Working Papers / Documents de travail

Longevity, Age-Structure, and Optimal Schooling

Noël Bonneuil
Raouf Boucekkine

WP 2014 - Nr 49

Longevity, age-structure, and optimal schooling

Noël BONNEUIL*

Institut national d'études démographiques,

and École des hautes études en sciences sociales

Raouf BOUCEKKINE†

Université d'Aix-Marseille

(Aix-Marseille School of Economics),

Cnrs and Ehess, France,

senior member, *Institut universitaire de France*

February 2015

Abstract

The mechanism stating that longer life implies larger investment in human capital, is premised on the view that individual decision-making

*bonneuil@ined.fr

†Raouf.Boucekkine@uclouvain.be

governs the relationship between longevity and education. This relationship is revisited here from the perspective of optimal period school life expectancy, obtained from the utility maximization of the whole population characterized by its age structure and its age-specific fertility and mortality. Realistic life tables such as model life tables are mandatory, because the age distribution of mortality matters, notably at infant and juvenile ages. Optimal period school life expectancy varies with life expectancy and mortality. Applications to stable population models and then to French historical data from 1806 to nowadays show that the population age structure has indeed modified the relationship between longevity and optimal schooling.

keywords: longevity, schooling, school life expectancy, age structure.

JEL C02, C65, C80, D90, H50, I26, I28, J10, J24

1 Introduction

Ben-Porath (1967) suggested a "mechanism" according to which longer life spans imply larger investment in human capital. This mechanism is central to several new growth theories in the field of unified growth theory. For example, Boucekkine et al. (2002, 2004) argued that early increases in life expectancy, mediated by the Ben-Porath mechanism, are at the origin of modern growth: longer lives would induce longer schooling and higher rates of human capital accumulation, giving rise to a new growth regime which allows the escape from the Malthusian trap.

French historical data are accurate enough to give us an insight into the relationship between longevity and education, at the moment of both the schooling and the mortality transitions. From the accurate estimates of female life expectancy at birth by French *départements* from 1806-10 to 1901-05 of Bonneuil (1997a) and from French schooling rates in 1837, 1850, 1867, and 1876 (Bonneuil, 2014), we regressed the time series of the growth rate of female life expectancy (tested to be stationary) on the growth rate of the female schooling rate (also tested to be stationary) in 1837-50, 1850-67, and 1867-76. On 82 French *départements*, these regressions yield 11 positive correlations (for *départements* located erratically on the territory), 6 negative, and 65 non significant. Then, in spite of the scarcity in time, the absence of clear correlation raises doubts that the relationship between longevity and schooling would be unequivocal.

In a quantity-quality trade-off model *à la* Becker (1991), Hazan and Zoabi (2006) show that the Ben-Porath mechanism may not always work because an increase in longevity affects not only the return to schooling (quality of children), but also the return to quantity or the optimal total number of children. The latter effect mitigates the Ben Porath mechanism and can in principle negate it. Under homothetic preferences, Hazan and Zoabi (2006) find that when fertility is endogenous, an exogenous increase in children's longevity has no effect on schooling.

Acemoglu and Johnson (2007) used empirical arguments to show the existence of a causal effect from health and disease environments on economic growth. Exploiting the wave of innovations in health worldwide from 1940 onward, they found no significant effects from changes in life expectancy on either total GDP or human capital. Using the available data on schooling from 1960 onward, they found no evidence of any Ben-Porath-like mechanism, and they concluded that "the most likely reason why the increase in life expectancy did not translate into greater education during this episode is that the affected countries faced bottlenecks in their education systems." Bloom, Canning, and Fink (2014), relaxing the assumption that initial health and income do not affect economic growth, used the same data to show that, on the contrary, the health transition should induce higher income levels.

Hazan (2010) questions the Ben-Porath mechanism explicitly. He starts from the cohort model of Boucekkine et al. (2002): all individuals of all cohorts are identical and make decisions about their lifetime consumption, schooling, and work time. Attending school for a longer time has a cost in terms of foregone labor income but this schooling time also induces a gain because longer schooling means higher wages in the labor market. In the case of a perfectly rectangular survival function, increased longer longevity leads to longer schooling only if the total expected number of hours spent at work during one's lifetime also rises. Hazan (2010) tested this property on US data for consecutive 10-year cohorts born between 1840 and 1970, to find that the total number of hours worked did not increase, and to conclude that the Ben-Porath mechanism was not relevant for the US during this period. Cervellati and Sunde (2013), however, argued that the connection between the total number of hours spent at work during an individual's lifetime and the Ben-Porath mechanism does not hold for non-perfectly rectangular survival functions.

All the studies referred to so far, except Acemoglu and Johnson (2007), are based on individual decision-making where agents decide about their optimal consumption stream and their lifetime accumulation of human capital over lifetime for a given ad-hoc survival function. As we mentioned, Hazan and Zoabi (2006) used a two-period overlapping generation model *à la* Becker. Hazan (2010) and Cervellati and Sunde (2013) used a continuous time (homogeneous) cohort model similar to the perpetual youth model *à la* Yaari-Blanchard, incorporating a schooling time decision model like Boucekkine et al. (2002, 2004).[1] However,

[1] Introducing within-cohort heterogeneity adds formidable complications (Boucekkine et al.,

decisions of education are not individual; on the contrary, at least in continental Europe, education is mainly run by the State. We innovate in relying on the alternative criterion of optimal period schooling, which is the only one so far to involve the whole age structure.

A basic formulation of the optimal period schooling equivalent could be the following: given the age structure of the population, its current fertility and mortality levels, would a planner seeking to maximize social welfare (say with respect to the Benthamite criterion) lengthen schooling in response to rising life expectancy? In contrast to the individual-cohort perspective adopted in the related literature, the key component of the period schooling optimum problem is that the decisions have to be taken on the basis of the overall demographic structure. Therefore, the current age structure of the population determines schooling decisions, whereas it is ignored in the literature on individual schooling decisions. Put simply, the main reason supporting the argument that longer life implies longer education is that the proportion of people above the maximum school completion age increases as mortality decreases. But this holds true only if the proportion of people under that age does not increase faster. Following a cohort, as did Hazan (2010) or Cervellati and Sunde (2013), is equivalent to fixing fertility at a constant value over time. The conditions of the moment however vary with fertility, so that if the proportion of people in school increases, compared to the proportion out of school, greater longevity may not be enough

2007), which are not germane to our purpose here.

to offset higher fertility, leading to *a decrease* in schooling, at the optimum. The relationship between length of life and schooling will thus depend on the balance between the additional young from higher fertility and the additional old from improved survival.

We shall then work with the concept of the *fictitious cohort*, where the forces of change affecting people of successive ages living at the same time are applied to a fictitious group of people born at a same time and who would experience these age-specific forces successively as they age. The tool of fictitious cohort is common in mathematical demography (Bonneuil, 1997b). The associated concept to the fictitious cohort is "school life expectancy", which is defined as "the total number of years of schooling (primary to tertiary) that a child can expect to receive, assuming that the probability of his or her being enrolled in school at any particular future age is equal to the current enrolment ratio at that age"[2] The *period* school life expectancy is then the sum over all ages of the probabilities at a given date to remain in school.

The conditions of the moment consist in the expected age of the oldest schooled individual at time t, the consumption schedule by age at t, the leisure schedule by age at t, the fertility schedule by age at t, and the life table at t. The optimal planner decides from the knowledge of these conditions of the moment, not from the knowledge of the stocks of population or from the school experience of each cohort present at t. In demography, family policies are based on

[2] https://www.cia.gov/library/publications/the-world-factbook/fields/2205.html

period fertility rates, often gathered in a total fertility rate. This rate is used in orientations of population policies, no matter the total number of children born from each cohort of women at time t. Similarly, period life expectancy gathers the mortality rates of the moment in a fictitious cohort, and insurance policies are decided from period life expectancy, not from cohort life expectancies. The optimal planner cannot modify past variables and is aware that, if the conditions at date t were maintained for a long enough time, the population would converge to the stable population associated with the fertility schedule and the life table at t ("ergodicity" theorem (Cohen, 1979)). Consequently, the optimal planner, taking decisions from the conditions of the moment, and not from the stocks inherited from the past, prepares the future.

We then move away from the Ben Porath framework, yet we still raise the question of the relationship between schooling and longevity. We shall study the optimal schooling rule adopted by a social planner and see how this rule responds to increased longer longevity. For the sake of clarity, we shall assume that the planner fixes the income transfer rates from working to non-working ages (people at school and pensioners) in such a way that the budget is never in deficit. Allowing the planner to borrow (subject to a no-Ponzi game condition) complicates the technicalities, but handling the age structure is already onerous (leading to nonlinear functional integral equations). So, we impose zero deficit at any date. By doing so, we offer a symmetrical setting to the one adopted in the cohort-individual literature. In the latter, the individual (representing the fictitious cohort) is followed-up over his/her lifetime (inter-temporal optimization). In contrast to Acemoglu and Johnson (2007), who criticized the Ben-Porath mechanism with a supply side argument (bottlenecks in the education sector modify the relationship between longevity and schooling), we focus on the demand side, with emphasis on the population age structure, to identify the cases where this mechanism fails to be optimal, even when education supply is not at stake.

In Section 2, we describe optimal period schooling from conditions of the moment through a maximization program, where the fictitious cohort has the structure of stable population associated with the fertility and the mortality of the moment. Stable populations correspond to constant mortality and fertility; they are also used to characterize the conditions of a current period, because a given population exposed to the conditions of the moment for a long time would converge to the stable population corresponding to the fertility and mortality schedules of the period then prevailing (property of strong ergodicity). In Section 3, we show how the period optimal schooling is modified when life expectancy or fertility are varied. We shall specify the model with realistic model life tables, and, in Appendix, present analytical yet unrealistic cases (constant, linear, and Gompertz mortality) to show the difficulty of obtaining analytical formulas. We shall situate empirical cases of French *départements* in the space defined by fertility, life expectancy, and schooling. Finally, we draw the trajectory of France as a whole in this space and discuss optimal schooling with demographically realistic assumptions. Our main result is that the relationship between longevity and schooling is mediated by infant mortality and by fertility, so that the the link between longer life and longer schooling is optimal only in countries with low mortality and low fertility.

2 Period optimal schooling, through the fictitious cohort

2.1 The problem

Instead of tracking individuals' life-cycle decisions, our schooling variable is not individual, but designates the *total* schooling time which is optimal at a given date. The program consists of the optimization at each date of welfare:

$$\max_{c(t,.),l(t,.),A_S(t)} V(t) \quad = \quad \int_0^{\bar{A}} p(t,a)u(c(t,a),\ell(t,a))\,da, \qquad \forall t, \tag{1}$$

where, at age a and time t, $p(t,a)$ represents the total number of people, $c(t,a)$ consumption per head; each individual of age a has one unit of time available, shared between $\ell(t,a)$ in leisure and $1-\ell(t,a)$ in labor supply for active or in time spent in school, u is the continuously differentiable strictly concave utility function, assumed additively separable in consumption and leisure; $\bar{A}$ is the maximal life span, $A_S(t)$ is the age of the oldest schooled individual at time t, and $S(t)$ is the total amount of schooling time at t. In contrast to the literature focusing on individual schooling decisions, the schooling variable $S(t)$ measures aggregate time at school spent across the whole economy at time t. We assume that for given total schooling S, productivity at work is an increasing function of S, say $\exp(\theta(S))$, where $\theta(.)$ is an increasing and concave production function. This mimics the specification in the literature with individual maximization. The exact timing of the relationship between schooling and productivity is made explicit here below. The constraints under which the optimization is solved are:

$$\int_0^{\bar{A}} p(t,a)c(t,a)\,da = \int_{A_S(t)}^{A_R} (1-\ell(t,a))p(t,a)w(t,a)\,da, \tag{2}$$

where $w(t,a)$ denotes wages at age a and time t. Equation (2) equalizes total consumption of individuals at every age with labor income in the stable population associated with the conditions of fertility and mortality of the moment. The age at retirement A_R is not a control variable, but taken here as constant for the sake of simplicity, because we focus on schooling. The school life expectancy is:

$$S(t) := \int_0^{A_S(t)} (1-\ell(t,b))\,db. \tag{3}$$

It corresponds to no cohort, but to a fictitious one, which at age b would have spent $(1-\ell(t,b))$ time units in school. This is the same calculus used in the total fertility rate or in the period life expectancy. It measures the conditions of schooling of the moment t. A particular case of wages, dependent of the school period life expectancy, is:

$$w(t,a) = \iota(a)\exp(\Theta(S(t)), \tag{4}$$

where ι describes a schedule of wages across age, after controlling for date. The specification of (4) highlights the presence of age in wages, for the sake of realism, but it is not necessary to study the influence of the age structure in the relationship between longevity and schooling.

State variables are the age-specific population $p(t,a)$ and the total schooling time $S(t)$; control variables are consumption $c(t,a)$, leisure $\ell(t,a)$, and the age $A_S(t)$ of the oldest schooled individual.

The Lagrangian is built from (1) and (2). The first-order condition obtained by differentiating with respect to $A_S(t)$ yields the relationship between schooling time and population:

$$\frac{1}{\theta'(S(t))} = (1-\ell(t,A_S(t))) \int_{A_S(t)}^{A_R} \frac{p(t,a)}{p(t,A_S(t))} \frac{1-\ell(t,a)}{1-\ell(t,A_S(t))} \, da. \tag{5}$$

It is also obtained from the first-order conditions obtained by differentiating with respect to $\ell(t,a)$:

$$\begin{array}{llcl} \text{for } a > A_S(t) & u'_\ell(c(t,a),\ell(t,a)) & = & \lambda e^{\theta(S(t))}\iota(a) \\ \text{for } a \leq A_S(t) & p(t,a)u'_\ell(c(t,a),\ell(t,a)) & = & \lambda \int_{A_S(t)}^{A_R} p(t,b)(1-\ell(t,b))e^{\theta(S(t))}\theta'(S(t))\iota(b)\, db, \end{array} \tag{6}$$

where λ is the Lagrangian multiplier, and equating λ in these two equations.

Equation (5) is close to Cervellati and Sunde (2013), who considered a single cohort to fix leisure time, time spent at school, and consumption from birth to death in a single maximisation program. These authors follow an individual, assimilated to a cohort, and used *cohort* school life expectancy (but there is a single cohort) as control variable. By construction, they did not deal with population, but only with survival of the cohort along with age. The decisive difference with us is that (5) involves not only survival, but age pyramids, which mix survival *and* fertility.

Equation (5) is also:

$$(1-\ell(t,A_S(t)))\ \int_{A_S(t)}^{A_R} p(t,a)(1-\ell(t,a))\theta'(S(t))\,da = (1-\ell(t,A_S(t)))\ p(t,A_S(t)), \tag{7}$$

where the left-hand side is the marginal benefit from increasing the marginal cost of increasing the last age of the oldest schooled individual (that is, by increasing $A_S(t)$), which depends on the curvature of the return function $\theta(.)$ through $S(t)$ and on the age-structure of the active population. The right-hand side is the marginal social loss, which consists of the income forgone by postponing entry into the labor market until age $a = A_S(t)$. In contrast to the formulas derived in Boucekkine et al. (2002), here the schooling decision depends on the age-structure.

3 Stable population

A stable population is a population closed to migration and characterized by constant fertility and mortality. These forces determine the age structure. Conversely, a stable age structure characterizes given fertility and mortality flows.

The stable population associated with the conditions of the moment t, namely mortality $\mu(t,a)$ and fertility $\phi(t,a)$, has its population growth rate $\rho(t)$ given by the Lotka equation:

$$1 = \int_0^{\bar{A}} \sigma(t,a) e^{-\rho(t)a} \phi(t,a)\, da\,, \tag{8}$$

where $\sigma(t,a) = 1 - \exp(-\int_0^a \mu(t,u)\, du)$ is the period survival function at t. The stable population generated by $\mu(t,a)$ and $\phi(t,a)$ has its own kinematics, indexed by a time τ, but for this population, the time t at which mortality and fertility are taken, is constant. We denote the population growth rate $\rho(t)$, but it is in fact dependent on $e_0(t)$ and on $\phi(t,.)$.

At time τ, the total number of people

$$p(\tau,a) = P(t) e^{\rho(t)\tau} \frac{\sigma(t,a) e^{-\rho(t)a}}{\int_0^{\bar{A}} \sigma(t,b) e^{-\rho(t)b}\, db} \tag{9}$$

aged a in this stable population (e.g. Bonneuil, 1997) depends on the mortality $\mu(t,a)$ and the population growth rate $\rho(t)$, which themselves do not depend on the kinematics inherent in the stable population and indexed by the time τ. The population size $P(t)$ at time t is the initial population for the kinematics of the associated stable population $p(\tau,a)$, $\tau = t, \cdots$.

By replacing $p(t,a)$ by its expression in Eq. (9) for $\tau = t$, Eq. (5) becomes:

$$\frac{1}{\theta'(S(t))} = (1 - \ell(t, A_S(t))) \int_{A_S(t)}^{A_R} \frac{\sigma(t,b)}{\sigma(t, A_S(t))} \frac{1 - \ell(t,b)}{1 - \ell(t, A_S(t))} e^{-\rho(t)(b - A_S(t))}\, db. \tag{10}$$

At constant fertility, if the life expectancy increases, $\mu(t,a)$ decreases for every age a and $e^{-\int_a^b \mu(t,u)\, du}$ increases, but $\rho(t)$ also increases and $e^{-\rho(t)(b-a)}$ decreases. So the two effects are in opposite directions.

The effect of an increase in period life expectancy $e_0(t) := \int_0^\infty a\mu(t,a)\sigma(t,a)\, da =$

$\int_0^\infty \sigma(t,a)\,da$ on $\frac{1}{\theta'}$ depends on the sign of:

$$\begin{aligned}\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)} &= (1-\ell(t,A_S(t)))\int_{A_S(t)}^{A_R} \frac{1-\ell(t,b)}{1-\ell(A_S(t))} e^{-\int_{A_S(t)}^{b}\mu(t,u)\,du} e^{-\rho(t)(b-A_S(t))} \\ &\qquad (-\int_{A_S(t)}^{b} \frac{\partial\mu(t,u)}{\partial e_0(t)}\,du - (b-A_S(t))\frac{\partial\rho(t)}{\partial e_0(t)})\,db\,. \end{aligned} \tag{11}$$

or

$$\begin{aligned}\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)} &= (1-\ell(t,A_S(t)))\int_{A_S(t)}^{A_R} \frac{1-\ell(b))}{1-\ell(A_S(t))} e^{-\int_{A_S(t)}^{b}\mu(t,u)\,du} e^{-\rho(t)(b-A_S(t))} \\ &\qquad (-\int_{A_S(t)}^{b} \frac{\partial(\rho+\mu(t,u))}{\partial e_0(t)}\,du)\,db\,. \end{aligned} \tag{12}$$

From (8), we obtain $\frac{\partial\rho(t)}{\partial e_0(t)}$ from the Lotka equation:

$$\frac{\partial\rho(t)}{\partial e_0(t)} = \frac{\int_0^{\bar{A}}(-\int_0^a \frac{\partial\mu(t,v)}{\partial e_0(t)}\,dv)e^{-\rho(t)a}\sigma(t,a)\phi(t,a)\,da}{\int_0^{\bar{A}} ae^{-\rho(t)a}\sigma(t,a)\phi(t,a)\,da}, \tag{13}$$

which is positive, because $\frac{\partial\mu(e_0(t),u)}{\partial e_0(t)} < 0$: empirical observation as well as mortality models indicate that mortality decreases at all ages when life expectancy increases.

Substituting $\frac{\partial\rho(t)}{\partial e_0(t)}$ into Eq. (12) yields:

$$\begin{aligned}\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)} &= (1-\ell(t,A_S(t)))\int_{A_S(t)}^{A_R} \frac{1-\ell(t,b)}{1-\ell(t,A_S(t))} e^{-\int_{A_S(t)}^{b}\mu(t,u)\,du} e^{-\rho(t)(b-A_S(t))} \\ &\qquad (-\int_{A_S(t)}^{b} \frac{\int_0^{\bar{A}}(-\int_0^a \frac{\partial\mu(t,v)}{\partial e_0(t)}\,dv)e^{-\rho(t)a}\sigma(t,a)\phi(t,a)\,da + \frac{\partial\mu(t,u)}{\partial e_0(t)}\int_0^{\bar{A}} ae^{-\rho(t)a}\sigma(t,a)\phi(t,a)\,da}{\int_0^{\bar{A}} ae^{-\rho(t)a}\sigma(t,a)\phi(t,a)\,da})\,du\,db \\ &= (1-\ell(t,A_S(t)))\int_{A_S(t)}^{A_R} \frac{1-\ell(t,b)}{1-\ell(t,A_S(t))} e^{-\int_{A_S(t)}^{b}\mu(t,u)\,du} e^{-\rho(t)(b-A_S(t))} \\ &\qquad (-\int_{A_S(t)}^{b} \frac{\int_0^{\bar{A}}(\int_0^a (\frac{\partial\mu(t,u)}{\partial e_0(t)} - \frac{\partial\mu(t,v)}{\partial e_0(t)})\,dv)e^{-\rho(t)a}\sigma(t,a)\phi(t,a)\,da}{\int_0^{\bar{A}} ae^{-\rho(t)a}\sigma(t,a)\phi(t,a)\,da})\,du\,db \end{aligned} \tag{14}$$

Marchand and Thélot (1991) explain that, until the nineteenth century, usual daily work coincided with daylight, and that work hours had been remaining the same, be it in town or in countryside, varying only with seasons. They do not mention any age component, but their description is consistent with the fact that age had no effect on the duration at work. OECD publishes incidences of employment by usual weekly hours worked and by quinquennial age class only from 2000 onwards.[3] From these data, we computed the mean total number of worked hours by age class: for France, this number was 36.5 (sd=0.3) for men and women aged 20-24, 37.9 (sd=0.3) for the 25-29, 38.3 (sd=0.2) for the 30-34, 38.3 (sd=0.3) for the 35-39, the 40-44, and the 45-49 age classes. The proximity of these numbers allows us to assume that the age profile of labor supply in one individual's day is constant with age, then:

$$\forall b \in (A_S(t), A_R], \quad 1 - \ell(t,b) = 1 - \ell(t, A_S(t))\,. \tag{15}$$

Then $\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)}$ has the sign of:

$$\begin{aligned} f(e_0(t), \phi(t,.)) \quad = \quad & (1 - \ell(t, A_S(t))) \int_{A_S(t)}^{A_R} e^{-\int_{A_S(t)}^{b} \mu(t,u)\,du} e^{-\rho(t)(b - A_S(t))} \\ & \left(- \int_{A_S(t)}^{b} \frac{\int_0^{\bar{A}} (\int_0^a (\frac{\partial \mu(t,u)}{\partial e_0(t)} - \frac{\partial \mu(t,v)}{\partial e_0(t)})\,dv) e^{-\rho(t)a} \sigma(t,a)\phi(t,a)\,da}{\int_0^{\bar{A}} a e^{-\rho(t)a} \sigma(t,a)\phi(t,a)\,da} \right) du\,db\,. \end{aligned} \tag{16}$$

where $\phi(t,.))$ is the fertility pattern underlying the value of the population growth rate $\rho(t)$ with respect to Eq. (8).

Figure 1 shows the value of $\frac{\partial \mu(t,a)}{\partial e_0(t)}$ from Ledermann model life tables, which are very realistic,[4] contrary to analytical formulas, which at best capture only certain age intervals (after 40 years for a Gompertz curve, for example). Figure 1

[3] http://stats.oecd.org/Index.aspx?DataSetCode=AVE_HRS&Lang=fr

[4] Model life tables are of common use in demography. Ledermann model life tables are semi-parametric, resulting from regressions performed at each age on some 300 empirical life tables.

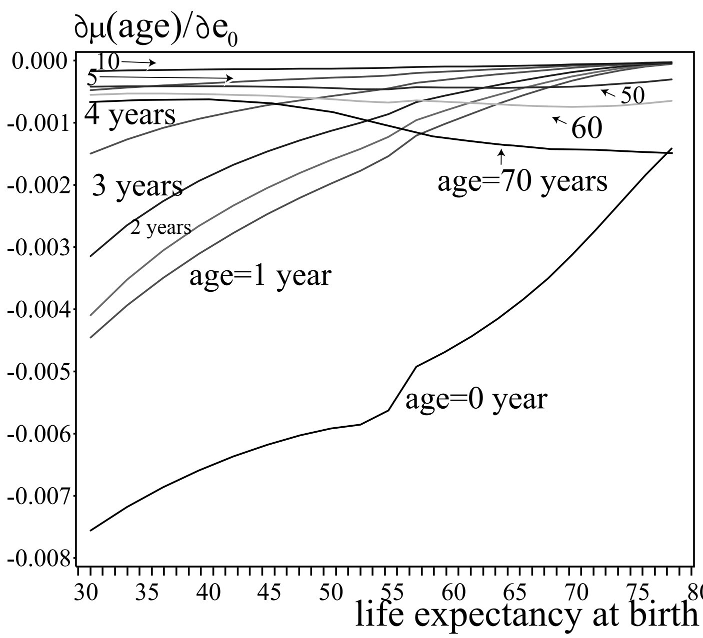

Figure 1: Value of $\frac{\partial \mu(t,a)}{\partial e_0(t)}$ with respect to life expectancy. Ledermann model life tables.

shows that infant ages, up to 4 years, are determinant in the computation of $\frac{\partial \mu(t,u)}{\partial e_0(t)} - \frac{\partial \mu(t,v)}{\partial e_0(t)}$: with increasing life expectancy, this importance of infant ages gradually vanishes, and the sign of $\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)}$ becomes more dependent on what occurs at old ages. In contrast to what happens at low life expectancy, $\frac{\partial \mu(t,u)}{\partial e_0(t)} - \frac{\partial \mu(t,v)}{\partial e_0(t)}$ for old ages dominates for high life expectancy. It is lower than values of this expression for younger ages, so that the sign of $\frac{\partial \mu(t,u)}{\partial e_0(t)} - \frac{\partial \mu(t,v)}{\partial e_0(t)}$ changes. Subsequently, the sign of $\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)}$ changes, too. *The shape of the life table is then determinant in the relationship between longevity and optimal schooling time.*

Eq. (16) shows that $\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)}$ depends on the improvement in mortality at young ages: when mortality decreases faster at young than at old ages:

$$\frac{\partial \mu(t,v)}{\partial e_0(t)} < \frac{\partial \mu(t,u)}{\partial e_0(t)} < 0 \quad \text{for } v < u, \tag{17}$$

which was historically the case at the beginning of the mortality transition, then from Eq. (16), $\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)} < 0$. The important decline of infant mortality in comparison with mortality at other ages implies that a larger proportion of children surviving and going to school, while juvenile mortality has not yet decreased greatly. The population includes more young children, but not many teenagers, whose mortality has decreased less. The consequence is that the average total number of schooling years *decreases*.

Conversely, in the post-transition era, infant mortality is already low and mortality at older ages has declined substantially, so that

$$\frac{\partial \mu(t,v)}{\partial e_0(t)} \geq \frac{\partial \mu(t,u)}{\partial e_0(t)} < 0 \quad \text{for } v < u \tag{18}$$

and $\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)} < 0$.

The sign of $\frac{\partial \mu(t,u)}{\partial e_0(t)} - \frac{\partial \mu(t,v)}{\partial e_0(t)}$ is modulated by $e^{-\rho(t)(t)a}\sigma(t,a)\phi(t,a)$ and by $e^{-\int_{A_S(t)}^{b} \mu(t,u)\,du} e^{-\rho(t)(b-A_S(t))}$: we propose to specify the fertility schedule and examine the relationship between longevity and schooling through their determinants which are fertility and mortality.

4 Empirical analysis

4.1 Specification

4.1.1 Mortality

Mortality studies have shown that mortality is regular with age only over short periods, say a year, subsequently across different cohorts living at the same time. Cohort mortality patterns on the contrary look irregular, because a cohort successively experiences uneven conditions. This is the reason why models of mortality are meaningful only for fictitious cohorts at a given period, and always unrealistic for actual cohorts followed-up over time.

In the Appendix, we present two toy cases: constant and linear mortality, to show that the relationship between longevity and schooling rapidly becomes complicated, then the more realistic case of Gompertz, because it is a well-known mortality model, realistic only after 40 years of age. However, as infant mortality plays a major role, only the case of model life tables is worth considering.

The model life tables we use (Ledermann, 1969) are semi-parametric, indexed by a single parameter $e(t)$, close to the life expectancy at birth:

$$\log_{10}(1 - e^{-\int_0^5 \mu(t,a+u)\,du}) = \gamma_a^1 + \gamma_a^2 \log_{10}(100 - e(t)) \tag{19}$$

where γ_a^1 and γ_a^2, $a = 0, 5, \cdots, 85$ were estimated by Ledermann (1969) from some three hundred empirical life tables. With this semi-parametric formula, we compute $\frac{\partial \mu(t,u)}{\partial e_0(t)}$ and $\frac{\partial \rho(t)}{\partial e_0(t)}$ numerically.

4.1.2 Fertility

We introduced the age structure through the simple stable population model. The population growth rate $\rho(t)$ then appears in Eq. (12) and (13). Rather than treating this population growth rate as exogenous, we gain better insight into the demographic processes at work by revealing its dependence on fertility with respect to Eq. (8).

Coale and Trussel (1974) calibrated the semi-parametric model of realistic five-year fertility schedules:

$$\phi(a) = M n(a) e^{m\nu(a)} \tag{20}$$

where $n(a)$ and $\nu(a)$ are given distributions by age. The parameter $M \geq 0$ represents a fertility level, $m \geq 0$ conditions the shape of the distribution: $m > 0$ reflects family limitation, $m = 0$ the absence of family limitation. We obtain the yearly fertility schedule by cubic spline on the cumulated function $\int_{15}^{a} \phi(u)\,du$.

4.2 Results

Figure 2 shows the computation of $f(e_0(t), \phi(t,.))$ for Lederman life tables (with $e_0(t)$ varying) and for Coale-Trussel fertility schedules $\phi(t,.)$ (with m and M varying). As we mentioned, both fertility and mortality patterns are both realistic, and the result no longer comes from a simplifying assumption on these schedules. The value of the total fertility rate (TFR:= $\int_0^{\bar{A}} \phi(t,a)\, da$) is given on each (m, M) to help situate the level of fertility.

Figure 2 shows that $f(e_0(t), \phi(t,.))$ and subsequently $\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)}$ change with respect to life expectancy and with fertility, For example, with $m \approx 0$ for nineteenth century France (Bonneuil, 1997a), we situate French *départements* over the course of the demographic transition. The population of the Calvados *département* was one of the earliest to embark on fertility decline and experienced an almost stationary fertility during the nineteenth century: TFR=3.11 and $e_0 = 47.5$ in 1806, TFR= 2.74, and $e_0 = 45.8$ in 1906 (Bonneuil, 1997a). At that point Calvados crosses the line $\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)} = 0$ from positive to negative values (Figure 2a).

At the other end of the spectrum of the transition, the French *département* of Finistère, whose population was one of the last to enter the fertility decline (TFR= 8.09, $e_0 = 28.7$ years in 1806, TFR= 4.11, $e_0 = 39.5$ years in 1906), $\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0}$ remains positive during this period (Figure 2a). So, for the same country, at the same epoch, namely nineteenth-century France, we find locations with positive signs, others with negative signs.

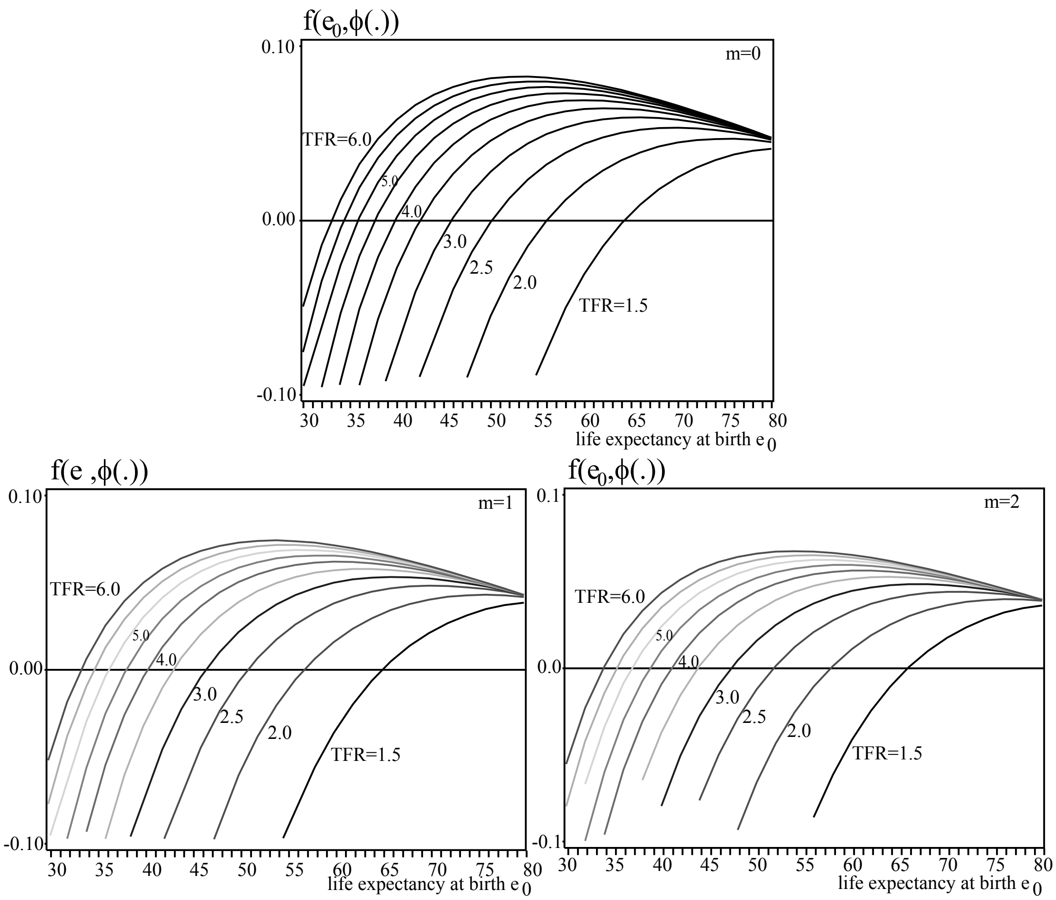

Figure 2: Value of $f(e_0(t), \phi(t, .))$, which has the same sign as $\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)}$, varying with life expectancy (Ledermann model life tables) and Coale-Trusssel fertility, m the parameter of family limitation, and TFR the total fertility rate. In this Figure, $\rho(t)$ is always positive.

In the problem of the relationship between longevity and schooling, one could ignore the role played by fertility, because children attending school are of a certain age. This would be a mistake, as it would mean using a mortality schedule which ignores the specific pattern of infant mortality. The relationship in fact depends on fertility as well as on mortality, because fertility determines the population growth rate and the age structure, which appears as $\frac{p(t,a)}{p(t,A_S(t))}$ in Eq. (10) for the general case or through $\rho(t)$ in Eq. (16) in the stable population case.

A very good way to capture the conditions of the moment in a real population is to fit the stable population associated with the fertility and the mortality schedules of this moment. Figure 3 shows the trajectory of $f(e_0, \phi(.))$ for these stable populations associated with the conditions of each year between 1806 and 1988 in the whole of France. The age-specific fertility distribution is known only from 1892 onwards (for the whole of the country). From 1806 to 1892, the age-specific schedule of 1892 is used with the TFR reconstructed by Bonneuil (1997a). Over the course of France's demographic transition, which took place during the nineteenth century, when life expectancy was still low, $f(e_0, \phi(.))$ is negative. With life expectancy increasing, this function becomes positive, with two abrupt returns to negative values during the World Wars, on account of the higher wartime mortality.

After the effect of declining fertility, we again find the key role played by infant mortality, which declined continuously (except during crises such as the 1870-71 Franco-Prussian war or the World Wars) between 1806 and 1988. The

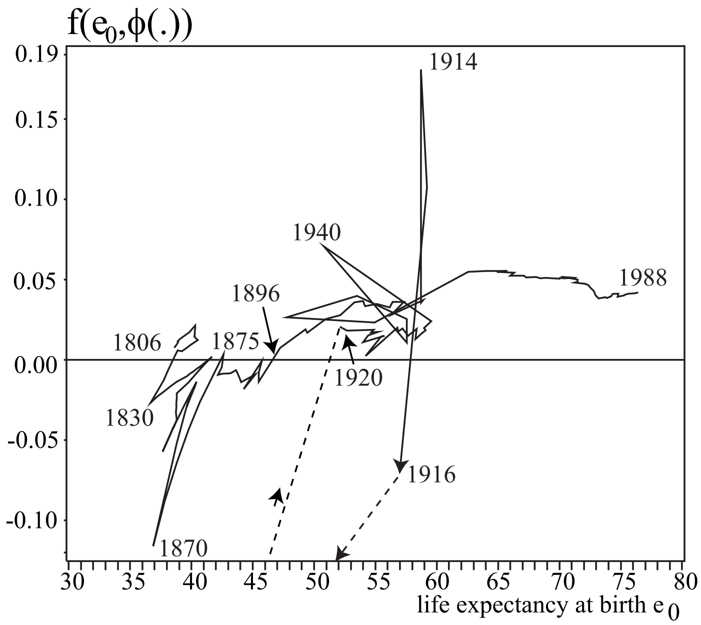

Figure 3: Value of $f(e_0(t), \phi(t,.))$, which has the same sign as $\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)}$, for the whole of France (*Data*: Bonneuil (1997a) for data from 1896 to 1892 and Ined for data after 1892.)

zero line of $f(e_0(t), \phi(t,.))$ was crossed when the variations of the mortality rate μ became dominated by the changes at old ages, and specifically in Figure 3, for period life expectancies over 46 years.

So, from Figures 2 and 3, in the case of θ strictly concave, the Ben-Porath proposition that schooling grows with longevity appears to be optimal only in post-transitional countries, and in conditions of sufficiently low mortality and low fertility. In countries currently experiencing the transition, especially in those where mortality is still high and fertility not high enough to attain the complete replacement of the population, the age structure is relevant, and longer life is no longer associated with schooling in a Ben Porath-like manner.

5 Conclusion

To address period optimal schooling, we proposed to solve the program (1). We obtained Eq. (12) and (13), from which the effect of an increase in life expectancy on schooling is not clear-cut and is influenced by the age structure. The realistic case of model life tables associated with model fertility schedules showed that the relationship between longevity and schooling depends on both life expectancy and fertility, in a non linear manner displayed in Figure 2. With varying fertility and mortality, notably during the demographic transition, the direction of this relationship may change or not. We took the French *départements* as empirical case studies. Finally we showed the trajectory of France as a whole from

1806 to 1988, and concluded that the Ben-Porath mechanism is optimal only in countries with low mortality and low fertility.

References

Acemoglu, D. and S. Johnson (2007): "Disease and Development: The Effect of Life Expectancy on Economic Growth," [Journal of Political Economy, 2007, 115()6, 925-985.

Ben-Porath, Y. (1967): "The production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 75, 352-365.

Becker, G. (1991): A Treatise on the Family, Cambridge: Harvard University Press.

Bloom, D., D. Canning, and G. Fink (2014): "Disease and Development Revisited," *Journal of Political Economy* 122(6), 1355-1366.

Bonneuil, N. (1997a): Transformation of the French Demographic Landscape, 1806-1906. Oxford: Clarendon Press.

Bonneuil N. (1997b): *Introduction à la modélisation démographique.* Paris: Armand Colin.

Bonneuil N. (2014) Morphological Transition of Schooling in 19th century France, *Journal of Mathematical Sociology* 38: 95-114.

Boucekkine, R., D. de la Croix, and O. Licandro (2002): "Vintage Human Capital, Demographic Trends, and Endogenous Growth," *Journal* *of Economic Theory*, 104(2), 340-375.

——————– (2004): “Modelling Vintage structures with DDEs: principles and applications,” *Mathematical Population Studies*, 11, 151-179.

Boucekkine, R., D. de la Croix, and D. Peters (2007): “Early literacy achievements, population density, and the transition to modern growth,” The Journal of the European Economic Association 5, 183-226.

Cervellati, M., and U. Sunde (2013): “Life Expectancy, Schooling, and Lifetime Labor Supply: Theory and Evidence Revisited” *Econometrica*, 81(5), 2055-2086, 2013.

Coale, A. J., and T. J. Trussel (1974): “Model fertility schedules: variations in the age structure of childbearing in human populations,” *Population Index*, 40, 185-258.

Cohen, J. (1979): “Ergodic theorems in demography,” *Bulletin of the American Mathematical Society*, 1, 275-190.

Hazan, M. and and Zoabi, H. (2006): “Does longevity cause growth? A theoretical critique”, Journal of Economic Growth 11(4), 363-376.

Hazan, M. (2009): “Longevity and Lifetime Labour Supply: Evidence and Implications,” *Econometrica*, 77(6), 1829-1863.

Ledermann, S. (1969): Nouvelles tables-types de mortalité. Paris: PUF/Ined.

Marchand, O. and Thélot, C. (1991): *Deux siècles de travail en France.* Paris: Insee.

Appendix: constant, linear, and Gompertz mortality

Constant mortality: $\forall u$, $\frac{\partial \mu(t,u)}{\partial e_0(t)} = k$ **constant** From Eq. (13),

$$\frac{\partial \rho(t)}{\partial e_0(t)} = -k \tag{21}$$

and from Eq. (8):

$$\frac{\partial (\frac{1}{\theta'(S(t))})}{\partial e_0(t)} = 0\,, \tag{22}$$

the increase in life expectancy has no effect on schooling.

In particular, if $\mu(t,u) = \mu$, $e_0(t) = \frac{1}{\mu}$, $\frac{\partial \mu(t,u)}{\partial e_0(t)} = -\frac{1}{e_0(t)^2}$ is constant with age and Eq. (22) holds true.

Linear mortality: $\mu(t,x) = Bx$ The linear case is not realistic, but is of interest here because it is the simplest departure from the case of mortality constant with age. Then

$$\begin{aligned} \sigma(t,a) &= e^{-\frac{B}{2}a^2} \\ e_0(t) &= \textstyle\int_0^{\bar{A}} \sigma(t,a)\,da = (\frac{\pi}{2B})^{\frac{1}{2}} \end{aligned} \tag{23}$$

which yields:

$$\begin{aligned} \tfrac{\partial \mu(t,u)}{\partial e_0(t)} &= -\tfrac{\pi}{e_0(t)^3} u \\ \tfrac{\partial \rho(t)}{\partial e_0(t)} &= \tfrac{\pi}{2e_0(t)^3} \tfrac{V_m + A_m^2}{A_m} \end{aligned} \tag{24}$$

where A_m and V_m are the period mean age at procreation and the variance of the age at procreation:

$$
\begin{array}{rcl}
A_m & = & \int_0^{\bar{A}} a e^{-\rho(t)(t)a} \sigma(t,a)\phi(t,a)\, da \\
V_m & = & -A_m^2 + \int_0^{\bar{A}} a^2 e^{-\rho(t)(t)a} \sigma(t,a)\phi(t,a)\, da
\end{array}
\qquad (25)
$$

Finally:

$$
\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)} = (1-\ell(t,A_S(t)))\frac{\pi}{2e_0(t)^3}\int_{A_S(t)}^{A_R}(b-A_S(t))(b+A_S(t)-A_m-\frac{V_m}{A_m})\sigma(t,b)e^{-\rho(t)(b-A_S(t))}\, db \qquad (26)
$$

Eq. (26) shows that for $A_S(t)$ high enough, $\frac{\partial(\frac{1}{\theta'(S(t))})}{\partial e_0(t)} > 0$, which would validate the Ben-Porath mechanism. The problem is that the decrease of $\frac{\partial \mu(t,u)}{\partial e_0(t)}$ with age described in Eq. (24) is contrary to real-world experience.

Gompertz mortality The Gompertz law of mortality is a close approximation of the observed force of mortality after 40 years of age (Pollard, 1973). It relies on two parameters A for the level and γ for the increase with age:

$$
\mu(t,a) = A\gamma^a \qquad \text{with } \gamma > 1 \text{ and } A < 1 \qquad (27)
$$

The survival function is then an exponential of an exponential: in no circumstances can a simple exponential portray an empirical human survival function.

Differentiating with respect to life expectancy $e_0(t)$ yields:

$$
\frac{\partial \mu(t,u)}{\partial e_0(t)} = (\frac{\partial \ln(A)}{\partial e_0(t)} + u\frac{\partial \ln(\gamma)}{\partial e_0(t)})\mu(t,u) \qquad (28)
$$

which is negative because $\frac{\partial \ln(A)}{\partial e_0(t)} < 0$ and $\frac{\partial \ln(\gamma)}{\partial e_0(t)} < 0$. After integration by parts:

$$
\int_0^b \frac{\partial \mu(t,u)}{\partial e_0(t)}\, du = \frac{A}{\ln\gamma}(\frac{\partial \ln A}{\partial e_0(t)}(\gamma^b - 1) + \frac{\partial \ln \gamma}{\partial e_0(t)}(b\gamma^b - \frac{1}{\ln\gamma}(\gamma^b - 1))) \qquad (29)
$$

Hence:

$$\int_{A_S(t)}^{b} \frac{\partial \mu(t,u)}{\partial e_0(t)}\,du = \frac{A}{\ln\gamma}\Big(\frac{\partial \ln A}{\partial e_0(t)}(\gamma^b - \gamma^{A_S(t)}) + \frac{\partial \ln\gamma}{\partial e_0(t)}(b\gamma^b - A_S(t)\gamma^{A_S(t)} - \frac{1}{\ln\gamma}(\gamma^b - \gamma^{A_S(t)}))\Big) \tag{30}$$

and

$$\frac{\partial \rho(t)}{\partial e_0(t)} = \frac{1}{A_m}\frac{A}{\ln\gamma}\int_0^{\bar{A}} \Big(\frac{\partial \ln A}{\partial e_0(t)}(1-\gamma^a) - \frac{\partial \ln\gamma}{\partial e_0(t)}(a\gamma^a - \frac{1}{\ln\gamma}(\gamma^a - 1))\Big)\sigma(t,a)\phi(t,a)e^{-\rho(t)a}\,da \tag{31}$$

with again $A_m = \int_0^{\bar{A}} ae^{-\rho(t)a}\sigma(t,a)\phi(t,a)\,da$. Eq. (30) and (31) introduced into Eq. (12) yield a computable formula. It remains however hard to interpret because the relationship of A and γ as functions of e_0 are not straightforward. This is better to work with model life tables, which are obtained from empirical data, and for which there is no need to compute which values of the parameters correspond to a given life expectancy. The difficulty arises because Gompertz, although good at describing mortality after 40 years of age, gives a poor fit before that age, especially at infant and juvenile ages.