On Aggregating Human Capital Across Heterogeneous Cohorts[*]

Jakub Growiec[†]   Christian Groth[‡]

February 21, 2013

Abstract. Based on a general framework for computing the aggregate human capital stock under heterogeneity across population cohorts, the paper derives the "true" functional relationship between the aggregate human capital stock, average years of schooling, and average work experience. Under the scenarios considered here, the macro-Mincer (log-linear) functional relationship between these variables is obtained only in cases which are inconsistent with heterogeneity or based on empirically implausible demographic survival laws. Our numerical results indicate that the macro-Mincer equation is nevertheless an empirically reasonable approximation of the true relationship, at least under standard calibrations.

Keywords: human capital, aggregation, heterogeneity, population cohort, Mincer equation

JEL Classification Numbers: J24, O47.

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, Lochner, Todd, 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, Canning and Sevilla, 2004). This *macro-Mincer* aggregative approach has been criticized from a number of standpoints, though. In particular, making such a micro–macro analogy has been shown to rest on the following questionable simplifying assumptions:

[†]Warsaw School of Economics, Institute of Econometrics, and National Bank of Poland, Economic Institute. E-mail: jakub.growiec@sgh.waw.pl. Corresponding author. Mailing address: Instytut Ekonometrii SGH, al. Niepodległości 162, 02-554 Warszawa, Poland.

[‡]University of Copenhagen, Department of Economics. E-mail: chr.groth@econ.ku.dk.

- the macro-Mincer approach requires perfect substitutability between unskilled and skilled labor (Pandey, 2008; Jones, 2011a, 2011b),

- 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, 2011b),

- it considers years of schooling as an exogeneous variable and thus neglects individuals' optimal decisions on the duration of their schooling (e.g., Jones, 2011a),

- it neglects the fact that maintaining a constant aggregate level of human capital in the society 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.[1]

The current paper elaborates on the last of the above points. Our analysis maintains the assumption that skill levels are perfectly substitutable and that there is no intra-cohort heterogeneity of tasks or skills. On the other hand, we 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 a framework without intra-cohort heterogeneity, we attempt to isolate the effects coming from the heterogeneity of human capital due to demographics alone.

The contribution of the current article to the literature is twofold. First, having clarified the outstanding problems related to the definitions of the aggregate human capital stock,[2] aggregate years of schooling, and aggregate work experience under heterogeneity across population cohorts, we have carried out a theoretical study leading to the conclusion that even if the micro-Mincer functional relationship does hold in a cross-section of individuals, the macro-Mincer (log-linear) equation generally does not. The only exceptions 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 where there are positive returns from work experience, the macro-Mincer equation requires the demographical survival law to have the "perpetual youth" property (Blanchard, 1985), which is empirically implausible. Under the scenario where people also retire at a certain age, the macro-Mincer equation cannot be recovered under any admissible survival law if there are positive returns from work experience. For the cases where the macro-Mincer equation does not hold, we derive the true aggregate relationships.[3]

---

[1]Some authors have also argued 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).

[2]We consider human capital as a one-dimensional stock of productive skills embodied in an individual, accumulated via schooling and on-the-job learning. By doing so, we set aside all the conceptual criticisms related to such definition of human capital (see, e.g., Hanushek and Woessmann, 2008).

[3]One can also demonstrate (see the working paper version of this article, Growiec and Groth, 2012) an important difference in aggregation results whether human capital stocks in the *whole population* or in the *labor force* are

Second, we add a quantitative edge to our analytical results by means of a numerical Monte Carlo study based on the general theoretical model. We find that, although theoretically misspecified, the macro-Mincer (log-linear) equation can nevertheless be perceived as an empirically reasonable approximation of the true relationship between the aggregate human capital stock, average years of schooling, and average work experience. In particular, this is the case if returns to schooling and work experience are assumed to be equal in the considered group of countries (or regions), with the observed heterogeneity coming only from differences in the number of years of schooling, retirement age, or demographical survival laws. It ceases to be the case if one (wrongfully) omits the work experience term from the macro-Mincer equation, though. We also find numerically that human capital heterogeneity across population cohorts decreases average social returns to schooling in the labor force as compared to private returns by about 1 pp, regardless of the assumed magnitude of returns to work experience.

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. Details and extensions have been relegated to the Appendix. Section 4 concludes.

2   Aggregation of human capital across population cohorts

2.1   Framework

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. Our results are obtained under the following assumptions.

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

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

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 accomodate other uses of time, such as leisure or childrearing. We thus also allow for

---

considered. In particular, the macro-Mincer (log-linear) functional relationship can only be obtained (under additional restrictions) for the latter case – on which we concentrate in the current paper – but not for the former. 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 strongly suggests that these concepts should not be used interchangeably.

retirement. We say that these alternative possibilities are exercised when $\ell_h(j, \tau) + \ell_Y(j, \tau) < 1$.[4]

Integrating equation (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:

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

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 eqations (cf. Heckman, Lochner, Todd, 2003), does not appear here because in equation (1) we have assumed human capital accumulation to be linear and not concave in work experience.[5]

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 $m(0) = 1, \lim_{\tau \to \infty} m(\tau) = 0$ and with $m(\tau)$ weakly decreasing in its whole domain. The survival probability does not depend on calendar time t.*

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. Accomodating 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*

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

The total labor force at time t is computed as

$$L(t) = \int_0^\infty P(t, \tau)\ell_Y(t - \tau, \tau)d\tau = \int_0^\infty bN(t - \tau)m(\tau)\ell_Y(t - \tau, \tau)d\tau. \tag{4}$$

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:

$$\frac{P(t, \tau)}{N(t)} = bm(\tau)\frac{N(t - \tau)}{N(t)} = bm(\tau)e^{-(b-d)\tau}, \qquad \text{independently of } t. \tag{5}$$

---

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

[5]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.

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:

$$\frac{L(t)}{N(t)} = \frac{\int_0^\infty bN(t - \tau)m(\tau)\ell_Y(\tau)d\tau}{N(t)} = \int_0^\infty be^{-(b-d)\tau}m(\tau)\ell_Y(\tau)d\tau. \tag{6}$$

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:

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

- **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:

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

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

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

2.2 Aggregation across cohorts

To be able to aggregate human capital stocks, years of schooling as well as work experience across heterogeneous population cohorts meaningfully, one needs to ensure that all the respective aggregative concepts are appropriately defined. This is particularly important in our current case because

5

certain analogies between micro- and macro-level variables are quite misleading here. The general framework, building on Assumptions 1–3, is consistent with the following definitions.

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

$$H_{LF}(t) = \int_0^\infty P(t,\tau)\ell_Y(\tau)h(\tau)d\tau. \tag{10}$$

Labor services provided by individuals of all ages are perfectly substitutable. The average human capital stock in the labor force is $h_{LF}(t) = \frac{H_{LF}(t)}{L(t)}$.

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

$$Q_{LF}(t) = \int_0^\infty P(t,\tau)\ell_Y(\tau)\left(\int_0^\tau \ell_h(s)ds\right)d\tau. \tag{11}$$

The average number of years of schooling in the labor force is thus $q_{LF}(t) = \frac{Q_{LF}(t)}{L(t)}$.

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

$$X_{LF}(t) = \int_0^\infty P(t,\tau)\ell_Y(\tau)\left(\int_0^\tau \ell_Y(s)ds\right)d\tau. \tag{12}$$

Average work experience in the labor force is thus $x_{LF}(t) = \frac{X_{LF}(t)}{L(t)}$.

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

Definition 4 *The* macro-Mincer equation *for the labor force takes the following form:*

$$h_{LF}(t) = h_0 \exp\left(\alpha q_{LF}(t) + \beta x_{LF}(t)\right). \tag{13}$$

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 could be rewritten for the whole population alive at t instead of just the labor force by dropping the $\ell_Y(\tau)$ term in the integrals. One could then also define the macro-Mincer equation for the 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), such an alternative case is also even less likely to reproduce the macro-Mincer result than the case considered here.

We shall now present our results under two specific survival laws $m(\tau)$, and then provide more general considerations relating to the (im)possibility or (im)plausibility of obtaining the macro-Mincer relationship as presumed in much of the related empirical literature.

2.3 Results under the "perpetual youth" survival law

Apart from Assumptions 1–3, let us now also assume the Blanchard (1985) simple "perpetual youth" survival law $m(\tau) = e^{-d\tau}$, where d is directly the aggregate death rate. Under this condition, the stationary age structure satisfies $\frac{P(t,\tau)}{N(t)} = be^{-b\tau}$. The results are presented in Table 1.

In the scenario "S+W", $H_{LF}(t)$ 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", $H_{LF}(t)$ 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 $H_{LF}(t)$. With a fixed share of time spent on work irrespective of individuals' age, it is however clear that $H_{LF}(t) = \bar{\ell}_Y H_{POP}(t)$, so that the qualitative results for both aggregates are identical up to a multiplicative constant. Also, the share of the working population is naturally $\frac{L(t)}{N(t)} = \bar{\ell}_Y$, and thus $h_{LF}(t) = h_{POP}(t)$ in this scenario.

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".

Proposition 1 (Sufficient conditions for macro-Mincer) *Let Assumptions 1–3 hold with $\mu \in [0, b)$ and assume the "perpetual youth" survival law. Then the macro-Mincer equation holds for the labor force (but not the whole population):*

- *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 equals the individual rate of return to education λ, while the Mincerian experience coefficient is zero. Apart from these two cases, the macro-Mincer equation does not hold.

Table 1: Average human capital, years of schooling, work experience, and the verification of the macro-Mincer equation under the "perpetual youth" survival law.

Scenario	S+W	S+W+R	Fix
$h_{LF}(t)$	$\frac{bh_0}{b-\mu}e^{\lambda S}$	$\frac{bh_0 e^{(\lambda-\mu)S}}{b-\mu}\frac{e^{(\mu-b)S}-e^{(\mu-b)R}}{e^{-bS}-e^{-bR}}$	$\frac{bh_0}{b-\lambda\bar{\ell}_h-\mu\bar{\ell}_Y}$
$q_{LF}(t)$	S	S	$\frac{\ell_h}{b}$
$x_{LF}(t)$	$\frac{1}{b}$	$\frac{1}{b} - \frac{R-S}{e^{b(R-S)}-1}$	$\frac{\ell_Y}{b}$
M-M, $\mu = 0$	YES, coefficient λ	YES, coefficient λ	NO
M-M, $\mu > 0$	YES, coefficients λ and 0	NO	NO

Note: "M-M" refers to the macro-Mincer functional relationship.

So 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 is very limited. It is obtained only if retirement is absent or if accumulated work experience does not affect workers' human capital stocks.[6]

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 ($\mu = 0$):[7]

Proposition 2 (Sufficient conditions for macro-Mincer) *Let Assumptions 1–3 hold and assume that the individuals have a fixed lifetime T. Then the macro-Mincer equation holds for the labor force (but not the whole population):*

- *under the "S+W" scenario with $\mu = 0$,*

- *under the "S+W+R" scenario with $\mu = 0$.*

In both cases $h_{LF}(t) = h_0 e^{\lambda S}$, that is, the Mincerian schooling coefficient equals the individual rate of return to education λ, whereas the Mincerian experience coefficient is zero. Apart from these two cases, the macro-Mincer functional relationship does not hold.

---

[6]As shown by Growiec and Groth (2012), the macro-Mincer equation is never obtained for the whole population if the survival law has the "perpetual youth" property.

[7]In the case $\lambda = b - d$, the formula $\frac{bh_0}{b-d-\lambda}\left(1 - e^{(\lambda-(b-d))S}\right)$ should be replaced by $bh_0 S$. Furthermore, if $\mu = b - d$, then the formula $\frac{bh_0}{b-d-\mu}\left(e^{(\mu-(b-d))S} - e^{(\mu-(b-d))T}\right)$ should be replaced by $bh_0(T - S)$, and the formula $\frac{bh_0}{b-d-\mu}\left(e^{(\mu-(b-d))S} - e^{(\mu-(b-d))R}\right)$ by $bh_0(R - S)$. This applies both to the case of the "perpetual youth" survival law and the case of fixed lifetimes.

Table 2: Average human capital, years of schooling, work experience, and the verification of the macro-Mincer equation under fixed lifetimes.

Case	S+W	S+W+R	Fix
		$T \neq 1/b \qquad b \neq d$	
$h_{LF}(t)$	$\dfrac{(b-d)h_0 e^{(\lambda-\mu)S}}{b-d-\mu}\dfrac{e^{(\mu-(b-d))S}-e^{(\mu-(b-d))T}}{e^{-(b-d)S}-e^{-(b-d)T}}$	$\dfrac{(b-d)h_0}{b-d-\mu}\dfrac{e^{(\lambda-(b-d))S}-e^{(\lambda-\mu)S+(\mu-(b-d))R}}{e^{-(b-d)S}-e^{-(b-d)R}}$	$\dfrac{bh_0}{b-d-\lambda\bar{\ell}_h-\mu\bar{\ell}_Y}\left(1-e^{(\lambda\bar{\ell}_h+\mu\bar{\ell}_Y-(b-d))T}\right)$
$q_{LF}(t)$	S	S	$\dfrac{\ell_h(1-dT)}{b-d}$
$x_{LF}(t)$	$\dfrac{1}{b-d}-\dfrac{T-S}{e^{(b-d)(T-S)}-1}$	$\dfrac{1}{b-d}-\dfrac{R-S}{e^{(b-d)(R-S)}-1}$	$\dfrac{\ell_Y(1-dT)}{b-d}$
M-M, $\mu=0$	YES, coefficient λ	YES, coefficient λ	NO
M-M, $\mu>0$	NO	NO	NO
		$T = 1/b \qquad b = d$	
$h_{LF}(t)$	$\dfrac{h_0 e^{\lambda S}}{\mu(T-S)}\left(e^{\mu(T-S)}-1\right)$	$\dfrac{h_0 e^{\lambda S}}{\mu(R-S)}\left(e^{\mu(R-S)}-1\right)$	$\dfrac{h_0}{T(\lambda\bar{\ell}_h+\mu\bar{\ell}_Y)}\left(e^{(\lambda\bar{\ell}_h+\mu\bar{\ell}_Y)T}-1\right)$
$q_{LF}(t)$	S	S	$\dfrac{\ell_h T}{2}$
$x_{LF}(t)$	$\dfrac{T-S}{2}$	$\dfrac{R-S}{2}$	$\dfrac{\ell_Y T}{2}$
M-M, $\mu=0$	YES, coefficient λ	YES, coefficient λ	NO
M-M, $\mu>0$	NO	NO	NO

Note: "M-M" refers to the macro-Mincer functional relationship.

2.5 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 3 (Sufficient condition for macro-Mincer) *Let Assumptions 1–3 hold and assume $\mu = 0$. Then under the "S+W" and "S+W+R" scenarios, the macro-Mincer equation holds for the labor force $h_{LF}(t)$ regardless of the underlying survival law $m(\tau)$. The Mincerian schooling coefficient is equal to the individual rate of return to education λ.*

Proof. Using equations (5)–(6), under the "S+W" scenario we have:

$$
\begin{aligned}
h_{LF}(t) &= \int_0^\infty \frac{P(t,\tau)}{L(t)} \ell_Y(t-\tau,\tau) h(t-\tau,\tau) d\tau = \int_S^\infty h_0 e^{\lambda S} b e^{-(b-d)\tau} m(\tau) \frac{N(t)}{L(t)} d\tau = \\
&= h_0 e^{\lambda S} \frac{\int_S^\infty b e^{-(b-d)\tau} m(\tau) d\tau}{\int_S^\infty b e^{-(b-d)\tau} m(\tau) d\tau} = h_0 e^{\lambda S}.
\end{aligned} \tag{14}
$$

Using equations (5)–(6) again, under the "S+W+R" scenario we have:

$$
\begin{aligned}
h_{LF}(t) &= \int_0^\infty \frac{P(t,\tau)}{L(t)} \ell_Y(t-\tau,\tau) h(t-\tau,\tau) d\tau = \int_S^R h_0 e^{\lambda S} b e^{-(b-d)\tau} m(\tau) \frac{N(t)}{L(t)} d\tau = \\
&= h_0 e^{\lambda S} \frac{\int_S^R b e^{-(b-d)\tau} m(\tau) d\tau}{\int_S^R b e^{-(b-d)\tau} m(\tau) d\tau} = h_0 e^{\lambda S}. \blacksquare
\end{aligned} \tag{15}
$$

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.

2.6 Necessary conditions for the macro-Mincer equation

Having obtained some positive results for very specific and empirically implausible survival laws, one should ask the converse, much more general question: For which survival law $m(\tau)$ will the macro-Mincer functional relationship be recovered 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:

$$h_{LF}(t) = h_0 \exp\left(\alpha q_{LF}(t)\right), \tag{16}$$

be obtained from the micro-level Mincerian equation?

The importance of the last question stems from the fact that the related applied literature[8] 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_{LF} 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 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 4 (Necessary conditions for macro-Mincer with S+W) *Let Assumptions 1–3 hold with $\mu \in (0, b)$ and assume that the simplified macro-Mincer equation holds for the labor force. 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 $h_{LF}(t) = \frac{b h_0}{b - \mu} e^{\lambda S}$.*

Proof. 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:

$$h_{LF}(t) = \int_S^\infty h(t - \tau, \tau) \frac{P(t, \tau)}{L(t)} d\tau = h_0 e^{(\lambda - \mu)S} \frac{\int_S^\infty b e^{(\mu - (b-d))\tau} m(\tau) d\tau}{\int_S^\infty b e^{-(b-d)\tau} m(\tau) d\tau}. \tag{17}$$

We shall use the notation:

$$\varphi(S) = \frac{\int_S^\infty b e^{(\mu - (b-d))\tau} m(\tau) d\tau}{\int_S^\infty b e^{-(b-d)\tau} m(\tau) d\tau} \tag{18}$$

which implies $h_{LF}(t) = \varphi(S) \cdot h_0 e^{(\lambda - \mu)S}$. 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

---

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

obtain:

$$\lim_{S \to +\infty} \frac{\varphi(S)}{e^{\mu S}} = \lim_{S \to +\infty} \frac{\int_S^\infty e^{(\mu-(b-d))\tau} m(\tau) d\tau}{e^{\mu S} \int_S^\infty e^{(-(b-d))\tau} m(\tau) d\tau} = \tag{19}$$

$$= \frac{1}{1 - \lim_{S \to +\infty} \frac{\mu \int_S^\infty e^{-(b-d)\tau} m(\tau) d\tau}{m(S) e^{-(b-d)S}}} = \frac{1}{1 + \lim_{S \to +\infty} \frac{\mu}{\frac{m'(S)}{m(S)} - (b-d)}}.$$

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 (17), 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

$$\lim_{S \to +\infty} \frac{\varphi(S)}{e^{\mu S}} = \lim_{S \to +\infty} Ge^{(H-\mu)S}. \tag{20}$$

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. Positing $\varphi(S) = Ge^{HS}$ and rearranging in (18) yields:

$$\int_S^\infty be^{(\mu-(b-d))\tau} m(\tau) d\tau = Ge^{HS} \int_S^\infty be^{-(b-d)\tau} m(\tau) d\tau. \tag{21}$$

Equation (21) is a *functional identity* and thus it holds for all $S \geq 0$. It is also possible to differentiate both sides of (21) with respect to S. Doing this twice and rearranging terms, we obtain:

$$\frac{m'(S)}{m(S)} = \frac{(\mu - H - b + d)e^{(\mu-H)S} - G(d - b + H)}{G - e^{(\mu-H)S}}. \tag{22}$$

Consider first the case $H > \mu$. In such a case we obtain $\lim_{S \to +\infty} Ge^{(H-\mu)S} = +\infty$. Coupled with equation (19), this implies:

$$\lim_{S \to +\infty} \frac{m'(S)}{m(S)} = b - d - \mu. \tag{23}$$

Comparing (23) and (22) we obtain:

$$\lim_{S \to +\infty} \frac{m'(S)}{m(S)} = b - d - H = b - d - \mu, \tag{24}$$

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 (22) and simplifying we obtain:

$$\frac{m'(S)}{m(S)} = -\frac{(G-1)(-b+d) + G\mu}{G-1} = (b-d) - \frac{G\mu}{G-1}. \tag{25}$$

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:

$$m(S) = \exp\left(\left((b-d) - \frac{G\mu}{G-1}\right)S\right), \qquad \forall\, (S \geq 0). \tag{26}$$

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 equation (26) 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:

$$N(t) = \int_{-\infty}^{t} bN(s)m(t-s)ds = N_0 e^{(b-d)t}. \tag{27}$$

From (26) and (27) it follows that

$$\int_{-\infty}^{t} b \exp\left(\left(\frac{G}{1-G}\right)\mu(t-s)\right)ds = 1. \tag{28}$$

Computing the last integral reveals that $G = \frac{b}{b-\mu} > 1$. Plugging this into (26), we obtain $m(\tau) = e^{-d\tau}$. Also, $\varphi(S) = \frac{b}{b-\mu}e^{\mu S}$ and thus $h_{LF}(t) = \frac{bh_0}{b-\mu}e^{\lambda S}$ so that the macro-Mincer equation holds with the Mincerian coefficient λ. ∎

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 unconditional probability of dying next year. According to empirical evidence (cf. e.g., Boucekkine, de la Croix and Licandro, 2002), this is clearly not the case, not even approximately.[9]

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 5 (Macro-Mincer impossible with S+W+R) *Let Assumptions 1–3 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. 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:

$$h_{LF}(t) = \int_{S}^{R} h(t-\tau, \tau)\frac{P(t, \tau)}{L(t)}d\tau = h_0 e^{(\lambda-\mu)S}\frac{\int_{S}^{R} be^{(\mu-(b-d))\tau}m(\tau)d\tau}{\int_{S}^{R} be^{-(b-d)\tau}m(\tau)d\tau}. \tag{29}$$

We shall use the notation:

$$\varphi(S) = \frac{\int_{S}^{R} be^{(\mu-(b-d))\tau}m(\tau)d\tau}{\int_{S}^{R} be^{-(b-d)\tau}m(\tau)d\tau} \tag{30}$$

which implies $h_{LF}(t) = \varphi(S) \cdot h_0 e^{(\lambda-\mu)S}$. 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

---

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

obtain:

$$\lim_{S \to R} \frac{\varphi(S)}{e^{\mu S}} = \lim_{S \to R} \frac{\int_S^R e^{(\mu - (b-d))\tau} m(\tau) d\tau}{e^{\mu S} \int_S^R e^{(-(b-d))\tau} m(\tau) d\tau} = \tag{31}$$

$$= \frac{1}{1 - \lim_{S \to R} \frac{\mu \int_S^R e^{-(b-d)\tau} m(\tau) d\tau}{m(S) e^{-(b-d)S}}} = 1.$$

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 (29), 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

$$\lim_{S \to R} \frac{\varphi(S)}{e^{\mu S}} = Ge^{(H-\mu)R}, \tag{32}$$

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 (30) yields:

$$\int_S^\infty be^{(\mu-(b-d))\tau} m(\tau) d\tau = Ge^{HS} \int_S^\infty be^{-(b-d)\tau} m(\tau) d\tau. \tag{33}$$

Equation (33) is a *functional identity* and thus it holds for all $S \in [0, R)$. It is also possible to differentiate both sides of (33) with respect to S. Doing this twice and rearranging terms, we obtain:

$$\frac{m'(S)}{m(S)} = \frac{(\mu - H - b + d)e^{(\mu-H)S} - G(d - b + H)}{G - e^{(\mu-H)S}}. \tag{34}$$

Solving (34) under the assumption $H < \mu$ we obtain:

$$m(S) = e^{(b-d-H)S} \left(\frac{e^{(\mu-H)R} - 1}{e^{(\mu-H)R} - e^{(\mu-H)S}} \right)^{\frac{\mu-2H}{\mu-H}}, \qquad \forall \, (S \in [0, R)). \tag{35}$$

Please note that the denominator necessarily tends to infinity as $S \to R_-$. 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 $\lim_{\tau \to R_-} m(\tau) = +\infty$, 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. ∎

This theoretical result further restricts the practical applicability of the (simplified) macro-Mincer relationship between average human capital and average years of schooling.

3 The macro-Mincer equation as an approximation

Having obtained a range of theoretical results, with rather negative conclusions for the 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 true relationship between average human capital and average years of schooling and work experience, despite being theoretically 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 comprehensive results based on a Monte Carlo study. Details and extensions have been relegated to the Appendix.

It turns out that, in spite of the theoretical misspecification, the approximation precision of the macro-Mincer equation is remarkably good. 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_{LF}, and especially so if the underlying experience coefficient μ is large.[10] In any case, we also find notable discrepancies between macro- and micro-level returns to schooling and work experience, however.

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 (but not within) cohorts, coupled with a realistic survival law put forward by Boucekkine, de la Croix and Licandro (2002) and further discussed by Azomahou, Boucekkine and Diene (2009). This survival law is given by the following function $m : [0, T^*] \to [0, 1]$:

$$m(\tau) = \frac{e^{-\beta\tau} - \alpha}{1 - \alpha}, \qquad \alpha > 1, \beta < 0. \tag{36}$$

The maximum lifetime of an individual is given by $T^* = -\frac{\ln \alpha}{\beta}$, whereas individuals' life expectancy is calculated as $E = \frac{1}{\beta} + \frac{\alpha \ln \alpha}{(1-\alpha)\beta}$.

We assume 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.

For every parameter configuration, we are going to compute the "true" average human capital stock in the population[11] h_{POP} and in the labor force h_{LF}, as well as respective measures of cumulative years of schooling q_{POP} and q_{LF}, and cumulative work experience x_{POP} and x_{LF}, based

---

[10]This last result is valid also when one controls for (exponential) human capital depreciation with $\mu > \delta > 0$. Please consult the Appendix.

[11]For theoretical derivations, please refer to Growiec and Groth (2012).

on our analytical framework. We shall identify each parameter configuration with a "country", assuming that the micro-Mincer equation holds exactly in every country 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 homogenous 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 assume that these parameters are equidistributed along a predefined interval. In the 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 regressions:

$$\ln h_{POP,i} = a_0 + a_1 q_{POP,i} + a_2 x_{POP,i} + \varepsilon_i, \qquad (37)$$
$$\ln h_{LF,i} = b_0 + b_1 q_{LF,i} + b_2 x_{LF,i} + \eta_i \qquad (38)$$

with ordinary least squares, based on the artificial data computed from the "true" model. The "goodness-of-fit" of the macro-Mincer equation to the "true" model will be assessed by comparing the R^2 of the regressions as well as within-sample mean absolute percentage errors (MAPE). We shall also compare our macro estimates of returns to schooling a_1, b_1 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 a_2, b_2 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 directly, or perhaps some adjustment is necessary.

Concurrently, we shall also report the respective numerical results for *simplified* macro-Mincer equations, obtained by omitting the experience variable, i.e., setting $a_2 = b_2 = 0$ in equations (37)–(38). Comparing the estimates for these simplified specifications with their counterparts from the fully specified macro-Mincer equations, we shall assess the magnitude of omitted variable bias incurred in the estimation of the simplified equations.[12]

The outcomes of our numerical exercises will be presented graphically in the form of plots with four panels (except for self-explanatory Fig. 2). In the upper panels we shall report actual and fitted country-level human capital averages. We shall also provide the parametric form of the estimated regression equations there (both full and simplified). In the lower panels we shall

---

[12]In the Appendix, we also consider numerically a case where returns to education and work experience are allowed to vary across countries. In line with our prior expectations, we find that the approximation quality deteriorates very quickly with increasing cross-country heterogeneity in returns to schooling or work experience. Such a case does not correspond directly to the issue of aggregating human capital across heterogeneous cohorts, though. Also, 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. On the other hand, since in known datasets (e.g., Psacharopoulos and Patrinos, 2004), returns to schooling tend to be largely divergent across countries, this auxiliary study indirectly implies that the empirical validity of the macro-Mincer equation should be rather limited unless the sample is conveniently restricted by the researcher.

report the residuals (actual minus fitted values) together with the upper and lower bounds of their respective 99% confidence intervals.

3.2 Calibration

The baseline calibration for parameters used in our numerical exercise is the following: (a) following Boucekkine, de la Croix and Licandro (2002), we assume $\alpha = 5.44, \beta = -0.0147$, implying a life expectancy of 73 years and maximum lifespan of 115 years; (b) the population 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 annum[13]; (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.

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

As demonstrated in the Appendix, 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, α and β (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. This applies particularly strongly to the macro-Mincer relationship in the labor force. 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 conclude that if returns to education λ and work experience μ are fixed across countries,[14] 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 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 α, are drawn randomly from Gaussian distributions. The parameter R is generated independently of the two other variables, whereas S and α 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.[15]

---

[13]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%).

[14]The alternative case is considered in the Appendix.

[15]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, α, and β, as long as the rates of return λ and μ are kept fixed.
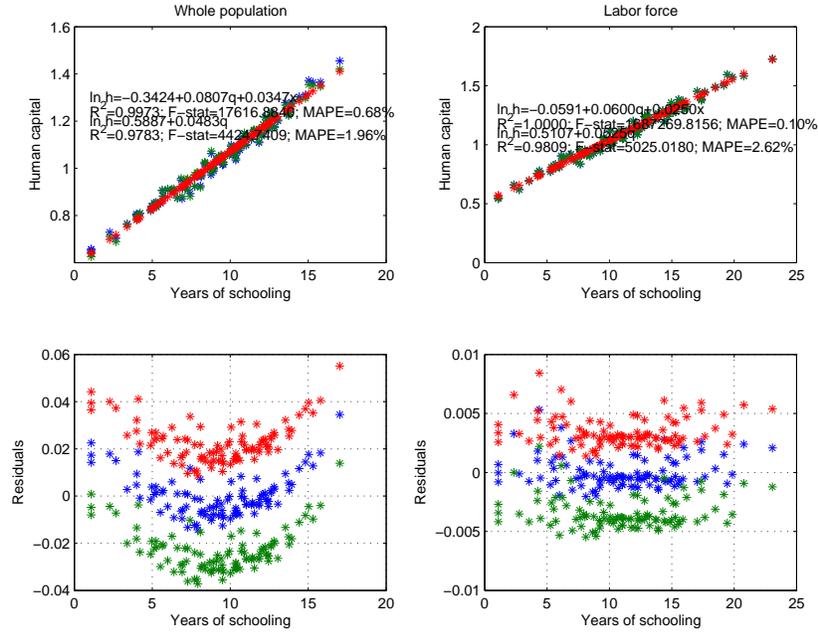
Figure 1: Quality of approximation of average human capital with the macro-Mincer equation: the case of randomly varying S, R, α.

As shown in Fig. 1, the macro-Mincer equation fits the "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. These two parameters are rather imprecisely estimated in the case of the whole population, though.

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 the Appendix), then – extrapolating from Proposition 3 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, especially for the labor force. 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: the macro-Mincer equation fits the true relationship between average human capital h_{LF}, average years of schooling q_{LF}, and average work experience x_{LF} in the labor force 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
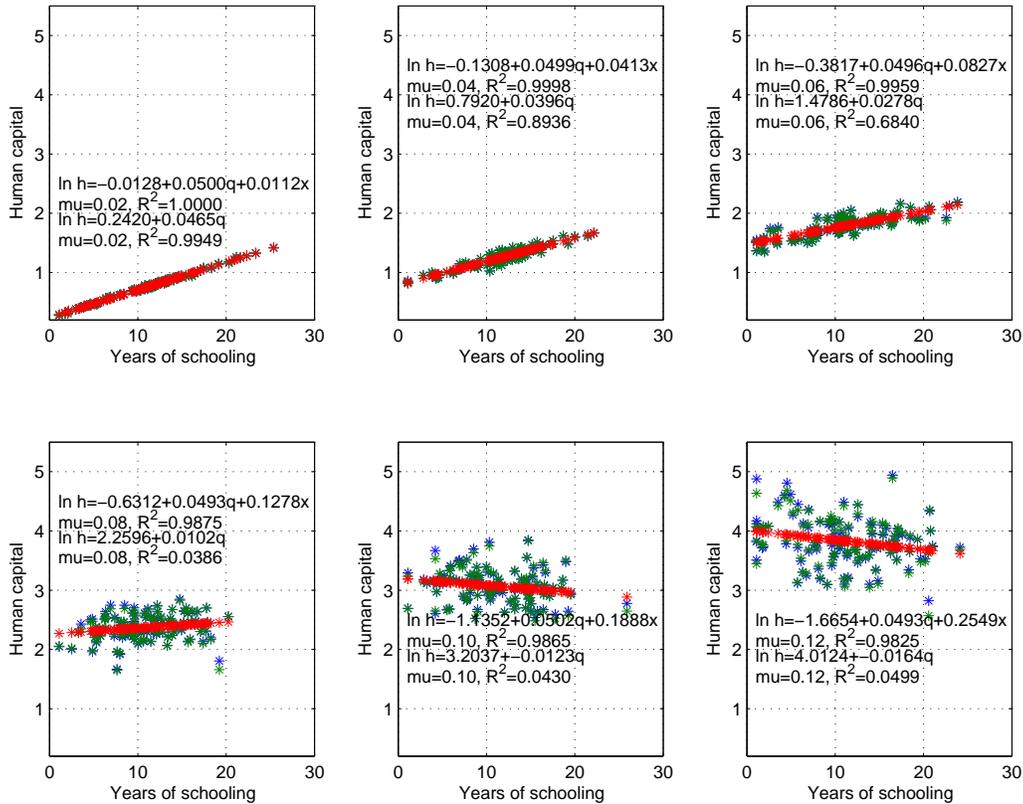
Figure 2: Quality of approximation of average human capital in the labor force h_{LF} with the macro-Mincer equation: dependence on the coefficient of returns to work experience μ.

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 empricial applications, it is of great importance that the Mincerian experience coefficient is not restricted to zero.

Table 3: Results of the Monte Carlo study – dependence on returns to work experience (parameter μ).

	R^2_{POP}	R^2_{LF}	M_{POP} [%]	M_{LF} [%]	a_0	a_1	a_2	b_0	b_1	b_2
					0	0.06	μ	0	0.06	μ
				Micro-level benchmarks						
				$\mu = 0.02$						
Mean	0.9971	1.0000	1.1929	0.0427	-0.2912	0.0640	0.0198	-0.0140	0.0500	0.0112
S.D.	0.0006	0.0000	0.1293	0.0062	0.0389	0.0015	0.0014	0.0009	0.0000	0.0000
				$\mu = 0.04$						
Mean	0.9686	0.9996	1.2203	0.2261	-0.8265	0.0885	0.0647	-0.1355	0.0501	0.0414
S.D.	0.0076	0.0001	0.1166	0.0290	0.0877	0.0032	0.0032	0.0090	0.0001	0.0004
				$\mu = 0.06$						
Mean	0.8629	0.9960	1.7334	0.4455	-1.8534	0.1333	0.1335	-0.3887	0.0504	0.0825
S.D.	0.0345	0.0014	0.1732	0.0581	0.1991	0.0074	0.0072	0.0262	0.0003	0.0012
				$\mu = 0.08$						
Mean	0.8710	0.9889	2.2289	0.6348	-3.2757	0.1953	0.2211	-0.7556	0.0510	0.1332
S.D.	0.0325	0.0034	0.2226	0.0767	0.3537	0.0129	0.0129	0.0510	0.0006	0.0024
				$\mu = 0.10$						
Mean	0.8765	0.9862	2.6302	0.7774	-4.9101	0.2673	0.3196	-1.2023	0.0517	0.1912
S.D.	0.0310	0.0036	0.2636	0.0959	0.5556	0.0205	0.0201	0.0771	0.0010	0.0036
				$\mu = 0.12$						
Mean	0.8769	0.9860	2.9170	0.8789	-6.6400	0.3445	0.4234	-1.6842	0.0526	0.2536
S.D.	0.0300	0.0034	0.2927	0.1101	0.7789	0.0287	0.0282	0.1091	0.0014	0.0051

Notes: M denotes MAPE. The regression parameters a_0 through b_2 are defined in equations (37)–(38).

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, α 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.[16] 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 3.

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. The macro-Mincer equation estimated for the whole population fits the true functional relationship visibly worse. As discussed in the previous subsection, the same (though not reported) can be said for the simplified macro-Mincer equation.

Apart from the general finding of the very good fit of the macro-Mincer equation to our data, it also stands out in Table 3 that the estimated value of $b_1 < \lambda$, and the difference between this macro-Mincer point estimate and its micro-level counterpart is remarkably stable across various values of μ. For this reason, we conclude that human capital heterogeneity across population cohorts – even though it also theoretically invalidates the log-linear functional form of the macro-Mincer equation – contributes to decrease *average* social returns to schooling in the labor force as compared to private returns by about 1 pp, at least under the baseline calibration of other model parameters. We also see that the estimated macro-level experience coefficient b_2 systematically diverges from its micro-level counterpart μ, downwards if $\mu < 0.04$ and upwards if $\mu > 0.04$.

4 Conclusion

The contribution of the current article to the literature is twofold. First, based on a general framework for computing the aggregate human capital stock under heterogeneity across (but not within) cohorts, building on Growiec (2010), we have shown that the macro-Mincer (log-linear) functional relationship between human capital, years of schooling and work experience is generally lost upon aggregation. It can in fact be obtained only in very special cases, and all of the cases which we were able to identify 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 – they require the demographical survival law to have the "perpetual youth" property (Blanchard,

---

[16]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.

1985), which is empirically implausible. For the cases where the macro-Mincer equation does not hold, we have derived the true aggregate relationships.

Second, we have shown numerically that the macro-Mincer equation can be perceived nevertheless as a quantitatively reasonable approximation of the true relationship between average human capital stocks, years of schooling, and work experience, with the observed heterogeneity coming from differences in the number of years of schooling, retirement age, or demographical survival laws, unless the work experience term is (wrongfully) omitted from the equation. We have also computed that human capital heterogeneity across population cohorts contributes to decrease average social returns to schooling in the labor force as compared to private returns by about 1 pp, regardless of the magnitude of individual returns to work experience.

References

[1] Azomahou, Théophile T., Raouf Boucekkine, and Bity Diene (2009) A closer look at the relationship between life expectancy and economic growth. International Journal of Economic Theory 5, 201-244.

[2] Bils, Mark and Peter J. Klenow (2000) Does schooling cause growth? American Economic Review 90(5), 1160–1183.

[3] Blanchard, Olivier (1985) Debt, deficits, and finite horizons. Journal of Political Economy 93, 223-247.

[4] Bloom, David E., David Canning, and Jaypee Sevilla (2004) The effect of health on economic growth: A production function approach. World Development 32 (1), 1-13.

[5] Boucekkine, Raouf, David de la Croix, and Omar Licandro (2002) Vintage human capital, demographic trends and endogenous growth. Journal of Economic Theory 104, 340-375.

[6] Caselli, Francesco (2005) Accounting for cross-country income differences. In: Philippe Aghion, Steven Durlauf (eds), Handbook of Economic Growth. Elsevier, Amsterdam.

[7] Caselli, Francesco and Wilbur J. Coleman (2006) The world technology frontier. American Economic Review 96 (3), 499-522.

[8] Growiec, Jakub (2010) Human capital, aggregation, and growth. Macroeconomic Dynamics 14 (2), 189-211.

[9] Growiec, Jakub and Christian Groth (2012) On aggregating human capital across heterogeneous cohorts. NBP Working Paper 134.

[10] Hamlen, Susan S. and William A. Hamlen (2012) The inconsistency of the quadratic Mincer equation: a proof. Theoretical Economics Letters 2(2), 115-120.

[11] Hanushek, Eric A. and Ludger Woessmann (2008) The role of cognitive skills in economic development. Journal of Economic Literature 46(3), 607-668.

[12] Heckman, James J., Lance J. Lochner, and Petra E. Todd (2003) Fifty years of Mincer earnings regressions. NBER Working Paper 9732.

[13] Hsieh, Chang-Tai, Peter J. Klenow (2010) Development accounting. American Economic Journal: Macroeconomics 2(1), 207-223.

[14] Jones, Benjamin F. (2011a) The human capital stock: a generalized approach. NBER Working Paper 17487.

[15] Jones, Benjamin F. (2011b) The knowledge trap: human capital and development reconsidered. Mimeo, Northwestern University.

[16] Klenow, Peter J. and Andrés Rodríguez-Clare (1997) The neoclassical revival in growth economics: has it gone too far? [In:] NBER Macroeconomics Annual, Vol. 12, ed. Ben Bernanke and Julio J. Rotemberg, 73-103. Cambridge, MA: MIT Press.

[17] Krueger, Alan B. and Mikael Lindahl (2001) Education for growth. Why and for whom? Journal of Economic Literature 39 (4), 1101-1136.

[18] Mincer, Jacob (1974) Schooling, Experience, and Earnings. New York: Columbia University Press.

[19] Pandey, Manish (2008) Human capital aggregation and relative wages across countries. Journal of Macroeconomics 30 (4), 1587-1601.

[20] Psacharopoulos, George and Harry A. Patrinos (2004) Returns to investment in education: a further update. Education Economics 12(2), 111-134.

A   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.

A.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 4 (Modification of Assumption 1) *Human capital of an individual τ years old, born at time j, is accumulated using a* linear *production function:*

$$\frac{\partial}{\partial \tau}h(j,\tau) = [\lambda \ell_h(j,\tau) + \mu \ell_Y(j,\tau) - \delta]h(j,\tau) \tag{39}$$

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 $h(j,0) \equiv h_0 > 0$.

Equation (39) can be straightforwardly integrated, yielding:

$$h(t-\tau,\tau) = h_0 \exp\left[\lambda \underbrace{\int_0^\tau \ell_h(t-\tau,s)ds}_{\text{years of schooling}} + \mu \underbrace{\int_0^\tau \ell_Y(t-\tau,s)ds}_{\text{work experience}} - \delta\tau\right]. \tag{40}$$

We shall keep all other features of our framework unchanged.

A.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 for the labor force 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 for the labor force 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 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 for the labor force 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 6 (Sufficient condition for macro-Mincer) *Let Assumptions 2–4 hold and assume $\mu = \delta$. Then under the "S+W" and "S+W+R" scenarios, the macro-Mincer equation holds for the labor force $h_{LF}(t)$ 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.

A.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 7 (Necessary conditions for macro-Mincer with S+W) *Let Assumptions 2–4 hold with $\mu \neq \delta$ and $b > \mu - \delta$. Assume that the simplified macro-Mincer equation holds for the labor force. 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 $h_{LF}(t) = \frac{bh_0}{b-\mu+\delta}e^{(\lambda-\delta)S}$.*

Proposition 8 (Macro-Mincer impossible with S+W+R) *Let Assumptions 2–4 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.

B Details of numerical experiments

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

B.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. 3. 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. 3 are the individual-level time profiles underlying our general aggregation framework discussed in Section 2.
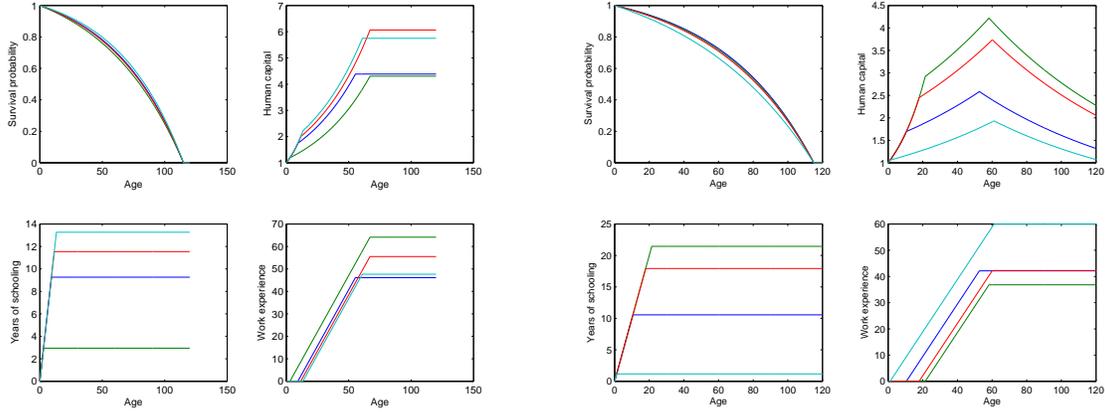
Figure 3: Time profiles of selected variables with no human capital depreciation (left panel) and with human capital depreciation at a rate $\delta = 0.01$ (right panel).

B.2 Heterogeneity in years of schooling

The first stylized numerical experiment 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. 4 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. This applies particularly strongly to the macro-Mincer relationship in the labor force. 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.

In the case of h_{POP}, the estimated macro-level return to schooling is higher than the micro-level return λ, and so is the macro-level return to work experience as compared to the micro-level return μ. The relationships are reversed in the case of h_{LF}. In the simplified macro-Mincer equation, despite the arguably good empirical fit, returns to schooling are systematically underestimated.

B.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. 5. 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, both in the whole population and in the labor force. Furthermore, the estimated log-linear equation – although quantitatively close to the "true" relationship – cannot match its curvature, leading to systematic errors.
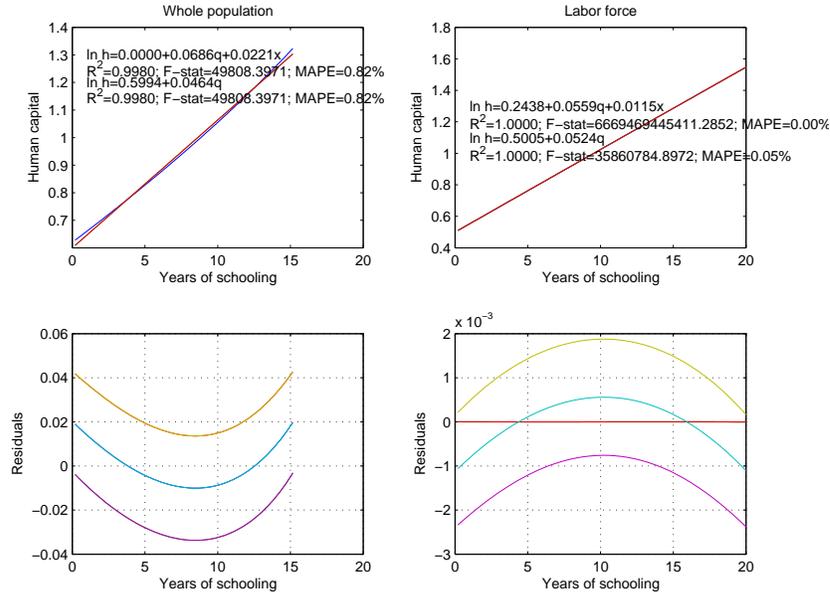
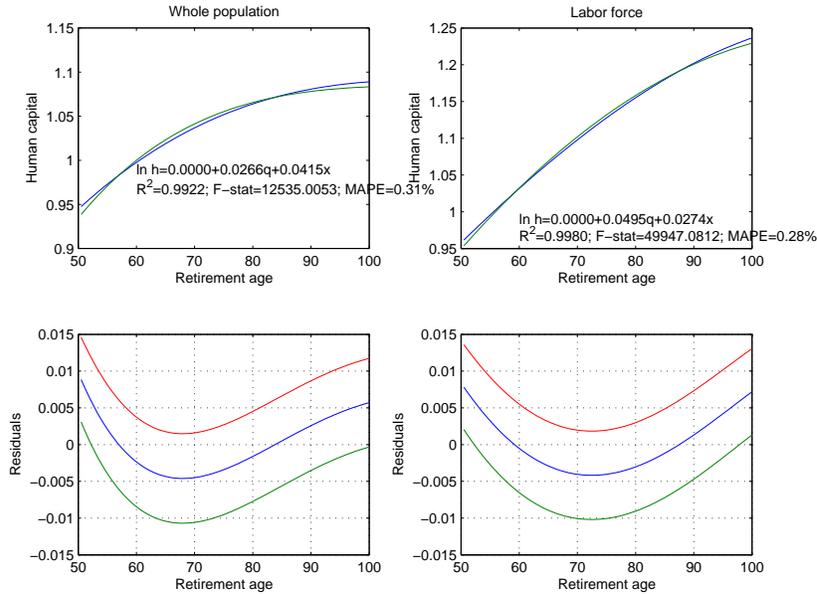Figure 4: Quality of approximation of average human capital with the macro-Mincer equation: the case of varying S.

Figure 5: Quality of approximation of average human capital with the macro-Mincer equation: the case of varying R.

B.4 Heterogeneity in life expectancy

As it is visible in Fig. 6, if the only source of cross-country heterogeneity is located in the parameters of the survival law, α and β, mapping uniquely into the measures of life expectancy E and the

maximum lifespan T^*, the macro-Mincer equation fits the data almost perfectly again.[17] The bias in estimates of macro-level returns to schooling and work experience is very small in the current case, especially when the macro-Mincer equation for the labor force is considered.
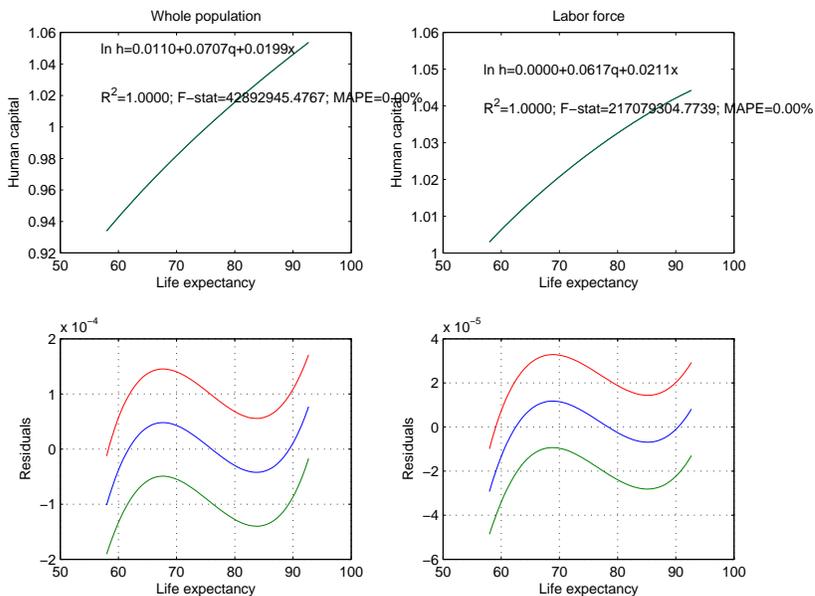

Figure 6: Quality of approximation of average human capital with the macro-Mincer equation: the case of varying E for a fixed T^*.

B.5 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.

In a related stylized example, we have also 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 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*

---

[17]Fig. 6 considers the case where the parameters α and β are varied simultaneously in a way that the maximum lifespan T^* is kept constant. The alternative where α is varied keeping β constant (and thus both E and T^* are variable) is available upon request. The results are qualitatively and quantitatively similar in both cases.
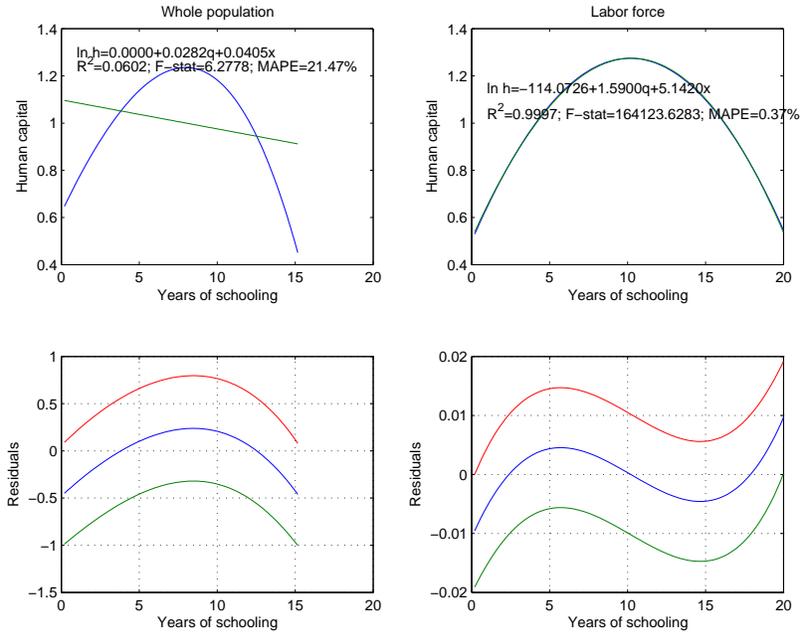
Figure 7: Quality of approximation of average human capital with the macro-Mincer equation: the case of varying S, perfectly negatively correlated with a varying λ.

functional relationship between them: the higher is S, the lower is λ.[18]

The aggregation results obtained under such circumstances are presented in Fig. 7. As we see, the macro-Mincer equation for the whole population misses the true human capital level completely. The macro-Mincer equation for the labor force fits the data very well, though, 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 4 below).

B.6 Accounting for human capital depreciation

With or without human capital depreciation, the goodness of fit of the macro-Mincer equation is vastly different whether the returns parameters λ and μ are allowed to vary across countries or not. 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-

---

[18]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.

Mincer regression depends crucially on the magnitude of variation of λ and μ in the sample.

Fig. 8 illustrates that allowing for human capital depreciation does not overturn the conclusion that the macro-Mincer equation fits the data remarkably well if rates of return λ and μ are equal across countries. 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. 9, 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 rather bad.

In both numerical examples, we are maintaining that S, R and α are randomly drawn for each country, like in section 3.4 of the article. In the current numerical exercise, all five parameters $S, R, \alpha, \lambda, \mu$ are generated independently. Again, the individual impact of human capital depreciation on the goodness-of-fit statistics of the macro-Mincer equation is rather small.
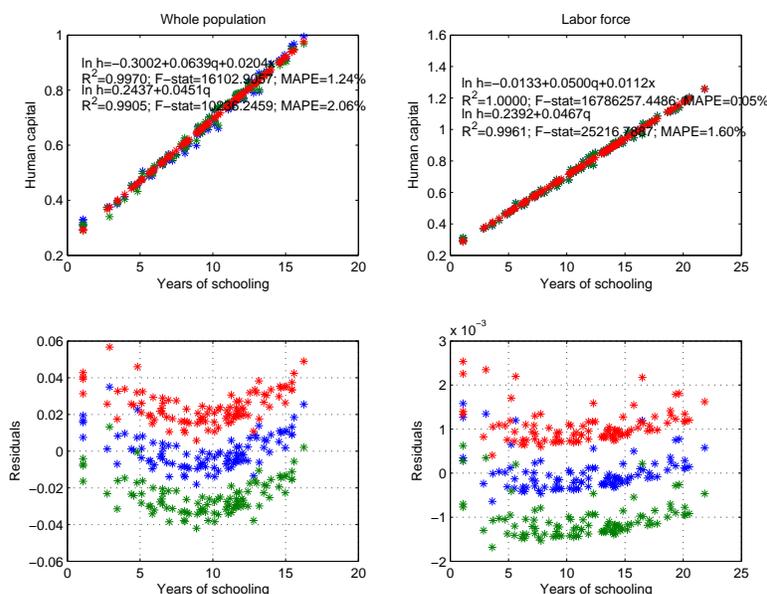

Figure 8: Quality of approximation of average human capital with the macro-Mincer equation: the case of randomly varying S, R, α, controling for human capital depreciation.

B.7   Monte Carlo study

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 3) 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.
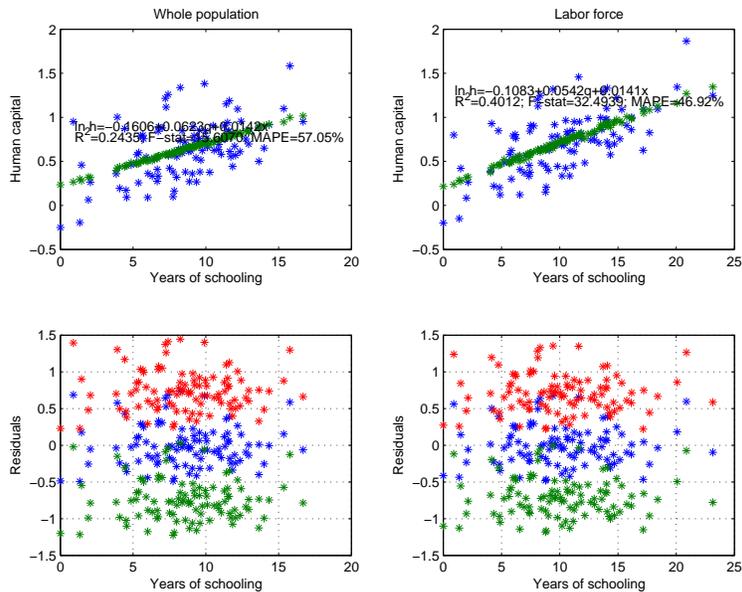
Figure 9: Quality of approximation of average human capital with the macro-Mincer equation: the case of randomly varying $S, R, \alpha, \lambda, \mu$, controling for human capital depreciation.

Quantitative results, for rather large $(S.D.(\lambda) = S.D.(\mu) = 0.01)$ 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 4. It must be noted that even our case of "large" 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.

These numerical results lead us to the conclusion that if the true cross-country variability in returns to schooling and/or work experience are in fact substantial, then the macro-Mincer equation is a poor approximation of the true relationship between aggregate human capital, average years of schooling, and average work experience. This result would remain quantitatively important even if existing datasets (e.g., Psacharopoulos and Patrinos, 2004) in fact overestimated the variability of these returns by a factor of four.

This result, however, is not really a consequence of human capital heterogeneity across the aggregated cohorts, but primarily of the imposition of a single value of the Mincerian schooling coefficient as well as the Mincerian experience coefficient to a heterogenous group of countries. As such, this result remains outside of the scope of applications of the theoretical model.

Table 4: Results of the Monte Carlo study – the case of heterogeneous returns to schooling and work experience.

	R^2_{POP}	R^2_{LF}	M_{POP} [%]	M_{LF} [%]	a_0	a_1	a_2	b_0	b_1	b_2
				Micro-level benchmarks						
					0	0.06	0.02	0	0.06	0.02
			$S.D.(\lambda) = S.D.(\mu) = 0.01$							
Mean	0.3072	0.4725	158.8161	516.4286	-0.3481	0.0664	0.0228	-0.0208	0.0499	0.0122
S.D.	0.0809	0.0751	1561	18556	0.9008	0.0335	0.0330	0.3915	0.0073	0.0172
			$S.D.(\lambda) = S.D.(\mu) = 0.001$							
Mean	0.9744	0.9887	3.6953	3.0938	-0.2888	0.0639	0.0197	-0.0129	0.0500	0.0112
S.D.	0.0047	0.0022	0.3417	0.2936	0.0970	0.0036	0.0035	0.0403	0.0008	0.0018
			$S.D.(\lambda) = S.D.(\mu) = 0$							
Mean	0.9971	1.0000	1.1929	0.0427	-0.2912	0.0640	0.0198	-0.0140	0.0500	0.0112
S.D.	0.0006	0.0000	0.1293	0.0062	0.0389	0.0015	0.0014	0.0009	0.0000	0.0000

Notes: M denotes MAPE. The regression parameters a_0 through b_2 are defined in equations (37)–(38).