Journal of Population Economics
https://doi.org/10.1007/s00148-019-00762-5

ORIGINAL PAPER

Endogenous education and the reversal in the relationship between fertility and economic growth

Alberto Bucci[1,2,3] · **Klaus Prettner**[4]

Received: 25 January 2019 / Accepted: 26 November 2019/[redacted]

Abstract
To reconcile the predictions of research and development (R&D)-based growth theory regarding the impact of population growth on productivity growth with the available empirical evidence, we propose a tractable, continuous-time, multisector, R&D-based growth model with endogenous education and endogenous fertility. As long as the human capital dilution effect is sufficiently weak, faster population growth may lead to faster aggregate human capital accumulation, to faster technological progress, and, thus, to a higher growth rate of productivity. By contrast, when the human capital dilution effect becomes sufficiently strong, faster population growth slows down aggregate human capital accumulation, dampens the rate of technical change, and, thus, reduces productivity growth. Therefore, the model can account for the possibly negative correlation between population growth and productivity growth in R&D-based models depending on the strength of the human capital dilution effect.

Keywords Human capital · Endogenous fertility · R&D-driven productivity growth · Non-monotonic growth effects of fertility

JEL classification I25 · J13 · O30 · O41

Responsible Editor: Alessandro Cigno

✉ Alberto Bucci
alberto.bucci@unimi.it

[1] Department of Economics, Management, and Quantitative Methods (DEMM), University of Milan, Milan, Italy

[2] CEIS (Centre for Economic and International Studies), University of Rome-'Tor Vergata', Rome, Italy

[3] RCEA (Rimini Centre for Economic Analysis), Wilfrid Laurier University, Waterloo, Canada

[4] University of Hohenheim, Stuttgart, Germany

1 Introduction

Since the seminal contribution of Romer (1990), models of endogenous technological change have enhanced our understanding of the forces driving long-run economic growth. Such models, indeed, have unveiled the crucial role for growth played by research and development (R&D) activities that, in turn, are determined by such factors as the individual incentives to do research, the size of the potential market, the future rewards that successful innovations accrue in terms of higher prospective monopolistic profits, and the availability of well-trained scientists. First-generation models of endogenous R&D-based growth (notably, Romer 1990; Grossman and Helpman 1991a; Aghion and Howitt 1992) suffered from a strong scale effect, i.e., the prediction that a larger population leads to faster economic growth and that a growing population implies hyper-exponential growth. Later contributions neutralized the strong scale effect by assuming either lower intertemporal knowledge spillovers in the production of new ideas (Jones 1995), increasing difficulty of R&D that comes with the accumulation of further knowledge (Kortum 1997; Segerstrom 1998), or a diluting effect of product proliferation on the resources that can be spent on quality-improving innovations (Peretto 1998; Howitt 1999). Nevertheless, even in these frameworks, a weak scale effect is still present, i.e., faster population growth unambiguously raises long-run economic growth.

In contrast to these theoretical implications, empirical analyses do not support a positive relation between population growth and long-run economic growth. Quite the contrary, most growth regressions maintain that the association between population growth and economic development is negative (see, for example, Brander and Dowrick 1994; Kelley and Schmidt 1995; Ahituv 2001; Li and Zhang 2007; Herzer et al. 2012).

Recent developments in economic growth theory have aimed at reconciling these empirical findings with the theoretical literature. This has been done either by modeling explicitly, in addition to the evolution of R&D, individuals' choice of also investing in human capital or by focusing on the so-called child quantity/quality tradeoff. According to the human capital channel (notable early contributions include Dalgaard and Kreiner 2001 and Strulik 2005), a higher birth rate slows down the accumulation of per capita human capital because it becomes more difficult to educate newborns in such a way that they attain the same education level as the rest of the population. This is the canonical human capital dilution-effect of population growth.

According to the child quantity/quality tradeoff channel, instead, if people choose to have more children, the overall resources that are available for education per child decrease such that there is a substitution of quantity for quality at the household level (cf. Becker 1960; Becker et al. 1990; Galor and Weil 2000; Galor 2005 and 2011). If taxes are used to finance public education, then this tradeoff is also present because more children imply a higher pupil-teacher ratio for a given size of the adult cohort and, thus, lower resources spent on education per child. This literature shows that under reasonable assumptions, the negative education effect of higher fertility overcompensates the positive quantity effect of higher fertility on the accumulation of *aggregate* human capital. This implies that faster population growth slows down aggregate human capital accumulation, decreases the flow of resources into R&D, and thereby reduces technological progress and economic growth (Strulik et al. 2013; Prettner 2014; Hashimoto and Tabata 2016; Baldanzi et al. 2019). This strand of the literature focuses on discrete-time overlapping generations models because it is very difficult to operationalize the quantity-quality substitution within a continuous-time setting. A notable exception is Chu et al. (2013)'s paper that uses a continuous-time setting with endogenous fertility and endogenous human capital accumulation to analyze the effects of patent protection in a model with endogenous quality-improving innovations. Their main finding is that strengthening patent protection increases technological progress, raises fertility, reduces human capital accumulation, and has an overall ambiguous effect on economic growth.

We aim to contribute to the above-mentioned literature along the following lines. First, we propose a continuous-time, multisector, R&D-based growth model with horizontal innovations that includes endogenous education and endogenous fertility decisions by agents. To our knowledge, we are the first to propose such a model in a continuous-time, analytically tractable setup. Unlike Bucci (2008), where the sign of the correlation between population growth and economic growth depends on the direction of technical change, and Bucci (2013, 2015), where an expansion in intermediate-inputs variety generates not only benefits (more specialization) but also costs (more complexity in assembling the larger set of available varieties of intermediates), in the present paper, population growth is endogenous and we are interested in modeling neither the possible effect of technical change on human capital investment nor the possible trade-off between potential benefits and potential costs related to intermediate-inputs proliferation. Instead, our goal here is to emphasize a totally new mechanism through which the relationship between population growth and economic growth might be non-uniform in sign either over time or across countries.

The new mechanism that we describe is based on the intensity of the negative human capital dilution effect, i.e., the adverse impact that a faster population growth rate bears on per capita human capital accumulation in conjunction with the quality/quantity substitution effect. To be more precise, unlike other papers, we do not postulate that the human capital dilution effect occurs only in a proportional manner. Instead, in our setting, all else equal, faster population growth may slow down per capita human capital investment either proportionally, less than proportionally, or more than proportionally. In other words, we share with the existing literature the view that, in the wake of an increase in the population, it becomes more difficult to accumulate human capital on a per capita basis. However, we extend this literature by showing that the strength with which such a dilution effect hits the accumulation of per capita human capital plays itself a fundamental role in determining the sign of the correlation between population growth and economic growth along a balanced growth path (BGP, hereafter).

Since our framework allows for a very general, non-monotonic relationship between population growth and economic growth, the model we put forward has the potential to reconcile the R&D-based growth literature with the existing empirical evidence. This is the second contribution of our paper. According to our results, as long as the human capital dilution effect is sufficiently weak, faster population growth may lead to faster aggregate human capital accumulation, faster technological progress, and, thus, a higher productivity growth rate in the long run. By contrast, if the human capital dilution effect becomes sufficiently strong, then faster population growth leads unambiguously to a negative impact on education. In this situation, faster population growth definitely slows down aggregate human capital accumulation, dampens technical change, and, thus, reduces the productivity growth rate in the long run. Remarkably, the threshold level of human capital dilution above which population growth slows down economic growth is below one, which implies that in our model a negative correlation between population growth and economic growth may arise even if faster population growth hits individual human capital investment less than proportionally.

Overall, this suggests that our setting can provide an additional and complementary explanation of the so-called ‘population-productivity reversal’ in R&D-based growth theory[1]: we account for the possibly negative correlation between population growth and economic growth by way of an amplification of the strength with which the per capita human capital dilution effect eventually operates over time and/or across countries.

At this stage, it is crucial to point out that we do not propose a Unified Growth model in which the reversal in the relation between fertility and economic growth (and, therefore, the transition across possibly different growth regimes) occurs endogenously. Doing so would, in fact, go beyond the scope of the present paper and would require a rather different modeling strategy based on the presence at some point of a corner-solution for educational investment and the resulting emergence of a demographic transition (as described in more detail by Strulik et al. 2013, for example).

Our paper is organized as follows. In “The model,” we describe the basic assumptions and the structure of the model. In “General equilibrium and BGP analysis,” we characterize its long-run BGP equilibrium, present our main results, and discuss them in detail. In “Concluding remarks and future research,” we draw our overall conclusions and present some ideas for future research.

2 The model

Consider a closed economy in which any individual purposefully invests in human capital and chooses how many children to have. The production side consists of three separate sectors. The research sector is characterized by free entry. Here, firms employ human capital and the existing number of non-rival ideas to engage in innovative activity that results in the invention of new blueprints for firms operating in the intermediate sector. The intermediate sector is composed of monopolistically competitive firms. There is a distinct firm producing each single variety of intermediates and holding a perpetual monopoly power over its sale. Finally, in the competitive final output sector, atomistic firms produce a homogeneous final good by employing human capital and all the available varieties of intermediates.

[1] This term was introduced first by Strulik et al. (2013) to describe that particular phenomenon according to which over the very long run, population growth is first positively and then negatively correlated with R&D-based productivity growth. According to them, in the course of time the increase of individual human capital overcompensates the associated decline of population growth in a way that leads to an ultimate rise of the aggregate human capital stock. Given that human capital is the driving force of R&D activity, this causes at some point higher R&D output and, hence, higher R&D-based economic growth. Unlike Strulik et al. (2013), in our paper the emphasis is on the intensity of the individual human capital dilution effect as a possible alternative source of the *population-productivity reversal* along a BGP.

2.1 Production

A representative firm produces the homogeneous final output through the following technology (see Ethier 1982; Romer 1987, 1990):

$$Y_t = n_t^{\overline{\alpha}} H_{Yt}^{1-Z} \int_0^{n_t} (x_{it})^Z di \quad \overline{\alpha} \geq 0, \quad 0 < Z < 1 \tag{1}$$

In (1), Y denotes production of the final good (the *numeraire* in the model), while x_i and H_Y are, respectively, the quantity of the ith intermediate input and the amount of human capital employed in this sector. The number of ideas existing at time t (n_t) coincides with the number of intermediate input varieties available in the economy at the same time. We assume that having a greater number of intermediate input varieties does not lead to any detrimental effect on total factor productivity ($\overline{\alpha} \geq 0$). As a whole, the production function (1) displays constant returns to scale with respect to the two private inputs (H_Y and x_i) and diminishing marginal returns to each of them, with $(1-Z)$ and Z corresponding to their respective shares in total GDP.[2] The fact that $Z \in (0;1)$ implies that human capital and intermediates are both necessary inputs in the production of final output. The inverse demand function for the ith intermediate reads as:

$$p_{it} = Z n_t^{\overline{\alpha}} H_{Yt}^{1-Z} (x_{it})^{Z-1} \tag{2}$$

The price of the generic ith intermediate input equals its marginal productivity in the production of final output. This is a consequence of the fact that the industry producing the final output is perfectly competitive.

In the intermediate sector firms engage in monopolistic competition. Each firm produces one (and only one) horizontally differentiated durable. Following Grossman and Helpman (1991b, Ch. 3), we assume that local intermediate monopolists have access to the same one-to-one technology:

$$x_{it} = h_{it}, \quad \forall i \in [0; n_t], \quad n_t \in [0; \infty), \tag{3}$$

where h_i is the amount of human capital required in the production of the ith durable, whose output is x_i. For given n, Eq. (3) implies that the total amount of human capital used in the intermediate sector at time t(*i. e.*, H_{It}) is:

$$\int_0^{n_t} (x_{it}) di = \int_0^{n_t} (h_{it}) di \equiv H_{It}. \tag{4}$$

Maximization of the generic ith intermediate firm's instantaneous profit and simultaneous use of Eq. (2) lead to the usual constant-markup-rule:

$$p_{it} = \frac{1}{Z} w_{It} = \frac{1}{Z} w_t = p_t, \quad \forall i \in [0; n_t], \quad n_t \in [0; \infty). \tag{5}$$

[2] Since final output is produced competitively under constant returns to scale with respect to its rival inputs, at equilibrium H_Y and x_i are rewarded according to their own marginal products. Hence, Z is the share of GDP accruing to intermediates, and $(1-Z)$ is the share of Y going to human capital.

The first part of Eq. (5) states that the price of intermediate good i is equal to a constant markup, $1/Z > 1$, over the marginal cost of production. The second part of Eq. (5) suggests that this price is the same across the different varieties of intermediates. To explain why, notice that in this economy the available human capital (H) is employed at any time to produce consumption goods (H_Y), intermediate inputs (H_I), and new ideas (H_n). Since it is assumed to be perfectly mobile across sectors, at equilibrium, human capital (the only input in the production of intermediates) will be rewarded according to the same wage rate, $w_t \equiv w_{Yt} = w_{nt} = w_{It}$.[3] Under symmetry (i.e., p and x equal across i), Eq. (4) leads to

$$x_{it} = H_{It}/n_t = x_t, \qquad \forall i \in [0; n_t] \tag{4.1}$$

$$\pi_{it} = n_t^{\overline{\alpha}-Z}\left[Z(1-Z)H_{Yt}^{1-Z}H_{It}^{Z}\right] = \pi_t \forall i \in [0; n_t]. \tag{6}$$

Thus, each intermediate firm will decide at any time t to produce the same quantity of output (x), to sell it at the same price (p), so earning the same instantaneous profit (π). The symmetry across durables is a direct consequence of the fact that each intermediate firm uses the same production technology (3), and faces the same demand function (see 2 and 5). Moreover, under symmetry, Eq. (1) can be recast as

$$Y_t = n_t^{R}\left(H_{Yt}^{1-Z}H_{It}^{Z}\right), \quad R \equiv 1 + \overline{\alpha} - Z > 0, \tag{1.1}$$

where R measures the degree of returns to specialization (Benassy 1998). In the present paper, it is immediate to verify that $R > 0$. This implies that in our setting the impact on total factor productivity of having a greater available number of intermediate input varieties is always positive (Ethier 1982, pp. 391–392).[4]

The aggregate production function (1) exhibits constant returns to H_Y and H_I together, but either increasing ($R > 1$), or decreasing ($0 < R < 1$), or else constant ($R = 1$) returns to an expansion of variety, while holding the quantity employed of each other input fixed. Hence, unlike other contributions (notably Devereux et al. 1996a, 1996b, 2000[5]), we allow for the possibility that the returns to specialization might be decreasing.

2.2 R&D activity

There is a large number of small competitive firms undertaking R&D activity. These firms produce ideas (n) taking the form of new varieties of intermediate

[3] The variables w_{Yt}, w_{It}, and w_{nt} denote the wage paid to a generic unit of human capital employed in the final output sector, the intermediate sector, and the research sector, respectively.

[4] So, unlike Bucci (2013), we rule out here the possibility that the returns to specialization might be negative. A negative R means that an increase in n would lead to some sort of inefficiency in the economy as, following a rise of the number of intermediate good varieties, total GDP would ceteris paribus decline in this case.

[5] In these papers, if all intermediates are hired in the same amount the returns to specialization are either increasing or at most constant.

inputs. A representative R&D firm uses only human capital to develop new ideas:

$$\dot{n} = \psi_t H_{nt}, n(0) > 0. \tag{7}$$

In Eq. (7), H_n is the number of researchers attempting to discover new ideas and ψ is the rate at which any single researcher can generate a new idea. Since the representative R&D-firm is small enough with respect to the whole sector, it takes ψ as given. So, Eq. (7) suggests that R&D activity is conducted under constant returns to scale to human capital (H_n). We postulate that the arrival rate ψ has the following specification:

$$\psi_t = \frac{1}{\chi} \frac{H_{nt}^{\mu-1}}{H_t^{\Phi}} n_t^{\eta}, \quad \chi > 0, \quad \mu > 0, \quad \Phi \lesseqgtr 0, \quad \eta < 1\,. \tag{7.1}$$

Using together (7) and (7.1), the production-function of new ideas reads as

$$\dot{n}_t = \frac{1}{\chi} \frac{H_{nt}^{\mu}}{H_t^{\Phi}} n_t^{\eta}, \mu \neq \Phi. \tag{8}$$

In the equations written above, χ is a strictly positive productivity parameter and H is the aggregate amount of human capital available in the economy. The rate at which a researcher can generate a new idea (ψ_t) is related to three different effects. The parameter η measures the traditional *intertemporal spillover effect* arising from the available stock of disembodied knowledge, n_t: $\eta < 0$ reflects the case where the rate at which a new innovation arrives declines with the number of ideas already discovered ("fishing-out effect"); if $0 < \eta < 1$, previous discoveries raise the productivity of current research effort ("standing-on-shoulders effect")[6]; $\eta = 0$ represents the situation in which the arrival rate of new ideas is independent of the existing stock of disembodied knowledge. The case $\eta = 1$ is ruled out from the analysis to avoid possible scale effects.

The parameter μ captures the effect on the arrival rate of a new innovation produced by the actual size of the R&D process (as measured by the number of units of skilled labor input devoted to it). A value $\mu = 0$ would imply that H_n is not an input to R&D-activity (Eq. 8). We ignore this unrealistic case by assuming that research human capital is indispensable to the discovery of new designs and that its contribution to the production of new ideas is always positive (i.e., $\mu > 0$). If $\mu = 1$, doubling the number of researchers H_n would not affect the arrival rate of a new idea in Eq. (7.1) and would therefore lead to exactly double the production of innovations per unit of time (Eq. 8); if $\mu \in (0; 1)$, due to the existence of congestion/duplication externalities ("stepping-on-toes effect"), increasing the number of researchers leads to a reduction in the rate at which each of them can discover a new idea (Eq. 7.1) and to a simultaneous increase (but less than proportional) in the total number of

[6] For a detailed discussion of the "*fishing-out*" and "*standing-on-shoulders*" effects, see Jones (1995, 2005).

innovations produced per unit of time (Eq. 8).[7] In accordance with Jones (2005), Eq. 16, p. 1074, we keep our analysis as general as possible by imposing no upper bound on μ.

According to Eq. (8), inventing the latest idea requires a skilled labor input equal to $H_n = (\chi H^{\Phi}/n^{\eta})^{1/\mu}$, which can change over time either because of the growth of n (*intertemporal knowledge-spillover effect*), or because of the growth of H, or for both reasons simultaneously. Suppose that an increase in population size raises the aggregate stock of human capital, H. If $\Phi > 0$, an expansion of the population will ultimately lead to a decrease of research human capital productivity (an increase in H_n). The hypothesis that the productivity of human capital employed in research may ultimately fall due to an increase in the population size can be justified by the fact that it becomes increasingly difficult to successfully introduce new varieties of (intermediate) goods in a more crowded market (*R&D difficulty* can rise also with the size of the population, as already suggested by Dinopoulos and Segerstrom 1999). In Eq. (8) the absolute value of Φ measures the strength of this effect: all the rest equal, a higher (and positive value of) Φ implies a more pronounced decline in the R&D-human capital productivity once an increase in the size of the population (and hence in H) has taken place.[8] The contrary happens when $\Phi < 0$: in this case, an increase in the population size that results in a rise of H, by reducing H_n, would ceteris paribus contribute to foster research human capital productivity (with the absolute value of Φ measuring again the strength of the effect). Such a possible outcome could be explained, for example, by the fact that a growing population is compatible with an increase in the ease of exchanging/diffusing ideas across people and/or creating research networks among scientists. In our model, all possibilities regarding the sign of the parameter Φ are left open. Notice that the Jones' (2005) formulation of the R&D process does not allow taking these important features of the inventive activity into account.[9]

The R&D sector is competitive and there is free entry. A representative R&D-firm has instantaneous profits equal to:

$$\text{R\&D-firm profits} = \underbrace{\left(\frac{1}{\chi}\frac{H_{nt}^{\mu}}{H_t^{\Phi}} n_t^{\eta}\right)}_{\dot{n}_t} V_{nt} - w_{nt} H_{nt}, \tag{9}$$

where

$$V_{nt} = \int_t^{\infty} \pi_{i\tau} e^{-\int_t^{\tau} r(s)ds} d\tau, \quad \tau > t. \tag{10}$$

In the last two equations, V_n denotes the market value of the generic ith intermediate firm (the one that has the exclusive right to produce the ith variety of capital goods); $\pi_{i\tau}$

[7] Likewise, if $\mu > 1$, increasing the number of researchers would imply an increase (more than proportional) in the total number of innovations produced per unit of time (Eq. 8).

[8] From $H_n = (\chi H^{\Phi}/n^{\eta})^{1/\mu}$, it is immediate to observe that $\partial H_n/\partial H = (1/\mu)(H_n/H)\Phi$. For given H and n, and for given parameters μ, χ, and η, if Φ is positive the larger the size of this parameter the bigger $\partial H_n/\partial H$.

[9] When $\Phi = 0$, Eq. (8) becomes: $\dot{n}_t = \frac{1}{\chi} H_{nt}^{\mu} n_t^{\eta}$, with $\chi > 0$, $\mu > 0$, and $\eta < 1$. This specification coincides with that employed by Jones (2005, Eq. 16). This means that $\mu > \Phi$ is the working assumption in this class of models. We will use this assumption below (see *Proposition 2* and *Remark 1*).

is the flow of instantaneous profits accruing to the same ith intermediate firm at date τ; $exp\left[-\int_t^{\tau} r(s)ds\right]$ is a present value factor that converts a unit of profit at time τ into an equivalent unit of profit at time t; r is the instantaneous interest rate (the real rate of return on households' asset holdings); and w_n is the wage rate going to one unit of research human capital. Equation (9) states that profits of a representative R&D-firm are equal to the difference between total R&D-revenues (R&D output, $\dot{n}$, times the price of an idea, V_n) minus total R&D costs related to *rival* inputs (human capital employed in research, H_n, times the wage accruing to one unit of this input, w_n). Equation (10), instead, reveals that the price of the generic ith idea is equal to the present discounted value of the returns resulting from the production of the ith variety of capital goods by profit-making intermediate firm i.

Combination of Eqs. (9) and (7.1) implies

$$w_{nt} = \frac{1}{\chi}\frac{H_{nt}^{\mu-1}}{H_t^{\Phi}} n_t^{\eta} V_{nt} = \psi_t V_{nt}. \tag{9.1}$$

This equation represents the standard research arbitrage condition stating that the compensation for the research activity (the left-hand side) should be equal to the value of the additional innovation that the marginal researcher is able to generate.

2.3 Households

The economy consists of many structurally identical households. Although the total number of households is constant through time (and normalized to unity), the size of the representative household may increase over time at the rate of population growth, $g_{L,t} \equiv \dot{L}_t / L_t$. Each member of the representative household can purposefully invest in human capital and has a Millian-type intertemporal utility function:[10]

$$U \equiv \int_0^{\infty} \left[log(c_t) + \upsilon \bullet log\left(g_{L,t}\right)\right] e^{-\rho t} dt, \tag{11}$$

where $\rho > 0$ is the subjective time discount rate, $c_t \equiv C_t / L_t$ is individual consumption, and $\upsilon > 0$ is the utility weight that an individual attaches to the number of children (see Prettner 2013). The flow budget constraint is

$$\dot{a}_t = r_t a_t + u_t h_t w_t - \left(1 + \Omega g_{L,t}\right) c_t, \quad a(0) > 0, \quad \Omega \geq 0, \tag{12}$$

[10] As it is well known, the difference between a *Millian-type* and a *Benthamite-type* intertemporal utility function is that in the former case the optimizing agent maximizes average utility of the dynasty, whereas in the latter case s/he maximizes aggregate utility of the dynasty. The *Benthamite* formulation would lead to serious complications in the presence of endogenous fertility decisions given that the endogenous fertility rate appeared in the exponent of the discount factor. However, the *Millian* formulation with a positive utility of the number of children captures the same effects as the *Benthamite* formulation such that the involved tradeoffs to the household are the same. Our choice of using a logarithmic specification for the instantaneous utility function serves the scope of making the problem even more tractable analytically.

where $a \equiv A/L$ denotes individual asset holdings, r is the real interest rate, u is the fraction of time spent on working, $h \equiv H/L$ is the individual human capital stock, w is the wage rate per unit of human capital, and Ω represents the resource-cost of rearing children, measured in terms of foregone consumption. Equation (12) states that per capita investment in asset holdings (the left-hand side) equals per capita saving (the right-hand side). Per capita saving, in turn, is equal to the difference between per capita total income (the sum of interest income, ra, and human capital income, uhw) and per capita consumption, $(1+\Omega g_L)c$, which includes also the (consumption–)cost of raising children. In what follows we shall assume $\Omega > 0$. However, as a special case of our model, in Appendix A, we analyze what happens when the child-rearing cost is zero ($\Omega = 0$).

From (13) it follows that aggregate asset holdings (A) evolve according to

$$\dot{A}_t = r_t A_t + (u_t H_t) w_t - \left(1 + \Omega g_{L,t}\right) C_t + A_t g_{L,t}, \qquad A(0) > 0, \tag{12.1}$$

where $uH \equiv H_E = H_Y + H_I + H_n$ is the share of the available aggregate human capital stock employed in production activities (namely, the production of consumption goods and intermediate inputs and the discovery of new ideas).

At each time $t \geq 0$, the fraction $(1-u) \in [0; 1]$ of the available stock of human capital is devoted to the accumulation of new human capital. Per capita human capital accumulates as

$$\dot{h}_t = [\sigma(1-u_t) - \xi g_{Lt}] h_t, \quad \sigma > 0, \quad \xi > 0, \quad h(0) > 0, \tag{13}$$

where σ and ξ are positive parameters. The first measures the productivity of the education sector, whereas the second is an indicator of the strength of the negative effect played by the growth rate of the population (the birth rate) on the growth rate of per capita human capital, i.e., the human capital dilution-effect.[11]

In Eq. (13) the returns to human capital per capita in the production of (new) human capital are assumed to be constant. Even though this assumption is empirically hard to defend (see Jones 2005, Section 6.2; Trostel 2004, among others), it is employed to make sure that, in the present context, human capital continues to be a growth engine in the very long run (namely, in a BGP equilibrium, to be defined in a moment).[12] With less than constant returns to human capital in Eq. (13), the results that we derive in this paper would still be present during the transition toward the BGP.

Using (13), the evolution of the aggregate stock of human capital, $H \equiv h \cdot L$, can be obtained as

[11] It is possible to show that introducing additional time costs of fertility in Eq. (13) leaves the key results of the model unchanged (see Appendix C for details).

[12] One can easily verify that, under the assumption of decreasing returns to human capital per capita in the production of (new) human capital, in the very long run human capital would no longer be a growth engine since economic growth would be wholly driven by population growth (i.e., fertility).

$$\dot{H}_t = \left[\sigma(1-u_t) + (1-\xi)g_{L,t}\right]H_t, \quad H(0) > 0. \tag{13.1}$$

Equation (13) states that, while the per capita human capital growth rate increases with the time spent on education, $(1-u)$, as long as $\xi > 0$, it also decreases with the population growth rate, g_L. To explain this point in more detail, first of all, observe what happens when $\xi = 1$: in this case, there exists a one-to-one dilution effect of population growth on per capita human capital investment.[13] The canonical explanation of this effect is that, since newborns enter the world uneducated, they reduce the existing stock of human capital per capita (h). Therefore, the speed at which this variable may accumulate over time decreases, ceteris paribus, in a proportional manner.[14] This effect, however, is not present in the original Lucas (1988, Eq. 14) formulation. Indeed, Lucas (1988) assumes that newborns enter the work-force endowed with a skill level proportional to the level already attained by older members of the family, so population growth per se does not reduce the current skill level of the representative worker. This assumption is based on the *social nature* of human capital accumulation, which, according to Lucas, has no counterpart in the accumulation of physical capital and of any other form of tangible assets. Note that, if $\xi = 0$, Eq. (13) would immediately be able to recover the view of Lucas (1988) as a very special result. However, since ξ represents the focus of our paper, we will purposefully ignore this case. A value of $\xi \in (0; 1)$ would represent, instead, an intermediate situation between the two described so far: in this case, a 1% increase in population growth would reduce, all the rest equal, the growth of per capita human capital by less than 1%. On the other hand, with $\xi > 1$, a 1% increase in population growth would reduce the growth of per capita human capital by more than 1%.

To our knowledge, in the growth literature, there exists no precise point estimate of ξ. However, research in the field of the economics of education, although not conclusive, can be used to gain at least some insights on the sign of ξ. Indeed, within this wide literature, the overwhelming majority of peer-reviewed publications supports the belief that there is a positive effect of a smaller class size on an individual student's achievement (in this regard, see, among others, the recent survey by Mathis 2016, p. 3). If one gives full credit to this finding, then the main implication of having a positive ξ in Eq. (13) is immediately clear: A lower population growth rate (a proxy for a reduced class size) is conducive to a faster rate of per capita human capital growth (a proxy for a better student's school performance).

Equation (13), however, also says that it becomes increasingly more difficult to accumulate new human capital on a per capita basis when ξ rises (faster population growth would hamper individual human capital accumulation less than proportionally

[13] The presence of g_L on the right hand side of Eq. (12) reflects also a sort of dilution-effect that, following agents' choice of having more children, may ultimately hit per capita asset investment in the form of an additional (consumption-)cost. More precisely, it reveals the cost (in terms of foregone consumption) of bringing the amount of per capita assets of the newcomers up to the average level of the existing population. Therefore, this formulation implies that, as long as $\Omega > 0$, population growth tends to reduce the speed of asset accumulation (by the average individual in the population).

[14] When $\xi = 1$, $\dot{h}/h = \dot{H}/H - g_L = \sigma(1-u) - g_L$: "*...population growth operates like depreciation of human capital* per capita" (Strulik 2005, Eq. 2, p. 135). Unlike Strulik (2005), we set the rate of physical obsolescence of human capital equal to zero. This is done just for the sake of simplicity.

if $0 < \xi < 1$; proportionally if $\xi = 1$; and more than proportionally if $\xi > 1$). Within the class-size debate, this appears to be consistent with the evidence that the effect of a class-size reduction (especially in the first years of schooling) is different across distinct groups of children, and is generally greater for those children coming from minorities, or other specific disadvantaged communities.[15]

In Hattie's (2005, p. 388) words the whole class-size debate can be condensed as follows: "…The major arguments in this review are that a synthesis of meta-analyses and other studies of class size demonstrate a typical effect-size of about 0.1-0.2, which…could be considered 'small'...".

In the light of the above, a priori we do not put any arbitrary upper bound on ξ and will present our results for the full attainable range of this parameter, i.e., $\xi > 0$ (even though the class-size debate summarized above supports a value of ξ which is both positive and tiny).

A representative agent in this economy chooses the optimal path of per capita consumption (c), the share of human capital to be devoted to production activities (u), and the birth rate, g_L (i.e., the number of children). As it is clear from the description of the household side, we treat the representative agent as infinitely lived to keep the model tractable and the mechanisms involved as clear as possible. Thus, we abstract from the analysis of population aging driven by changing mortality and we also abstract from the analysis of retirement decisions. For contributions that focus on these aspects and the insights that emerge from this literature, see Bloom et al. (2007), Prettner and Canning (2014), Cipriani (2014, 2018), Kuhn et al. (2015), and Cipriani and Pascucci (2019). Since the representative household in the economy is infinitely lived, there are also no interactions between changes in mortality and endogenous fertility choices in our framework. For papers that analyze the interactions between mortality and fertility decisions in detail see Cigno (1998), Kalemli-Ozcan (2003), Doepke (2005), Fanti and Gori (2014), and Cipriani and Fioroni (2019).

3 General equilibrium and BGP analysis

Since human capital is fully employed and perfectly mobile across sectors, the following market-clearing conditions must hold at equilibrium:

$$H_{E,t} \equiv u_t H_t = H_{Y,t} + H_{It} + H_{nt}, \quad \forall t \geq 0, \tag{14}$$

$$w_{I,t} = w_{nt}, \forall t \geq 0, \tag{15}$$

[15] See Zyngier (2014). According to Bressoux et al. (2009, p. 560): "…The effect of class size is shown to be significant and negative: a smaller class size improves student achievement. The impact is evaluated as being between 2.5% and 3% of a standard deviation of the scores… It is worth noting that the effect of class size seems more beneficial to low-achieving students within classes. The effect is particularly large for classes in priority education areas… This finding shows the complexity of the education production function and proves that it is essential to study how resources impact different students differently…".

$$W_{I,t} = W_{Y,t}, \forall t \geq 0. \tag{16}$$

Equation (14) states that aggregate labor demand (the right-hand side) should equal the fraction of the available human capital stock employed in production and R&D activities (the left-hand side). Equations (15) and (16) together state that, for the previous equality to be met, wages should adjust in such a way that the salary earned by one unit of skilled labor in the intermediate sector equals the salary earned by the same unit of skilled labor if employed either in research or in the production of final goods.

Moreover, since, in this economy, household's asset holdings are equal to the aggregate value of firms, the following equation should be also satisfied at equilibrium:

$$A_t = n_t V_{n,t}, \tag{17}$$

where V_{nt} is given by Eq. (10) and satisfies

$$\dot{V}_{nt} = r_t V_{nt} - \pi_t. \tag{17.1}$$

In the model, the ith idea allows the ith intermediate firm to produce the ith variety of durables. This explains why, in Eq. (17), total assets (A) equal the number of profit-making intermediate firms (n) times the market value (V_n) of each of them (equal, in turn, to the price of the corresponding idea). On the other hand, Eq. (17.1) states that the return on the value of the ith intermediate firm (rV_n) must be equal to the sum of the instantaneous monopoly profit accruing to the ith intermediate-input producer (π) and the capital gain or loss matured on V_n during the time interval dt, $\dot{V}_n$.

We are now able to move to a formal definition and characterization of the model's BGP equilibrium.

3.1 Definition: BGP equilibrium

A BGP equilibrium in this economy is a long-run equilibrium path along which

(i) All variables depending on time grow at constant exponential rates;
(ii) The sectoral shares of human capital employment ($s_j = H_j/H$, j = Y, I, n) are constant.

From this definition, Proposition Proposition 1 follows immediately:

Proposition 1 In the BGP equilibrium, the following results hold

$$\frac{g_L \equiv \frac{\dot{L}_t}{L_t} = -[\xi + \rho\Omega(1-\upsilon)(1+\sigma)] + \sqrt{[\xi + \rho\Omega(1-\upsilon)(1+\sigma)]^2 + 4\boldsymbol{\xi\Omega\rho\upsilon}(1+\sigma)}}{2\xi\Omega}, \tag{18}$$

$$1-u = \left(\frac{\sigma-\rho}{\sigma}\right), \tag{19}$$

$$\frac{\dot{H}_{Y,t}}{H_{Y,t}} = \frac{\dot{H}_{I,t}}{H_{I,t}} = \frac{\dot{H}_{n,t}}{H_{n,t}} = \frac{\dot{H}_t}{H_t} \equiv \gamma_H = \sigma(1-u) + (1-\xi)g_L = (\sigma-\rho) + (1-\xi)g_L, \quad (20)$$

$$\frac{\dot{h}_t}{h_t} \equiv \gamma_h = \sigma(1-u) - \xi g_L = (\sigma-\rho) - \xi g_L, \quad (21)$$

$$\frac{\dot{n}_t}{n_t} \equiv \gamma_n = \Upsilon[(\sigma-\rho) + (1-\xi)g_L] = \Upsilon\gamma_H, \quad (22)$$

$$\frac{\dot{w}_{nt}}{w_{nt}} = \frac{\dot{w}_{I,t}}{w_{I,t}} = \frac{\dot{w}_{Y,t}}{w_{Y,t}} \equiv \gamma_w = R\gamma_n = R\Upsilon[(\sigma-\rho) + (1-\xi)g_L] = R\Upsilon\gamma_H, \quad (23)$$

$$r = -\xi g_L + \sigma + \gamma_w = \sigma + (\sigma-\rho)R\Upsilon - [\xi(1+R\Upsilon) - R\Upsilon]g_L, \quad (24)$$

$$\gamma_y \equiv \frac{\dot{y}_t}{y_t} = \gamma_c \equiv \frac{\dot{c}_t}{c_t} = r - \rho = \gamma_a \equiv \frac{\dot{a}_t}{a_t} = \gamma_H + R\gamma_n - g_L$$
$$= (\sigma-\rho)(1+R\Upsilon) - [\xi(1+R\Upsilon) - R\Upsilon]g_L, \quad (25)$$

$$s_n = \frac{Z(1-Z)\gamma_n}{(1-Z+Z^2)[r + (1-R)\gamma_n - \gamma_H] + Z(1-Z)\gamma_n} \cdot u, \quad (26)$$

$$s_I = \left(\frac{Z^2}{1-Z+Z^2}\right)(u-s_n), \quad (27)$$

$$s_Y = \left(\frac{1-Z}{1-Z+Z^2}\right)(u-s_n), \quad (28)$$

$$\frac{H_t^{\mu-\Phi}}{n_t^{1-\eta}} = \frac{\chi}{s_n^{\mu}}\gamma_n \quad (29)$$

$$R \equiv 1 + \overline{\alpha} - Z > 0, \qquad \Upsilon \equiv \frac{\mu-\Phi}{1-\eta}$$

Proof: See Appendix A. ■

Equation (18) gives the growth rate of the population (the endogenous birth rate) in the BGP equilibrium. We see that g_L is certainly positive, increases with the parental preferences for children (υ), and decreases with the consumption cost of each child (Ω) –see Appendix B for a graphical analysis and a detailed economic interpretation of these (and other) effects. Equation (19) gives the allocation of the available stock of human capital between production (u) and non-production ($1-u$) activities along a BGP. Equation (20) shows the BGP equilibrium growth rate of the economy's human capital stock (H) and the growth rates of human capital employment in the final output (H_Y), intermediate inputs (H_I), and research (H_n) sectors. Equations (21) and (22) provide the BGP equilibrium growth rates of per capita human capital (h) and the economy's stock of disembodied knowledge (n). Equation (23) shows the growth rate of the (common) wage accruing to one unit of human capital employed in the three productive sectors of the economy in the BGP equilibrium; Eq. (24) provides the equilibrium real rate of return on asset holdings (r). According to Eq. (25), per capita income (y), per capita consumption (c), and per capita asset holdings (a) all grow at the same constant rate along a BGP. The BGP equilibrium shares of the existing human capital stock devoted to the production of ideas ($s_n \equiv H_n/H$), the production of intermediate inputs ($s_I \equiv H_I/H$), and the production of consumption goods ($s_Y \equiv H_Y/H$) are displayed in Eqs. (26), (27), and (28), respectively. Finally, Eq. (29) expresses the relation that holds in the long run among some function of the ratio of H_t to n_t, the growth rate of the number of ideas, γ_n, and the share of the available human capital stock devoted to R&D-activity, s_n. From this equation, it is evident that the restriction $\mu \neq \Phi$ (see Eq. 8) prevents, ceteris paribus, γ_n to be independent of H_t. It is also apparent from (29) that, with $\gamma_n > 0$ and $s_n > 0$, the ratio $H_t^{\mu-\Phi}/n_t^{1-\eta}$ is always positive.

In Appendix A, we show that (a) The shares of human capital allocated to production activities in the BGP equilibrium (s_n, s_Y, and s_I) are all between zero and one when $\gamma_n > 0$, $u \in (0;1)$ and the inequality $r > \gamma_H - (1-R)\gamma_n$ is fulfilled. This inequality allows V_{nt} to be positive at any time $t \geq 0$ in the BGP equilibrium; (b) The two transversality conditions, $\lim_{t\to+\infty} \lambda_{at} a_t = 0$ and $\lim_{t\to+\infty} \lambda_{ht} h_t = 0$, are simultaneously satisfied when $\sigma \cdot u > 0$.

Notice that, while some parameters of the model (namely, ξ, σ, and ρ) affect the BGP growth rate of the economy (γ_y) both directly and indirectly (i.e., through their impact on the birth rate, g_L), other parameters (i.e., Ω and υ) influence economic growth (γ_y) only indirectly through their sole effect on g_L. Finally, there is a third set of technological parameters ($\overline{\alpha}$, Z, μ, Φ, and η)—contributing to define R and Υ—that show a direct impact on γ_y, while at the same time having no effect on g_L.

Given the results stated in Proposition 1, reasonable parameter constellations exist such that realistic values for g_L, u, γ_H, γ_h, r, γ_y, γ_n, and γ_w can be attained. To provide such a numerical example, which we do not intend as a full-fledged calibration exercise because we derive our crucial results analytically, we use the following parameters. For the discount rate we choose $\rho = 0.038$, which lies between the values used by Jones (1995) and Chu et al. (2013). For the elasticity of final output with respect to physical capital (Z), we assume a value of 0.33 such that the labor share attains a value of 2/3 (cf. Jones 1995). For the parameter ξ that measures the negative impact of population growth on individual human capital investment, we choose a value of 0.75, such that class size reduces teaching efficiency but by less than one-for-one (which acknowledges the literature that is skeptical regarding a strong impact of class size on education outcomes). Finally, we assume $\Omega = 0.5$ such that the consumption expenditures of a child are half of those of an adult. Given these values, we adjust the remaining parameters that are much more difficult to estimate ($\overline{\alpha}$, μ, Φ, υ, σ, and η) such that we get realistic values for population growth, human capital growth, the share of time spent at work, per capita GDP growth, technological progress, wage growth, and the real rate of return on capital.[16]

[16] With these parameter values, we also have: $r - \gamma_H + (1-R)\gamma_n = 0.031 > 0$.

In Table 1, we report the results of this simulation. The first column of the table contains the variable names, and the second column, the values obtained for the USA as an average over the past decades based on DeLong and Magin (2009), Burkhauser et al. (2012), Psacharopoulos and Patrinos, (2018), Bureau of Labor Statistics (2019), and the data of The World Bank (2019a, 2019b). The detailed calculations and time periods used are described in the note to the table. The third column contains the outcomes of the model simulations for a baseline case. We observe that all endogenous variables are reasonably close to the obtained values for the USA. Finally, the forth column contains the values of the endogenous variables for an alternative simulation run in which we increase the parameter υ such that population growth increases slightly. For this simulation, we observe that the growth rate of per capita GDP decreases such that the model predicts an inverse relation between population growth and economic growth as found empirically for modern economies (Brander and Dowrick 1994; Kelley and Schmidt 1995; Ahituv 2001; Li and Zhang 2007; Herzer et al. 2012).

data of the World Bank (2019a)'s World Development Indicators over the time period 1970 to 2017. For the growth of human capital, we transform the increase in the average years of total schooling (age 15+) between 1970 and 2010 from the World Bank (2019b)'s Education Statistics into a measure of human capital according to a Mincerian specification with a rate of return to schooling of 9.08% (which is the average value for the USA as reported by Psacharopoulos and Patrinos 2018). To get the share of time at work rather than in education (u), we assume that a working life lasts for 45 years (from age 20 to age 65) and divide average years of total schooling by 45 years plus average years of total schooling. For the real rate of return on capital, we take plausible values reported by DeLong and Magin (2009). For the growth rate of total factor productivity, we take the value reported by the Bureau of Labor Statistics (2019) for the available time period 1987–2018. For wage growth we proxy by the household size-adjusted post-tax post-transfer income that includes health insurance as reported by Burkhauser et al. (2012) for the available time period 1979–2007. Since our model does not consider a changing composition of the labor force and different additional compensation components (such as health insurance), post-tax post-transfer income including health insurance is the closest data-based proxy for the expression of wages in the model (w) that we can get.

Proposition 2 analyzes the interaction between population growth and economic growth in this model economy's BGP equilibrium.

Proposition 2 Assume $\Omega > 0$ and $\Upsilon > 0$ (which implies $\mu > \Phi$). The sign of the relation between population growth and economic growth in the BGP equilibrium crucially depends on the magnitude of ξ.

- When $\xi \in (0; 1)$, then
- $\frac{\partial \gamma_y}{\partial g_L} > 0 \quad if \quad 0 < \xi < \bar{\xi} \equiv \frac{R\Upsilon}{1 + R\Upsilon} < 1;$
- $\frac{\partial \gamma_y}{\partial g_L} = 0 \quad if \quad 0 < \xi = \bar{\xi} \equiv \frac{R\Upsilon}{1 + R\Upsilon} < 1;$
- $\frac{\partial \gamma_y}{\partial g_L} < 0 \quad if \quad 0 < \bar{\xi} \equiv \frac{R\Upsilon}{1 + R\Upsilon} < \xi < 1;$
- When $\xi \geq 1$, then $\frac{\partial \gamma_y}{\partial g_L} < 0$.

Proof: The proof uses the fact that in the model there are two parameters (Ω and υ) that influence economic growth (γ_y) only indirectly, i.e., through their effect on g_L. Given this fact, suppose that a change in υ occurs such that g_L varies while, at the same time, R, Υ, σ, ρ, and ξ do not change (clearly, one can use a similar argument if the initial change regards Ω, instead of υ). Then,

$$\frac{\partial\gamma_y}{\partial g_L}=\frac{\frac{\partial\gamma_y}{\partial\upsilon}}{\frac{\partial g_L}{\partial\upsilon}}=\frac{-[\xi(1+R\Upsilon)-R\Upsilon]\frac{\partial g_L}{\partial\upsilon}}{\frac{\partial g_L}{\partial\upsilon}}=-[\xi(1+R\Upsilon)-R\Upsilon].$$

Therefore, $\frac{\partial\gamma_y}{\partial g_L}>0$ if $0<\xi<\overline{\xi}\equiv\frac{R\Upsilon}{1+R\Upsilon}<1$,

$$\frac{\partial\gamma_y}{\partial g_L}=0 \quad if \quad 0<\xi=\overline{\xi}\equiv\frac{R\Upsilon}{1+R\Upsilon}<1,$$

$$\frac{\partial\gamma_y}{\partial g_L}<0 \quad if \quad 0<\overline{\xi}\equiv\frac{R\Upsilon}{1+R\Upsilon}<\xi<1.$$

Table 1 Observed data and a comparison to the simulated outcomes for a baseline scenario and an alternative scenario with faster population growth

Variable	Data	Baseline simulation	Alternative simulation
g_L	0.0101	0.0104	0.0109
u	0.7735	0.7525	0.7525
r	0.0500–0.0700	0.0569	0.0567
γ_h	0.0055	0.0047	0.0043
γ_H	0.0156	0.0151	0.0152
γ_y	0.0176	0.0189	0.0186
γ_n	0.9000	0.9056	0.9133
γ_w	1.3100	1.4217	1.4339

For the first column, we compute average growth rates of per capita GDP and the population based on the data of the World Bank (2019a)'s World Development Indicators over the time period 1970 to 2017. Forthe growth of human capital, we transform the increase in the average years of total schooling (age 15+)between 1970 and 2010 from the World Bank (2019b)'s Education Statistics into a measure of humancapital according to a Mincerian specification with a rate of return to schooling of 9.08% (which is theaverage value for the United States as reported by Psacharopoulos and Patrinos 2018). To get the share oftime at work rather than in education (u), we assume that a working life lasts for 45 years (from age 20 toage 65) and divide average years of total schooling by 45 years plus average years of total schooling. Forthe real rate of return on capital, we take plausible values reported by DeLong and Magin (2009). For thegrowth rate of total factor productivity, we take the value reported by the Bureau of Labor Statistics (2019)for the available time period 1987-2018. For wage growth we proxy by the household size-adjusted posttaxpost-transfer income that includes health insurance as reported by Burkhauser et al. (2012) for theavailable time period 1979-2007. Since our model does not consider a changing composition of the laborforce and different additional compensation components (such as health insurance), post-tax post-transferincome including health insurance is the closest data-based proxy for the expression of wages in the model(w) that we can get.

Notice that $-[\xi(1+R\Upsilon)-R\Upsilon]<0$ for any $\xi\geq 1$. Thus, $\frac{\partial\gamma_y}{\partial g_L}<0$ if $\xi\geq 1$. ■.

Proposition 2 states that when $\xi\in(0;1)$—this is the case that the empirical class-size debate concisely abridged above would support—the relation between economic growth and population growth may be non-monotonic in the long run. An intuitive explanation for the result goes as follows. After combining Eqs. (20), (22), and (25), it is possible to conclude that

$$\gamma_y=\gamma_H+R\gamma_n-g_L=\gamma_H-g_L+R\Upsilon\gamma_H=(1+R\Upsilon)\gamma_H-g_L=\underbrace{\left(\frac{1+R\Upsilon}{\Upsilon}\right)}_{\substack{>0,\text{ according to}\\ \text{our assumptions}}}\gamma_n-g_L. \tag{30}$$

Looking at Eq. (30) it is immediate to see that an increase in the population growth rate (obtained, for example, through a change either in υ, or in Ω, or else in both) yields two different effects on the growth rate of per capita income:

- The first (i.e., the term $-g_L$ in Eq. 30) is direct. This is the canonical negative *dilution effect*: when newborns enter the world, they reduce, ceteris paribus, the existing stock of per capita human capital. So, in order to equip every single member of the rising population (including the newborns) with the same amount of human capital, extra-resources need to be explicitly used for this purpose, which finally hampers long-run economic growth;
- The second, instead, is indirect and describes the impact that the same increase in the population growth rate has, in turn, on the economy's growth rate of human capital (γ_H) and therefore on the growth rate of ideas (γ_n). As long as $\xi<1$, this effect is always positive (see Eqs. 20 and 22).

Thus,

- If ξ is large enough ($\xi\geq 1$), both the direct and indirect effects mentioned above are negative, and $\partial\gamma_y/\partial g_L$ is negative too;

- If ξ is small enough ($0<\xi<1$) these two effects are both operative but opposing in sign, with the first being negative and the second positive. As a consequence, the sign of $\partial\gamma_y/\partial g_L$ may well be ambiguous in this case. Our results (Proposition 2) show that, when $\xi\in(0;1)$, the sign of $\partial\gamma_y/\partial g_L$ crucially depends on whether ξ is below ($\partial\gamma_y/\partial g_L>0$), above ($\partial\gamma_y/\partial g_L<0$), or else equal ($\partial\gamma_y/\partial g_L=0$) to a threshold, $\overline{\xi}\equiv\frac{R\Upsilon}{1+R\Upsilon}=\frac{(\overline{\alpha}+1-Z)(\mu-\Phi)}{(\overline{\alpha}+1-Z)(\mu-\Phi)+(1-\eta)}\in(0;1)$, which is ultimately influenced by the underlying parameters $\overline{\alpha}$, Z, μ, Φ, and η.

The following remark analyzes in more detail the way in which $\overline{\xi}$ depends specifically on each of these five crucial parameters.[17]

Remark 1 Assume $\Upsilon>0$ (which implies $\mu>\Phi$). Then we have the following dependence of the threshold $\overline{\xi}$ on the underlying parameters of the model:

[17] We are particularly grateful to a referee for suggesting us to reflect more thoughtfully on this issue.

- $$\frac{\partial\overline{\xi}}{\partial\mu}=\frac{(1-\eta)\left(\overline{\alpha}+1-Z\right)}{\left[(1-\eta)+\left(\overline{\alpha}+1-Z\right)(\mu-\Phi)\right]^{2}}>0$$
- $$\frac{\partial\overline{\xi}}{\partial\Phi}=-\frac{(1-\eta)\left(\overline{\alpha}+1-Z\right)}{\left[(1-\eta)+\left(\overline{\alpha}+1-Z\right)(\mu-\Phi)\right]^{2}}<0$$
- $$\frac{\partial\overline{\xi}}{\partial\eta}=\frac{(\mu-\Phi)\left(\overline{\alpha}+1-Z\right)}{\left[(1-\eta)+\left(\overline{\alpha}+1-Z\right)(\mu-\Phi)\right]^{2}}>0$$
- $$\frac{\partial\overline{\xi}}{\partial\overline{\alpha}}=\frac{(\mu-\Phi)(1-\eta)}{\left[(1-\eta)+\left(\overline{\alpha}+1-Z\right)(\mu-\Phi)\right]^{2}}>0$$
- $$\frac{\partial\overline{\xi}}{\partial Z}=-\frac{(\mu-\Phi)(1-\eta)}{\left[(1-\eta)+\left(\overline{\alpha}+1-Z\right)(\mu-\Phi)\right]^{2}}<0$$

The following table shows in a compact way the relation between a change in the parameters $\overline{\alpha}$, Z, μ, Φ, and η and the expected sign of the effect of a variation in the endogenous population growth rate on the endogenous growth rate of per capita income. In the table, a formal distinction is made between countries being in a *Post-Malthusian Regime* (where $\partial\gamma_y/\partial g_L > 0$) and countries being in a *Modern Growth Regime* (where $\partial\gamma_y/\partial g_L < 0$).[18]

Proposition 2 and Table 2 show that growth of per capita GDP depends positively on population growth for sufficiently low levels of the negative effect exerted by population growth on schooling outcomes (that is for levels of ξ lower than a threshold $\overline{\xi}$), whereas the reverse holds true when ξ is sufficiently high. Intuitively, ξ is likely to be low enough in countries that are at an early stage of their economic development (e.g., in countries being in a *Post-Malthusian regime*) and where the amount of available embodied knowledge (i.e., human capital) is rather small. In particular, because these economies have a stock of human capital that is comparatively low, probably class size does not significantly affect the process of skill acquisition in these countries. This could be due, for example, to the fact that, in this group of economies, for acquiring (typically basic and general) skills, it does not matter that much if a teacher delivers her lecture in front of many or a few students. By contrast, if the stock of embodied knowledge available in the economy is on average already high and specialized (e.g., in rich countries being in a *Modern Growth Regime*), class size may have a pronounced negative impact on human capital accumulation: the background abilities already available to students are more qualified in such a case, and attaining further (specialized) skills would require a more direct interaction between professors and students, which is probably not necessary in the first group of countries). This overall narrative could explain the differential sign in the relation between per capita GDP growth and population growth in low—as opposed to high—income countries with the former exhibiting a comparatively smaller ξ than the latter. As for the evolution over time of the same relation, one could endogenize ξ in such a way that it increases with

[18] See Galor and Weil (2000), and Galor (2005, 2011) for detailed descriptions of the *Post-Malthusian* and *Modern Growth regimes*, and for an exhaustive analysis of the driving forces of economic growth, population growth, and human capital accumulation in each of the two different regimes.

the level of the available human capital stock. In this case, the relation between population growth and economic growth would certainly be positive at earlier (and negative at later) stages of economic development.

Table 2 is very much specific about the possible relation among the values of certain underlying parameters of the model (namely, $\overline{\alpha}$, Z, μ, Φ, and η), the magnitude of $\overline{\xi}$ (and, hence, of the relative ratio $\overline{\xi}/\xi$), the sign of $\partial\gamma_y/\partial g_L$, and the possible growth regime of a country. Unlike the shares of capital and (skilled) labor (Z and $1-Z$, respectively, that are regularly computed and internationally compared either across countries or over time), the other parameters ($\overline{\alpha}$, μ, Φ, and η) are, as already recognized above, much more difficult to measure and estimate both across countries (at the same or at a different stage of economic development) and/or over time. This notwithstanding, we can provide a simple qualitative interpretation of the results summarized in Table 2. To this end, consider those economies in which there is a negative relation between population growth and economic growth (these are the economies that the literature identifies as *modern economies* or as economies settled in a *Modern Growth regime*). As we are going to argue below, it is not difficult to think of these countries as those in which one might find (if properly measured or estimated) the smallest values of η, μ, and $\overline{\alpha}$ and the largest values of Z and Φ.

In the model, $\eta < 1$ measures the strength of the intertemporal spillover occurring in the production of new ideas and arising from the available stock of disembodied knowledge, n_t. So, in a sense, η is a measure of the intensity of the already invented ideas (n_t) in the production of the next ones. This externality may be either positive ($0 < \eta < 1$) or negative ($\eta < 0$) or else equal to zero ($\eta = 0$). It is not hard to believe that, in advanced economies, η might, indeed, be very low or even negative. In the latter case ($\eta < 0$), the typical *fishing-out effect* takes place: the rate at which a new innovation arrives declines with the number of ideas already discovered (i.e., inventing the latest idea becomes more difficult in those environments that are already rich of ideas).

In our setting $\mu > 0$ represents, instead, the intensity of research human capital in the production of new ideas. If $\mu \in (0; 1)$, then—as a result of the presence of some *congestion* or *duplication externality*—increasing the number of researchers would lead, ceteris paribus, to an increase (but only less than proportional) in the total number of innovations produced per unit of time. It is rather straightforward to think that this would be the most plausible case to observe in modern advanced economies where a comparatively larger number of researchers is already allocated to R&D.[19]

In contrast to η, the parameter $\overline{\alpha}$ reveals the extent of a different source of externality related to the existing number of ideas, n_t. The externality represented by $\overline{\alpha}$ exhibits its effect directly in the sector that produces goods, rather than in that producing innovations. In Ethier (1982, p. 392)'s words "…these economies reflect not an increased plant size but rather a greater division of labor; they are what Balassa (1967, ch. 5) refers to as 'horizontal specialization'…and were the subject of my earlier paper (1979), where they were called 'international' returns to

[19] *"The number of scientists and engineers engaged in R&D in the United States has grown dramatically over time, from under 500,000 in 1965 to nearly 1 million in 1989. Other advanced countries have experienced even larger increases in R&D employment…"* (Segerstrom 1998, p. 1290).

scale…". In brief, $\overline{\alpha}$ measures the force of a particular example of externality that is totally exterior to any individual firm producing final goods and that still arises from the number of ideas already circulating in an economy as a whole. Though, in our model, the role of $\overline{\alpha}$ cannot be understood disjointedly from the one played by $(1-Z)$. Both these two terms, in fact, define $R \equiv \overline{\alpha} + (1-Z)$ that represents "…the degree to which society benefits from 'specializing' production between a larger number of intermediates" (Benassy 1998, p. 63). The lower $\overline{\alpha}$ and $(1-Z)$, the lower is R. In economies where the degree of "horizontal specialization" n is already very high (i.e., in modern economies), it is reasonable to conjecture that ulteriorly specializing production among an even higher number of varieties would not generate social benefits greater than those generated, all the rest equal, in different settings where n is comparatively smaller. In other words, it makes sense to think that R, if appropriately measured or estimated, would be decidedly lower in modern (as opposed to Post-Malthusian) economies. As a consequence, one can reasonably maintain that $\overline{\alpha}$ and $(1-Z)$ might well be smaller in the first (as opposed to the second) group of countries.[20]

The presence of Φ in Eq. (8) is meant to capture the possible impact that an increase in the total stock of human capital has on the ability of a country to continue inventing new ideas. When $\Phi > 0$, a rise of the total stock of human capital, H, causes, ceteris paribus, researchers' productivity to decline. This result is consistent with the belief (Ha and Howitt 2007, Eq. 5, p. 740; Bucci and Raurich 2017, p. 190) that having more human capital, by resulting in more horizontal innovations, may ultimately lead to thinning any given aggregate research effort over a disproportionately large amount of small and separate R&D projects. This eventually results in a general economy-wide loss of productivity in doing research. This effect, which is associated with aggregate human capital, H, is (although linked) actually different from the congestion/duplication externality effect related, instead, to the total number of people devoted to R&D, H_n. In rich economies, where it is reasonable to presume that the aggregate amount of human capital is comparatively larger, the size of Φ is more likely to be greater in comparison with poorer countries.

All these aspects considered, it is reasonable to refer to those economies characterized by the smallest values of η, μ, $\overline{\alpha}$, and the (skilled-)labor share $1-Z$ and the largest value of Φ, as economies settled in a *Modern Growth regime*. For these countries, our model suggests that the threshold $\overline{\xi}$ would be sufficiently small. Combined with the fact that high-income countries are also those in which a comparatively greater value of ξ can most likely be found (see above), this discussion leads us to believe that $\overline{\xi} < \xi$ would be the case more probable to happen in these economies. In the end, this would

[20] While (as already said) there seemingly exists no available point estimate of the parameter $\overline{\alpha}$ (and, more generally, of the degree of returns to specialization, R), there are instead many studies that document very well the decline of the labor share of income, especially in the richest economies (see Fig. 2, p. 71, and Fig. 3, p. 73, in Karabarbounis and Neiman 2014; ILO-OECD 2015). In this regard, the OECD (2012) claims that over the period from 1990 to 2009 the share of labor compensation in national income declined in 26 out of 30 countries for which data were available, and calculated that the median (adjusted) labor share of national income across these economies fell from 66.1% to 61.7%. More recent OECD calculations find that the average adjusted labor share in G20 countries went down by about 0.3 percentage points per year between 1980 and the late 2000s. Similar downward trends of the labor share have also been documented by other international institutions (IMF 2007; European Commission 2007; BIS 2006; ILO 2012).

Table 2 The relation between the parameters $\overline{\alpha}$, Z, μ, Φ, and η and the expected sign of the impact of a change in the endogenous population growth rate on the endogenous growth rate of per capita income: *Post-Malthusian* vs. *Modern Regime countries* ($\mu > \Phi$ is assumed in the table)

$\uparrow\overline{\alpha}$ $\downarrow Z$ $\uparrow(1-Z)$	$\uparrow\overline{\xi}$	$\overline{\xi} > \xi$ (more likely)	$\frac{\partial\gamma_y}{\partial g_L} > 0$	Post-Malthusian Regime
$\uparrow\mu$ $\downarrow\Phi$ $\uparrow\eta$	$\uparrow\overline{\xi}$	$\overline{\xi} > \xi$ (more likely)	$\frac{\partial\gamma_y}{\partial g_L} > 0$	Post-Malthusian Regime
$\downarrow\overline{\alpha}$ $\uparrow Z$ $\downarrow(1-Z)$	$\downarrow\overline{\xi}$	$\overline{\xi} < \xi$ (more likely)	$\frac{\partial\gamma_y}{\partial g_L} < 0$	Modern Growth Regime
$\downarrow\mu$ $\uparrow\Phi$ $\downarrow\eta$	$\downarrow\overline{\xi}$	$\overline{\xi} < \xi$ (more likely)	$\frac{\partial\gamma_y}{\partial g_L} < 0$	Modern Growth Regime

explain why in theory one should observe a negative long-run correlation between population growth and per capita income growth in *Modern* (as opposed to *Post-Malthusian*) economies.

4 Concluding remarks and future research

Available evidence shows that (1) across countries, high R&D-driven productivity growth co-exists with high growth of human capital and low/negative growth of the population; (2) over time, and within the group of countries at the knowledge frontier, productivity growth increases when fertility decreases. These two observations are difficult to explain by means of the typical R&D-based growth theory.

In this paper, we have proposed a continuous-time, analytically tractable, R&D-driven growth model that includes both endogenous education and fertility decisions with the objective of emphasizing a new mechanism that is able to yield a negative association between population growth and productivity growth over the very long run (i.e., in a BGP equilibrium). This new mechanism is based on the intensity of the negative human capital dilution effect, that is, the adverse impact that faster population growth bears on per capita human capital accumulation due to the fact that when the population gets larger it becomes more difficult to keep on accumulating human capital on a per capita basis. According to our results, as long as the human capital dilution effect is sufficiently strong, a faster rate of population growth definitely slows down aggregate human capital accumulation, dampens the rate of technical change, and, thus, reduces the productivity growth rate. In this way, our model allows to reconcile the R&D-based growth literature with the existing empirical evidence according to which, when disembodied knowledge is advanced by educated scientists, high productivity growth can be sustained also when fertility is below its replacement level and the population is declining.

For future research, three promising routes are arguably worth taking. First of all, from a theoretical point of view, it would be interesting to complement our analysis with the study of the long-run adjustment dynamics of the model toward its balanced growth path. This would allow to examine the mutual evolution of population, technological progress, human capital investment, and per capita income along an economic and demographic transition leading to a long-run equilibrium characterized by low fertility and mortality rates, positive human capital formation, and persistent growth both in technology and in per capita income. Second, from an applied point of view, it would be fruitful to obtain reliable country-based point estimates for such parameters as $\overline{\alpha}$, μ, Φ, and η. Indeed, these parameters ultimately determine, along with the capital share Z, the threshold-value of ξ (i.e., $\overline{\xi}$). Given its own $\overline{\xi}$, it is then feasible for any country to design and implement pro-growth policies leading to a positive causation among per capita skill acquisition, demographic change (population growth), and real income per person. Finally, it might be worth augmenting the endogenous growth framework we proposed in this article with the inclusion of an endogenous demographic transition mechanism (along the lines illustrated by Strulik et al. 2013). This would allow to derive an R&D-based Unified Growth model in which the reversal in the relation between fertility and economic growth occurs endogenously across different economic-demographic regimes.

Acknowledgments We are indebted to three anonymous referees and to the editors of this journal (in particular Prof. Alessandro Cigno and Prof. Madeline Zavodny) for providing constructive comments and valuable suggestions.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

APPENDICES

APPENDIX A: DERIVATION OF EQS. (18)–(29), AND ANALYSIS OF THE SPECIAL CASE WHERE $\Omega = 0$

The *Hamiltonian function* (J_t) related to the inter-temporal optimization problem (11)–(12)–(13) in the text reads as:

$J_t = [\log(c_t) + \upsilon \log(g_{Lt})]e^{-\rho t} + \lambda_{at}[r_t a_t + u_t h_t w_t - (1 + \Omega g_{L,\,t})c_t] + \lambda_{ht}[\sigma(1 - u_t) - \xi g_{Lt}]h_t$, where λ_{at} and λ_{ht} are the co-state variables associated with the two state variables, a_t and h_t, respectively. The (necessary) first order conditions are:

$$\frac{\partial J_t}{\partial c_t} = 0 \Leftrightarrow \frac{e^{-\rho t}}{c_t} = \lambda_{at}\left(1 + \Omega g_{L,t}\right) \tag{A1}$$

$$\frac{\partial J_t}{\partial g_{Lt}} = 0 \quad \Leftrightarrow \quad \frac{\upsilon e^{-\rho t}}{g_{Lt}} - \Omega \lambda_{at} c_t - \xi \lambda_{ht} h_t = 0 \tag{A2}$$

$$\frac{\partial J_t}{\partial u_t} = 0 \quad \Leftrightarrow \quad \lambda_{at} = \sigma \frac{\lambda_{ht}}{w_t} \tag{A3}$$

$$\frac{\partial J_t}{\partial a_t} = -\dot{\lambda}_{at} \quad \Leftrightarrow \quad \lambda_{at} r_t = -\dot{\lambda}_{at} \tag{A4}$$

$$\frac{\partial J_t}{\partial h_t} = -\dot{\lambda}_{ht} \quad \Leftrightarrow \quad \lambda_{at} u_t w_t + \lambda_{ht}[\sigma(1-u_t) - \xi g_L] = -\dot{\lambda}_{ht}, \tag{A5}$$

along with the two transversality conditions:

$$\lim_{t \to +\infty} \lambda_{at} a_t = 0, \qquad \lim_{t \to +\infty} \lambda_{ht} h_t = 0,$$

and the initial conditions:

$$a(0) > 0, \qquad h(0) > 0.$$

From (A1) it follows that:

$$e^{-\rho t} = \lambda_{at}\left(1 + \Omega g_{L,t}\right) c_t. \tag{A6}$$

Plugging this expression into (A2) yields:

$$c_t \lambda_{at}\left[\frac{\upsilon}{g_{Lt}} - \Omega(1-\upsilon)\right] = \xi \lambda_{ht} h_t. \tag{A7}$$

Employing (A3) into (A7) delivers:

$$\left[\frac{\upsilon}{g_{Lt}} - \Omega(1-\upsilon)\right] = \frac{\xi}{\sigma} \frac{h_t w_t}{c_t}. \tag{A8}$$

Using (A3) in (A5), instead, gives:

$$\frac{\lambda^{\bullet}{}_{ht}}{\lambda_{ht}} = (\xi g_{Lt} - \sigma). \tag{A9}$$

Along a BGP all variables depending on time (including L_t) grow at constant, possibly positive, exponential rates. Therefore, in a BGP equilibrium (where $g_{Lt} = g_L$, $\forall\, t \geq 0$) Eq. (A9) suggests that $\lambda^{\bullet}{}_{ht}/\lambda_{ht}$ is also constant. Eqs. (A3) and (A4) imply, respectively:

$$\frac{\lambda^{\bullet}{}_{at}}{\lambda_{at}} = \frac{\lambda^{\bullet}{}_{ht}}{\lambda_{ht}} - \frac{w^{\bullet}{}_t}{w_t}, \tag{A10}$$

$$\frac{\lambda^{\bullet}{}_{at}}{\lambda_{at}} = -r_t \tag{A11}$$

Combination of (A9), (A10) and (A11) leads to:

$$r_t = \sigma + \frac{\dot{w}_t}{w_t} - \xi g_L. \tag{A12}$$

Since human capital is perfectly mobile across sectors, at equilibrium it will be rewarded according to the same wage, *i.e.*:$w_{Yt} = w_{It} = w_{nt} \equiv w_t$. Moreover, along the BGP this common wage will grow at a constant exponential rate, implying that $\frac{\dot{w}_{Yt}}{w_{Yt}} = \frac{\dot{w}_{It}}{w_{It}} = \frac{\dot{w}_{nt}}{w_{nt}} \equiv \frac{\dot{w}_t}{w_t}$ is constant. Accordingly, Eq. (A12) implies that in the BGP equilibrium the real rate of return on asset holdings, r , will be constant, as well.

From (A1) and (A4) together:

$$\frac{\dot{c}_t}{c_t} = r - \rho, \quad c_t \equiv \frac{C_t}{L_t}. \tag{A13}$$

By combining (A7), (A13), (A4), (A9) and (13) in the main text, after some simple algebra it is finally possible to obtain:

$$u_t = \frac{\rho}{\sigma} = u, \quad \forall t \geq 0, \quad \Leftrightarrow \quad 1 - u = \left(\frac{\sigma - \rho}{\sigma}\right). \tag{A14}$$

Eq. (A14) suggests that along a BGP the allocation of human capital between productive and non-productive activities is also constant.

Making use of Eqs. (6) and (10) in the main text, we find that along a BGP:

$$V_{nt} = Z(1-Z)\left(\frac{H_{Yt}}{n_t}\right)^{1-Z}\left(\frac{H_{It}}{n_t}\right)^{Z}\frac{n_t^R}{[r + (1-R)\gamma_n - \gamma_H]}, \quad R \equiv 1 + \overline{\alpha} - Z$$
$$> 0; \quad \frac{\dot{n}_t}{n_t} \equiv \gamma_n; \quad \frac{\dot{H}_t}{H_t} \equiv \gamma_H \tag{A15}$$

Notice that, for any $0 < Z < 1$, $H_{Yt} > 0$, $H_{It} > 0$, and $n_t > 0$, V_{nt} is always positive as long as:

$$r > \gamma_H - (1-R)\gamma_n. \tag{A15'}$$

Given V_{nt}, from Eq. (9') in the main text:

$$w_{nt} = \frac{Z}{\chi}(1-Z)s_n^{\mu-1}H_t^{\mu-1-\Phi}n_t^{\eta}\left(\frac{H_{Yt}}{n_t}\right)^{1-Z}\left(\frac{H_{It}}{n_t}\right)^{Z}\frac{n_t^R}{[r + (1-R)\gamma_n - \gamma_H]}, \tag{A16}$$

where $s_n \equiv H_{nt}/H_t$ is (by definition) constant in a BGP-equilibrium.

We can now use Eqs. (5), (2) and (4’) in the main text, obtaining:

$$w_{It} = Z^2 \left(\frac{H_{Yt}}{n_t}\right)^{1-Z} \left(\frac{H_{It}}{n_t}\right)^{Z-1} n_t^R. \tag{A17}$$

From Eq. (15) in the main text, by equalizing (A16) and (A17) in this appendix one gets:

$$s_I \equiv \frac{H_{It}}{H_t} = \frac{Z\chi}{(1-Z)} \frac{[r + (1-R)\gamma_n - \gamma_H]}{s_n^{\mu-1}} \frac{n_t^{1-\eta}}{H_t^{\mu-\Phi}}. \tag{A18}$$

Combining Eqs. (1) and (4’) in the text yields:

$$w_{Yt} \equiv \frac{\partial Y_t}{\partial H_{Yt}} = (1-Z)\left(\frac{H_{Yt}}{n_t}\right)^{-Z} \left(\frac{H_{It}}{n_t}\right)^{Z} n_t^R. \tag{A19}$$

From (16) in the main text and (A18) above, equalization of (A17) and (A19) in this appendix delivers:

$$s_Y \equiv \frac{H_{Yt}}{H_t} = \left(\frac{1-Z}{Z^2}\right) s_I = \frac{\chi}{Z} \frac{[r + (1-R)\gamma_n - \gamma_H]}{s_n^{\mu-1}} \frac{n_t^{1-\eta}}{H_t^{\mu-\Phi}}. \tag{A20}$$

Along a BGP all variables depending on time grow at constant (possibly positive) exponential rates and the sectoral shares of human capital employment are also constant. Therefore, from Eq. (8) in the main text it follows that:

$$\frac{\dot{n}_t}{n_t} \equiv \gamma_n = \left(\frac{\mu-\Phi}{1-\eta}\right)\gamma_H. \tag{A21}$$

If $\mu - \Phi = 1 - \eta$, we have a very special case of the model in which human capital and technology grow at the same rate $\gamma_n = \gamma_H \equiv \gamma$ in the long-run BGP equilibrium. Here we allow for the most general possible case in which $\mu \neq \Phi \neq \Phi + 1 - \eta$.

Using Eqs. (A16), (A17), (A19) and (A21), we observe that along a BGP wages grow at a common and constant rate:

$$\frac{\dot{w}_{nt}}{w_{nt}} = \frac{\dot{w}_{It}}{w_{It}} = \frac{\dot{w}_{Yt}}{w_{Yt}} \equiv \frac{\dot{w}_t}{w_t} = R\gamma_n. \tag{A22}$$

From (17) in the text and (A15) in this appendix we conclude that along a BGP:

$$\frac{\dot{a}_t}{a_t} \equiv \gamma_a = \gamma_H + R\gamma_n - g_L, \quad a_t \equiv A_t/L_t, \quad g_L \equiv \dot{L}_t/L_t. \tag{A23}$$

Merging (12) in the main text and (A11) in this appendix yields:

$$\frac{\dot{\lambda}_{at}}{\lambda_{at}} = -\gamma_a + u\frac{h_t w_t}{a_t} - (1 + \Omega g_L)\frac{c_t}{a_t}. \tag{A24}$$

Similarly, from the combination of (13) in the body-text and (A9) in this appendix we get:

$$\frac{\dot{\lambda}_{ht}}{\lambda_{ht}} = -\gamma_h - \sigma u, \quad \gamma_h \equiv \dot{h}_t/h_t, \quad h_t \equiv H_t/L_t. \tag{A25}$$

Note that (A25) and (A9), taken together, confirm that in this economy: $\dot{h}_t = [\sigma(1-u) - \xi g_L]h_t$, see Eq. (13) in the text.

Combination of Eqs. (A9), (A10), (A22), (A23), (A24) and Eq. (13) in the body-text, finally leads to:

$$\frac{c_t}{a_t} = \frac{u}{(1 + \Omega g_L)}\left[\sigma + \frac{h_t w_t}{a_t}\right], \tag{A26}$$

where $\gamma_h = \gamma_H - g_L$ has been used.

By employing, again, Eq. (13) in the main text, Eq. (A22), Eq. (A23) and the fact that $\gamma_h = \gamma_H - g_L$, from (A26) one immediately concludes that:

$$\frac{c_t}{a_t} = \frac{u}{(1+\Omega g_L)}\left[\sigma + \underbrace{\frac{h(0)w(0)}{a(0)}}_{\equiv \Gamma(0)>0}\right], \tag{A27}$$

where $h(0) > 0$, $a(0) > 0$ and $w(0) > 0$ are the given initial values (*i.e.*, at $t = 0$) of h_t, a_t and w_t, respectively.

Eqs. (A27) implies that along a BGP (where u and g_L are constant):

$$\gamma_c = \gamma_a. \tag{A28}$$

Eq. (A26) can be recast as:

$$h_t w_t = -\sigma a_t + \left(\frac{1 + \Omega g_L}{u}\right)c_t. \tag{A29}$$

Similarly, Eq. (A8) can be rewritten as:

$$h_t w_t = \frac{\sigma}{\xi}\left[\frac{\upsilon}{g_L} - \Omega(1-\upsilon)\right]c_t. \tag{A8'}$$

After equating (A29) and (A8'), we are able to compute the endogenous value of c_t/a_t:

$$\frac{c_t}{a_t}=\frac{\sigma}{\left(\frac{1+\Omega g_L}{u}\right)-\frac{\sigma}{\xi}\left[\frac{\upsilon}{g_L}-\Omega(1-\upsilon)\right]}. \tag{A30}$$

Eq. (A30) confirms that, along a BGP, c_t and a_t grow at a common rate, see (A27) and (A28) above. This happens because the ratio of these two variables is constant in the BGP equilibrium. Finally, we equate (A30) and (A27) and get:

$$\Gamma(0)\xi\Omega g_L^2+\{\Gamma(0)[\xi+\rho\Omega(1-\upsilon)]+\rho\sigma\Omega(1-\upsilon)\}g_L-\rho\upsilon[\sigma+\Gamma(0)]=0. \tag{A31}$$

Eq. (A31) gives the endogenous value of g_L along a BGP-equilibrium. If we simplify the analysis by normalizing all the relevant initial conditions to one, in such a way that $h(0)=w(0)=a(0)=1>0$, and therefore $\Gamma(0)\equiv\frac{h(0)w(0)}{a(0)}=1$, then Eq. (A31) easily becomes:

$$\xi\Omega g_L^2+[\xi+\rho\Omega(1-\upsilon)(1+\sigma)]g_L-\rho\upsilon(1+\sigma)=0. \tag{A31'}$$

The solution to this equation is:

$$g_L=\frac{-[\xi+\rho\Omega(1-\upsilon)(1+\sigma)]\pm\sqrt{[\xi+\rho\Omega(1-\upsilon)(1+\sigma)]^2+4\xi\Omega\rho\upsilon(1+\sigma)}}{2\xi\Omega} \tag{A31''}$$

One root is positive and the other one is negative. The positive root is:

$$g_L=\frac{-[\xi+\rho\Omega(1-\upsilon)(1+\sigma)]+\sqrt{[\xi+\rho\Omega(1-\upsilon)(1+\sigma)]^2+4\xi\Omega\rho\upsilon(1+\sigma)}}{2\xi\Omega}.$$

Using Eqs. (1) and (4') in the body-text and the definitions of $y_t\equiv\frac{Y_t}{L_t}$, $R\equiv1+\overline{\alpha}-Z>0$, and $\dot{L}_t/L_t\equiv g_L$, we obtain the growth rate of real per capita output along a BGP:

$$\gamma_y\equiv\frac{\dot{y}_t}{y_t}=\frac{\dot{Y}_t}{Y_t}-g_L=\gamma_H+R\gamma_n-g_L=\gamma_a=\gamma_c, \tag{A32}$$

see Eqs. (A23) and (A28).

Eq. (A32) is important because it says that along a BGP, per capita income (y), per capita asset holdings (a) and per capita consumption (c) all grow at the same rate.

We are now able to compute the BGP-equilibrium values of s_n, s_I and s_Y. Eq. (14) in the main text suggests that:

$$u=s_Y+s_I+s_n. \tag{A33}$$

Using the fact that (see A20) $s_Y = \left(\frac{1-Z}{Z^2}\right)s_I$ into the expression above we obtain:

$$s_I = \left(\frac{Z^2}{1-Z+Z^2}\right)(u-s_n). \tag{A34}$$

Hence:

$$s_Y = \left(\frac{1-Z}{1-Z+Z^2}\right)(u-s_n). \tag{A35}$$

According to (A20), however, it is also true that:

$$s_Y \equiv \frac{H_{Yt}}{H_t} = \frac{\chi}{Z}\frac{[r+(1-R)\gamma_n-\gamma_H]}{s_n^{\mu-1}}\frac{n_t^{1-\eta}}{H_t^{\mu-\Phi}}.$$

Equating the last expression to (A35) yields:

$$\frac{H_t^{\mu-\Phi}}{n_t^{1-\eta}} = \frac{\chi\left(1-Z+Z^2\right)}{Z(1-Z)}\frac{[r+(1-R)\gamma_n-\gamma_H]}{s_n^{\mu-1}(u-s_n)}. \tag{A36}$$

From Eq. (8) in the body-text we get:

$$\frac{H_t^{\mu-\Phi}}{n_t^{1-\eta}} = \frac{\chi}{s_n^{\mu}}\gamma_n. \tag{A37}$$

Equalization of (A36) and (A37) finally leads to:

$$s_n = \frac{Z(1-Z)\gamma_n}{\left(1-Z+Z^2\right)[r+(1-R)\gamma_n-\gamma_H]+Z(1-Z)\gamma_n}\cdot u. \tag{A38}$$

Given s_n above, it is now possible to compute the BGP ratio $\frac{H_t^{\mu-\Phi}}{n^{1-\eta}}$ (by using either Eq. A37 or Eq. A36), along with s_I and s_Y (see Eqs. A34 and A35).

Note that, when $\gamma_n > 0$, $Z \in (0; 1)$, and Eq. (A15') is satisfied, then the following inequality does hold:

$$0 < \Xi \equiv \frac{Z(1-Z)\gamma_n}{\left(1-Z+Z^2\right)[r+(1-R)\gamma_n-\gamma_H]+Z(1-Z)\gamma_n} < 1.$$

With $u \in (0; 1)$, this implies that:

$$0 < s_n \equiv \Xi u < 1, \qquad \text{and} \qquad 0 < (u-s_N) = u - \Xi u = u(1-\Xi) < 1.$$

The last result allows S_Y and S_I to be also strictly between zero and one along a BGP. In Eq. (A37) it is evident that, with $\chi > 0$, $\gamma_n > 0$ and $s_n \in (0; 1)$, the ratio $H_t^{\mu-\Phi}/n_t^{1-\eta}$ is always positive.

Finally, using Eqs. (A9), (A10), (A22), (A23), (13') in the body-text, and the definition of $h_t \equiv H_t/L_t$, it can be easily shown that along a BGP the two transversality conditions:

$$\lim_{t\to+\infty} \lambda_{at} a_t = 0 \text{ and } \lim_{t\to+\infty} \lambda_{ht} h_t = 0,$$

are simultaneously checked when:

$$\sigma \cdot u > 0,$$

for any $\lambda_a(0) > 0$, $\lambda_h(0) > 0$, $a(0) > 0$, and $h(0) > 0$.

The condition $\sigma \cdot u > 0$ always holds in our model, as $\sigma \cdot u = \rho > 0$ (see A14).

In the model, as a particular case, Ω can also be equal to zero (see Eq. 12). When $\Omega = 0$ the law of motion of per capita asset holdings becomes:

$$\dot{a}_t = (r_t a_t + u_t h_t w_t) - c_t, \qquad a(0) > 0.$$

Economically this means that, following an agent's choice of having more children, no dilution effect of population growth (in the form of an additional consumption cost) would hit per capita asset investment. In other words, as to the speed at which investment in asset-holdings by an average individual in the population occurs, a change in the population size would play no role in this specific case.

When $\Omega = 0$, from Eq. (18) in the main text, one can easily observe that:

$$\lim_{\Omega\to 0} g_L = \frac{0}{0}.$$

Since all the major endogenous variables of the model depend on g_L (see Proposition 1 in the text), such variables would inevitably take an indeterminate value when g_L does so. As a consequence, studying the behavior of our model when $\Omega = 0$ appears to be extremely relevant from an economic as well as an algebraic point of view. In this regard, we apply *de l'Hôpital's rule* to Eq. (18) in the text. More formally, call by $B(\Omega)$ and $E(\Omega)$ the numerator and the denominator of Eq. (18) in the body-text, respectively. Then, since $\lim_{\Omega\to 0} g_L \equiv \frac{B(\Omega)}{E(\Omega)} = \frac{0}{0}$, by the *l'Hôpital's rule*:[21]

$$\lim_{\Omega\to 0} g_L = \lim_{\Omega\to 0} \frac{B'(\Omega)}{E'(\Omega)} = \frac{\rho\upsilon(1+\sigma)}{\xi}.$$

[21] We are indebted to a referee for suggesting us this condensed proof.

This means that an equilibrium growth rate of the population (g_L) does exist, is finite, and is strictly positive when $\Omega = 0$.

The following Lemma introduces explicit constraints on the (relation among the feasible) values of the model's parameters such that the resultant endogenous variables are economically meaningful when $\Omega = 0$.

LEMMA

Assume $\Omega = 0$. Then,

- γ_H, γ_h, γ_n, γ_w, $\gamma_y = \gamma_a = \gamma_c$, and r are all positive;
- $0 < (1-u) < 1$;
- $r > \gamma_H - (1-R)\gamma_n$, (see A15')

if the following assumptions on the parameters' values hold true:

(i) $0 < \xi < 1$;
(ii) $\Upsilon > \frac{1}{1-\xi}$;
(iii) $(\sigma - \rho) > 0$;
(iv) $0 < \upsilon < \frac{(\sigma-\rho)}{\rho(1+\sigma)} < \frac{\sigma+(\sigma-\rho)R\Upsilon}{\rho(1+\sigma)(1+R\Upsilon)}$.

Proof: Immediate from Eqs. (19)–(25) in the main text when $\Omega = 0$ and $g_L = \frac{\rho\upsilon(1+\sigma)}{\xi}$. ■

From Proposition 2 in the text, $\xi \in (0; 1)$ seems the most interesting situation to be analyzed. This is the reason why we still focus on this case here (assumption *i* above). Also notice that supposing $\Upsilon > \frac{1}{1-\xi} > 0$ (assumption *ii*) is equivalent to postulating $\mu - \Phi > \frac{1-\eta}{1-\xi} > 0$, which again implies $\mu > \Phi$ (exactly as in Proposition 2 in the text).

The next Proposition analyzes the interaction between population growth and economic growth in the BGP equilibrium of this simpler model ($\Omega = 0$).

PROPOSITION

Assume $\Omega = 0$ and $\Upsilon > \frac{1}{1-\xi} > 0$ (which still implies $\mu > \Phi$). The sign of the relation between population growth and economic growth in the BGP equilibrium crucially continues to depend on the magnitude of ξ. Again, we observe that:

- $$\frac{\partial\gamma_y}{\partial g_L} > 0 \quad \text{if} \quad 0 < \xi < \bar{\xi} \equiv \frac{R\Upsilon}{1+R\Upsilon} < 1;$$
- $$\frac{\partial\gamma_y}{\partial g_L} = 0 \quad \text{if} \quad 0 < \xi = \bar{\xi} \equiv \frac{R\Upsilon}{1+R\Upsilon} < 1;$$
- $$\frac{\partial\gamma_y}{\partial g_L} < 0 \quad \text{if} \quad 0 < \bar{\xi} \equiv \frac{R\Upsilon}{1+R\Upsilon} < \xi < 1.$$

Proof: Again, suppose that there is a change in υ such that g_L ultimately varies while at the same time R, Υ, σ, ρ, and ξ do not. Then:

$$\frac{\partial \gamma_y}{\partial g_L} = \frac{\frac{\partial \gamma_y}{\partial \upsilon}}{\frac{\partial g_L}{\partial \upsilon}} = \frac{-[\xi(1+R\Upsilon)-R\Upsilon]\frac{\partial g_L}{\partial \upsilon}}{\frac{\partial g_L}{\partial \upsilon}} = -[\xi(1+R\Upsilon)-R\Upsilon].$$

$$\text{Therefore :} \quad \begin{array}{l} \frac{\partial \gamma_y}{\partial g_L} > 0 \quad \text{if} \quad 0 < \xi < \overline{\xi} \equiv \frac{R\Upsilon}{1+R\Upsilon} < 1, \\ \frac{\partial \gamma_y}{\partial g_L} = 0 \quad \text{if} \quad 0 < \xi = \overline{\xi} \equiv \frac{R\Upsilon}{1+R\Upsilon} < 1, \\ \frac{\partial \gamma_y}{\partial g_L} < 0 \quad \text{if} \quad 0 < \overline{\xi} \equiv \frac{R\Upsilon}{1+R\Upsilon} < \xi < 1. \end{array}$$

This proposition (obtained when $\Omega = 0$) confirms the results of the most general case of the model ($\Omega > 0$, Proposition 2 in the text): when $\xi > 0$ population growth operates like a form of depreciation in the per capita human capital investment function. In addition, if ξ is sufficiently small ($0 < \xi < 1$ – see Proposition 2 in the text and the proposition above), a threshold level of this parameter, $\overline{\xi} \equiv [R\Upsilon/(1+R\Upsilon)] \in (0;1)$, does exist such that if ξ is below (respectively, above or equal to) this threshold, then economic growth depends positively (respectively, negatively, or does not depend at all) on population growth.

APPENDIX B: g_L AS A FUNCTION OF Ω, υ, ρ, σ, AND ξ

The graph of:

$$g_L = \frac{-[\xi + \rho\Omega(1-\upsilon)(1+\sigma)] + \sqrt{[\xi + \rho\Omega(1-\upsilon)(1+\sigma)]^2 + 4\xi\Omega\rho\upsilon(1+\sigma)}}{2\xi\Omega},$$

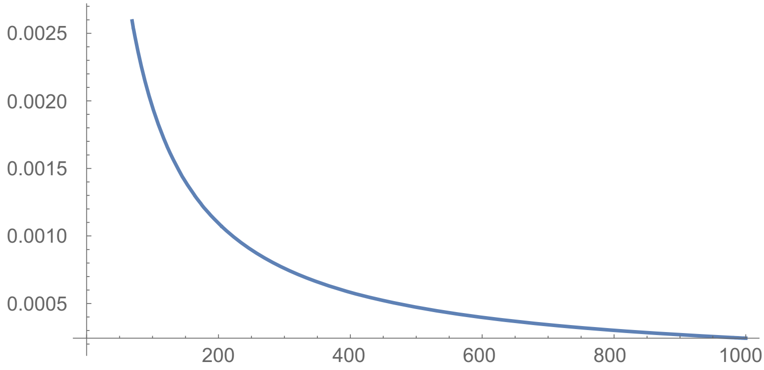
as a function of Ω, υ, ρ, σ, and ξ, is reported below. In what follows, we use:

$$\begin{array}{l} \rho = 0.038 \\ \sigma = 0.0505 \\ \xi = 0.75 \\ \Omega = 0.5 \\ \upsilon = 0.2 \end{array}$$

Under this parameter-constellation, realistic values for the main endogenous variables of the model can be obtained (see Table 1 in the text).

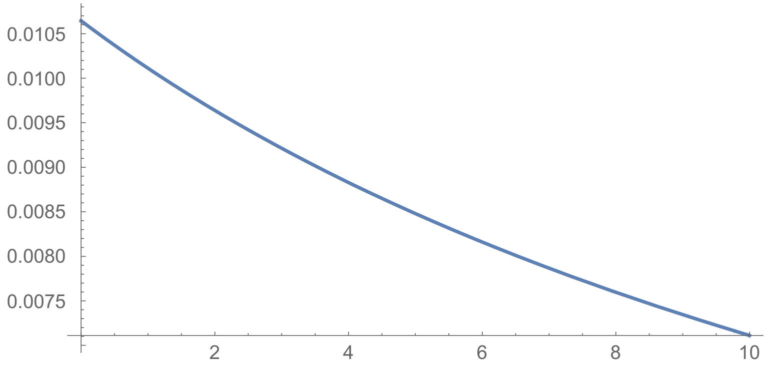
It is apparent from the graphs above that in a BGP-equilibrium the endogenous birth rate, g_L, increases with the preference of parents for children (υ), but decreases with the consumption-cost of each child (Ω), as stated in the main text.

Fig. 1 *The graph of* g_L *(vertical axis) as a function of* Ω *(horizontal axis),* $\Omega \in [0.000000000001; 1000]$



The intuition for these results goes as follows. If Ω increases, this raises the costs of each child in terms of foregone consumption such that individuals would choose to reduce fertility (Figures 1 and 2); if υ increases, parents derive more utility per child and therefore fertility rises (Figures 3 and 4); if ρ increases, parents are more impatient, educate themselves less and choose to have more children because this raises utility (Figure 5); if σ is higher, education for the children is easier and this mitigates the human capital dilution effect of population growth such that parents choose to have more children (Figure 6); finally, if ξ increases, the dilution effect of population growth on human capital accumulation becomes stronger such that parents substitute quality for quantity in their decision of having children and therefore reduce fertility (Figure 7).

Fig. 2 *The graph of* g_L *(vertical axis) as a function of* Ω *(horizontal axis),* $\Omega \in [0.000000000001; 10]$



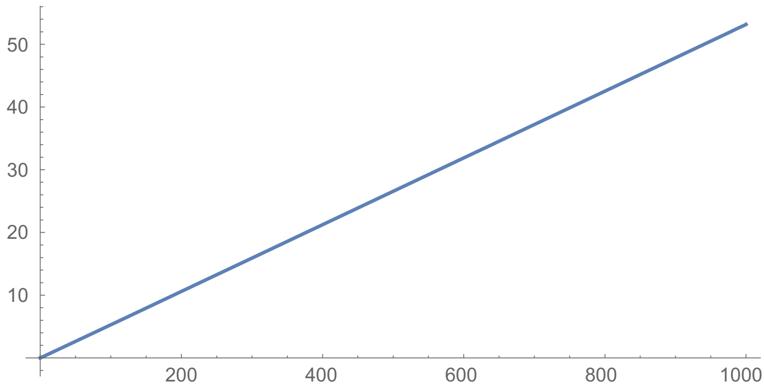


Fig. 3 *The graph of g_L (vertical axis) as a function of υ (horizontal axis),* $\upsilon \in [0.000000000001; 1000]$

APPENDIX C: INTRODUCING CHILD–COSTS IN TERMS OF PARENTAL TIME

The objective of this appendix is to see what happens when one introduces additional time-costs of fertility. At this aim, assume that the original law of motion of per capita human capital (Eq. 13 in the text) becomes:

$$\dot{h}_t = [\sigma(1-\varphi g_{Lt}-u_t)-\xi g_{LT}]h_t, \qquad \sigma > 0, \qquad \xi > 0, \qquad \varphi > 0 \qquad (13a)$$

where the term φg_{Lt} identifies another source of costs (in terms of time subtracted to the activity of skill acquisition, $1-u_t$) related to the choice of having more children. As in the original version of the paper, here we continue to assume that:

Fig. 4 *The graph of g_L (vertical axis) as a function of υ (horizontal axis),* $\upsilon \in [0.000000000001; 1]$

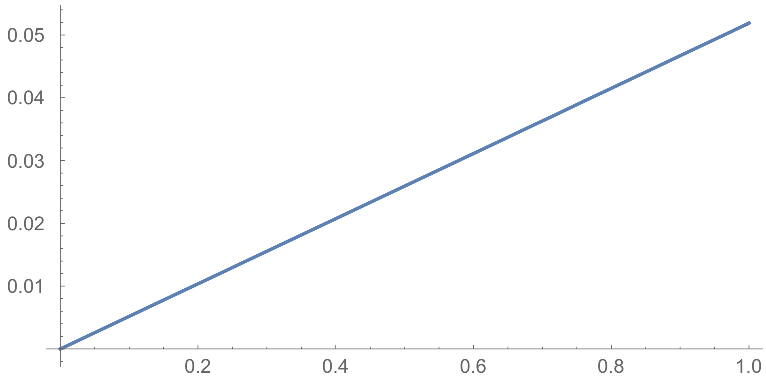


Fig. 5 *The graph of g_L (vertical axis) as a function of ρ (horizontal axis), $\rho \in [0.000000000001; 1]$*

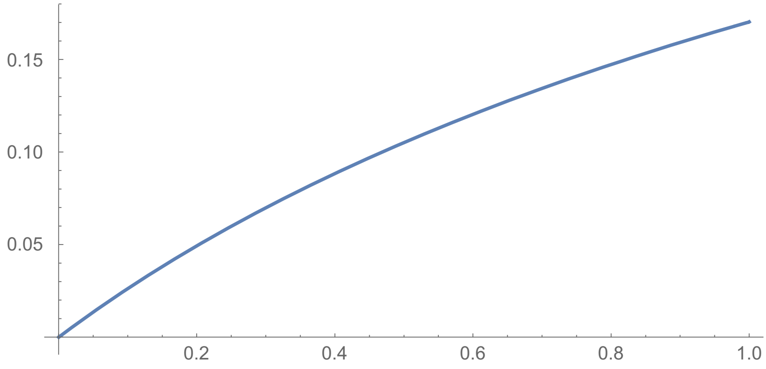
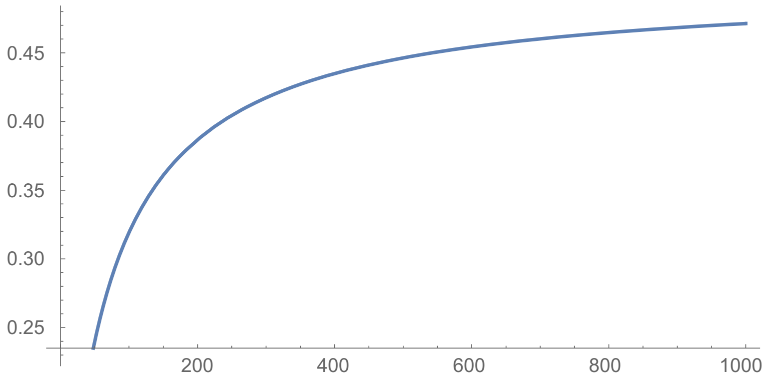
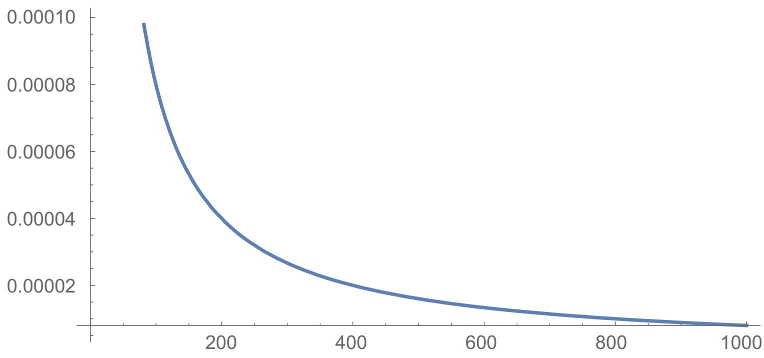


Fig. 6 *The graph of g_L (vertical axis) as a function of σ (horizontal axis), $\sigma \in [0.000000000001; 1000]$*



?

Fig. 7 *The graph of g_L (vertical axis) as a function of ξ (horizontal axis), $\xi \in [0.000000000001; 1000]$*



- Individuals have a Millian-type intertemporal utility function of the form:

$$U \equiv \int_0^{\infty} [\log(c_t) + \upsilon \log(g_{Lt})] e^{-\rho t} dt, \quad \rho > 0, \quad \upsilon > 0.$$

- The flow budget constraint is:

$$\dot{a}_t = r_t a_t + u_t h_t w_t - \left(1 + \Omega_{L,t}\right) c_t, \qquad \Omega \geq 0.$$

Finally, we continue to use the same definition of *Balanced Growth Path* (BGP) *equilibrium* provided in the main text (Section 3).

The *Hamiltonian function* (J_t) of the dynamic, intertemporal problem now reads as:

$$J_t = [\log(c_t) + \upsilon \log(g_{Lt})] e^{-\rho t} + \lambda_{at}\left[r_t a_t + u_t h_t w_t - \left(1 + \Omega g_{L,t}\right) c_t\right]$$
$$+ \lambda_{ht}[\sigma(1 - \varphi g_{Lt} - u_t) - \xi g_{Lt}] h_t,$$

where λ_{at} and λ_{ht} continue to be the co-state variables associated with the two state variables (a_t and h_t, respectively). The necessary first order conditions are:

$$\frac{\partial J_t}{\partial c_t} = 0 \quad \Leftrightarrow \quad \frac{e^{-\rho t}}{c_t} = \lambda_{at}\left(1 + \Omega g_{L,t}\right) \tag{C1}$$

$$\frac{\partial J_t}{\partial g_{Lt}} = 0 \quad \Leftrightarrow \quad \frac{\upsilon e^{-\rho t}}{g_{Lt}} - \Omega \lambda_{at} c_t - \lambda_{ht}(\sigma\varphi + \xi) h_t = 0 \tag{C2}$$

$$\frac{\partial J_t}{\partial u_t} = 0 \quad \Leftrightarrow \quad \lambda_{at} = \sigma \frac{\lambda_{ht}}{w_t} \tag{C3}$$

$$\frac{\partial J_t}{\partial a_t} = -\dot{\lambda}_{at} \quad \Leftrightarrow \quad \lambda_{at} r_t = -\dot{\lambda}_{at} \tag{C4}$$

$$\frac{\partial J_t}{\partial h_t} = -\dot{\lambda}_{ht} \quad \Leftrightarrow \quad \lambda_{at} u_t w_t + \lambda_{ht}[\sigma(1 - \varphi g_{Lt} - u_t) - \xi g_{Lt}] = -\dot{\lambda}_{ht}, \tag{C5}$$

along with the two transversality conditions:

$$\lim_{t \to +\infty} \lambda_{at} a_t = 0, \qquad \lim_{t \to +\infty} \lambda_{ht} h_t = 0,$$

and the initial conditions:

$$a(0) > 0, \qquad h(0) > 0.$$

From (C1) it follows that:

$$e^{-\rho t} = \lambda_{at}\left(1 + \Omega g_{L,t}\right)c_t. \tag{C6}$$

Plugging this expression into (C2) yields:

$$c_t\lambda_{at}\left[\frac{\upsilon}{g_{Lt}} - \Omega(1-\upsilon)\right] = \lambda_{ht}(\sigma\varphi + \xi)h_t. \tag{C7}$$

Employing (C3) into (C7) delivers:

$$\left[\frac{\upsilon}{g_{Lt}} - \Omega(1-\upsilon)\right] = \frac{(\sigma\varphi + \xi)}{\sigma}\frac{h_t w_t}{c_t}. \tag{C8}$$

Using (C3) in (C5), instead, gives:

$$\frac{\lambda_{ht}}{\lambda_{ht}} = (\sigma\varphi + \xi)g_{Lt} - \sigma. \tag{C9}$$

Along a BGP all variables depending on time (including L$_t$) grow at constant (possibly positive) exponential rates. Therefore, in a BGP equilibrium (where $g_{Lt} = g_L, . \forall t \geq 0$) Eq. (C9) suggests that $\lambda^{\bullet}{}_{ht}/\lambda_{ht}$ is also constant. Eqs. (C3) and (C4) imply, respectively:

$$\frac{\lambda^{\bullet}{}_{at}}{\lambda_{at}} = \frac{\lambda^{\bullet}{}_{ht}}{\lambda_{ht}} - \frac{w^{\bullet}{}_{t}}{w_t}, \tag{C10}$$

$$\frac{\lambda^{\bullet}{}_{at}}{\lambda_{at}} = -r_t. \tag{C11}$$

Combining together (C9), (C10), and (C11) leads immediately to:

$$r_t = \sigma + \frac{w_t}{w_t} - (\sigma\varphi\xi)g_L. \tag{C12}$$

Since human capital is perfectly mobile across sectors, at equilibrium it will be rewarded according to the same wage, *i.e.*: $w_{Yt} = w_{It} = w_{nt} \equiv w_t$. Moreover, along the BGP this common wage will grow at a constant exponential rate, implying that $\frac{w^{\bullet}{}_{Yt}}{w_{Yt}} = \frac{w^{\bullet}{}_{It}}{w_{It}} = \frac{w^{\bullet}{}_{nt}}{w_{nt}} \equiv \frac{w^{\bullet}{}_{t}}{w_t}$ is constant. Accordingly, Eq. (C12) implies that in a BGP equilibrium the real rate of return on asset holdings, r , will be constant, as well.

From (C1) and (C4) together:

$$\frac{\dot{c_t}}{c_t} = r - \rho, \quad c_t \equiv \frac{C_t}{L_t}. \tag{C13}$$

By combining (C7), (C13), (C4), (C9), and (13a) in this appendix, after some simple algebra it is finally possible to obtain:

$$u_t = \frac{\rho}{\sigma} = u, \quad \forall t \geq 0, \tag{C14}$$

Eq. (C14) suggests that along a BGP the fraction of time allocated to productive activities is also constant. Making use of Eqs. (6) and (10) in the main text, we find that along a BGP:

$$V_{nt} = Z(1-Z)\left(\frac{H_{Yt}}{n_t}\right)^{1-Z}\left(\frac{H_{It}}{n_t}\right)^{Z}\frac{n_t^R}{[r+(1-R)\gamma_n - \gamma_H]}, \quad R \equiv 1 + \overline{a} - Z$$
$$> 0; \quad \frac{\dot{n}_t}{n_t} \equiv \gamma_n; \quad \frac{\dot{H_t}}{H_t} \equiv \gamma_H. \tag{C15}$$

Given V_{nt}, from Eq. (9') in the main text:

$$w_{nt} = \frac{Z}{\chi}(1-Z)s_n^{\mu-1}H_t^{\mu-1-\Phi}n_t^{\eta}\left(\frac{H_{Yt}}{n_t}\right)^{1-Z}\left(\frac{H_{It}}{n_t}\right)^{Z}\frac{n_t^R}{[r+(1-R)\gamma_n - \gamma_H]}, \tag{C16}$$

where $s_n \equiv H_{nt}/H_t$ is (by definition) constant in a BGP-equilibrium.

We can now use Eqs. (5), (2) and (4') in the main text, obtaining:

$$w_{It} = Z^2\left(\frac{H_{Yt}}{n_t}\right)^{1-Z}\left(\frac{H_{It}}{n_t}\right)^{Z-1}n_t^R. \tag{C17}$$

From Eq. (15) in the main text, by equalizing (C16) and (C17) in this appendix one gets:

$$s_I \equiv \frac{H_{It}}{H_t} = \frac{Z\chi}{(1-Z)}\frac{[r+(1-R)\gamma_n - \gamma_H]}{s_n^{\mu-1}}\frac{n_t^{1-\eta}}{H_t^{\mu-\Phi}}. \tag{C18}$$

Combining Eqs. (1) and (4') in the body-text yields:

$$w_{Yt} \equiv \frac{\partial Y_t}{\partial H_{Yt}} = (1-Z)\left(\frac{H_{Yt}}{n_t}\right)^{-Z}\left(\frac{H_{It}}{n_t}\right)^{Z}n_t^R. \tag{C19}$$

From (16) in the main text and (C18) above, equalization of (C17) and (C19) in this appendix delivers:

$$s_Y \equiv \frac{H_{Yt}}{H_t} = \left(\frac{1-Z}{Z^2}\right) s_I = \frac{\chi}{Z} \frac{[r + (1-R)\gamma_n - \gamma_H]}{s_n^{\mu-1}} \frac{n_t^{1-\eta}}{H_t^{\mu-\Phi}}. \tag{C20}$$

Along a BGP all variables depending on time grow at constant (possibly positive) exponential rates, and the sectorial shares of human capital employment are also constant. Therefore, from Eq. (8) in the main text it follows that:

$$\frac{\dot{n_t}}{n_t} \equiv \gamma_n = \left(\frac{\mu-\Phi}{1-\eta}\right) \gamma_H. \tag{C21}$$

If $\mu - \Phi = 1 - \eta$, we have a very special case of the model in which human capital and technology grow at the same rate $\gamma_n = \gamma_H \equiv \gamma$ in the long-run BGP equilibrium. Here we continue to allow for the most general possible case in which: $\mu \neq \Phi \neq \Phi + 1 - \eta$.

Using Eqs. (C16), (C17), (C19) and (C21), we observe that along a BGP wages grow at a common and constant rate:

$$\frac{\dot{w_{nt}}}{w_{nt}} = \frac{\dot{w_{It}}}{w_{It}} = \frac{\dot{w_{Yt}}}{w_{Yt}} \equiv \frac{\dot{w_t}}{w_t} = R\gamma_n. \tag{C22}$$

From (17) in the main text and (C15) in this appendix we conclude that along a BGP:

$$\frac{\dot{a}_t}{a_t} \equiv \gamma_a = \gamma_H + R\gamma_n - g_L, \quad a_t \equiv A_t/L_t, \quad g_L \equiv \dot{L}_t/L_t. \tag{C23}$$

Merging (12) in the main text and (C11) in this appendix yields:

$$\frac{\dot{\lambda_{at}}}{\lambda_{at}} = -\gamma_a + u\frac{h_t w_t}{a_t} - (1 + \Omega g_L)\frac{c_t}{a_t}. \tag{C24}$$

Similarly, from the combination of (13a) and (C9) in this appendix we get:

$$\frac{\dot{\lambda_{ht}}}{\lambda_{ht}} = -\gamma_h - \sigma u, \quad \gamma_h \equiv \dot{h}_t/h_t, \quad h_t \equiv H_t/L_t. \tag{C25}$$

Note that (C25) and (C9), taken together, confirm that in this economy:
$\dot{h}_t = [\sigma(1-\varphi g_L - u) - \xi g_L] h_t$, see Eq. (13a) above.
Combination of Eqs. (C9), (C10), (C22), (C23), (C24) and (13a) leads to:

$$\frac{c_t}{a_t} = \frac{u}{(1 + \Omega g_L)} \left[\sigma + \frac{h_t w_t}{a_t}\right], \tag{C26}$$

where $\gamma_h = \gamma_H - g_L$ has been used.

By employing again Eqs. (C22) and (C23), along with the fact that $\gamma_h = \gamma_H - g_L$, from (C26) one immediately concludes that:

$$\frac{c_t}{a_t} = \frac{u}{(1+\Omega g_L)}\left[\sigma + \underbrace{\frac{h(0)w(0)}{a(0)}}_{\equiv \Gamma(0)>0}\right], \tag{C27}$$

where $h(0) > 0$, $a(0) > 0$ and $w(0) > 0$ are the given initial values (*i.e.*, at $t = 0$) of h_t, a_t and w_t, respectively.

Eqs. (C27) implies that along a BGP (where u and g_L are constant):

$$\gamma_c = \gamma_a. \tag{C28}$$

Eq. (C26) can be recast as:

$$h_t w_t = -\sigma a_t + \left(\frac{1+\Omega g_L}{u}\right) c_t. \tag{C29}$$

Similarly, Eq. (C8) can be rewritten as:

$$h_t w_t = \frac{\sigma}{(\sigma\varphi + \xi)}\left[\frac{\upsilon}{g_L} - \Omega(1-\upsilon)\right] c_t. \tag{C8'}$$

After equating (C29) and (C8'), we are able to compute the endogenous value of c_t/a_t:

$$\frac{c_t}{a_t} = \frac{\sigma}{\left(\frac{1+\Omega g_L}{u}\right) - \left(\frac{\sigma}{\sigma\varphi + \xi}\right)\left[\frac{\upsilon}{g_L} - \Omega(1-\upsilon)\right]}. \tag{C30}$$

Eq. (C30) confirms that, along a BGP, c_t and a_t grow at a common rate, see Eqs. (C27) and (C28) above. This happens because the ratio of these two variables is constant in the BGP equilibrium. Finally, we equate (C30) and (C27) and (after some algebra) get:

$$\begin{aligned} &\Gamma(0)(\sigma\varphi + \xi)\Omega g_L^2 \\ &+ \{\Gamma(0)[(\sigma\varphi + \xi) + \rho\Omega(1-\upsilon)] + \rho\sigma\Omega(1-\upsilon)\} g_L - \rho\upsilon[\sigma + \Gamma(0)] \\ &= 0. \end{aligned} \tag{C31}$$

Eq. (C31) gives the endogenous value of g_L (as a function of some of the model's parameters and initial conditions) along a BGP-equilibrium. If we simplify the analysis by normalizing all the relevant initial conditions to one, in such a way that $h(0) = w(0) = a(0) = 1 = \frac{h(0)w(0)}{a(0)} \equiv \Gamma(0) > 0$, then Eq. (C31) easily becomes:

$$(\sigma\varphi+\xi)\Omega g_L^2+[(\sigma\varphi+\xi)+\rho\Omega(1-\upsilon)(1+\sigma)]g_L-\rho\upsilon(1+\sigma)=0. \quad \text{(C31')}$$

The solution to this equation is:

$$g_L=\frac{-[(\sigma\varphi+\xi)+\rho\Omega(1-\upsilon)(1+\sigma)]\pm\sqrt{[(\sigma\varphi+\xi)+\rho\Omega(1-\upsilon)(1+\sigma)]^2+4(\sigma\varphi+\xi)\Omega\rho\upsilon(1+\sigma)}}{2(\sigma\varphi+\xi)\Omega}. \quad \text{(C31'')}$$

One root is positive and the other one is negative. The positive root is:

$$g_L=\frac{-[(\sigma\varphi+\xi)+\rho\Omega(1-\upsilon)(1+\sigma)]+\sqrt{[(\sigma\varphi+\xi)+\rho\Omega(1-\upsilon)(1+\sigma)]^2+4(\sigma\varphi+\xi)\Omega\rho\upsilon(1+\sigma)}}{2(\sigma\varphi+\xi)\Omega}$$
$$=\frac{-[\hat{\xi}+\rho\Omega(1-\upsilon)(1+\sigma)]+\sqrt{[\hat{\xi}+\rho\Omega(1-\upsilon)(1+\sigma)]^2+4\hat{\xi}\Omega\rho\upsilon(1+\sigma)}}{2\hat{\xi}\Omega}, \qquad \hat{\xi}\equiv(\sigma\varphi+\xi)$$

Clearly, in the expression above we are considering the specific case where $\Omega>0$.[22] This expression coincides exactly with Eq. (18) in the main text whenever the parameter ξ is re-scaled to take explicitly into account also the time-costs of fertility (in the production of per capita human capital), that is when $\hat{\xi}$ is defined as: $\hat{\xi}\equiv(\sigma\varphi+\xi)$.

Using Eqs. (1) and (4') in the main text and the definitions of $y_t\equiv\frac{Y_t}{L_t}$, $R\equiv1+\overline{\alpha}-Z>0$, and $\dot{L}_t/L_t\equiv g_L$, we obtain the growth rate of real per capita output along a BGP:

$$\gamma_y\equiv\frac{\dot{y}_t}{y_t}=\frac{\dot{Y}_t}{Y_t}-g_L=(\gamma_H-g_L)+R\gamma_n=\gamma_h+R\gamma_n=\gamma_a=\gamma_c, \quad \text{(C32)}$$

see Eqs. (C23) and (C28) above.

Eq. (C32) is important because it says that along a BGP, per capita income (y), per capita asset holdings (a), and per capita consumption (c) all grow at the same rate.

Using (13a), (C14) and (C21) in this appendix, along with the definition of $\gamma_H=\gamma_h+g_L$, it is possible to re-state (C32) as:

$$\gamma_y=\gamma_H+R\gamma_n-g_L=\gamma_H-g_L+R\Upsilon\gamma_H=(1+R\Upsilon)\gamma_H-g_L=\underbrace{\left(\frac{1+R\Upsilon}{\Upsilon}\right)}_{>0,\text{ according to our assumptions}}\gamma_n-g_L\cdot \quad \text{(C33)}$$

where Υ continues to be defined as (see the main text):

$$\Upsilon=\frac{\mu-\Phi}{1-\eta}.$$

[22] If $\Omega=0$, then it is easy to see that an indeterminate form (of the type: $g_L=0/0$) would arise.

Notice that, while some parameters of the model (namely, ξ, σ, ρ, and now also φ) affect the BGP growth rate of the economy (γ_y) both directly and indirectly (*i.e.*, through their impact on the birth rate, g_L), other parameters (υ and Ω) influence economic growth (γ_y) only indirectly through their sole effect on g_L. Finally, there exists a third set of technological parameters (more precisely: $\overline{\alpha}$, Z, μ, Φ, and η – contributing to define R and Υ) that show a direct impact on γ_y, while at the same time having no effect on g_L. Using the fact that υ and Ω influence economic growth (γ_y) only indirectly (*i.e.*, through their effect on g_L), suppose that there occurs a change in υ[23] such that g_L ultimately varies while at the same time R, Υ, σ, ρ, ξ, and φ do not change at all. Then, it is possible to observe that:

$$\frac{\partial \gamma_y}{\partial g_L} = \frac{\frac{\partial \gamma_y}{\partial \upsilon}}{\frac{\partial g_L}{\partial \upsilon}} = \frac{-[(\sigma\varphi + \xi)(1 + R\Upsilon) - R\Upsilon]\frac{\partial g_L}{\partial \upsilon}}{\frac{\partial g_L}{\partial \upsilon}} = -[(\sigma\varphi + \xi)(1 + R\Upsilon) - R\Upsilon].$$

In the model $R \equiv 1 + \overline{\alpha} - Z > 0$. If we continue to assume (as we do in the main text) that $\Upsilon > 0$, then we conclude:

- $\frac{\partial \gamma_y}{\partial g_L} > 0 \quad \text{if} \quad 0 < \xi < \overline{\xi} \equiv \frac{R\Upsilon}{1 + R\Upsilon} - \sigma\varphi < 1,$
- $\frac{\partial \gamma_y}{\partial g_L} = 0 \quad \text{if} \quad 0 < \xi = \overline{\xi} \equiv \frac{R\Upsilon}{1 + R\Upsilon} - \sigma\varphi < 1,$
- $\frac{\partial \gamma_y}{\partial g_L} < 0 \quad \text{if} \quad 0 < \overline{\xi} \equiv \frac{R\Upsilon}{1 + R\Upsilon} - \sigma\varphi < \xi < 1,$

After noticing that $-[(\sigma\varphi + \xi)(1 + R\Upsilon) - R\Upsilon] < 0$ for any $\xi \geq 1$, we also have:

- $\frac{\partial \gamma_y}{\partial g_L} < 0 \cdot \quad \textbf{\textit{if}}\ \xi \geq 1.$

These results are qualitatively similar to those presented and discussed in the main text (see Proposition 2, where φ has been taken exactly equal to zero).[24]

References

Aghion P, Howitt P (1992) A model of growth through creative destruction. Econometrica 60(2):323–351

Ahituv A (2001) Be fruitful or multiply: on the interplay between fertility and economic development. J Popul Econ 14(1):51–71

Balassa B (1967) Trade liberalization among industrial countries. McGraw-Hill, New York

[23] Clearly, one can use the same argument if the initial change has to do with Ω, instead of υ.

[24] At this stage it is worth observing that, while $\overline{\xi} \equiv \frac{R\Upsilon}{1+R\Upsilon} - \sigma\varphi$ is always lower than one, $\overline{\xi} \equiv \frac{R\Upsilon}{1+R\Upsilon} - \sigma\varphi > 0$ whenever $0 < \varphi < \frac{1}{\sigma}\left(\frac{R\Upsilon}{1+R\Upsilon}\right)$, that is when φ is positive but sufficiently small. This is the assumption that we implicitly use in the comparative statics results reported above (also in the light of the fact that, as far as we know, reliable point estimates of φ do not exist).

Baldanzi A, Bucci A, Prettner, K (2019) Children's health, human capital accumulation, and R&D-based economic growth. Macro Dyn (forthcoming)
Becker GS (1960) *An economic analysis of fertility*. In: In: National Bureau of Economic esearch (Eds.), *Demographic and economic change in developed countries*. Princeton University press, Princeton, pp 209–231
Becker GS, Murphy KM, Tamura R (1990) Human capital, fertility and economic growth. J Polit Econ 98(5): 12–37
Benassy J-P (1998) Is there always too little research in endogenous growth with expanding product variety? Eur Econ Rev 42(1):61–69
BIS (Bank for International Settlements) 2006 76th Annual Report. Basel
Bloom DE, Canning D, Mansfield RK, Moore M (2007) Demographic change, social security systems, and savings. J Monet Econ 54(1):92–114
Brander JA, Dowrick S (1994) The role of fertility and population in economic growth. J Popul Econ 7(1):1–25
Bressoux P, Kramarz F, Prost C (2009) Teachers' training, class size and students' outcomes: learning from administrative forecasting mistakes. Econ J 119(536):540–561
Bucci A (2008) Population growth in a model of economic growth with human capital accumulation and horizontal R&D. J Macro 30(3):1124–1147
Bucci A (2013) Returns to specialization, competition, population, and growth. J Econ Dyn Control 37(10): 2023–2040
Bucci A (2015) Product proliferation, population, and economic growth. J Hum Capital 9(2):170–197
Bucci A, Raurich X (2017) Population and economic growth under different growth engines. Ger Econ Rev 18(2):182–211
Bureau of Labor Statistics (2019) Sources of growth in real output in the private business sector, 1987–2018. Downloadable at: https://www.bls.gov/mfp/outputpb.pdf [accessed on July 15, 2019]
Burkhauser RV, Larrimore J, Simon KI (2012) A "second opinion" on the economic health of the American middle class. Natl Tax J 65(1):7–32
Chu AC, Cozzi G, Liao C-H (2013) Endogenous fertility and human capital in a Schumpeterian growth model. J Popul Econ 26(1):181–202
Cigno A (1998) Fertility decisions when infant survival is endogenous. J Popul Econ 11(1):21–28
Cipriani GP (2014) Population aging and PAYG pensions in the OLG model. J Popul Econ 27(1):251–256
Cipriani GP (2018) Aging, retirement, and pay-as-you-go pensions. Macro Dyn 22(5):1173–1183
Cipriani GP, Fioroni T (2019) Health spending, education, and endogenous demographics in an OLG model. In: Bucci A, Prettner K, Prskawetz A (eds) *Human Capital and Economic Growth: The Impact of Health, Education and Demographic Change*. Palgrave-MacMillan, Basingstoke *forthcoming*
Cipriani GP, Pascucci F (2019) Pension policies in a model with endogenous fertility. J Pension Econ Finan, forthcoming
Dalgaard C-J, Kreiner CT (2001) Is declining productivity inevitable? J Econ Growth 6(3):187–203
DeLong JB, Magin K (2009) The U.S. equity return premium: past, present, and future. J Econ Perspect 23(1): 193–208
Devereux MB, Head AC, Lapham BJ (1996a) Monopolistic competition, increasing returns, and the effects of government spending. J Money Credit Bank 28(2):233–254
Devereux MB, Head AC, Lapham BJ (1996b) Aggregate fluctuations with increasing returns to specialization and scale. J Econ Dyn Control 20(4):627–656
Devereux MB, Head AC, Lapham BJ (2000) Government spending and welfare with returns to specialization. Scand J Econ 102(4):547–561
Dinopoulos E, Segerstrom PS (1999) A Schumpeterian model of protection and relative wages. Am Econ Rev 89(3):450–472
Doepke M (2005) Child mortality and fertility decline: does the Barro-Becker model fit the facts? J Popul Econ 18(2):337–366
Ethier WJ (1982) National and international returns to scale in the modern theory of international trade. Am Econ Rev 72(3):389–405
European Commission (EC, 2007) The labour income share in the European Union. In: Employment in Europe 2007, Directorate-general for employment, Social Affairs and Equal Opportunities, Brussels, pp. 237–72
Fanti L, Gori L (2014) Endogenous fertility, endogenous lifetime and economic growth: the role of child policies. J Popul Econ 27(2):529–564
Galor O (2005) From stagnation to growth: unified growth theory. In: Aghion P, Durlauf SN (eds) Handbook of economic growth. Elsevier-North Holland, Amsterdam, pp 171–293
Galor O (2011) Unified growth theory. Princeton University Press, Princeton

Galor O, Weil D (2000) Population, technology, and growth: from Malthusian stagnation to the demographic transition and beyond. Am Econ Rev 90(4):806–828

Grossman GM, Helpman E (1991a) Quality ladders in the theory of growth. Rev Econ Stud 58(1):43–61

Grossman GM, Helpman E (1991b) Innovation and growth in the global economy. MIT Press, Cambridge

Ha J, Howitt P (2007) Accounting for trends in productivity and R&D: a Schumpeterian critique of semi-endogenous growth theory. J Money Credit Bank 39(4):733–774

Hashimoto K, Tabata K (2016) Demographic change, human capital accumulation and R&D-based growth. Canadian Econ Rev 59(2):707–737

Hattie J (2005) The paradox of reducing class size and improving learning outcomes. Int J Educ Res 43(6): 387–425

Herzer D, Vollmer S, Strulik H (2012) The long-run determinants of fertility: one century of demographic change 1900–1999. J Econ Growth 17(4):357–385

Howitt P (1999) Steady endogenous growth with population and R&D inputs growing. J Polit Econ 107(4): 715–730

ILO (International Labour Office) (2012) Global wage report 2012/13: wages and equitable growth, Geneva

ILO (International Labour Office) – OECD (Organisation for Economic Cooperation and Development) (2015) *The Labour Share in G20 Economies*. Report prepared for the G20 Employment Working Group - Antalya, Turkey, 26-27 February 2015. Downloadable at: https://www.oecd.org/g20/topics/employment-and-social-policy/The-Labour-Share-in-G20-Economies.pdf. Accessed date: September 5, 2019.

IMF (International Monetary Fund) (2007) *The globalization of labor*. In: In World Economic Outlook. April: Spillovers and cycles in the world economy, Washington, pp 161–192

Jones CI (1995) R&D-based models of economic growth. J Polit Econ 103(4):759–783

Jones CI (2005) Growth and ideas. In: Aghion P, Durlauf SN (eds) Handbook of economic growth. Elsevier-North Holland, Amsterdam, pp 1063–1111

Kalemli-Ozcan S (2003) A stochastic model of mortality, fertility, and human capital investment. J Dev Econ 70(1):103–118

Karabarbounis L, Neiman B (2014) The global decline of the labor share. Quarterly J Econ 129(1):61–103

Kelley AC, Schmidt R (1995) Aggregate population and economic growth correlations: the role of the components of demographic change. Demography 32(4):543–555

Kortum S (1997) Research, patenting and technological change. Econometrica 65(6):1389–1419

Kuhn M, Wrzaczek S, Prskawetz A, Feichtinger G (2015) Optimal choice of health and retirement in a life-cycle model. J Econ Theory 158(A):186–212

Li H, Zhang J (2007) Do high birth rates hamper economic growth? Rev Econ Stat 89(1):110–117

Lucas RE (1988) On the mechanics of economic development. J Mon Econ 22(1):3–42

Mathis WJ (2016) The effectiveness of class size reduction. National Education Policy Center and School of Education, University of Colorado Boulder, Boulder Retrieved from: https://nepc.colorado.edu/sites/default/files/publications/Mathis%20RBOPM-9%20Class%20Size.pdf. Accessed date: September 5, 2019

OECD (Organisation for Economic Cooperation and Development) (2012) Employment outlook 2012, Paris

Peretto PF (1998) Technological change and population growth. J Econ Growth 3(4):283–311

Prettner K (2013) Population aging and endogenous economic growth. J Popul Econ 26(2):811–834

Prettner K (2014) The non-monotonous impact of population growth on economic prosperity. Econ Lett 124(1):93–95

Prettner K, Canning D (2014) Increasing life expectancy and optimal retirement in general equilibrium. Economic Theory 56(1):191–217

Psacharopoulos G, Patrinos AH (2018) Returns to investment in education: a decennial review of the global literature. Educ Econ 26(5):445–458

Romer PM (1987) Growth based on increasing returns due to specialization. Am Econ Rev 77(2):56–62

Romer PM (1990) Endogenous technological change. J Polit Econ 98(5):71–102

Segerstrom PS (1998) Endogenous growth without scale effects. Am Econ Rev 88(5):1290–1310

Strulik H (2005) The role of human capital and population growth in R&D-based models of economic growth. Rev Int Econ 13(1):129–145

Strulik H, Prettner K, Prskawetz A (2013) The past and future of knowledge-based growth. J Econ Growth 18(4):411–437

Trostel PA (2004) Returns to scale in producing human capital from schooling. Oxford Econ Pap 56(3):461–484

World Bank (2019a) World development indicators. Downloadable at: http://datatopics.worldbank.org/world-development-indicators/ [Accessed on May 16, 2019]

World Bank (2019b) Education statistics. Downloadable at: http://datatopics.worldbank.org/education/ [Accessed on May 16, 2019]

Zyngier D (2014) Class size and academic results, with a focus on children from culturally, linguistically and economically disenfranchised communities. *Evid Base*, issue 1, pp. 1-23. Retrieved from: https://www.exeley.com/exeley/journals/evidence_base/2014/1/pdf/10.21307_eb-2014-001.pdf. Accessed date: September 5, 2019