defined in Eq. (16).

literature (e.g., [Bils and Klenow, 2000](#), [Caselli, 2005](#)), primary education tends to yield higher returns than secondary education, and higher still than tertiary education. A stylized representation of these findings within our framework is to impose a strict negative correlation between these two variables. As an example, we have considered the case where there is a *linear* functional relationship between them: the higher is S , the lower is λ .¹⁵

The aggregation results obtained under such circumstances are presented in [Fig. 10](#). The macro-Mincer equation fits the data very well, yielding an R^2 just marginally short of unity, and very small residuals. This is however not a general result but only a coincidence of functional forms, for two reasons. First, the estimated parameters are two orders of magnitude away from their micro-level counterparts, $\lambda = 0.06$ and $\mu = 0.02$. This is because the estimated macro-Mincer equation tries to incorporate the years of schooling-returns to schooling tradeoff in the macro-level returns to schooling and work experience directly, which is at odds with the true model. Second, further numerical experiments (available upon request) indicate that if the relationship between S and λ were nonlinear, the goodness of fit of the macro-Mincer equation would fall considerably, aligning again both with the intuition and the results of our general Monte Carlo study (see [Table 5](#)).

D.2. Accounting for heterogeneity in returns to education or work experience along with human capital depreciation

Allowing for human capital depreciation may be a realistic assumption but it is not conducive to the resulting goodness of fit of the macro-Mincer equation to the theoretically derived “true” data. We find that, with or without human capital depreciation, this fit is vastly different whether the returns parameters λ and μ are allowed to vary across countries or not. As opposed to the cases considered in the main text, if such possibility is allowed, the macro-Mincer equation cannot match the assumed heterogeneity, leaving a very large part of human capital variation unexplained. In result, the R^2 of the macro-Mincer regression depends crucially on the magnitude of variation of λ and μ in the sample.

¹⁵ In cross-country data on years of schooling and returns ([Psacharopoulos and Patrinos, 2004](#)), this correlation is indeed negative, but not as strict: it amounts to -0.37 .

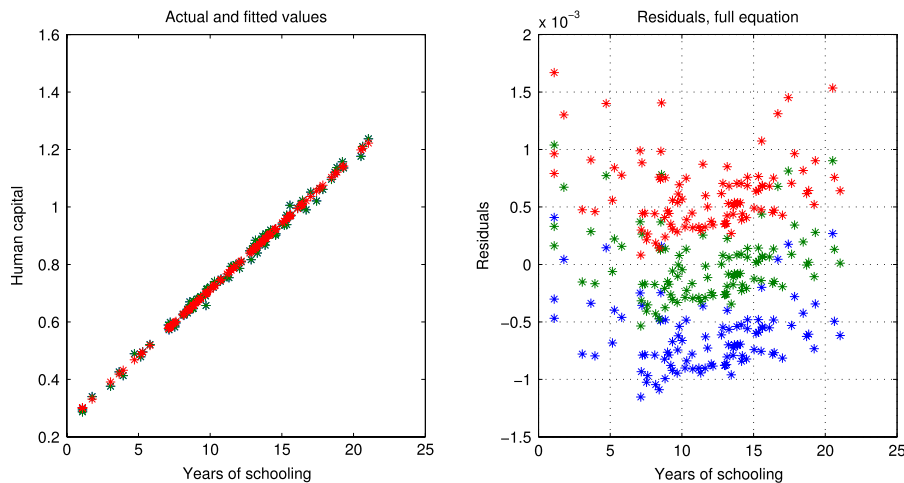


Fig. 11. Quality of approximation of average human capital with the macro-Mincer equation: the case of randomly varying S, R, a , controlling for human capital depreciation. *Notes.* Estimated equations: $\ln h_i = -0.0133 + 0.0500q_i + 0.0112x_i$, with $R^2 \approx 1$ and MAPE = 0.05%; $\ln h_i = 0.2392 + 0.0467q_i$ with $R^2 \approx 0.996$ and MAPE = 1.60%. Left: blue asterisks denote true data points, green asterisks denote their macro-Mincer approximations, and red asterisks denote their simplified macro-Mincer approximations. Right: blue and red asterisks delineate the boundaries of a 95% confidence interval around green asterisks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

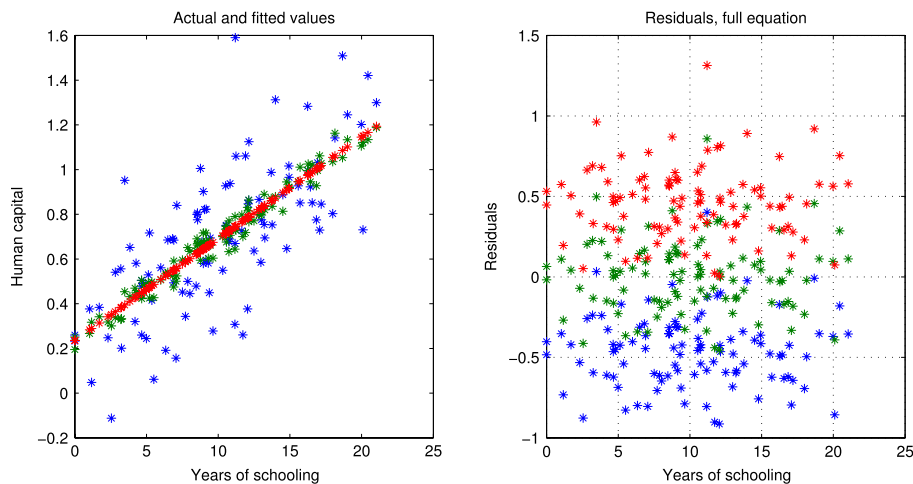


Fig. 12. Quality of approximation of average human capital with the macro-Mincer equation: the case of randomly varying S, R, a, λ, μ , controlling for human capital depreciation. *Notes.* Estimated equations: $\ln h_i = 0.2227 + 0.0389q_i + 0.0048x_i$, with $R^2 \approx 0.3339$ and MAPE = 115.5%; $\ln h_i = 0.3342 + 0.0053q_i$ with $R^2 \approx 0.3335$ and MAPE = 116.2%. Left: blue asterisks denote true data points, green asterisks denote their macro-Mincer approximations, and red asterisks denote their simplified macro-Mincer approximations. Right: blue and red asterisks delineate the boundaries of a 95% confidence interval around green asterisks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 11 illustrates that allowing for human capital depreciation itself (but keeping rates of return λ and μ equal across countries) does not overturn the conclusion that the macro-Mincer equation fits the data remarkably well. In fact, the individual impact of human capital depreciation on the goodness-of-fit statistics of the macro-Mincer equation is rather negligible. In Fig. 12, on the other hand, λ and μ are assumed to have a standard deviation of 0.01 which is substantial but still much lower than observed in real-world cross-country data. In such a case, the fit of the macro-Mincer equation is bad.

In both numerical examples, we are maintaining that S, R and a are randomly drawn for each country, like in Section 3.4 of the article. In the latter one, on top of that λ and μ are also allowed to vary (independently). The impact of the latter modification on the goodness-of-fit statistics of the macro-Mincer equation is overwhelming.

D.3. Monte Carlo study with heterogeneous returns to schooling and work experience

To strengthen our arguments about the case which allows λ and μ to be country-specific, we have repeated our Monte Carlo

study (cf. Table 4) for various levels of cross-country variability in the key variables of interest: returns to schooling λ and returns to work experience μ . It is confirmed that if these two parameters are known with certainty and are equal across countries, the macro-Mincer equation fits the data remarkably well, but its fit deteriorates rapidly once the returns are allowed to vary across countries. The reason is that such heterogeneity cannot be accounted for by a single macro-Mincer equation with equal coefficients for the whole sample of countries.

Quantitative results, for moderate ($S.D.(\lambda) = S.D.(\mu) = 0.01$) and very small ($S.D.(\lambda) = S.D.(\mu) = 0.001$) variability in λ and μ as well as the case of their zero variability, are collected in Table 5. It must be noted that even our case of “moderate” variation in returns remains rather conservative as compared to Psacharopoulos and Patrinos (2004) data. In their cross-country dataset, the (unweighted) average return on an additional year of schooling across the world is 9.6%, with a standard deviation of 4.3%, i.e., the estimated standard deviation is about four times larger than in our case of “large” variability of returns. Under such circumstances, the fit of the macro-Mincer equation to real-world cohort-specific data must be expected to be very poor. This result

remains outside of the scope of applications of the theoretical model, though.

References

- Azomahou, Théophile T., Boucekkine, Raouf, Diene, Bity, 2009. A closer look at the relationship between life expectancy and economic growth. *Int. J. Econ. Theory* 5, 201–244.
- Bils, Mark, Klenow, Peter J., 2000. Does schooling cause growth? *Amer. Econ. Rev.* 90 (5), 1160–1183.
- Blanchard, Olivier, 1985. Debt, deficits, and finite horizons. *J. Polit. Econ.* 93, 223–247.
- Bloom, David E., Canning, David, Sevilla, Jaypee, 2004. The effect of health on economic growth: A production function approach. *World Dev.* 32 (1), 1–13.
- Boucekkine, Raouf, de la Croix, David, Licandro, Omar, 2002. Vintage human capital, demographic trends and endogenous growth. *J. Econom. Theory* 104, 340–375.
- Caselli, Francesco, 2005. Accounting for cross-country income differences. In: Aghion, Philippe, Durlauf, Steven (Eds.), *Handbook of Economic Growth*. Elsevier, Amsterdam.
- Caselli, Francesco, Coleman, Wilbur J., 2006. The world technology frontier. *Amer. Econ. Rev.* 96 (3), 499–522.
- Growiec, Jakub, 2010. Human capital, aggregation, and growth. *Macroecon. Dyn.* 14 (2), 189–211.
- Growiec, Jakub, Groth, Christian, 2012. On aggregating human capital across heterogeneous cohorts. NBP Working Paper 134.
- Hamlen, Susan S., Hamlen, William A., 2012. The inconsistency of the quadratic Mincer equation: a proof. *Theoret. Econ. Lett.* 2 (2), 115–120.
- Heckman, James J., Lochner, Lance J., Todd, Petra E., 2003. Fifty years of Mincer earnings regressions. NBER Working Paper 9732.
- Hsieh, Chang-Tai, Klenow, Peter J., 2010. Development accounting. *Amer. Econ. J.: Macroecon.* 2 (1), 207–223.
- Jones, Benjamin F., 2014a. The human capital stock: a generalized approach. *Amer. Econ. Rev.* 104 (11), 3752–3777.
- Jones, Benjamin F., 2014b. The knowledge trap: human capital and development reconsidered. Mimeo, Kellogg School of Management, Northwestern University.
- Klenow, Peter J., Rodríguez-Clare, Andrés, 1997. The neoclassical revival in growth economics: has it gone too far? In: Bernanke, Ben, Rotemberg, Julio J. (Eds.), *NBER Macroeconomics Annual*. Vol. 12. MIT Press, Cambridge, MA, pp. 73–103.
- Krueger, Alan B., Lindahl, Mikael, 2001. Education for growth. Why and for whom? *J. Econom. Lit.* 39 (4), 1101–1136.
- Lagakos, David, Moll, Benjamin, Porzio, Tommaso, Qian, Nancy, Schoellman, Todd, 2012. Experience matters: human capital and development accounting. NBER Working Paper 18602.
- Mincer, Jacob, 1974. *Schooling, Experience, and Earnings*. Columbia University Press, New York.
- Pandey, Manish, 2008. Human capital aggregation and relative wages across countries. *Journal of Macroecon.* 30 (4), 1587–1601.
- Psacharopoulos, George, Patrinos, Harry A., 2004. Returns to investment in education: a further update. *Educ. Econ.* 12 (2), 111–134.