Network

Rimini Centre for Economic Analysis
Working Paper Series

wp 24-05

Human Capital-based Growth with Depopulation and Class-size Effects: Theory and Empirics

Alberto Bucci
Lorenzo Carbonari
Giovanni Trovato
Pedro Trivin

Human Capital-based Growth with Depopulation and Class-size Effects: Theory and Empirics

Alberto Bucci* Lorenzo Carbonari† Giovanni Trovato‡
Pedro Trivin§

April 28, 2024

Abstract

Building on Lucas (1988), we develop a model in which the impact of population dynamics on per capita GDP and human capital depends on the balance of intertemporal altruism effects towards future generations and class-size effects on an individual's education investment. We show that there is a critical level of the class-size effect that determines whether a decline in population growth will lead to a decrease or an increase in a country's long-run growth rate of real per capita income. We take the model to OECD data, using a semi-parametric technique. This allows us to classify countries into groups based on their long-term growth trajectories, revealing patterns not captured by previous studies on the topic.

JEL Classification*: J11, O11, O41.*

Keywords*: Long-run economic growth; Depopulation; Class-size effects; Human capital investment.*

*ICEA (International Centre for Economic Analysis), and University of Milan (Department of Economics, Management and Quantitative Methods, DEMM). E-mail: alberto.bucci@unimi.it.

†Università degli Studi di Rome "Tor Vergata"(Dipartimentio di Economia e Finanza). E-mail: lorenzo.carbonari@uniroma2.it.

‡Università degli Studi di Rome "Tor Vergata"(Dipartimentio di Economia e Finanza). E-mail: giovanni.trovato@economia.uniroma2.it.

§University of Milan (Department of Economics, Management and Quantitative Methods, DEMM). E-mail: pedro.trivin@unimi.it.

1 Introduction

This paper advances both a theoretical argument and a piece of empirical evidence to shed light on the possible mechanisms by which a reduction in population size may exert a favorable impact on economic growth, particularly through the enhancement of educational attainments.

While an increase in population growth lowers the long-run level and the short-run growth rate of real per capita GDP within the standard Solow (1956) model,[1] there exists a definitely positive correlation between the long-run growth rate of real per capita GDP and population size (*strong scale-effect*) in the first generation models of fully-endogenous growth, such as those proposed by Romer (1990), Grossman and Helpman (1993), and Aghion and Howitt (1992). However, it is now widely recognized that this strong scale effect is not supported by empirical evidence.

As a consequence, in the second-generation models of endogenous growth, exemplified by Jones (1995), Kortum (1997), and Segerstrom (2000), it is population growth (as opposed to population size) to be posited as the ultimate engine of (semi-endogenous) economic growth (*weak scale-effect*). The class of semi-endogenous growth models is characterized by the common assumption of diminishing returns of knowledge to the production of new ideas, and therefore it posits the need for a positive population growth rate for the long-run economic growth rate to be maintained positive, as well. Diminishing returns imply that an increasing effort is required to continue to innovate (or to enhance a product) at a given rate. Hence, with a constant fraction of the population engaged in research, population growth is ultimately the only source of sustained economic growth.

Most recently, Sasaki and Hoshida (2017) have revived the debate on the economic growth implications of a changing population by introducing a negative population growth rate within the canonical semi-endogenous growth setting by Jones (1995). They were among the first to show that with negative (and exogenous) population growth, the rate of technological change falls to zero while the (semi-endogenous) growth rate of per capita output may still be positive in the long-run. Although interesting, Sasaki and Hoshida (2017)'s analysis seems empirically implausible (Jones, 2022, p. 3492). Jones (2022, Result 1, p. 3494) finds that in a fully-endogenous growth framework, a negative and exogenous population growth is compatible only with stagnating knowledge and living standards in the very long-run (the so-called *empty-planet result*). In the second part of his analysis, Jones (2022) extends the

[1] Christiaans (2011), using a modified Solow-type growth model, shows that the correlation between population growth and real per capita income growth may become non-monotonous when population growth is negative.

canonical R&D–based, semi-endogenous growth setup with exogenous population growth to the case where fertility choices are endogenous. In this extension two different equilibria are characterized and analyzed. In the decentralized equilibrium, the endogenous rate of growth of the population can be negative, as it ultimately depends on several parameters. In this framework, negative (and endogenous) population growth rates still lead the economy toward the empty-planet result. Instead, in the optimal equilibrium (which differs from the decentralized one for the only fact that it explicitly takes into account that a larger population may raise the future rate of technological progress), the economy reaches in the long-run a state where economic growth is positive provided that the (endogenous) population growth rate is positive itself. Given that both the decentralized– and the optimal–equilibrium endogenous population growth rates depend on structural policy parameters, Jones (2022) concludes that it is important to put in place policies that allow switching to the optimal allocation that guarantees the sustained exponential growth in population, and therefore in knowledge and in living standards.[2] What makes our analysis fundamentally different from Jones (2022) and the other related contributions mentioned above is the crucial role that in the current paper individuals' educational choices (as opposed to their fertility ones) may play in the presence of a population that declines over time at an exogenous rate.

In the light of all this, we believe that our paper is most closely related to recent works by Feichtinger et al. (2023), Boikos et al. (2023), Bucci (2023), and Siskova et al. (2023). Feichtinger et al. (2023) include endogenous education investments and human capital accumulation in the vein of Lucas (1988) into the model of Jones (2022). They show that the basin of attraction of the Jones (2022) solution with depopulation and economic decline shrinks (and, for reasonable parameter values, may even completely vanish) when investing in education becomes a possibility. In other words, the Jones (2022) empty-planet result may disappear in the presence of human capital investment.[3] The empty-planet result of

[2]Bucci and Prettner (2020) also study the long-run correlation between population growth and economic growth under endogenous fertility choices. In addition to endogenous fertility, their model also includes endogenous education investments within an algebraically-tractable multi-sector, horizontal R&D-driven growth model. Their analysis is capable, too, of yielding a negative association between population growth and productivity growth in the long run. The mechanism that they describe in their work to obtain this result is, however, based on the existence of an explicit quality/quantity substitution effect between population (quantity) and human capital accumulation (quality). With respect to Bucci and Prettner (2020), the present paper does not hinge upon any quality/quantity trade-off and investigates the theoretical conditions under which in a human capital-based growth model, without R&D activity and where individuals' education decisions are characterized by the presence of class-size effects (as empirical evidence seems mostly to suggest – see the next section of this paper), a negative and exogenously-given population growth rate may be compatible with a positive long-run growth rate of real income per capita.

[3]Using a discrete-time model, Strulik (2022) makes a similar point in arguing that human capital accu-

Jones (2022) describes a situation in which, with a population that declines over time, the stock of knowledge and the level of GDP per capita would eventually stabilize as time goes to infinity. While current data seem to corroborate that TFP is already stagnating, this is not the case for GDP per capita. To explain this 'puzzle', Boikos et al. (2023) include human capital accumulation in the Jones (2022) model and show that TFP stagnation may be compatible with an increasing level of GDP per capita in the very long-run. In other words, they demonstrate that adding human capital to the Jones (2022) model with negative (and exogenous) population growth may contribute to reconcile the long-term no-stagnation-result in GDP per capita with the long-term TFP-stagnation-result. Bucci (2023) analyzes, within an R&D-based economy that invests also in human capital, the conditions (related to the size of a crucial parameter affecting the law of motion of per capita human capital) under which in the very long-run a negative population growth rate may be conducive to positive growth rates of ideas, per capita income, and per capita human capital, respectively. The main difference between the current paper and all the works briefly summarized above is that in the current article we completely abstract from R&D-activity and focus our attention solely on the role that human capital accumulation may (under particular conditions) eventually play for the growth rate of real per capita GDP to continue to be positive even under a shrinking population size.

In a recent stimulating work, Siskova et al. (2023) empirically quantify the changes in aggregate human capital, given the observed fertility (and population) declines. Overall, they find that declining fertility is only partly compensated by increasing education and health investments per capita when all countries are included in their regressions. Indeed, according to them, the elasticity of individual human capital with respect to fertility is about -0.124 %. This elasticity is further reduced when the focus is just on those countries that face population declines and when migrations are also included as a control variable of the regressions. This implies that countries subject to population decline find themselves in a more difficult position to compensate for the human capital effects of declining fertility. Unlike our contribution, Siskova et al. (2023), are not interested in grasping the possible theoretical mechanisms hiding behind the relationship between human capital and growth in the presence of declining population.

Our paper aims to rationalize the relationship between depopulation and educational attainments, clarifying some observed empirical patterns. Population dynamics is governed by two primary determinants: fertility and mortality rates. Owing to the issue of reverse causality, the task of distinguishing the individual contributions of fertility and mortality

mulation may sustain in the very long-run both the process of ideas-creation and the process of economic growth in spite of a declining population, even under decreasing returns to education.

rates to educational attainment poses a significant empirical challenge. In the vast majority of countries, empirical evidence indicates that individuals with higher education levels exhibit lower mortality rates, and this trend extends to their offspring, who exhibit higher survival probabilities. Worldwide, it has been observed that women with advanced education levels tend to have fewer children.[4]

Table 1: Percentiles for Population Growth in OECD countries (1950-2019)

Percentile	value
1%	-0.529
5%	-0.205
10%	0.003
25%	0.325
50%	0.706
75%	1.251
90%	1.972
95%	2.390
99%	3.208

United Nations's (2022) projections indicate that the populations of 61 countries or areas are expected to decline by at least 1 percent between 2022 and 2050, especially due to persistently low fertility rates and, in certain instances, high emigration levels. This demographic trend underscores the necessity for policymakers to anticipate and manage the potential socioeconomic impacts of a shrinking population base. Table 1 illustrates that, during the period from 1950 to 2019, the lowermost 5 % percentile of the population growth rate distribution among OECD countries registered negative values.

The well-documented negative correlation between population growth and educational attainment[5] serves as a foundation for our analysis, which further posits that this association is non-monotonic. We focus on OECD countries where the decrease in population size for the lowest five percent is linked to the highest (in absolute value) coefficients in the following simple (two fixed) effects quantile regressions:

$$Q(\text{Human capital}_{it}|\text{population growth rate}_{it}; q) = \alpha_i + \beta X_{it}, \tag{1}$$

[4]The impact of education on fertility rates is particularly pronounced in non-OECD countries, which display higher fertility figures (Lutz and Kc, 2011).

[5]See, e.g., Lutz et al. (2017).

in which the aggregate educational attainment X is captured: (i) by the gross secondary enrollment (source: World Bank) or, alternatively, (ii) by the Human Capital Index (source: Penn World Tables). In (1) the probability $P(\text{Human capital}_{it} \leq f(\hat{\alpha} + \hat{\beta} X_{it}) = q$, where q are the quantile of the population growth rate distribution. Figure 1 illustrates a notably negative impact on the indicators of educational attainment at the lower end of the population growth rate distribution.

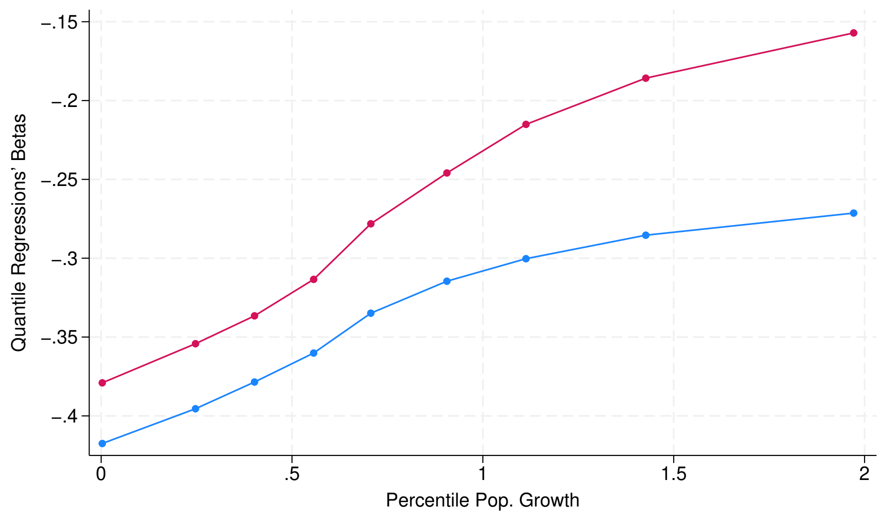

Figure 1: Quantile Regression.

Red line: $\hat{\beta}$ for Schooling. Blue line: $\hat{\beta}$ for Human Capital Index

In the first part of the paper we provide a rationale for such a figure. Building on Lucas (1988), we propose a model in which, in the long-run, the influence of a diminishing population on the growth of real per capita GDP and human capital is non-monotonic.[6] In particular, the main theoretical result of our model is that depending on whether the agents' degree of altruism toward future generations is larger or smaller than the class-size effect that characterizes any individual's human capital investment, a differential effect of an increasingly negative population growth rate on a country's long-term rate of real per capita income and human capital growth may be observed.[7]

[6]In an R&D-based growth model with diminishing technological opportunities in the sector that produces new ideas for new varieties of differentiated intermediates, Bucci et al. (2021) obtains a non-monotonic association between population growth and GDP growth.

[7]For an insightful analysis of the relevance of the intertemporal altruism effect in the relationship between economic and demographic growth, see Boucekkine and Fabbri (2013). In particular, they conclude that when the inverse of the intertemporal elasticity of substitution in consumption is larger than one, decreasing the intertemporal altruism parameter does favor the realization of the so-called *Parfit's Repugnant Conclusion*.

In the second part of the paper, we take the model to the data, contributing to the empirical literature on the association between population dynamics and human capital.[8] We employ a Finite Mixture Model, which facilitates addressing the unobserved heterogeneity that stems from the non-monotonic and non-uniform relationship between population change and human capital, as posited by our theoretical framework. This estimation technique also enables the execution of a cluster analysis, allowing for the categorization of countries into distinct groups according to the homogeneity in the conditional distribution of their long-term growth rates, which is in relation to the estimated latent factors.

Because one important ingredient of our theoretical model (as well as a fundamental departure of it from Lucas, 1988) is the recognition that the presence of class-size effects indeed characterizes the process of acquiring skills, the next section is devoted to an informed analysis of the vast body of empirical literature that has recently highlighted the relevance of such effects, at least in some particular contexts.

Outline The paper is structured as follows. The next section reviews the empirical literature on class-size effects in human capital investment. Section 3 lays down the model while Section 4 presents the data, the estimation method and results. Section 5 concludes. Proofs are in the Appendix.

2 Class-Size effects in human capital investment: A brief review of the empirical literature

This section reviews the empirical literature that suggests some correlation between larger class sizes and lower student achievement, encompassing cognitive and non-cognitive skills, abilities, and test scores.[9]

In an influential paper, Angrist and Lavy (1999) used the so-called *Maimonides' rule* to estimate the class-size effect. This rule exploits statutory limits on class size in Israel as a source of quasi-experimental variation. As first noted by the authors, Israeli schools face a

[8]The focus of prior research has primarily centered on life expectancy, whose change is generally found to exert a rather negligible influence on educational attainments, or on fertility rates, as widely discussed in Hazan and Zoabi (2006), Hazan (2009), Cervellati and Sunde (2013) and Cervellati and Sunde (2015).

[9]See, among others, Fredriksson et al. (2013), Konstantopoulos and Chung (2009), Bosworth (2014), Finn et al. (2005). As for the relation between class size and students' scores, while Finn et al. (2005) find that a small class is on average associated with a significantly higher graduation rate, Konstantopoulos and Chung (2009) notice that "...longer periods in small classes produce higher increases in achievement in later grades for all types of students...".

maximum class size of 40, so that, in principle, grade cohorts of 41 are split into two different classes, while slightly smaller cohorts of 39 may be taught in just one large class. Analyzing data on class average scores for the population of Israeli 4th and 5th graders tested in June 1991, the authors found a substantial return to class size reductions. They used 1327 class means in 625 schools and scores in verbal and math achievement. They estimated effect-sizes across classes of 0.18 (for 5th grade students) and 0.13 (for 4th grade students).

Many different applications of the Maimonides' rule research design in other settings also report statistically significant learning gains from smaller classes. An example is Urquiola (2006) who studied 10,018 third-grade students in Bolivia with the 608 different class sizes varying from very small up to 40. The effect-sizes were estimated between 0.18 and 0.23.

Wößmann and West (2006) examined the effect of class size on student achievement across numerous countries, using data from the Third International Mathematics and Science Study (TIMSS). This data enabled the researchers to rule out both between-school and within-school factors that influence student distribution. In the majority of countries in their sample, the impact of class size on student performance was not as significant as the effect found by Krueger (1999) in the Student/Teacher Achievement Ratio (STAR) project. However, the results varied significantly from country to country. On one end of the spectrum, Greece and Iceland showed substantial class-size effects. On the other hand, any significant class-size effect was dismissed for Canada, Portugal, Singapore, and Slovenia. The findings of Wößmann and West (2006) align with an explanation rooted in the quality of the teaching force: smaller classes have a noticeable positive effect on student achievement primarily in those countries where the average teaching quality appears to be relatively low.

Dustmann et al. (2003) is one of the rare studies that not only measures the direct impact of class-size reductions (such as the likelihood of students staying in school, a key measure of overall educational achievement) but also the indirect impacts, specifically on future earnings. The authors used micro-data to investigate the effects of reducing class sizes on the decision to stay in school beyond the age of 16 for students in England and Wales, and on their future earnings. They based their analysis on several waves of the National Child Development Study (NCDS), a longitudinal survey of children born in one week in 1958. This data set provided a unique range of background variables, helping to avoid the common problem of omitted variable bias that is often seen in empirical research. They found a significant negative relationship between class size and the probability of staying in school at age 16. This finding remained robust across various specifications that controlled for different background variables, including past performance tests and types of schools. Furthermore, the decision to stay in school beyond the minimum leaving age had a considerable impact on wages at ages 33 and 42 (and for women, also at age 23). In conclusion, a smaller class size appears to have a significantly positive effect on the likelihood of students staying in school, as well as on their future earnings.[10]

While early studies primarily focused on the outcomes of students at primary and secondary education levels, more recent discussions have begun to explore how class size can also impact academic achievement at the tertiary education level.[11] In this context, Bandiera et al. (2010), using administrative records from a leading UK university, concluded the following: (i) The average effect size is -0.108, implying that if a student were to move from an average-sized class to a class that is one standard deviation larger, their test score would decrease by 0.108. (ii) The effect size is only significant for the smallest and largest ranges of class sizes, and it is zero for intermediate class sizes. This suggests that the impact of class size is non-linear. (iii) Students at the top of the test score distribution are more affected by changes in class size, especially when class sizes are very large. This indicates that the highest-achieving students would benefit the most from a reduction in class sizes, particularly when class sizes are exceedingly large. (iv) Lastly, the study found that the effect of class size does not change based on proxies for students' wealth.

Kara et al. (2021) extend the previous work of Bandiera et al. (2010) and estimate the effect of class size on academic (higher education) performance of university students, distinguishing between STEM (Science–Technology–Engineering–Mathematics) and non-STEM fields of study. Using administrative data from a large UK higher education institution, it is found that: (i) Larger classes are associated with significantly lower grades (the average effect size is –0.08); (ii) This average effect masks, however, considerable differences across academic fields of study, as a larger effect is observed in STEM subjects (–0.11) than in non-STEM subjects (–0.04); (iii) In terms of students' socio-economic status, ability, and gender, smaller classes seem to be particularly beneficial for students from a low socio-economic background, and (within the STEM fields of study) for higher ability and male students.[12]

This key evidence forms the basis for the theoretical model that follows. If one fully accepts these empirical findings, their primary implication for our forthcoming theory is straightforward: A more negative population growth rate (i.e., a smaller population size,

[10] Hence, "...it is worthwhile to investigate the indirect impact of school quality enhancements, such as class size reductions, over and above its direct impact on educational performance..." Dustmann et al. (2003).

[11] A synthesis of this literature can be found in Kara et al. (2021).

[12] According to Krueger and Whitmore (2001), the beneficial effects of small classes on college aspirations appear to be stronger for those students who received free or reduced-price lunch. For these low socio-economic-status (SES) students, Finn et al. (2005) find that "...the odds of graduating were 67% greater for students attending small classes for 3 years and almost 2.5 times greater for students attending small classes for 4 years...".

which serves as a proxy for smaller class sizes) should lead to a faster rate of per capita human capital growth (i.e., a higher level of per capita human capital stock, which serves as a proxy for improved student performance).

This selective and unavoidably brief review of empirical studies in the field of educational economics suggests that the number of students in a class (and, at a macroeconomic level, a country's population size and growth rate) significantly influences individual educational achievements. This includes the accumulation of students' human capital and the skills they acquire. Generally, students' academic performance improves when they are in smaller classes.

3 The model

Technology We build on Lucas (1988). The model is set in continuous time, with time running on the interval $[0, \infty)$. At any point in time, the production of a homogeneous consumption good Y_t is obtained by employing, under constant returns to scale, solely human capital H_{Yt} as an input:

$$Y_t = AH_{Yt}, \tag{2}$$

where $A > 0$ is the productivity of human capital employed in production, H_Y. In this economy, the available stock of human capital (H) can be employed either for producing (under perfect competition) a homogeneous consumption/final good or for *producing* new human capital. At any point in time, the share of human capital used for producing the final good is $u_t \in [0, 1]$, whereas the share of human capital used for producing new human capital is $1 - u_t$. For the sake of simplicity, we postulate that, once produced, total output Y_t is consumed, i.e., $Y_t = C_t$.[13] Let now N_t denote the existing population at time t. In per capita terms, (2) becomes:

$$y_t = A\left(u_t h_t\right) = c_t \tag{3}$$

where $y_t \equiv Y_t/N_t$, $h_t \equiv H_t/N_t$ and $c_t \equiv C_t/N_t$. We assume that population grows over time at an exogenous rate n such that

$$\frac{\dot{N}_t}{N_t} = n \quad \text{with} \quad n \gtreqless 0 \quad \text{and} \quad N(0) > 0. \tag{4}$$

Per capita human capital accumulation is given by

$$\dot{h}_t = \xi(1 - u_t)h_t - \delta h_t \quad \text{with} \quad h(0) > 0, \tag{5}$$

[13] See Jones (2022, p. 3498, Table 1).

where ξ is the efficiency of human capital investment and $\delta \in (0, 1)$ is the instantaneous depreciation rate of human capital per capita. Unlike Lucas (1988), here we account also for the possible presence of *class-size effects* in schooling investment. Specifically, we assume that the efficiency with which human capital may be augmented ξ increases when population size N_t (a proxy for a class size) shrinks and that the faster population size declines and the faster ξ rises over time, i.e.,

$$\xi = \sigma - \varepsilon \left(\frac{\dot{N}_t}{N_t} \right) \quad \text{with} \quad \sigma > 0 \text{ and } \varepsilon \geq 0. \tag{6}$$

Equation (6) tells us that there might exist (for any $\varepsilon > 0$) a negative correlation between the rate of population growth and the efficiency with which any single individual can augment her own stock of embodied knowledge (human capital).

In particular, when $\varepsilon > 0$ and $n > 0$, an increase in the population growth rate leads both to a faster rise in population size over time and to a quicker decline in the efficiency of human capital investment (ξ). All other factors being equal, this eventually results in a slower rate of per capita human capital accumulation. Instead, when $\varepsilon > 0$ and $n < 0$, a rise in the rate at which population decreases results in a faster decline in population size over time and in a swifter increase in the efficiency of human capital investment (ξ). When all other factors remain constant, this eventually prompts an acceleration in the speed at which an individual may acquire new human capital over time. In the remaining scenario where $n = 0$, our model simplifies to the one originally proposed by Lucas (1988), with $\xi = \sigma > 0$. The same outcome would occur if $\varepsilon = 0$.

Overall, the existence of a class size effect implies that a population that decreases faster might positively influence the efficiency of per capita human capital acquisition over time. Consequently, a smaller population size (which can be seen as a reflection of a smaller class size) ultimately results in improved student academic performance. In our model, the degree of the class size effect is therefore directly determined by the value of the parameter $\varepsilon > 0$.

Combine (5) and (6) to get

$$\dot{h}_t = (\sigma - \varepsilon n)(1 - u_t)h_t - \delta h_t \quad \text{with} \quad n \gtreqless 0, \tag{7}$$

which explicitly illustrates that, when $n < 0$, a stronger population decrease (i.e., a larger $|n|$) enhances the growth of per capita human capital by improving the efficiency of individual educational technology, for any $\varepsilon > 0$.

The specific point that the next section tries to examine is whether (and, eventually, under which condition(s)) the presence of a class-size effect in education may contribute to avoiding losses of economic growth in the long-run, following a steady decline in population size.

Households The economy is populated by many structurally-identical households. Therefore, we analyze the behavior of a representative infinitely-lived family with perfect foresight whose size coincides with the size of the whole population N_t. Each member of the representative household has a *Constant-Intertemporal-Elasticity of Substitution* (CIES) instantaneous felicity function of the form:

$$u(c_t) = \frac{c_t^{1-\theta} - 1}{1-\theta}, \quad \text{with } \theta > 0,\ \theta \neq 1. \tag{8}$$

The representative household allocates time-resources towards the acquisition of human capital and chooses the optimal path of the time-share of human capital to be devoted to production and educational activities. The problem faced by the head of the household is to maximize the household's discounted infinite lifetime utility. Using (3), (4), (7) and (8), the household's problem is:

$$\max_{\{c_t,u_t,h_t\}_{t=0}^{\infty}} \mathcal{U} \equiv \int_0^{\infty} u(c_t) N_t^{\mu} e^{-\rho t} dt = \int_0^{\infty} \frac{c_t^{1-\theta} - 1}{1-\theta} N(0)^{\mu} e^{-(\rho-\mu n)t} dt \tag{9}$$

s.t.

$$\begin{aligned}
&\frac{\dot{N}_t}{N_t} = n \gtreqless 0, \\
&c_t = A(u_t h_t) = y_t, \\
&\dot{h}_t = (\sigma - \varepsilon n)(1-u_t)h_t - \delta h_t, \\
&h(0) > 0, \quad N(0) > 0,
\end{aligned}$$

where $\rho > 0$ is the subjective discount rate and $\mu \in [0,1]$ is the intertemporal altruism parameter.[14] We assume $\rho - \mu n > 0$ and, for the sake of simplicity, normalize $N(0)$ to 1.

[14] In (9), $\mu = 0$ corresponds to the Millian intertemporal utility case while $\mu = 1$ corresponds to the Benthamite intertemporal utility case. Although it is difficult to have a precise point estimate of agents' degree of altruism toward future generations, we now have some support for the hypothesis that it is very small and most likely less than one. Indeed, Altonji et al. (1997) is, to our knowledge, one of the very few attempts at obtaining a direct estimate of agents' degree of altruism. They test for a specific form of altruism, namely that of parents who make money transfers to their children. According to the theory of inter-vivos transfers, we would face perfect altruism if an increase, say, of one dollar in the income of parents making transfers to a child, coupled with a simultaneous one-dollar decrease in that child's income, resulted in the parents' increasing their transfer to the same child by exactly one dollar. To test this hypothesis the authors use the 1968-89 Panel Study of Income Dynamics (PSID) data set and control for the principal theoretical determinants of money transfers (the current and permanent incomes of the parents, the child, and the child's siblings). Their findings suggest that redistributing one dollar from a recipient child to donor parents

In this economy, a balanced growth path (BGP) equilibrium is an equilibrium where variables depending on time grow at constant exponential rates. Using equation (7) we see immediately that along a BGP the fraction of human capital employed in production and educational activities is constant, that is $u_t = u \ \forall t \geq 0$. Moreover, it is possible to show that along a BGP equilibrium, the following results do hold:[15]

$$(1-u) = \frac{(\sigma-\rho) - \delta(1-\theta) + n(\mu-\varepsilon)}{\theta(\sigma-\varepsilon n)}, \tag{10}$$

and

$$\frac{\dot{y}_t}{y_t} = \frac{\dot{c}_t}{c_t} = \frac{\dot{h}_t}{h_t} \equiv \gamma = (\sigma-\varepsilon n)(1-u) - \delta = \frac{(\sigma-\rho-\delta) + n(\mu-\varepsilon)}{\theta}. \tag{11}$$

Notice that in the absence of any population change ($n = 0$), provided that the condition $\sigma > \rho + \delta$ holds,[16] the model would predict positive economic growth without any (weak or strong) scale effect.

Negative population growth and long-run economic growth We now focus on the specific case where $n < 0$ which is probably the most relevant one today, at least for OECD countries.[17]

From (11) we get:

$$\frac{\partial\gamma}{\partial n} = \frac{1}{\theta}(\mu-\varepsilon). \tag{12}$$

With $\theta > 0$, this implies that[18]

$$\text{Sign}\left(\frac{\partial\gamma}{\partial n}\right) = \text{Sign}(\mu-\varepsilon). \tag{13}$$

leads to only about a 13-cent increase in the parents' transfer to that child. Using panel data on bequests, rather than inter-vivos transfers from parents to children, Laitner and Ohlsson (2001) obtain similar results.

[15] See the Appendix A.1 for a complete derivation of the results of the model.

[16] See, e.g., Strulik (2005, p. 135).

[17] Bricker and Ibbitson (2019) contend that the world is on the brink of an under-population crisis. They report, indeed, decreasing rates of population growth in the 21st century, particularly in prosperous regions such as Europe and Japan, and forecast a comparable trend in nations like China, Brazil, and even in historically high-fertility regions like India and Sub-Saharan Africa. They claim that the global population will start its decline in approximately three decades, a transformation they deem to be irreversible. Jones (2022, p. 3489) points out that "the natural rate of population growth (i.e., births minus deaths, ignoring immigration) is already negative in Japan and in many European countries such as Germany, Italy, and Spain".

[18] For estimates of the intertemporal elasticity of substitution using macro data, see, e.g., Campbell (2003) and Yogo (2004). For estimates using microdata, see, e.g., Attanasio and Weber (1993) and Vissing-Jørgensen (2002).

If $\varepsilon > \mu$, then the class-size effect is quite strong (i.e., it is larger than the intertemporal altruism parameter). In this case, any further decrease in n would imply: (i) a faster decline in population size *but* also (ii) a higher real per capita GDP growth rate along the BGP. If instead, $\varepsilon < \mu$, then the class-size effect is rather negligible (i.e., it is smaller than the intertemporal altruism parameter). In this case, any further decrease in n would imply: (i) a faster decline in population size *and* (ii) a lower real per capita GDP growth rate along the BGP. Lastly, if $\varepsilon = \mu$ then any additional reduction in n would have no impact on the long-term economic growth rate of the economy.

4 Quantitative analysis

The model presented above has two fundamental testable predictions.

1. In the long-run (i.e., along a BGP), the consequences on real per capita GDP growth of a declining population may be non-monotonic.

2. There exists a threshold-value of the class-size effect above/below which we would observe a differential impact of a more negative growth rate of population on a country's long-run (BGP) growth rate of real per capita income.[19] In particular, following a more negative population growth rate, we would observe a higher growth rate of the economy when the class-size effect in educational achievement is sufficiently large, i.e., when $\varepsilon > \mu$.

To test these theoretical predictions against OECD data, we also account for the possibility that fundamental parameters differ across countries. Guided by the theoretical model, our parameter estimations enable us to identify situations where depopulation negatively impacts growth and where it does not.

Data Our empirical analysis relies on the Penn World Table (PWT version 10.01), a revered database widely employed in the study of macroeconomic dynamics. Covering 183 countries from 1950 to 2019, this dataset provides crucial information on relative levels of income, output, input, and productivity (Feenstra et al., 2015).

From this database, we use data on the Human Capital Index, population, and a proxy that measures the proportion of human capital allocated to the production of the final good. This proxy is derived from equation (3) as follows:

[19] In our simplified theoretical framework, this threshold coincides with the intertemporal altruism parameter, μ, which is itself very likely to be quite small.

$$u_{it} = \frac{Y_{it}}{TFP_{it} \times HC_{it}},$$

where Y, TFP, and HC represent, respectively, GDP per capita, Total Factor Productivity, and the Human Capital Index for country i and year t, all obtained from the PWT dataset.

Subsequently, we utilize u_{it} to compute the share of human capital dedicated to fostering new human capital $(1 - u_{it})$. Adhering to the theoretical framework, this variable is standardized within each country, ensuring its values fall within the range 0 to 1.

Due to the closer alignment of our theoretical model with developed economies, our analysis focuses on a selective set of 34 OECD countries observed between 1954 and 2019. Table A.1 in the Appendix A.2 provides descriptive statistics for the variables used in the analysis.[20]

Estimates Assume for simplicity $\delta = 0$. Then, using (11), equation (7) can be written as:

$$\frac{\dot{h}}{h_t} \equiv \gamma = \sigma(1 - u_t) - \varepsilon \left[n_t \times (1 - u_t) \right]. \tag{14}$$

The empirical counterpart of (14) is

$$\gamma_{it} = \zeta_0 + \zeta_1(1 - u_{it}) + \zeta_2 \left[n_{it} \times (1 - u_{it}) \right] + \zeta_3 n_{it} + \epsilon_{it}, \tag{15}$$

where γ_{it} is the annual growth rate of the Human Capital Index and ϵ_{it} is an iid error term.[21] Confronting (14) with (15), we see immediately that $\hat{\zeta}_1 = \hat{\sigma}$ and $\hat{\zeta}_2 = -\hat{\varepsilon}$. Notice also that in (15) the interaction between population growth (n_t) and the investment in education $(1 - u_t)$ is complete.

Table 2 presents the OLS Fixed Effects (FE), OLS Random Effect (RE), and Feasible GLS (FGLS) estimates, respectively, for (15). As predicted by the theory, for the interaction $n_t \times (1 - u_t)$, all three models show a negative association with the dependent variable γ. The FE and RE models show similar coefficient estimates of -0.19 and -0.22, respectively, both significant at a 1% level. The FGLS model suggests a slightly smaller coefficient of -0.09, also significant at a 1% level. As predicted by the theory, the variable $1 - u_t$ is positively

[20] To mitigate potential selection biases, countries that joined the OECD after 2010 (such as Latvia-2016, Lithuania-2018, Colombia-2018, and Costa Rica-2021) are excluded. Additionally, the choice to start our sample period in 1954 is due to missing TFP values in earlier years.

[21] Our model predicts that, along the BGP, the stock of human capital expands at the same rate as real GDP, as stated in (11). Since our primary interest lies in the former, we utilize the growth rate of the Human Capital index (HC) as a proxy for γ. However, empirical data show that, despite a strong correlation, $\dot{y}/y$ and $\dot{h}/h$ are not necessarily identical. Employing the growth rate of real GDP as the dependent variable in (15) does not change the core findings of our analysis. However, this approach yields less precise estimates.

associated with γ in all models, with the FE model showing the largest coefficient estimate (0.79), and the FGLS showing the smallest (0.24), but again with a remarkably low standard error. Lastly, the variable n_t, which has been added to account for the complete interaction, also has a positive relationship with γ, with an approximate coefficient value of 0.1 across all models. Finally, the R^2 values indicate that the FGLS model has a considerably better fit ($R^2 = 0.74$) to the data compared to the FE ($R^2 = 0.17$) and RE ($R^2 = 0.05$) models. Time and individual controls were included in all models.

Table 2: OLS,GLS Estimation

	Dependent variable: γ		
	(FE)	(RE)	(FGLS)
$n_t \times (1 - u_t)$	-0.19^{***}	-0.22^{***}	-0.09^{***}
	(0.05)	(0.05)	(0.00)
$(1 - u_t)$	0.79^{***}	0.51^{***}	0.24^{***}
	(0.13)	(0.05)	(0.00)
n_t	0.09^{*}	0.10^{*}	0.10^{***}
	(0.04)	(0.04)	(0.00)
Intercept		0.41^{***}	0.00
		(0.05)	(0.00)
Time and individual controls	✓	✓	✓
R^2	0.17	0.05	0.74
Adj. R^2	0.13	0.05	
Num. obs.	2039	2039	2005
σ_idios		0.45	
σ_id		0.23	

Note: *p<0.1; **p<0.05; ***p<0.01
σ_idios: Variance Idiosyncratic; σ_id: Variance individual, Standard Errors in Brackets.

FGLS model offers the most precise and best-fitting model for the data, given its low standard errors and high R^2 value. Notice also that in the OLS models, the inference appears to be biased due to non-normal residuals. This is evident from the Shapiro-Wilk test, which rejects the assumption of normality with a test statistic of 0.988 and a highly significant p-value of 0.000. This suggests that the residuals don't follow a normal distribution, which can lead to biased and inefficient estimates in the OLS models. On the other hand, the FGLS model, which allows for heteroscedasticity or different variances in the residuals, presents a more robust approach in this case. It also takes into account the potential correlations among the covariates.[22]

In the following, we apply a Finite Mixture Model (FMM) to the equation (14) to cluster countries according to the unobserved heterogeneity. In general, Mixture Models allow to test the behavior of human capital growth dynamics, under the assumption that unobserved heterogeneity affects parameters estimation. In particular, we employ a FMM, relaxing the hypothesis of iid residuals and allowing for correlated random terms.[23] In such a model, the random component captures the impact of unobserved country-specific variables, limiting the effects of the omitted variable bias. FMM allows dealing with the unobserved heterogeneity due to the non-monotonic, non-uniform relationship between regressors and response, as predicted by our theory. Moreover, through this estimation procedure, we are able to perform a cluster analysis: we sort countries into groups based on the homogeneity of the conditional distribution of their long-run growth rates with respect to the estimated unobservable factors.

According to the Generalized Linear Models framework (McCullagh and Nelder, 1989), equation (14) can be written as:

$$\gamma_{it} = \zeta_{0i} + \zeta_{1i}(1 - u_{it}) + \zeta_{2i}\left[n_{it} \times (1 - u_{it})\right] + \zeta_{3i} n_{it} + \epsilon_{it}, \qquad (16)$$

where parameters $\zeta_i = \zeta_{0i}, \zeta_{1i}, \zeta_{2i}, \zeta_{3i}$ are now country specific. In particular, they capture the country-specific unobserved factors that affect human capital growth, through population change, the share of human capital dedicated to fostering new human capital, and its interaction with population growth.

We assume that parameters in ζ_i can be empirically described by random variables, with unspecified probability function, and cluster-specific variances σ_i. In this way, eqn. (16) takes explicitly into account the between countries random terms correlation.

The nonparametric maximum likelihood estimator (NPMLE) of the distribution is discrete,[24] with a finite number of locations and masses. This implies that the country-specific latent variables are modeled as measures of the difference between country *i-th*'s covariates and their sample mean. We assume that γ_{it} is a conditionally independent realization of the

[22] This flexibility makes the FGLS model more suitable for data sets where the assumption of homoscedasticity (equal variances) in the residuals and independence of covariates is violated.

[23] Assuming that some of the fundamental covariates were not included in the model specification and that their joint effects can be accounted for by adding latent variables to the linear predictor, it is possible to relax the assumption of iid residuals. See, e.g., Aitkin (1997) and McLachlan and Peel (2000).

[24] See, e.g., Laird (1978) and Heckman and Singer (1984).

potential human capital growth, given the set of random factors, which varies over countries and accounts for both individual variation and dependence among country-specific rates of growth.

Table 3 presents the results, while Table 4 provides the corresponding country classification. From the previous specification, we obtain a classification of countries into 5 clusters. The first noteworthy result is the positive impact of the coefficient $1 - u_t$ across the different clusters. Consistently with the theory, investment in human capital is positively associated with the HC index's growth rate, ranging from 0.30 (Cluster 5) to a maximum of 1.91 (cluster 3). This variance suggests that, despite the consistent positive relationship between investment in human capital and its growth rate, this relationship is heterogeneous across clusters, underscoring the influence of unobserved factors.

A distinct pattern emerges when examining the interaction term capturing the class-size effect (ε). Here, we observe three distinct patterns. Only Cluster 5 exhibits a positive effect. All the remaining clusters share a comparably negative coefficient, ranging from -0.26 to -0.42, with the exception of Cluster 3 which displays a significantly stronger negative relationship, with a coefficient of -0.76. The negative sign aligns with the theoretical model, where a negative relationship exists between population growth and the efficiency of human capital (see equation (6)). However, the positive sign found for one group contradicts the class-size effect, indicating that for some countries, larger population growth is associated with more efficient human capital. Below we will try to better understand this effect by studying the composition of the different groups.

When considering the impact of the coefficient linked to the population growth variable, a careful interpretation is required. Unlike the previous two coefficients, this one is included in the model to account for empirical interactions but does not have a direct implication derived from the theoretical framework.[25] Finally, although frequently not statistically significant, most clusters exhibit a negative intercept. This indicates a tendency towards a lower baseline growth rate of the HC index.

Overall, these estimates suggest that the relationship between population growth, human capital investment, and the growth rate of the HC index is complex and varies across clusters. The interaction effect is particularly varied, with some clusters showing a positive relationship and others negative. This could reflect different demographic dynamics or varying effectiveness of human capital investment in different sub-populations or contexts represented by the clusters. The direct effect of human capital investment is consistently positive across all clusters, emphasizing its importance in the growth of the HC index. The varied signs and magnitudes of the population growth rate indicate that its impact on the

[25]See Brambor et al. (2006) for a detailed explanation of the interpretation of interaction models.

HC index's growth may depend on other factors captured by the cluster distinctions.

We seek a deeper understanding of our findings by examining the countries grouped into various clusters in our analysis, as detailed in Table 4. Among the 34 OECD countries in our sample, there is an uneven distribution of countries across groups. The largest contingent falls into Cluster 5 (12 countries), followed by Cluster 2 (10 countries) and Cluster 4 (6 countries). The remaining groups (Cluster 1 and Cluster 3) consist of 3 countries each. Initial observations make it challenging to discern any consistent patterns within the classification.[26] Indeed, we find no substantial differences in terms of investment of human capital or economic growth across groups during the period under study, and although there are some differences in terms of the HC index, there is not a direct relationship with the interaction coefficient from Table 4.

Remarkably, clusters with higher population growth (Cluster 3 and Cluster 4) exhibit more pronounced class-size effects. In Cluster 3 we find Luxembourg ($\partial\gamma/\partial n = -0.649$), Mexico ($\partial\gamma/\partial n = -1.284$) and Republic of Korea ($\partial\gamma/\partial n = -0.649$).[27] In 2021, the OECD reports that Korea and Mexico have some of the largest class sizes in both primary (with an average of 22.45 and 22.57 students per teacher, respectively) and secondary education (with an average of 26.49 and 26.17 students per teacher, respectively).[28] In 2022, instead, Luxembourg had one of the highest proportions of young adults aged 18-24 in education or training, either full-time or part-time, despite having one of the lower average student-to-teacher ratios in both primary and secondary schools.[29] Among the countries grouped in Cluster 5, those with average class sizes in primary and secondary school that were above the OECD mean were Australia ($\partial\gamma/\partial n = 0.251$), France ($\partial\gamma/\partial n = -0.076$), and Hungary ($\partial\gamma/\partial n = -0.375$).[30] Average class sizes in Australia have steadily risen since the 1980s. Caps that were 25 in high schools and 22 in technical schools have since expanded. Even primary schools, which now cap early years below 26, have seen an increase from the low point in 1981 (Zyngier, 2014). Similarly, in France, average class sizes have begun to decline only gradually since 2017 in primary school, aligning with the new government's class size reduction policy.[31]

[26] It is worth noting that our empirical method classifies countries based on unobserved factors, thus a clear-cut classification based on observables is not expected.

[27] Table A.3 presents, for each country, the marginal effects of population growth rate on the GDP growth rate, using the same country classification as Table 4 and the average values of $(1 - u_{it})$.

[28] Source: `https://stats.oecd.org/`.

[29] See OECD (2023).

[30] Source: `https://stats.oecd.org/`.

[31] Initially focused on high-priority educational areas (REP+), the policy capped first-grade classes at 12 pupils starting in the 2017-2018 school year and expanded to second-grade classes in the 2018-2019 school year (Bressoux et al., 2019).

The documented relationship between population growth and class size appears to be non-linear. Indeed, despite having the third-highest average growth rate of the population, Cluster 5 exhibits a positive coefficient in the interaction term. Countries in line with the theoretical model, showing a negative effect of the interaction term are the majority (22 out of 34), with a large heterogeneity among them. Within these groups, we find the Mediterranean and Middle-Eastern countries, most Nordic countries in the sample (except for Norway), along with the Anglo-Saxon countries (Canada, UK, and US) outside Oceania, Chile, and Japan. In examining the group of countries deviating from our theoretical predictions, we find a cluster comprising mostly European nations, Australia, New Zealand, and Norway. It is worth noting that the European block in this cluster excludes Germany, Poland, and Slovakia, making it more difficult to find a common pattern across the clusters.

In summary, our results underscore the presence of a class-size effect in certain countries, while for others, population growth appears not to impact human capital efficiency. Regrettably, the country classification doesn't elucidate the underlying mechanisms. Nevertheless, it is evident that the relationship is heterogeneous and warrants further research, as this question holds significance for future economic policy considerations.

Table 3: Finite Mixture Model

	Dependent variable: γ				
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
$n_t \times (1 - u_t)$	-0.31**	-0.26***	-0.76***	-0.42*	0.36***
	(0.10)	(0.07)	(0.11)	(0.17)	(0.10)
$(1 - u_t)$	0.65***	0.77***	1.91***	1.47***	0.30*
	(0.15)	(0.11)	(0.20)	(0.19)	(0.13)
n_t	0.31***	0.14**	0.37***	0.34*	-0.31***
	(0.08)	(0.05)	(0.09)	(0.13)	(0.07)
Intercept	0.03	-0.24	-0.42*	-0.90***	-0.22
	(0.14)	(0.13)	(0.18)	(0.17)	(0.14)
Time Fixed Effects: ✓					
Observations	2039				
σ_u	0.404				

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Standard Errors in brackets

Table 4: Country Classification, model in Table 3

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	Greece	Chile	Luxembourg	Canada	Australia
	Portugal	Estonia	Mexico	Czech Republic	Austria
	Turkey	Finland	Republic of Korea	Germany	Belgium
		Iceland		Sweden	Denmark
		Israel		United Kingdom	France
		Italy		United States	Hungary
		Japan			Ireland
		Poland			Netherlands
		Slovakia			New Zealand
		Spain			Norway
					Slovenia
					Switzerland
Avg. $\partial\gamma_t/\partial n_t$	0.012	-0.093	-0.797	0.048	-0.061
n_t	0.933	0.8734	1.537	0.688	0.704
$1-u_t$	0.627	0.578	0.637	0.579	0.634
h_t	0.753	0.808	0.693	0.822	0.699
GDP growth	0.026	0.027	0.028	0.027	0.312

Robustness Table A.4 presents a modified version of the model, adding as a control variable the one-period lagged value of the HC index, denoted by h_{t-1}. Under this specification, the fixed parameter related to it is fixed and exhibits a significant negative effect on the dependent variable. The random parameters show different effects across clusters. Compared to our main specification, the estimates for these parameters are generally less precise. Moreover, the parameter for population growth is never significant across groups. The parameter related to the interaction $n_t \times (1-u_t)$ shows a negative effect on γ in Cluster 1 and Cluster 3, while it has a positive effect in Cluster 2 and Cluster 4. These findings indicate that the influence of the random parameters on the dependent variable varies across different clusters, emphasizing the heterogeneous nature of the data and the value of using a mixture model approach.

The intercept terms also demonstrate a significant positive effect on γ across all clusters, reinforcing the robustness of the model. These estimates confirm the complexity of the relationship between the predictors and the dependent variable, across different clusters.

Table A.5 provides more information about the country classification. We find the strongest estimated class-size effect in countries belonging to Cluster 3 ($\partial\gamma/\partial n = -0.603$). Qualitatively, this result is consistent with our findings from the main analysis using the model in Table 3 for all countries within this cluster, with the sole exception of Portugal.

Cluster 1 is characterized by a negligible negative effect of population growth on the growth rate of the HC index (-0.037), indicating a near-neutral relationship for this group of countries. This finding aligns well with the result from our main analysis for countries like Denmark, France, Hungary, Ireland, Italy, Mexico, Norway, Poland, Slovenia, Spain, and Turkey.

The cases of Luxembourg and the Republic of Korea, now grouped in Cluster 2, are the more problematic since for these two countries the estimates suggest a positive effect of population on γ, which contradicts our previous findings. Controlling for the lagged value of the HC index appears to be important for these countries.

Overall, these estimates corroborate our earlier findings, revealing varied economic dynamics among the clusters. They indicate that although investment in human capital remains high across groups, the marginal impact of population growth on the growth of the HC index varies markedly. This variation may be attributed to interactions with other factors that are not observable.

5 Conclusion

In this paper, we presented a novel perspective on the relationship between population growth and the growth of real per capita GDP and human capital. Drawing on the work of Lucas (1988), we developed a model that illustrates a non-monotonic long-run influence of diminishing population on these growth metrics. Our main theoretical finding is that the impact of a negative population growth rate on a country's long-term real per capita income and human capital growth can vary, depending on whether the altruism towards future generations exceeds or falls short of the class-size effect on human capital investment.

We applied our model to OECD data. We use a Finite Mixture Model to address the unobserved heterogeneity arising from the non-monotonic and non-uniform relationship between population change and human capital. This method also allowed us to perform a cluster analysis, categorizing countries based on their long-term growth rates' homogeneity in relation to the estimated latent factors.

The Finite Mixture Model reveals that parameter heterogeneity is a significant issue when analyzing the macroeconomic impact of the class-size effect. Our findings, which appear to be robust to different specifications, contribute to the extensive literature on this topic by providing empirical evidence on the relationship between population dynamics, human capital accumulation, and economic development using publicly available cross-country data.

References

Philippe Aghion and Peter Howitt. A model of growth through creative destruction. *Econometrica*, 60(2):323–351, 1992. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/2951599.

Murray Aitkin. The calibration of p-values, posterior bayes factors and the aic from the posterior distribution of the likelihood. *Statistics and Computing*, 7(4):253–261, 1997.

J.K. Altonji, F. Hayashi, and L.J. Kotlikoff. Parental altruism and inter vivos transfers: Theory and evidence. *Journal of Political Economy*, 105(6):1121–1166, 1997.

J.D. Angrist and V. Lavy. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics*, 114(2):533–575, 1999.

Orazio P Attanasio and Guglielmo Weber. Consumption growth, the interest rate and aggregation. *The Review of Economic Studies*, 60(3):631–649, 1993.

O. Bandiera, V. Larcinese, and I. Rasul. Heterogeneous class size effects: New evidence from a panel of university students. *Economic Journal*, 120(549):1365–1398, 2010.

Spyridon Boikos, Alberto Bucci, and Tiago Neves Sequeira. May human capital rescue the empty planet? *Economics Letters*, 232:111352, 2023.

R. Bosworth. Class size, class composition, and the distribution of student achievement. *Education Economics*, 22(2):141–165, 2014.

Raouf Boucekkine and Giorgio Fabbri. Assessing Parfit's repugnant conclusion within a canonical endogenous growth set-up. *Journal of Population Economics*, 26:751–767, 2013.

Thomas Brambor, William Roberts Clark, and Matt Golder. Understanding interaction models: Improving empirical analyses. *Political analysis*, 14(1):63–82, 2006.

Pascal Bressoux, Laurent Lima, and Christian Monseur. Reducing the number of pupils in French first-grade classes: Is there evidence of contemporaneous and carryover effects? *International Journal of Educational Research*, 96:136–145, 2019.

Darrell Bricker and John Ibbitson. *Empty planet: the shock of global population decline*. Hachette UK, 2019.

Alberto Bucci. Can a negative population growth rate sustain a positive economic growth rate in the long run? *Mathematical Social Sciences*, 122:17–28, 2023.

Alberto Bucci and Klaus Prettner. Endogenous education and the reversal in the relationship between fertility and economic growth. *Journal of Population Economics*, 33(3):1025–1068, 2020.

Alberto Bucci, Lorenzo Carbonari, and Giovanni Trovato. Variety, competition, and population in economic growth: theory and empirics. *Macroeconomic Dynamics*, 25(5):1303–1330, 2021.

John Y Campbell. Consumption-based asset pricing. *Handbook of the Economics of Finance*, 1:803–887, 2003.

Matteo Cervellati and Uwe Sunde. Life expectancy, schooling, and lifetime labor supply: theory and evidence revisited. *Econometrica*, 81(5):2055–2086, 2013.

Matteo Cervellati and Uwe Sunde. The effect of life expectancy on education and population dynamics. *Empirical Economics*, 48:1445–1478, 2015.

Thomas Christiaans. Semi-endogenous growth when population is decreasing. *Economics Bulletin*, 31(3):2667–2673, 2011.

C. Dustmann, N. Rajah, and A. van Soest. Class size, education, and wages. *Economic Journal*, 113(485):F99–F120, 2003.

Robert C Feenstra, Robert Inklaar, and Marcel P Timmer. The next generation of the Penn World Table. *American Economic Review*, 105(10):3150–3182, 2015.

Gustav Feichtinger, Dieter Grassb, Michael Kuhnb, Klaus Prettnerc, and Alexia Prskawetza. Resurrecting economic growth? Human capital accumulation as an antidote to depopulation. 2023.

J.D. Finn, S.B. Gerber, and J. Boyd-Zaharias. Small classes in the early grades, academic achievement and graduating from high school. *Journal of Educational Psychology*, 97(2): 214–223, 2005.

P. Fredriksson, B. Öckert, and H. Oosterbeek. Long-term effects of class size. *Quarterly Journal of Economics*, 128(1):249–285, 2013.

Gene M Grossman and Elhanan Helpman. *Innovation and growth in the global economy*. MIT press, 1993.

Moshe Hazan. Longevity and lifetime labor supply: Evidence and implications. *Econometrica*, 77(6):1829–1863, 2009.

Moshe Hazan and Hosny Zoabi. Does longevity cause growth? A theoretical critique. *Journal of Economic Growth*, 11:363–376, 2006.

James Heckman and Burton Singer. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica: Journal of the Econometric Society*, pages 271–320, 1984.

Charles I Jones. R & D-based models of economic growth. *Journal of political Economy*, 103(4):759–784, 1995.

Charles I Jones. The end of economic growth? Unintended consequences of a declining population. *American Economic Review*, 112(11):3489–3527, 2022.

E. Kara, M. Tonin, and M. Vlassopoulos. Class size effects in higher education: Differences across stem and non-stem fields. *Economics of Education Review*, 82:102104, 2021.

S. Konstantopoulos and V. Chung. What are the long-term effects of small classes on the achievement gap? Evidence from the lasting benefits study. *American Journal of Education*, 116(1):125–154, 2009.

Samuel S Kortum. Research, patenting, and technological change. *Econometrica: Journal of the Econometric Society*, pages 1389–1419, 1997.

A. Krueger and D. Whitmore. The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project STAR. *Economic Journal*, 111(468):1–28, 2001.

A.B. Krueger. Experimental estimates of education production functions. *Quarterly Journal of Economics*, 114(2):497–532, 1999.

Nan Laird. Nonparametric maximum likelihood estimation of a mixing distribution. *Journal of the American Statistical Association*, 73(364):805–811, 1978.

J. Laitner and H. Ohlsson. Bequest motives: A comparison of Sweden and the United States. *Journal of Public Economics*, 79(1):205–236, 2001.

Robert E Lucas. On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42, 1988.

Wolfgang Lutz and Samir Kc. Global human capital: Integrating education and population. *Science*, 333(6042):587–592, 2011.

Wolfgang Lutz, William P Butz, and KC ed Samir. World population & human capital in the twenty-first century: An overview. 2017.

Peter McCullagh and John Ashworth Nelder. Binary data. In *Generalized linear models*, pages 98–148. Springer, 1989.

Geoffrey J McLachlan and David Peel. Mixtures of factor analyzers. In *Proceedings of the seventeenth international conference on machine learning*, pages 599–606, 2000.

OECD. *Education at a Glance 2023*. 2023. doi: https://doi.org/https://doi.org/10.1787/e13bef63-en. URL https://www.oecd-ilibrary.org/content/publication/e13bef63-en.

Paul M Romer. Endogenous technological change. *Journal of political Economy*, 98(5, Part 2):S71–S102, 1990.

Hiroaki Sasaki and Keisuke Hoshida. The effects of negative population growth: An analysis using a semiendogenous R&D growth model. *Macroeconomic Dynamics*, 21(7):1545–1560, 2017.

Paul S Segerstrom. The long-run growth effects of r&d subsidies. *Journal of Economic Growth*, 5:277–305, 2000.

Martina Siskova, M Kuhn, K Prettner, and A Prskawetz. Does human capital compensate for population decline? *The Journal of the Economics of Ageing*, 26:100469, 2023.

Robert M Solow. A contribution to the theory of economic growth. *The quarterly journal of economics*, 70(1):65–94, 1956.

H. Strulik. The role of human capital and population growth in R&D-based models of economic growth. *Review of International Economics*, 13(1):129–145, 2005.

Holger Strulik. Long-run economic growth despite population decline, 2022. URL https://www.holger-strulik.org/my_papers/growth_pop_decline.pdf.

United Nations. World population prospects 2022. 2022. URL https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf.

M. Urquiola. Identifying class size effects in developing countries: Evidence from rural Bolivia. *Review of Economics and Statistics*, 88(1):171–177, 2006.

Annette Vissing-Jørgensen. Limited asset market participation and the elasticity of intertemporal substitution. *Journal of Political Economy*, 110(4):825–853, 2002.

L. Wößmann and M. West. Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review*, 50(3):695–736, 2006.

Motohiro Yogo. Estimating the elasticity of intertemporal substitution when instruments are weak. *Review of Economics and Statistics*, 86(3):797–810, 2004.

D. Zyngier. Class size and academic results, with a focus on children from culturally, linguistically and economically disenfranchised communities. *Evidence Base*, pages 1–23, 2014. Downloadable at: https://www.classsizematters.org/wp-content/uploads/2014/05/EvidenceBase2014Issue1.pdf.

A Appendix

A.1 Solution of the Model and Conditions on its Parameter-Values

With the initial population size normalized to one, $N(0) = 1$, The Hamiltonian function $\mathcal{H}_t$ related to the consumer's maximization problem (9) in the main text, reads as:

$$\mathcal{H}_t = \left[\frac{(Au_th_t)^{1-\theta} - 1}{1-\theta}\right] e^{-(\rho-\mu n)t} + \lambda_t \left[(\sigma - \varepsilon n)(1-u_t)h_t - \delta h_t\right], \quad (\rho - \mu n) > 0, \tag{A.1}$$

where λ_t is the co-state variable associated with the law of motion of human capital (it is the shadow price of human capital). The (necessary) first-order conditions read as:

$$\frac{\partial \mathcal{H}_t}{\partial u_t} = 0 \Rightarrow (Au_th_t)^{-\theta} Ae^{-(\rho-\mu n)t} = \lambda_t(\sigma - \varepsilon n), \tag{A.2}$$

$$\frac{\partial \mathcal{H}_t}{\partial h_t} = -\dot{\lambda}_t \Rightarrow (Au_th_t)^{-\theta} Au_te^{-(\rho-\mu n)t} + \lambda_t \left[(\sigma - \varepsilon n)(1-u_t) - \delta\right] = -\dot{\lambda}_t. \tag{A.3}$$

along with the transversality condition

$$\lim_{t\to\infty} \lambda_t h_t = 0, \tag{17}$$

and the initial condition

$$h(0) > 0. \tag{18}$$

Remember that along a BGP: $u_t = u$, $\forall t \geq 0$. By merging (A.2) and (A.3), we immediately obtain

$$\lambda_t(\sigma - \varepsilon n)u + \lambda_t \left[(\sigma - \varepsilon n)(1-u_t) - \delta\right] = -\dot{\lambda}_t. \tag{A.4}$$

Equation (A.4) boils down to:

$$-\frac{\dot{\lambda}_t}{\lambda_t} = \sigma - \varepsilon n - \delta. \tag{A.5}$$

Take logs and derive with respect to time (A.2) and obtain

$$-\frac{\dot{\lambda}_t}{\lambda_t} = \theta\frac{\dot{h}_t}{h_t} + (\rho - \mu n) \tag{A.6}$$

Plugging (7) into (A.6) leads to

$$-\frac{\dot{\lambda}_t}{\lambda_t} = \theta(\sigma - \varepsilon n)(1-u) - \theta\delta + (\rho - \mu n). \tag{A.7}$$

Equalization of (A.5) and (A.7) finally implies

$$(1-u) = \frac{(\sigma-\rho)-\delta(1-\theta)+n(\mu-\varepsilon)}{\theta(\sigma-\varepsilon n)}, \tag{A.8}$$

which implies that along a BGP

$$\frac{\dot{y}_t}{y_t} = \frac{\dot{c}_t}{c_t} = \frac{\dot{h}_t}{h_t} \equiv \gamma = \frac{(\sigma-\rho-\delta)+n(\mu-\varepsilon)}{\theta}. \tag{A.9}$$

We now find conditions on the parameter-values such that the model can be economically meaningful. First of all, since we are interested in the determinants of positive long-run growth rates of per capita income, consumption, and human capital, the common growth rate of the economy (A.9) must be positive. This requires

$$(\sigma-\rho) > \delta + n(\varepsilon-\mu). \tag{A.10}$$

At the same time, the share of human capital employed in education $(1-u)$ must be strictly between zero and one. This leads to the following three further restrictions[32]

$$(\sigma-\rho) > \delta(1-\theta)+n(\varepsilon-\mu), \tag{A.11}$$

$$(\sigma-\delta)(1-\theta)-\rho+n\left[\mu-\varepsilon(1-\theta)\right] < 0, \tag{A.12}$$

$$\sigma > \varepsilon n. \tag{A.13}$$

It is easy to show that (A.12) also allows the transversality condition to be checked along a BGP. By putting together (A.10) and (A.11), one observes that the following condition must hold

$$\begin{aligned} &(\sigma-\rho) > \max\left\{\left[\delta+n(\varepsilon-\mu)\right], \left[\delta(1-\theta)+n(\varepsilon-\mu)\right]\right\} \\ \Rightarrow\ &(\sigma-\rho) > \delta+n(\varepsilon-\mu). \end{aligned} \tag{A.14}$$

If conditions (A.12), (A.13) and (A.14) are simultaneously met, then (i) the common growth rate of the economy (A.9) is positive; (ii) the share of human capital employed in educational activities (A.8) is strictly between zero and one; (iii) the transversality condition is checked along a BGP.

[32]Notice that if $\theta = 1$, then the instantaneous utility function of the representative agent is logarithmic in consumption, and the condition (A.12) would be always met, as $(\rho-\mu n) > 0$.

A.2 Tables

Table A.1: Descriptive Statistics, OECD Sample

Variable	N	Mean	St. Dev.	Min	Max
h_t	2,039	2.825	0.564	1.173	3.892
γ_t	2,039	0.687	0.515	−4.301	4.110
n_t	2,039	0.845	0.792	−2.103	4.976
$1 - u_t$	2,039	0.609	0.295	0.000	1.000

Table A.2: LogLikelihoods and Penalized Criteria for model in Table 3

Cluster	lik	aic	bic	icl
1	$-1,381.402$	$2,902.804$	$3,296.219$	$3,296.219$
2	$-1,192.881$	$2,535.763$	$2,957.279$	$2,957.450$
3	$-1,137.134$	$2,434.267$	$2,883.885$	$2,884.557$
4	$-1,112.217$	$2,394.435$	$2,872.153$	$2,874.344$
5	$-1,089.701$	$2,359.402$	$2,865.222$	$2,866.210$

Table A.3: Population growth and GDP growth, model in Table 3

Cluster 1	$\partial\gamma_t/\partial n_t$	Cluster 2	$\partial\gamma_t/\partial n_t$	Cluster 3	$\partial\gamma_t/\partial n_t$	Cluster 4	$\partial\gamma_t/\partial n_t$	Cluster 5	$\partial\gamma_t/\partial n_t$
Greece	0.261	Chile	-0.273	Luxembourg	-0.452	Canada	-0.250	Australia	0.251
Portugal	0.300	Estonia	0.256	Mexico	-1.284	Czech Republic	0.288	Austria	-0.179
Turkey	-0.223	Finland	0.022	Republic of Korea	-0.649	Germany	0.225	Belgium	-0.165
		Iceland	-0.186			Sweden	0.125	Denmark	-0.163
		Israel	-0.529			United Kingdom	0.153	France	-0.076
		Italy	0.043			United States	-0.109	Hungary	-0.375
		Japan	-0.008					Ireland	-0.031
		Poland	0.066					Netherlands	-0.047
		Slovakia	0.117					New Zealand	0.151
		Spain	-0.056					Norway	-0.058
								Slovenia	-0.262
								Switzerland	0.005

Table A.4: Finite Mixture Model II

	Dependent variable: γ			
	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Fixed Parameter:				
h_{t-1}	-1.03***	-1.03***	-1.03***	-1.03***
	(0.07)	(0.07)	(0.07)	(0.07)
Random Parameters:				
Intercept	0.79***	0.79***	0.77***	0.62***
	(0.14)	(0.17)	(0.19)	(0.14)
$n_t \times (1 - u_t)$	-0.17**	0.63***	-0.82***	0.20*
	(0.06)	(0.13)	(0.12)	(0.09)
$(1 - u_t)$	0.18	-0.02	1.35***	0.77***
	(0.11)	(0.17)	(0.21)	(0.11)
n_t	0.13	-0.37	0.39	-0.17
	(0.04)	(0.10)	(0.09)	(0.20)
Time Fixed Effects: ✓				
Observations	2039			
σ_u	0.403			

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table A.5: Country Classification, model in Table A.4

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
	Chile	Israel	Australia	Canada
	Denmark	Luxembourg	Austria	Czech Republic
	France	Republic of Korea	Belgium	Estonia
	Hungary		Netherlands	Finland
	Iceland		New Zealand	Germany
	Ireland		Portugal	Greece
	Italy		Switzerland	Japan
	Mexico			Slovakia
	Norway			Sweden
	Poland			United Kingdom
	Slovenia			United States
	Spain			
	Turkey			
Avg. $\partial\gamma_t/\partial n_t$	-0.037	0.184	-0.603	0.007
n_t	0.986	0.880	1.212	0.550
$1 - u_t$	0.632	0.615	0.653	0.560
h_t	0.744	0.422	1.230	0.657
GDP growth	0.029	0.024	0.043	0.028