amse
école d'économie d'aix-marseille
aix-marseille school of economics

Working Papers / Documents de travail

Optimal Population Growth as an Endogenous Discounting Problem: The Ramsey Case

Raouf Boucekkine
Blanca Martínez
José Ramón Ruiz-Tamarit

WP 2017 - Nr 31

Optimal population growth as an endogenous discounting problem: The Ramsey case*

R. Boucekkine† B. Martínez‡ J. R. Ruiz-Tamarit §

August 29, 2017.

Abstract

This paper revisits the optimal population size problem in a continuous time Ramsey setting with costly child rearing and both intergenerational and intertemporal altruism. The social welfare functions considered range from the Millian to the Benthamite. When population growth is endogenized, the associated optimal control problem involves an endogenous *effective* discount rate depending on past and current population growth rates, which makes preferences intertemporally dependent. We tackle this problem by using an appropriate maximum principle. Then we study the stationary solutions (balanced growth paths) and show the existence of two admissible solutions except in the Millian case. We prove that only one is optimal. Comparative statics and transitional dynamics are numerically derived in the general case.

Keywords: Optimal population size, Population ethics, Optimal growth, Endogenous discounting, Optimal demographic transitions

JEL classification: C61, C62, J1, O41.

*Authors acknowledge the support of the Belgian research programmes ARC on Sustainability, the financial support from the Spanish Ministerio de Ciencia e Innovación, Project ECO2013-48884-C3-1-P and from the Spanish Ministerio de Economía, Industria y Competitividad, Project ECO2016-76818-C3-3-P.

†Corresponding author. Aix-Marseille Univ., CNRS, EHESS, Centrale Marseille, AMSE and IMRA. Senior member, IUF. raouf.boucekkine@univ-amu.fr

‡Department of Economics, Universidad Complutense de Madrid, Spain and Instituto Complutense de Análisis Económico (ICAE). blmartin@ccee.ucm.es

§Department of Economic Analysis, Universitat de València, Spain, and IRES Department of Economics, Université Catholique de Louvain, Belgium. ramon.ruiz@uv.es

1 Introduction

There is a growing body of the economic and population ethics literatures concerned with the demographic dimension of the sustainable growth debate (see for example, Arrow et al., 2004). Clearly, if sustainability implies coping with the needs of current generations without compromising those of the far future generations (i.e. the so-called Brundtland criterion), the question of sustainable demographic paths, namely those which can achieve the latter intergenerational fairness goal, comes easily into the story. Of course this question can be easily connected to current hot environment-oriented issues: *ceteris paribus*, larger populations are likely to pollute more and to deplete more quickly natural resources, which is a further strong threat on sustainable development (see Boucekkine, Martinez and Ruiz-Tamarit, 2014). Here, we strictly stick to the intergenerational fairness problem outlined by Arrow et al. (2004), and abstract from the environmental ingredients.

The key question is indeed a basic and recurrent one in population ethics: what is the optimal population size? By which social welfare ordering can we argue that the current European or world population is suboptimal or not? Clearly this question can be asked with or without the global warming threat. As recently outlined by Dasgupta (2005), the question of optimal population size traces back to antiquity. For example, Plato concluded that the number of citizens in the ideal city-state is 5,040, arguing that it is divisible by every number up to ten and have as many as 59 divisors, which would allow for the population to "... suffice for purposes of war and every peacetime activity, all contracts for dealings, and for taxes and grants" (cited in Dasgupta, 2005).

The early related economic literature is due to Edgeworth (1925): he claimed that the use of total utilitarianism (that's the Benthamite social welfare function) is highly problematic as it leads to choose a bigger population size (compared for example to the Millian social welfare function or average utilitarianism) with quite lower standard of living. Some recent inspections into this issue have reached the same conclusion. In particular, Nerlove, Razin and Sadka (1985), who examined the robustness of Edgeworth's claim to parental altruism within a simple static model, found that the claim still holds when the utility function of adults is increasing in the number of children and/or the utility of children. Interestingly enough, the latter economic literature was contemporaneous (and probably sympathetic) to a masterpiece of the population ethics literature, the 1984 Parfit's *Reasons and Persons* book. In particular, Parfit explicitly attributed to total utilitarianism the same unpleasant implication as Edgeworth 60 years ago, and he called it a *repugnant conclusion*.

The use of dynamic frameworks to assess the *repugnant conclusion* and its correlates traces back essentially to the 90s. Intriguingly, the settings considered were rooted in the endogenous growth literature, a strongly rising stream at that time. Palivos and Yip (1993) and Razin and Yuen (1995) are two excellent representatives of this literature. In particular, Palivos and Yip showed that Edgeworth's claim cannot hold in the framework of endogenous growth driven by an AK production function. The determination of the optimal population growth rate relies on the following trade-off: on one hand, the utility function depends explicitly on the demographic growth rate; on the other, the latter induces the standard linear dilution effect on capital accumulation, and therefore on economic growth. Palivos and Yip proved that in such a case the Benthamite criterion leads to a smaller population size and a higher growth rate of the economy provided the intertemporal elasticity of substitution is lower than one. More recently, Boucekkine and Fabbri (2013) have examined a more general AK framework with endogenous demographic growth allowing for any type of correlation between demographic and economic growth at equilibrium. This is made possible by considering a large class of dilution functions: in particular, following Blanchet (1988) who proved that such functions are nonlinear when accounting for the age structure of capital, Boucekkine and Fabbri found that the *repugnant conclusion* would more easily arise under non-monotonic dilution schemes.

In Boucekkine and Fabbri (2013) and Palivos and Yip (1993), human populations do not play any role in the production side of the economy since the assumed production function is AK. Clearly, if humans do not produce, a strong pro-natalist ingredient is lost. What if the size of population matters in the production function as in neoclassical growth? Things should be much more involved. In the extreme case where the production function is AN (that's only human capital matters), Boucekkine, Fabbri and Gozzi (2014) show that the results depend strongly on the humans' life span. If the life span is large enough, Parfit's *repugnant conclusion* for total utilitarianism does not hold: even more, all individuals of all generations will receive the same consumption, and therefore will enjoy the same welfare. If life spans are small enough, even the Benthamite social welfare function would legitimate finite time extinction.

This paper is concerned with the much more essential Ramsey version of the problem, that's the production function is neoclassical and both capital and labor (or human capital) are production factors. We shall not incorporate the finite life span assumption into the story as it has been already deeply explored in Boucekkine, Fabbri and Gozzi. Individuals have infinite lives and there are decreasing returns with respect to labor in the production sector. Time is continuous. To our knowledge, very few papers have tackled the optimal population problem within this frame. Perhaps the most popular contribution along this line is Barro and Sala-i-Martin (2004), section 9.2.2. Essentially, the very vast majority of papers dealing with endogenous fertility use overlapping generations and discrete time. The seminal contribution to this line of research is Barro and Becker (1989) and their fertility choice model within a Ramsey structure. Typically, agents live two periods (childhood and adulthood), the utility of children enters linearly the utility of parents but the degree of altruism is a decreasing function of the number of children. Barro and Sala-i-Martin's 2004 model (BSM hereafter) can be seen, roughly speaking, as a continuous time formulation of the latter seminal model. Accordingly, the counterpart of the number of children variable is the continuous growth rate of population, just like in the models surveyed above on the optimal population problem.

There are two main drawbacks in the BSM. First of all, BSM do not consider the case where the induced social welfare function is either Benthamite or Millian since their model is initially derived from Barro and Becker (1993) where the degree of altruism is a decreasing function of the number of children. Bringing the latter assumption to the continuous time model does not allow to tackle an important element of the optimal population debate. This simply reflects the fact that Barro and his coauthors have a clearly distinct focus. The second drawback is technical: as any Ramsey model, BSM's is not tractable. Moreover, endogenizing demographic growth brings more nonlinearity into the problem, favoring multiplicity of stationary equilibria and the like (a fact acknowledged, although not rigorously studied, in Barro and Becker, 1989, pages 489-490). These issues are left in the dark in BSM.

In this paper, we take seriously the specific implications of endogenous population growth. When we endogenize population growth decisions connecting the fertility choice with economic variables, the Ramsey growth model experiences indeed a drastic change in its structure. The standard discounted optimal control problem assumes that the instantaneous utility function depends on contemporary variables alone, and that the intertemporal welfare function discounts utility stream at a fixed exponential rate. However, the simple modification to an hyperbolic discount function causes systematic changes in decisions which are responsible of a time inconsistency in intertemporal choices. The same problem of time inconsistency appears when the intertemporal non-separability of preferences comes from an endogenous and variable discount function. In our model, the endogeneity of exponential population growth at a variable rate transforms the standard optimal control problem with a constant rate of time preference, into a new and nontrivial dynamic optimization problem, and one has to be cautious in the application of the Pontryagin's maximum principle. This is because the induced non-constant *effective* discount rate becomes endogenous, which makes preferences intertemporally dependent. From the literature on endogenous discounting (see for example, Obstfeld, 1990, or more recently, Marin-Solano and Navas, 2009), we can implement a mathematical solution by introducing a new state variable representing the accumulated stock of impatience. Then, we can solve the transformed problem within the standard optimal control approach. The transformed is however far nontrivial as it involves a problem in higher dimension with a pure state constraint. We handle it using the appropriate approach described for example in Sethi and Thompson (2000). Furthermore, we study two sets of questions. One is related to the existence and uniqueness of stationary solutions (balanced growth paths): we show that two admissible steady states exist provided the social welfare function is not Millian; however, only one is proved to be optimal. Last but not least, we give some insight into the short term dynamics of the model. In particular, we numerically study the *optimal demographic transitions* in line with the typical imbalance effect analysis as designed in Boucekkine, Martinez and Ruiz-Tamarit (2008).

The article is organized as follows. Section 2 describes the model economy. Section 3 describes the optimal growth problem and applies an appropriate maximum principle to derive the set of necessary conditions. Section 4 analyses the long-run dynamics and characterize the balanced growth paths as described above, including comparative statics. Section 5 is numerical, it derives in particular some useful transitional dynamics. Section 6 concludes.

2 The model

The model economy is a one sector closed economy. Output is obtained according to a neoclassical production function depending on the technological level, the physical capital stock and labor input. The latter, under the assumption of a fixed relation between labor supply and population, will be identified with the population stock. The aggregate production function, which comes from the direct summation of the individual production functions for many identical firms, is

$$Y(t) = A(t)^{1-\beta} K(t)^{\beta} N(t)^{1-\beta}. \tag{1}$$

In this function technical progress is assumed Harrod-neutral. Technological level, denoted by A, is exogenous and evolves according to the differential equation

$$\dot{A}(t) = xA(t). \tag{2}$$

This is the standard law of motion for technology in Neoclassical growth theory, where it is assumed that technical progress arrives at a constant growth rate $x \geq 0$. The solution to the above equation implies that A increases monotonically according to the exponential form

$$A(t) = A_0 \exp(x(t - t_0)), \tag{3}$$

where is $A_0 = A(t_0) > 0$ is the initial technological level.

The economy is populated by many identical and infinitely lived agents. In this context, there is no point for differentiating between parents and children. Households face an infinite planning horizon, representing an immortal extended family where each member can be seen as a dynasty. Consequently, given that we focus on the link between demography and economic growth, trying to endogenize the demographic growth in a continuous time Ramsey model, we shall adapt the fertility conceptual schema from demographic theories to the requirements of our own model. From an aggregate point of view, we only have to deal with two demographic variables: population level and its variation. Population stock, denoted by N, is endogenously determined and it evolves according to the linear differential equation

$$\dot{N}(t) = n(t) N(t), \tag{4}$$

where the rate of population growth $n(t)$ is a control variable.

The initial population stock is $N(t_0) = N_0 > 0$, and we assume that $n(t) \geq 0 \ \forall t$. With respect to the individual preferences we assume that they are represented by a twice continuously differentiable instantaneous utility function, which depends positively on the current per capita consumption, and positively on the rate of population growth. The structure of our model allows for the existence of a long-run balanced growth path, defined as an allocation in which consumption per capita grows at a positive constant rate and the population growth rate is constant. We assume that the particular instantaneous utility function is of the form

$$U(c(t), n(t)) = \frac{c(t)^{1-\Phi} n(t)^{\epsilon(1-\Phi)}}{1-\Phi}. \tag{5}$$

In this function, the parameter $\Phi \left(\equiv \frac{c \cdot U_{cc}}{-U_c}\right)$ represents the inverse of the *conventional* intertemporal elasticity of substitution coefficient, which is constant and it is allowed to take values above or below unity, $0 < \Phi \lesseqgtr 1$; the parameter $\epsilon \left(\equiv \frac{U_n}{U_c}\frac{n}{c}\right)$ represents the weight of population changes in utility relative to the weight of consumption, and it is assumed positive but lower than one, $0 < \epsilon < 1$. According to the above parameter configuration we get $U_c > 0$, $U_n > 0$ and $U_{cc} < 0$, while we need $\Phi > \frac{\epsilon - 1}{\epsilon}$ for $U_{nn} < 0$. However, the latter parameter constraint always holds for $\Phi > 0$ and $\epsilon < 1$. To ensure the strict concavity of the instantaneous utility function we assume that $\Phi > \frac{\epsilon}{1+\epsilon}$.[1]

In the present framework where population is endogenous because the stock N depends on the population growth rate n, which is currently decided by economic agents, we omit any population stock effect in the representation of individuals' preferences. We consider that people do not care about the population size N but only about the per capita number of offspring n. That is, the stock effect is not modeled entering the instantaneous utility function as a direct argument, but affecting other variables and functions in the model.

Finally, we introduce the aggregate resources constraint according to which output may be devoted to consumption, to capital accumulation or to rear population changes. Strictly speaking, here there are no parents rearing children but people looking after people. For the sake of simplicity we do not consider capital depreciation. Hence, net investment equals gross

[1] This parameter constraint is a sufficient condition for the determinant of the Hessian matrix to be positive, but also implies that $\Phi > \frac{\epsilon - 1}{\epsilon}$. Consequently, the Hessian matrix is negative definite, which corresponds to the standard sufficient condition for the utility function to be strictly concave.

investment and the capital stock is governed by the differential equation

$$C(t) + \dot{K}(t) + bn(t)K(t) = Y(t). \tag{6}$$

The initial capital stock is $K(t_0) = K_0 > 0$. Adapting from Barro and Sala-i-Martín (2004), we assume that the per capita rearing cost is $b_0 + b\frac{K(t)}{N(t)}$, where $b_0 \geqslant 0$ and $b \geqslant 0$. This cost includes either purchases of market goods and services or the opportunity cost of time devoted to population rearing. Then, if we consider the real number representing the change in population size, total resources allocated to them are $\left(b_0 + b\frac{K(t)}{N(t)}\right)\dot{N}(t) = b_0 n(t) N(t) + bn(t) K(t)$. To obtain (6) we have simplified by setting $b_0 = 0$.

3 The optimal growth problem

In the optimal growth problem, the benevolent planner has to consider the effect of population size on social welfare. In this setting, given that we are not particularly interested in the case $\Phi \to 1$ but rather in the most empirically relevant case in which $\Phi > 1$, we define the social welfare (which is the planner's objective function) as

$$W = \int_{t_0}^{+\infty} \frac{c^{1-\Phi} n^{\epsilon(1-\Phi)}}{1-\Phi} N^{\lambda} e^{-\rho(t-t_0)} dt \tag{7}$$

Parameter ρ is the positive social rate of discount or time preference. Parameter $\lambda \in [0,1]$ contributes to specify social preferences, which are represented using a Millian, an intermediate, or a Benthamite intertemporal welfare function. In one extreme, when $\lambda = 0$ (average utilitarianism), the central planner maximizes per capita utility (average utility of consumption per capita). In the other, when $\lambda = 1$ (classical utilitarianism), the central planner maximizes total utility (the addition across total population of utilities of per capita consumption).[2]

[2]The literature differentiates between two types of altruism depending on the two parameters ρ and λ. The first one is *intertemporal* altruism and depends on the discount rate applied to future population utility. The second one is *intergenerational* altruism and depends on the number of individuals which is taken into account each period. In particular, for representative and infinitely lived agent models, parameter λ controls for the degree of altruism towards total population including future generations. When agents are selfish the central planner maximizes W under $\lambda = 0$, and population size has no direct effect on the intertemporal utility. Instead,

The central planner's problem consists then in choosing the sequence $\{c(t), n(t), t \geq t_0\}$ that solves the optimization problem

$$\max_{\{K,N,c,n\}} \quad (7) \qquad s.t. \quad (1), (2), (4), \text{and } (6), \tag{8}$$

$$\text{given } A(t_0) = A_0 > 0,\ K(t_0) = K_0 > 0, \text{ and } N(t_0) = N_0 > 0.$$

Before solving the dynamic problem, we define the variables $\widetilde{c}(t) = \frac{c(t)}{A(t)}$ and $\widetilde{k}(t) = \frac{K(t)}{A(t)N(t)}$, which allow to write in per capita efficiency terms either the integrand and the dynamic resources constraint, $\dot{\widetilde{k}}(t) = \widetilde{k}(t)^{\beta} - \widetilde{c}(t) - (x + (1+b)\,n(t))\,\widetilde{k}(t)$.

In this context, solving equation (4) we get the following expression for the endogenous population size

$$N(t) = N_0 \exp\left(\int_{t_0}^{t} n(\tau)\,d\tau\right). \tag{9}$$

Hence, expressions (3) and (9) allow for the transformation

$$A(t)^{1-\Phi} N(t)^{\lambda} e^{-\rho(t-t_0)} = A_0^{1-\Phi} N_0^{\lambda} \exp\left(-\int_{t_0}^{t} (\rho - x(1-\Phi) - \lambda n(\tau))\,d\tau\right). \tag{10}$$

This term plays the role of a **variable discount factor** which also depends on past and current rates of population growth. So, adapting from Obstfeld (1990), Palivos et al. (1997), Ayong and Schubert (2007), and Schumacher (2011) who analyze optimal control problems extended to an **endogenous discounting** framework, we can define the accumulated stock of impatience as the non-negative

$$\Delta(t) = \int_{t_0}^{t} (\rho - x(1-\Phi) - \lambda n(\tau))\,d\tau = (\rho - x(1-\Phi))(t-t_0) - \lambda \int_{t_0}^{t} n(\tau)\,d\tau \geq 0, \tag{11}$$

where for obvious reasons $\Delta(t_0) = \Delta_0 = 0$.[3] This is a new state variable for which the motion equation reads

$$\dot{\Delta}(t) = \rho - x(1-\Phi) - \lambda n(t) \equiv \Theta(n(t)) \gtreqless 0. \tag{12}$$

when agents are altruistic the central planner maximizes W under $\lambda = 1$, and the intertemporal utility function includes total population as a determinant.

[3]The non-negativity of Δ might be replaced by a weaker constraint in line with Assumption 4 in Palivos et al. (1997) given the goal of a well-defined optimization problem.

That is, the **effective discount rate** is endogenous because of the endogeneity of population growth rates. Further, impatience is inversely proportional to the number of offspring, $\Theta'(\cdot) = -\lambda \leqslant 0$ and $\Theta''(\cdot) = 0$, except for the Millian case in which we recover the standard constant discount factor. This may happen because as population grows agents care more about the future, given that the increased population represents an investment having a positive impact on future welfare. The negative effect of population growth on the effective discount rate is greater as higher is the intergenerational altruism. Moreover, from (12) we get

$$\Theta(\cdot) \begin{array}{lll} > 0 & & 0 < n(t) < \frac{\rho - x(1-\Phi)}{\lambda} \\ = 0 & \text{whenever} & n(t) = \frac{\rho - x(1-\Phi)}{\lambda} \\ < 0 & & \frac{\rho - x(1-\Phi)}{\lambda} < n(t). \end{array} \tag{13}$$

A direct consequence of definition (11) is that the solution trajectory for population size may be rewritten as

$$N(t) = N_0 \exp\left(\frac{(\rho - x(1-\Phi))(t-t_0) - \Delta(t)}{\lambda}\right). \tag{14}$$

Overall, after introducing the new variable Δ, the intertemporal optimization problem becomes an **autonomous problem** without discounting and infinite planning horizon. According to Pittel (2002) and based on Marin-Solano and Navas (2009), due to the effective non-constant discount rate, the Pontryagin's maximum principle cannot be applied directly because intertemporally dependent preferences can create a time-consistency problem. We need this state variable to solve the problem within the standard optimal control approach, where it is no use distinguishing between present value and current value specifications. Here we follow Seierstad and Sydsaeter (1987), Chiang (1992), and Sethi and Thompson (2000).

Then, we can write the Hamiltonian function

$$\underset{\{\widetilde{c},n,q,\widetilde{k},\upsilon,\Delta\}}{H} = \frac{\widetilde{c}^{1-\Phi} n^{\epsilon(1-\Phi)} A_0^{1-\Phi} N_0^{\lambda} e^{-\Delta}}{1-\Phi} + q\left(\widetilde{k}^{\beta} - \widetilde{c} - (x + (1+b)n)\widetilde{k}\right) + \upsilon(\rho - x(1-\Phi) - \lambda n). \tag{15}$$

Here $q \geqslant 0$ and $\upsilon \geqslant 0$ are the co-states for $\widetilde{k}$ and Δ respectively.[4] If we ignore the constraints

[4]The multiplier ν represents the marginal shadow value of relaxing the constraint (12). That is, the shadow price of (the accumulated stock of) impatience. As previously said, because of empirical reasons, we focus on the case in which $\Phi > 1$ and, consequently, $U(\widetilde{c}, n, \Delta) < 0$. Here we choose to write H in its canonical form with a positive sign preceding the (positive) multiplier ν.

involving only control variables, $\widetilde{c} \geqslant 0$ and $n \geqslant 0$, the first order necessary conditions arising from Pontryagin's Maximum principle are

$$q = \widetilde{c}^{-\Phi} n^{\epsilon(1-\Phi)} A_0^{1-\Phi} N_0^{\lambda} e^{-\Delta}, \tag{16}$$

$$q\,(1+b)\,\widetilde{k} = \epsilon \widetilde{c}^{1-\Phi} n^{\epsilon(1-\Phi)-1} A_0^{1-\Phi} N_0^{\lambda} e^{-\Delta} - \upsilon\lambda, \tag{17}$$

$$\overset{\bullet}{\widetilde{k}} = \widetilde{k}^{\beta} - \widetilde{c} - (x + (1+b)\,n)\,\widetilde{k}, \tag{18}$$

$$\overset{\bullet}{q} = (x + (1+b)\,n)\,q - q\beta\widetilde{k}^{\beta-1}, \tag{19}$$

$$\overset{\bullet}{\Delta} = \rho - x\,(1-\Phi) - \lambda n, \tag{20}$$

$$\overset{\bullet}{\upsilon} = \frac{\widetilde{c}^{1-\Phi} n^{\epsilon(1-\Phi)}}{1-\Phi} A_0^{1-\Phi} N_0^{\lambda} e^{-\Delta}, \tag{21}$$

Finally, we also need the initial conditions A_0, N_0, K_0, $\widetilde{k}_0 = \frac{K_0}{A_0 N_0}$, and Δ_0, as well as the transversality conditions

$$\lim_{t\to+\infty} H\,(t) = 0, \tag{22}$$

$$\lim_{t\to+\infty} q\,(t) \geqslant 0 \quad \text{and} \quad \lim_{t\to+\infty} q\,(t)\,\widetilde{k}\,(t) = 0, \tag{23}$$

$$\lim_{t\to+\infty} \upsilon\,(t) \geqslant 0 \quad \text{and} \quad \lim_{t\to+\infty} \upsilon\,(t)\,\Delta\,(t) = 0. \tag{24}$$

The necessary conditions in the present dynamic optimization problem are also sufficient for a maximum because the Hamiltonian function satisfies the required concavity conditions [see Appendix A]. Looking at (15) we can also check that H is autonomous. Consequently, along the optimal path H is constant and, given that our transversality condition (22) says that H eventually converges to zero, we conclude that

$$H = 0 \quad \forall t. \tag{25}$$

Next, given the solution to equation (20) as shown in (11), which assumes $\Delta(t) \geq 0$, as well as the finite values of the strictly concave function $\frac{\widetilde{c}^{1-\Phi} n^{\epsilon(1-\Phi)}}{1-\Phi} A_0^{1-\Phi} N_0^{\lambda}$, we can integrate (21) to obtain the expression[5]

$$\upsilon(t) = \upsilon(t_0) + \int_{t_0}^{t} \frac{\widetilde{c}^{1-\Phi} + n^{\epsilon(1-\Phi)}}{1-\Phi} A_0^{1-\Phi} N_0^{\lambda} e^{-\Delta} d\tau.$$

Then, substituting into the transversality condition (24) we get

$$\left(\upsilon(t_0) + \int_{t_0}^{+\infty} \frac{\widetilde{c}^{1-\Phi} n^{\epsilon(1-\Phi)} A_0^{1-\Phi} N_0^{\lambda} e^{-\Delta}}{1-\Phi} d\tau\right) \int_{t_0}^{+\infty} \Theta(n(\tau))\, d\tau = 0.$$

This condition holds if and only if, for any $\lim\limits_{t\to+\infty} \Delta(t)$ different from zero,

$$\upsilon(t_0) = \int_{t_0}^{+\infty} -\frac{\widetilde{c}^{1-\Phi} n^{\epsilon(1-\Phi)}}{1-\Phi} A_0^{1-\Phi} N_0^{\lambda} e^{-\Delta} d\tau.$$

Consequently, we conclude that the multiplier ν takes the value

$$\upsilon(t) = \int_{t}^{+\infty} -\frac{\widetilde{c}^{1-\Phi} n^{\epsilon(1-\Phi)}}{1-\Phi} A_0^{1-\Phi} N_0^{\lambda} e^{-\Delta} d\tau \quad \forall t \geq t_0. \tag{26}$$

Then, the first order conditions reduce to (16)-(19) together with (11), (26), the transversality conditions

$$0 = \frac{\widetilde{c}^{1-\Phi} n^{\epsilon(1-\Phi)}}{1-\Phi} A_0^{1-\Phi} N_0^{\lambda} e^{-\Delta} + q\dot{\widetilde{k}} + \upsilon(\rho - x(1-\Phi) - \lambda n), \tag{27}$$

$$\lim_{t\to+\infty} q(t) \geqslant 0 \quad \text{and} \quad \lim_{t\to+\infty} q(t)\, \widetilde{k}(t) = 0, \tag{28}$$

$$0 < \left| \int_{t_0}^{+\infty} (\rho - x(1-\Phi) - \lambda n(t))\, dt \right|, \tag{29}$$

and the initial conditions A_0, N_0, and $\widetilde{k}_0 = \frac{K_0}{A_0 N_0}$.

Consider now equations (16) and (17). As we have seen, gross product may be allocated to consumption, investment, or offspring. On the margin, goods must be equally valuable if they are consumed or accumulated as new physical capital. Namely, the marginal utility of

[5] The convergence of the objective integral (7) is shown in Nairay (1984), even for the case in which $\Phi < 1$, by proving that such a limit value exists and is finite.

consumption today must be equal to the current shadow price qe^{Δ} of physical capital (consumption tomorrow). Moreover, at equilibrium the marginal utility of population growth must be equal to the sum of the current implicit value of the *full* (rearing and dilution) marginal cost of increasing population $qe^{\Delta}(1+b)\widetilde{k}$, plus the current shadow value of the accumulated impatience scaled by the weight of the increased population in social welfare $\upsilon e^{\Delta}\lambda$. Taken together, these equations give the tangency condition

$$\frac{\epsilon\widetilde{c}}{n} = (1+b)\widetilde{k} + \lambda\frac{\upsilon}{q}, \tag{30}$$

which describes the optimal allocation between consumption goods and children. The marginal rate of substitution between n and c must be equal to the *full* marginal cost of increasing population plus the degree of intergenerational altruism times the relative (shadow) prices of impatience and physical capital.

Moreover, differentiating (16) with respect to time and substituting (19) and (20), we get the corresponding adapted version of the Ramsey rule,

$$\frac{\dot{\widetilde{c}}}{\widetilde{c}} = \frac{1}{\Phi}\left(\beta\widetilde{k}^{\beta-1} - (x+n+bn) - (\rho - x(1-\Phi) - \lambda n) + \epsilon(1-\Phi)\frac{\dot{n}}{n}\right), \tag{31}$$

where $\frac{1}{\Phi} = \frac{-U_c}{cU_{cc}}$, $\beta\widetilde{k}^{\beta-1} = f'\left(\widetilde{k}\right)$, $\rho - x(1-\Phi) - \lambda n = \Theta(n)$, and $\epsilon(1-\Phi) = \frac{nU_{cn}}{U_c}$. The growth rate of per capita consumption in efficiency units depends: *i*) positively on the net marginal productivity[6] of per capita capital in efficiency units; *ii*) negatively on the effective discount rate; as well as *iii*) on the rate of change of the population growth rate. For $\Phi > 1$ we get $U_{cn} < 0$, which implies that c and n are gross substitutes in utility. In this case their corresponding rates of growth are inversely related to each other. All the three above arguments are endogenous because of the endogeneity of the rate of population growth.

Finally, for the purpose of facilitating comparison with the exogenous discount rate model, the above Ramsey rule may be written as

[6] Even if we do not consider capital depreciation, the exogenous technical progress and the increase in the population size are the cause of a marginal dilution effect which adds to the corresponding marginal rearing cost to determine the net marginal productivity.

$$\frac{\overset{\bullet}{\widetilde{c}}}{\widetilde{c}} = \frac{-U_c}{cU_{cc}} \left(f'\left(\widetilde{k}\right) - \rho - x\Phi - (1+b-\lambda)\, n + \frac{U_{cn}}{U_c}\overset{\bullet}{n} \right). \tag{32}$$

If the net return to capital exceeds the effective discount rate, agents would decide to invest now in physical capital leaving less resources to consumption and child rearing today. In the standard model with exogenous discount and constant rate of population growth, this would suffice to explain an increasing per capita consumption. However, in our model the expected increasing resources may allow for different combinations. Obviously, the additional future resources are available for the simultaneous growth of per capita consumption and population size, but this is not the only possibility given that the above expression still admits an increasing consumption with a decreasing population, or a decreasing consumption with an increasing population. In any case, all of them will produce more future welfare.

On the other hand, differentiating (17) with respect to time and substituting (18)-(21), we get the corresponding adapted version of the Meade rule (Dasgupta, 1969; Constantinides, 1988). This is a rule for the optimal population growth, which comes from the balance between the gains and losses due to the introduction of a new member into society.

$$\frac{\overset{\bullet}{n}}{n} = \frac{n}{\epsilon\,(1-\Phi)-1}\frac{(1+b)}{\epsilon\widetilde{c}}\left((1-\beta)\,\widetilde{k}^{\beta} - \widetilde{c}\right) + \frac{n}{\epsilon\,(1-\Phi)-1}\frac{\lambda}{\epsilon\,(1-\Phi)}$$

$$+\frac{\rho - x\,(1-\Phi) - \lambda n}{\epsilon\,(1-\Phi)-1} - \frac{(1-\Phi)}{\epsilon\,(1-\Phi)-1}\frac{\overset{\bullet}{\widetilde{c}}}{\widetilde{c}}, \tag{33}$$

where $\frac{1}{\epsilon(1-\Phi)-1} = \frac{U_c}{nU_{cn}-U_c} < 0$, $(1-\beta)\,\widetilde{k}^{\beta} = f\left(\widetilde{k}\right) - \widetilde{k}f'\left(\widetilde{k}\right)$, $\rho - x\,(1-\Phi) - \lambda n = \Theta\,(n)$, and $1-\Phi = \frac{U_c + cU_{cc}}{U_c} < 0$. The change in the population growth rate depends: *i*) negatively on the difference between the marginal product of an additional person and his consumption measured in efficiency units; *ii*) positively on the degree of intergenerational altruism; *iii*) negatively on the effective discount rate; and *iv*) negatively on the growth rate of per capita consumption in efficiency units. It is worth noticing that again the changes in c and n are inversely related to each other.

Writing in terms of a more general specification, the Meade rule takes the following form

$$\overset{\bullet}{n} = \frac{n\left(U_c\right)^2}{nU_{cn}U_n - U_cU_n} A\left(1+b\right)\left(f\left(\widetilde{k}\right) - \widetilde{k}f'\left(\widetilde{k}\right) - \widetilde{c}\right) + \frac{n\left(U_c\right)^2}{n\left(U_{cn}\right)^2 - U_cU_{cn}}\lambda$$

$$+\frac{nU_c}{nU_{cn} - U_c}\Theta\left(n\right) - \frac{n\left(U_c + cU_{cc}\right)}{nU_{cn} - U_c}\frac{\overset{\bullet}{\widetilde{c}}}{\widetilde{c}}. \tag{34}$$

According to our model, when the per capita consumption exceeds the marginal product of labor, there is an incentive for increasing the rate of population growth as well as the per capita consumption level. Moreover, for a given degree of altruism, if the effective discount rate is negative, the above incentive will be stronger.

4 The balanced growth path and comparative statics

In the previous section we have solved the model for any exogenous and constant rate of technical progress. However, in the present section we assume, for the sake of simplicity, a constant technological level, that is x=0. Otherwise, we could not find most of the analytical long-run results associated with the balanced growth path. Hereafter we characterize the long term equilibria identifying the balanced growth path along which $\widetilde{c}$ and n are constant. In steady state $\overset{\bullet}{\widetilde{k}} = 0$, but given the transversality condition (27) we observe that $\overset{\bullet}{\Delta} = \rho - \lambda n^* \equiv \Theta\left(n^*\right) > 0$, which is compatible with the constraint (29). This implies, from (11), that

$$\Delta^* = \left(\rho - \lambda n^*\right)\left(t - t_0\right), \tag{35}$$

which makes q non-stationary according to (16). Consequently, we introduce a new variable $p = qe^{\Delta}$ and, hence, $\frac{\overset{\bullet}{p}}{p} = \frac{\overset{\bullet}{q}}{q} + \overset{\bullet}{\Delta}$. Now, in steady state $\overset{\bullet}{\widetilde{k}} = \overset{\bullet}{p} = 0$ and equations (16)-(19) can be written as

$$p^* = \widetilde{c}^{*^{-\Phi}} n^{*^{\epsilon(1-\Phi)}} A_0^{1-\Phi} N_0^{\lambda}, \tag{36}$$

$$\left(1+b\right)\widetilde{k}^* = \frac{\epsilon\widetilde{c}^*}{n^*} - \left(\frac{\upsilon^*}{q^*}\right)\lambda, \tag{37}$$

$$\widetilde{k}^{*^{\beta}} = \widetilde{c}^* + \left(1+b\right)n^*\widetilde{k}^*, \tag{38}$$

$$\beta \widetilde{k}^{*\beta-1} = \rho + (1+b-\lambda)\, n^*, \tag{39}$$

Moreover, from (26) or (27) we get

$$v^* = \frac{-\widetilde{c}^{*1-\Phi} n^{*\epsilon(1-\Phi)} A_0^{1-\Phi} N_0^{\lambda} e^{-\Delta^*}}{(1-\Phi)(\rho-\lambda n^*)} = \frac{-\widetilde{c}^* q^*}{(1-\Phi)(\rho-\lambda n^*)}. \tag{40}$$

These expressions allow us to directly obtain the stationary values $\widetilde{c}^*$, n^*, and $\widetilde{k}^*$ corresponding to the balanced growth path and, by substitution, all the remaining endogenous variables of the model. After some algebraic manipulations we get

$$\widetilde{k}^* = \left(\frac{\beta}{\rho + (1+b-\lambda)\, n^*}\right)^{\frac{1}{1-\beta}}, \tag{41}$$

$$\widetilde{c}^* = \frac{\beta^{\frac{\beta}{1-\beta}} \left(\rho + ((1-\beta)(1+b)-\lambda)\, n^*\right)}{\left(\rho + (1+b-\lambda)\, n^*\right)^{\frac{1}{1-\beta}}}, \tag{42}$$

$$\widetilde{y}^* = \left(\frac{\beta}{\rho + (1+b-\lambda)\, n^*}\right)^{\frac{\beta}{1-\beta}}, \tag{43}$$

where n^* corresponds to the roots of the second degree polynomial equation with real coefficients

$$\Psi_a n^{*2} + \Psi_b n^* + \Psi_c = 0, \tag{44}$$

These coefficients depend on the structural parameters of the model in the following way

$$\Psi_a(\lambda, \Phi, b, \beta, \epsilon) = \lambda\left(((1-\beta)(1+b)-\lambda)(\epsilon(1-\Phi)-1) - \beta(1-\Phi)(1+b)\right), \tag{45}$$

$$\Psi_b(\lambda, \Phi, b, \beta, \epsilon, \rho) = \rho(1-\Phi)(1+b)(\beta - \epsilon(1-\beta)) + \lambda(\epsilon(1-\Phi)2\rho - \rho), \tag{46}$$

$$\Psi_c(\Phi, \epsilon, \rho) = -\rho^2 \epsilon (1-\Phi). \tag{47}$$

The roots are:

$$n_1^* = \frac{-\Psi_b + \sqrt{D}}{2\psi_a}, \tag{48}$$

$$n_2^* = \frac{-\Psi_b - \sqrt{D}}{2\psi_a}, \tag{49}$$

where $D = \Psi_b^2 - 4\Psi_a\Psi_c$ is the discriminant. In case $D \geq 0$ the roots $n_1^\star$ and $n_2^\star$ are both real. Moreover, from (40), (26), and (13) we get $\Theta(n^*) > 0$, that is

$$0 < n_i^* < \frac{\rho}{\lambda} \qquad \forall i = \{1, 2\}. \tag{50}$$

4.1 The case of the Millian welfare function: $\lambda = 0$

We first analyze the Millian case. When the central planner maximizes per capita utility (average utilitarianism), $\lambda = 0$ and population size has no direct effect on the intertemporal utility. It is easily checked that since $\Psi_a = 0$ for $\lambda = 0$, equation (44) has a unique solution given by,

$$n^* = \frac{\epsilon\rho}{(1+b)(\beta - \epsilon(1-\beta))} \tag{51}$$

which is positive for $\beta > \epsilon(1-\beta)$. Notice that under the Millian case equation (50) is always checked.

Substituting (51) in (41)-(43) we obtain the stationary values for $\widetilde{k}^*$, $\widetilde{c}^*$ and $\widetilde{y}^*$,

$$\widetilde{k}^* = \left(\frac{\beta - \epsilon(1-\beta))}{\rho(1+\epsilon)}\right)^{\frac{1}{1-\beta}}, \tag{52}$$

$$\widetilde{c}^* = \frac{-\rho(\epsilon - \beta(1+\epsilon))^{\frac{\beta}{1-\beta}}}{(-\rho(1+\epsilon))^{\frac{1}{1-\beta}}}, \tag{53}$$

$$\widetilde{y}^* = \left(\frac{\beta - \epsilon(1-\beta))}{\rho(1+\epsilon)}\right)^{\frac{\beta}{1-\beta}}, \tag{54}$$

Proposition 1 summarizes the associated comparative statics findings:

Proposition 1. When $\beta > \epsilon(1-\beta)$:

$\frac{\partial n^*}{\partial b} < 0$, $\frac{\partial n^*}{\partial \epsilon} > 0$, $\frac{\partial n^*}{\partial \rho} > 0$, $\frac{\partial n^*}{\partial \beta} < 0$, $\frac{\partial n^*}{\partial \Phi} = 0$

$\frac{\partial \widetilde{y}^*}{\partial b} = 0$, $\frac{\partial \widetilde{y}^*}{\partial \epsilon} < 0$, $\frac{\partial \widetilde{y}^*}{\partial \rho} < 0$, $\frac{\partial \widetilde{y}^*}{\partial \beta} > 0$, $\frac{\partial \widetilde{y}^*}{\partial \Phi} = 0$

Proposition 1 is trivially checked by taking the partial derivative with respect to the corresponding parameter in (51) and (54). In general, as we can see from the above two sets of partial derivative signs, the optimal long-run rate of population growth and the long-run per capita income level (measured in efficiency terms) are inversely correlated.

In particular, we can identify the following parameter-variable relationships. First, recall that b represents the opportunity cost of parental time devoted to child rearing. Then, an economy where parents experience a higher cost of offspring will optimally choose in the long-run a lower rate of population growth. Note that in equation (43) we can observe two different effects of a change in the per capita rearing costs on the per capita income level. First, an increase in b directly reduces the resources devoted to capital accumulation, which implies a lower long-run level of $\widetilde{y}^*$. Moreover, a higher b reduces the optimal population growth rate, which has an indirect positive effect on $\widetilde{y}^*$. These two effects are of opposite sign and, only in the Millian case, exactly compensate each other. Consequently, the per capita income level is independent of b.

Second, recall that ϵ represents the weight of children in utility relative to consumption. Then, a society with higher preference for children will optimally choose in the long-run a higher population rate of growth, and will experience a lower per capita income level. Moreover, recall that an economy showing a low ρ represents a patient society. Then, we find that in the long-run an impatient society will optimally choose a higher population rate of growth, and will reach a lower level of per capita income.

On the other hand, an economy with higher β is an economy with a technology implying a higher elasticity of output with respect to capital (higher capital share). Then, in the long-run, this economy will optimally choose a lower population rate of growth, and will have the benefit of a higher per capita income level. Finally, we observe that in the Millian case the inverse of the intertemporal elasticity of substitution in consumption, Φ, has no effect on the long run population growth rate or the level of income per capita.

4.2 The case $\lambda \neq 0$

When $\lambda \neq 0$, two distinct balanced growth paths emerge. We show that both are admissible in the sense that the two associated demographic growth rates are positive.

Proposition 2. Under the parameter constraints $0 < \lambda \leqslant 1$, $\Phi > 1$, and $\epsilon < \frac{\beta}{1-\beta}$, which imply $\Psi_c > 0$ and $\Psi_b < 0$, if $\Psi_a > 0$ then we get two real positive values n_1^* and n_2^*, which are different as long as $D > 0$.

Proof. We consider the most empirically relevant case in which $\Phi > 1$ and assume $\epsilon < \frac{\beta}{1-\beta}$. Under these assumptions we get $\Psi_c > 0$ and $\Psi_b < 0$. Given $D \geqslant 0$, if $\Psi_a > 0$, we have $\sqrt{D} < |\Psi_b|$, and we can express $\sqrt{D}$ as $\sqrt{D} = |\Psi_b| - \theta$, with $\theta > 0$. Taking into account all the above, it is straightforward to check that $n_1^\star = \frac{2|\Psi_b|+\theta}{2\Psi_a} > 0$ and $n_2^\star = \frac{\theta}{2\Psi_a} > 0$. ■

Remark 1. From Proposition 2 and using the expressions in equations (48) and (49) we get

$$n_1^* > n_2^* > 0, \tag{55}$$

$$2\Psi_a n_1^* + \Psi_b = \sqrt{D} > 0, \tag{56}$$

$$2\Psi_a n_2^* + \Psi_b = -\sqrt{D} < 0. \tag{57}$$

Multiplicity of balanced growth paths cannot be a surprise in a model with endogenous fertility. In the seminal Barro and Becker 1989 discrete time OLG model, multiplicity is possible in the case where the cost of rearing children is large enough. In our model, the existence of two distinct solutions is generated under much milder parametric assumptions not related to the cost of rearing children. This said, the optimality analysis of the steady state solutions in our model does allow to eliminate one of the two candidates, as demonstrated here below.

Proposition 3. Given the parameter constraints and the results shown in Proposition 2 and Remark 1, the two real positive values n_1^* and n_2^* are separated from each other by the root's limiting upper-bound, implying that condition (50) does not hold for both. That is, we get

$$0 < n_2^* < \frac{\rho}{\lambda} < n_1^*. \tag{58}$$

Proof. We consider the different combinations ordering the upper-bound and the two roots, and we conclude that only one of such orderings is feasible because it is the only that requires a compatible relationship between structural parameters.

First, $\frac{\rho}{\lambda} > n_1^* > n_2^*$. That is $\frac{\rho}{\lambda} > \frac{-\Psi_b+\sqrt{D}}{2\Psi_a}$, or $\lambda\sqrt{D} < 2\rho\Psi_a + \lambda\Psi_b$. Then, if $2\rho\Psi_a + \lambda\Psi_b \leqslant 0$ the previous inequality is incompatible because $D > 0$. Alternatively, if $2\rho\Psi_a + \lambda\Psi_b > 0$ we can square the two sides of the inequality getting the new inequality $-\lambda^2\Psi_c < \rho^2\Psi_a + \rho\lambda\Psi_b$. Using equations (45), (46), and (47) to transform into a constraint between structural parameters alone we get $0 < -\lambda(1-\beta)(1+b)$, which is incompatible.

Second, $n_1^* > n_2^* > \frac{\rho}{\lambda}$. That is $\frac{-\Psi_b-\sqrt{D}}{2\Psi_a} > \frac{\rho}{\lambda}$, or $-\lambda\sqrt{D} > 2\rho\Psi_a + \lambda\Psi_b$. Then, if $2\rho\Psi_a + \lambda\Psi_b \geqslant 0$ the previous inequality is incompatible because $D > 0$. Alternatively, if $2\rho\Psi_a + \lambda\Psi_b < 0$ we can square the two sides of the inequality getting the new inequality $-\lambda^2\Psi_c < \rho^2\Psi_a + \rho\lambda\Psi_b$. Using equations (45), (46), and (47) to transform into a constraint between structural parameters alone we get $0 < -\lambda(1-\beta)(1+b)$, which is incompatible.

Therefore, $n_2^* < \frac{\rho}{\lambda} < n_1^*$ is the only case which is compatible with the signs of the coefficients Ψ_a, Ψ_b, and Ψ_c. ■

Remark 2. Given that $\lim\limits_{\lambda\to 0}\Psi_a = 0$, $\lim\limits_{\lambda\to 0}\Psi_b = \rho(1-\Phi)(1+b)(\beta-\epsilon(1-\beta)) < 0$, $\lim\limits_{\lambda\to 0}\Psi_c = -\rho^2\epsilon(1-\Phi) > 0$, and consequently $\lim\limits_{\lambda\to 0}\sqrt{D} = \left|\lim\limits_{\lambda\to 0}\Psi_b\right| = -\rho(1-\Phi)(1+b)(\beta-\epsilon(1-\beta)) > 0$, we get

$$\lim_{\lambda\to 0} n_2^* = \lim_{\lambda\to 0}\frac{-\Psi_b-\sqrt{D}}{2\Psi_a} = \frac{\epsilon\rho}{(1+b)(\beta-\epsilon(1-\beta))} > 0. \tag{59}$$

which corresponds to the selfish case (see equation (51)).

Comparative statics results

Unfortunately, when $\lambda \neq 0$, the model becomes much more complex analytically speaking. Recall that in this case the endogenous discounting is active and preferences become intertemporally related. The comparative statics of the optimal steady state population growth rate, n_2^* become intractable in general. To get an immediate idea about it, let us consider the vector of parameters $\Omega = (\lambda, \Phi, b, \beta, \epsilon, \rho)$. Then, by successive differentiation of equation (44) with respect to the components of such a vector we get the general formula:

$$\frac{\partial n_2^*}{\partial \Omega} = \frac{-\left(n_2^*\right)^2 \frac{\partial \Psi_a}{\partial \Omega} - n_2^* \frac{\partial \Psi_b}{\partial \Omega} - \frac{\partial \Psi_c}{\partial \Omega}}{2\Psi_a n_2^* + \Psi_b}. \tag{60}$$

Therefore, even though one can sign the terms $\frac{\partial \Psi_i}{\partial \Omega}$ where $i \in \{a, b, c\}$, which is not always the case indeed, this might be unlikely to do the job. We can establish the following partial results

$$\frac{\partial \Psi_a}{\partial \lambda} = \Psi_a - (\epsilon (1 - \Phi) - 1) \lambda > 0,$$

$$\frac{\partial \Psi_a}{\partial \Phi} = \lambda (-\epsilon ((1 - \beta)(1 + b) - \lambda) + \beta (1 + b)) > 0,$$

$$\frac{\partial \Psi_a}{\partial b} = \lambda ((1 - \beta)(\epsilon (1 - \Phi) - 1) - \beta (1 - \Phi)) \lesseqqgtr 0 \text{ depending on whether } \beta\Phi \lesseqqgtr 1,$$

$$\frac{\partial \Psi_a}{\partial \beta} = -\lambda ((1 - \beta)(\epsilon (1 - \Phi) - 1) + (1 - \Phi)(1 + b)) > 0,$$

$$\frac{\partial \Psi_a}{\partial \epsilon} = \lambda ((1 - \beta)(1 + b) - \lambda)(1 - \Phi) < 0, \text{ because we assume } (1 - \beta)(1 + b) - 2\lambda > 0,$$

$$\frac{\partial \Psi_b}{\partial \lambda} = \epsilon (1 - \Phi) 2\rho - \rho < 0,$$

$$\frac{\partial \Psi_b}{\partial \Phi} = -\rho (1 + b)(\beta - \epsilon (1 - \beta)) - \lambda \epsilon 2\rho < 0,$$

$$\frac{\partial \Psi_b}{\partial b} = \rho (1 - \Phi)(\beta - \epsilon (1 - \beta)) < 0,$$

$$\frac{\partial \Psi_b}{\partial \beta} = \rho (1 - \Phi)(1 + b)(1 + \epsilon) < 0,$$

$$\frac{\partial \Psi_b}{\partial \epsilon} = -\rho (1 - \Phi)((1 - \beta)(1 + b) - 2\lambda) > 0, \text{ because we assume } (1 - \beta)(1 + b) - 2\lambda > 0,$$

$$\frac{\partial \Psi_b}{\partial \rho} = (1 - \Phi)(1 + b)(\beta - \epsilon (1 - \beta)) + \lambda \epsilon (1 - \Phi) 2 - \lambda < 0,$$

$$\frac{\partial \Psi_c}{\partial \Phi} = \rho^2 \epsilon > 0,$$

$$\frac{\partial \Psi_c}{\partial \epsilon} = -\rho^2 (1 - \Phi) > 0,$$

$$\frac{\partial \Psi_c}{\partial \rho} = -2\rho\epsilon (1 - \Phi) > 0.$$

Unfortunately, all these properties do not allow us to sign the derivative $\frac{\partial n_2^*}{\partial \lambda}$, which is one of the important tasks we have to accomplish as increasing λ allows to move from the Millian to the Benthamite social welfare function. However, on gets after the appropriate substitutions

$$\frac{\partial n_2^*}{\partial \lambda} = \frac{-(n_2^*)^2 \frac{\partial \Psi_a}{\partial \lambda} - n_2^* \frac{\partial \Psi_b}{\partial \lambda} - \frac{\partial \Psi_c}{\partial \lambda}}{2\Psi_a n_2^* + \Psi_b} = \frac{-n_2^*}{2\Psi_a n_2^* + \Psi_b} \left(n_2^* \frac{\partial \Psi_a}{\partial \lambda} + \frac{\partial \Psi_b}{\partial \lambda} \right)$$
$$= \frac{-n_2^*}{2\Psi_a n_2^* + \Psi_b} \left(n_2^* \Psi_a - \left(\epsilon \left(1 - \Phi\right) - 1\right) \left(n_2^* \lambda - \rho \right) + \rho \epsilon \left(1 - \Phi\right) \right). \tag{61}$$

Given (55), (57), and (58), we get

$$\frac{\partial n_2^*}{\partial \lambda} < 0 (> 0)$$

depending on whether

$$n_2^* \Psi_a < (>) \left(\epsilon \left(1 - \Phi\right) - 1\right) \left(n_2^* \lambda - \rho \right) - \rho \epsilon \left(1 - \Phi\right).$$

One can proceed in the same way for all the other comparative statics and identify sufficient conditions for the intended properties to hold. Unfortunately, it is not possible to extract sharp necessary and sufficient conditions. As a last example, consider the other altruism parameter ϵ and try to sign the derivative $\frac{\partial n_2^*}{\partial \epsilon}$, **which is expected to be strictly positive**:

$$\frac{\partial n_2^*}{\partial \epsilon} = \frac{-(n_2^*)^2 \frac{\partial \Psi_a}{\partial \epsilon} - n_2^* \frac{\partial \Psi_b}{\partial \epsilon} - \frac{\partial \Psi_c}{\partial \epsilon}}{2\Psi_a n_2^* + \Psi_b} = \frac{-n_2^*}{2\Psi_a n_2^* + \Psi_b} \left(n_2^* \frac{\partial \Psi_a}{\partial \epsilon} + \frac{\partial \Psi_b}{\partial \epsilon} + \frac{1}{n_2^*} \frac{\partial \Psi_c}{\partial \epsilon} \right)$$
$$= \frac{-n_2^*}{2\Psi_a n_2^* + \Psi_b} \left(\left(n_2^* \lambda - \rho \right) \left(1 - \Phi\right) \left(\left(1 - \beta\right) \left(1 + b\right) - 2\lambda \right) + \left(1 - \Phi\right) \frac{\left(n_2^* \lambda + \rho \right) \left(n_2^* \lambda - \rho \right)}{n_2^*} \right). \tag{62}$$

Consequently, if $(1-\beta)(1+b) - 2\lambda > 0$ (sufficient), then $\frac{\partial n_2^*}{\partial \epsilon} > 0$. Moreover, it also implies $(1-\beta)(1+b) - \lambda > 0$.

It appears clearly that when $\lambda \neq 0$, numerical exploration of the comparative statics properties is unavoidable. We shall do that together with the investigation of short-run dynamics.

5 Numerical experiments

In this section we complement the previous analytical results with some numerical exercises. Since the comparative statics are analytically ambiguous in the general case, we present some numerical results in section 5.1. Moreover, section 5.2 is devoted to give some insight into the short term dynamics of the model. We consider the following parameter values: $\rho = 0.05$, $\beta = 0.36$ and $\Phi = 2$, which roughly conform to the standard values used in the literature (Caballé and Santos, 1993; Cantón and Meijdam, 1997). Per capita rearing cost and the propensity to have children are given by $b = 1$, $\epsilon = 0.3$, according to Barro and Sala-i-Martin (2004) and de la Croix and Doepke (2003). Finally we assume A_0=1, $x = 0$ and fix $\lambda = 0.5$ as the benchmark value for the altruism parameter.

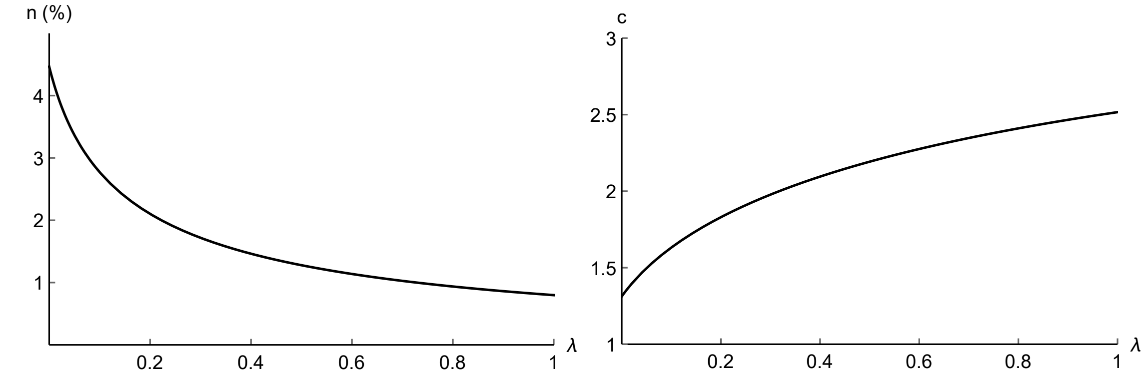
5.1 Comparative statics

We first analyze how the long run population growth rate and the per capita consumption change as the degree of intertemporal altruism increases (Figure 1). One can see that consistently with Palivos and Yip, the population growth rate is decreasing with λ, that's the Benthamite social welfare function delivers the lowest demographic growth in the long-run. The fact that consumption per capita is at the same time increasing with λ allows also to conclude that either in the AK or in the Ramsey case (with decreasing returns to both human or physical capital) no *repugnant conclusion* arises under Benthamite preferences[7]. The Millian case shows in contrast the largest growth rate of population and the lowest per capita consumption. Nonetheless, in our calibrated model, the quantitative differences between the two extreme cases, though significant, can hardly lead us to conclude for any opposite *repugnant conclusion*.

Figure 2 shows a negative relationship between the child rearing cost and the stationary rate of population growth. The same result is obtained in the selfish case. However, the effect of higher child rearing costs on the long run income per capita depends on the degree of intertemporal altruism. We proved that in the Millian case $\lambda = 0$, the long run income per capita is independent of b. When $\lambda \neq 0$, numerical experiments show that the steady state income per capita decreases with the child rearing cost.

[7]We also obtain the same result for a wide range of reasonable parameter values.

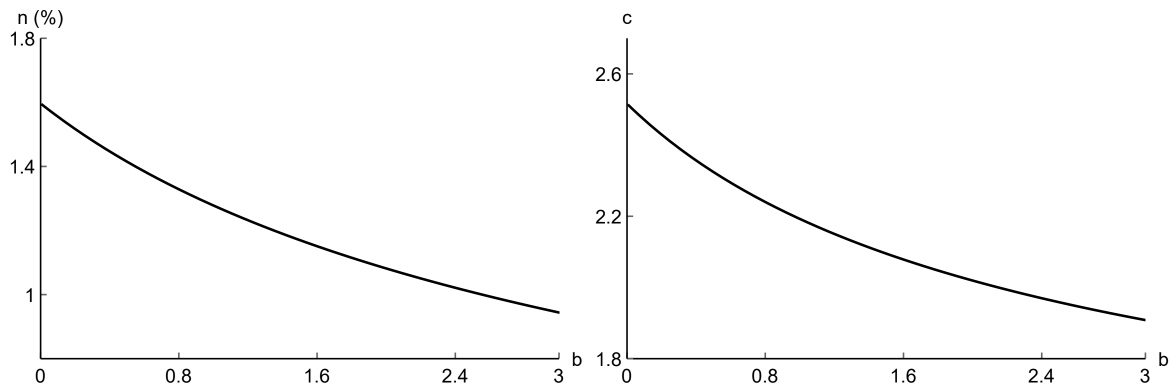
Figure 1: Population growth rate and consumption per capita values as the altruism parameter λ changes.



(a) long run n as λ changes

(b) long run c as λ changes

Figure 2: Population growth rate and consumption per capita values as rearing cost b changes.

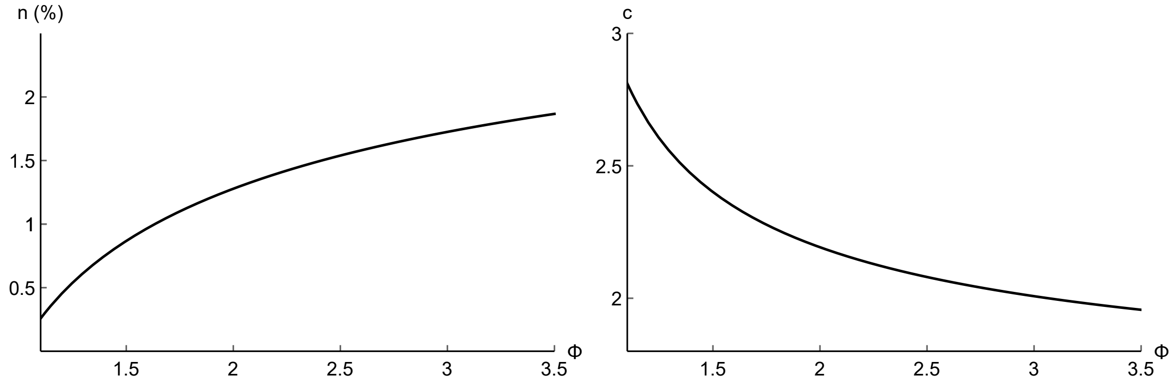


(a) long run n rate as b changes

(b) long run c as b changes

Finally we study the comparative statics with respect to the inverse of the intertemporal elasticity of substitution in consumption Φ. Under the Millian case (see Proposition 1), the long run values of the relevant variables of the model are independent of Φ. However, when $\lambda \neq 0$, an increase in the inverse of the intertemporal elasticity of substitution in consumption rises the population growth rate in the long run. As a consequence, income per capita decreases with Φ.

Figure 3: Population rate and consumption per capita values as the inverse of the intertemporal elasticity of substitution Φ changes.



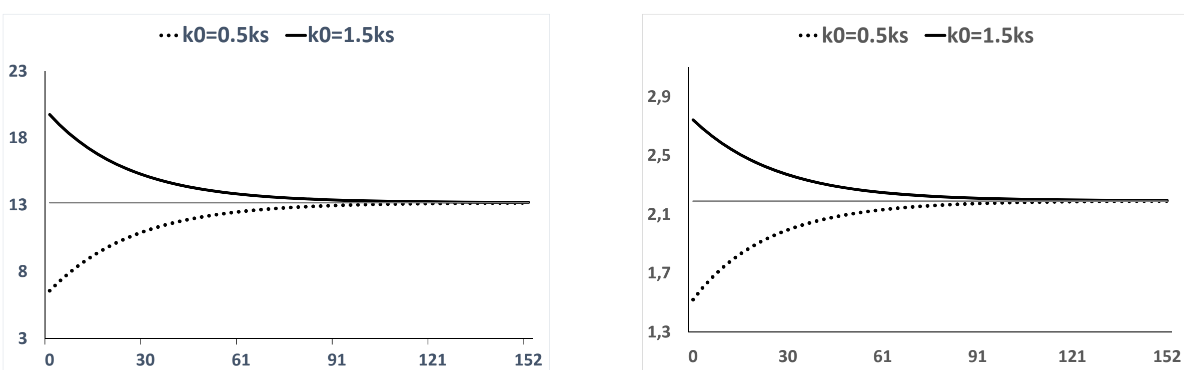
(a) long run n as Φ changes

(b) long run c as Φ changes

5.2 Short run dynamics

We numerically study the optimal paths focusing on the typical imbalance effect analysis. We induce a transition process choosing $k_0 \neq \widetilde{k}^*$, and analyze two different situations, depending on the position of the economy, below or above the long run value of the capital stock per capita. In particular we set the initial condition $k_0 = 0.5\widetilde{k}^*$ and $k_0 = 1.5\widetilde{k}^*$ (Figures 4 and 5).

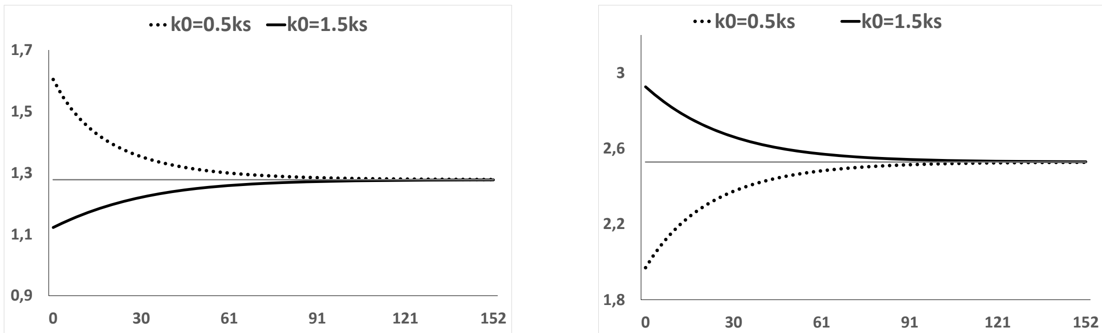
Figure 4: per capita physical capital stock and consumption



(a) physical capital stock

(b) per capita consumption

Figure 5: population growth rate and per capita income



(a) population growth rate

(b) per capita income

The paths obtained can be roughly interpreted as optimal demographic transitions. When capital per capita is below the stationary value, capital is relatively scarce with respect to labor (or human capital). The economy starts investing massively in capital, and capital is gradually substituted for labor. As the process of substitution proceeds forward, the optimal population growth rate goes down leading to a kind of demographic transition (decreasing population growth rate, increasing consumption per capita) until convergence to the stationary equilibrium. It's worth pointing out that these dynamics can be interpreted as imbalance effect dynamics as depicted in Barro and Sala-i-Martin (2004), chapter 5, or more recently in Boucekkine, Martinez and Ruiz-Tamarit (2008). Notice also the symmetry of optimal trajectories corresponding to initial relatively scarce capital and initial relatively abundant capital respectively. Symmetry here is granted because we have one production sector, the shape of imbalance effects are much less symmetric in two-sector models à la Lucas-Uzawa (see Boucekkine et al., 2008).

6 Conclusion

In this paper, we have studied an optimal population size problem of the Ramsey type in a continuous time framework. We have motivated this problem within the population ethics debate: the tricky question is not to derive optimal demographic paths but which social welfare functions are the most appropriate to cope with intergenerational fairness. We show that our framework does not go at odds with the early AK or AN literature dealing with these questions. In particular, the Benthamite criterion is shown to not deliver any *repugnant conclusion*, neither in the long run nor in the short run. Indeed, within the class of social welfare functions considered, the Benthamite criterion is the one which leads with the lowest stationary demographic growth and to the largest consumption per capita. Our contribution is also methodological. We show that this type of problems with endogenous demography can be unambiguously connected to the class of optimal control problems with endogenous discounting, and should be therefore treated accordingly, that's with an appropriate version of the maximum principle. It goes without saying that this proviso applies *a fortiori* to any other extension of this model, for example if one has in mind to incorporate ecological concerns into the problem. Also notice, as already put forward by Ayong le Kama and Schubert (2007), that integrating those concerns would imply an additional channel of endogenous discounting: indeed, the planner could choose to also discount with respect to the ecological state, assigning the maximal weight to the situations of ecological emergency.

7 Appendix A

According to the Mangasarian's Sufficiency Theorem the necessary conditions of the Maximum Principle for an optimum are also sufficient if the Hamiltonian function $H\left(\widetilde{c}, n, q, \widetilde{k}, \upsilon, \Delta\right)$ given in (15) is concave in $\left(\widetilde{c}, n, \widetilde{k}, \Delta\right)$ jointly, under the proviso that the transversality conditions (23) and (24) hold. Here, the Hessian matrix associated to the Hamiltonian function may be written as follows

$$H_{essian} = (1-\Phi)\, H_{\Delta\Delta} \begin{pmatrix} \frac{-\Phi}{\widetilde{c}^2} & \frac{\epsilon(1-\Phi)}{\widetilde{c}n} & 0 & \frac{-1}{\widetilde{c}} \\ \frac{\epsilon(1-\Phi)}{\widetilde{c}n} & \frac{\epsilon(\epsilon(1-\Phi)-1)}{n^2} & \frac{-(1+b)}{\widetilde{c}} & \frac{-\epsilon}{n} \\ 0 & \frac{-(1+b)}{\widetilde{c}} & \frac{-(1-\beta)\beta}{\widetilde{c}\widetilde{k}^{2-\beta}} & 0 \\ \frac{-1}{\widetilde{c}} & \frac{-\epsilon}{n} & 0 & \frac{1}{1-\Phi} \end{pmatrix}, \tag{A.1}$$

where $(1-\Phi)\, H_{\Delta\Delta} = \widetilde{c}^{1-\Phi} n^{\epsilon(1-\Phi)} A_0^{1-\Phi} N_0^{\lambda} e^{-\Delta}$ is nonnegative.

A necessary and sufficient condition for $H\left(\widetilde{c}, n, q, \widetilde{k}, \upsilon, \Delta\right)$ to be concave in $\left(\widetilde{c}, n, \widetilde{k}, \Delta\right)$ is that the Hessian matrix is negative semidefinite. Moreover, a necessary and sufficient condition for a negative semidefinite H_{essian} is that the sign of the determinants known as principal minors accommodate to the following sequence: $\widetilde{D}_1 \leqslant 0$, $\widetilde{D}_2 \geqslant 0$, $\widetilde{D}_3 \leqslant 0$, and $\widetilde{D}_4 \equiv |H_{essian}| \geqslant 0$.

Given the parameter constraints assumed in this model, in particular $\Phi > 1$ and $1 > \epsilon(1-\Phi)$, it is easy to show that the required concavity conditions on the Hamiltonian function are satisfied if

$$\frac{\widetilde{c}}{\widetilde{k}} \frac{\widetilde{k}^{\beta}}{\widetilde{k}} \frac{1}{n^2} \geqslant \frac{(1+b)^2}{\epsilon(1-\beta)\beta}. \tag{A.2}$$

This condition imposes a stronger requirement on the degree of concavity of the production function. The above inequality, given that $f\left(\widetilde{k}\right) = \widetilde{k}^{\beta}$, may be rewritten as

$$f''\left(\widetilde{k}\right) \leqslant -\frac{(1+b)^2}{\epsilon} \frac{n^2}{\widetilde{c}} < 0. \tag{A.3}$$

In particular, given (41), (42) and (58), a sufficient condition for the required concavity condition (A.2) to be satisfied is that $1 > \beta\epsilon$, which always holds because of the assumed parameter configuration of the model.

8 Appendix B: Volterra's derivatives

In what follows we adapt to our model the analysis from Pittel (2002), appendix to chapter 5. Recall that the particular instantaneous utility function is of the form

$$U\left(c\left(t\right),n\left(t\right)\right)=\frac{c\left(t\right)^{1-\Phi}n\left(t\right)^{\epsilon(1-\Phi)}}{1-\Phi} \tag{B.1}$$

whereas the welfare function takes the following intertemporal form

$$W\left(\widetilde{c}\left(t\right),n\left(t\right)\right)=\int_{t_0}^{+\infty}\frac{\widetilde{c}\left(t\right)^{1-\Phi}n\left(t\right)^{\epsilon(1-\Phi)}A_0^{1-\Phi}N_0^{\lambda}e^{-\Delta(t)}}{1-\Phi}dt \tag{B.2}$$

$$\Delta\left(t\right)=\int_{t_0}^{t}\Theta\left(n\left(s\right)\right)ds \tag{B.3}$$

$$\Theta\left(n\left(s\right)\right)=\rho-x\left(1-\Phi\right)-\lambda n\left(s\right) \tag{B.4}$$

Due to the structure of the exponential term, **intertemporal preferences are not time-additive**. Consequently, although the marginal utilities $U_c=c^{-\Phi}n^{\epsilon(1-\Phi)}$ and $U_n=\epsilon c^{1-\Phi}n^{\epsilon(1-\Phi)-1}$ in (B.1) represent the corresponding changes in utility at time t, with intertemporal preferences being **recursive** as in (B.2) we need the Volterra derivatives to determine the corresponding intertemporal marginal utilities. Changes in the determinants of W will have an impact on the current utility index but they can also affect the perception of future utility via the impact on the accumulated discount rates with which the future utility levels are discounted.

For the sake of simplicity we can write (B.2) in a more compact form

$$W=\int_{t_0}^{+\infty}F\left(\widetilde{c}\left(t\right),n\left(t\right)\right)e^{-\Delta(t)}dt=\int_{t_0}^{+\infty}F\left(\widetilde{c}\left(t\right),n\left(t\right)\right)\exp\left(-\int_{t_0}^{t}\Theta\left(n\left(s\right)\right)ds\right)dt \tag{B.5}$$

where

$$F\left(\widetilde{c}\left(t\right),n\left(t\right)\right)=\frac{\widetilde{c}\left(t\right)^{1-\Phi}n\left(t\right)^{\epsilon(1-\Phi)}A_0^{1-\Phi}N_0^{\lambda}}{1-\Phi} \tag{B.6}$$

On the other hand, from equation (26) we get

$$\upsilon(t)=\int_{t}^{+\infty}-F\left(\widetilde{c}(\tau),n(\tau)\right)e^{-\Delta(\tau)}d\tau=\int_{t}^{+\infty}-F\left(\widetilde{c}(\tau),n(\tau)\right)\exp\left(-\int_{t_0}^{\tau}\Theta\left(n(s)\right)ds\right)d\tau \tag{B.7}$$

and

$$\upsilon(t)e^{\Delta(t)}=\left(\int_{t}^{+\infty}-F\left(\widetilde{c}(\tau),n(\tau)\right)\exp\left(-\int_{t_0}^{\tau}\Theta\left(n(s)\right)ds\right)d\tau\right)\exp\left(\int_{t_0}^{t}\Theta\left(n(s)\right)ds\right) \tag{B.8}$$

Then, we define the new variable

$$\omega(t)=-\upsilon(t)e^{\Delta(t)}=\int_{t}^{+\infty}F\left(\widetilde{c}(\tau),n(\tau)\right)\exp\left(-\int_{t}^{\tau}\Theta\left(n(s)\right)ds\right)d\tau \tag{B.9}$$

The Volterra derivative is used to determine the derivatives of the functional W near time t, which is supplied in (B.9).

Volterra derivative with respect to $\widetilde{c}$:

$$\frac{\partial F\left(\widetilde{c}(t),n(t)\right)}{\partial\widetilde{c}}\exp\left(-\int_{t_0}^{t}\Theta\left(n(s)\right)ds\right)=\widetilde{c}(t)^{-\Phi}n(t)^{\epsilon(1-\Phi)}A_0^{1-\Phi}N_0^{\lambda}e^{-\Delta(t)} \tag{B.10}$$

which corresponds to the right hand side of the first order condition (16).

Volterra derivative with respect to n:

$$\left(\frac{\partial F\left(\widetilde{c}(t),n(t)\right)}{\partial n}-\frac{\partial\Theta\left(n(s)\right)}{\partial n}\omega(t)\right)\exp\left(-\int_{t_0}^{t}\Theta\left(n(s)\right)ds\right)=$$

$$=\frac{\epsilon\widetilde{c}(t)^{1-\Phi}n(t)^{\epsilon(1-\Phi)}A_0^{1-\Phi}N_0^{\lambda}e^{-\Delta(t)}}{n(t)}-\lambda\upsilon(t) \tag{B.11}$$

which corresponds to the right hand side of the first order condition (17).

References

[1] Arrow K., P. Dasgupta, L. Goulder, G. Daly et al. (2004). Are We Consuming Too Much?. *Journal of Economic Perspectives*, 18, 147-172.

[2] Ayong le Kama, A. and K. Schubert (2007). A note on the Consequences of an Endogenous Discounting depending on the Environmental Quality. *Macroeconomic Dynamics*, 11 (2), 272-289.

[3] Barro, R. and X. Sala-i-Martin (2004). Economic Growth. MIT Press.

[4] Barro, R., and G. Becker (1989). Fertility Choice in a Model of Economic Growth. *Econometrica*, 57, 481-501.

[5] Blanchet, D. (1988). Age Structure and Capital Dilution Effects in Neoclassical Growth Models. *Journal of Population Economics*, 1 (3), 183-194.

[6] Boucekkine, R., G. Fabbri and G. Gozzi (2014). Egalitarianism under Population Change: Age Structure Does Matter. *Journal of Mathematical Economics*, 55, 86-100.

[7] Boucekkine, R., and G. Fabbri (2013). Assessing Parfit's Repugnant Conclusion within a Canonical Endogenous Growth Set-up. *Journal of Population Economics*, 26, 751-767.

[8] Boucekkine, R., B. Martínez and J. R. Ruiz-Tamarit (2008). Note on Global Dynamics and Imbalance Effects in the Lucas-Uzawa Model. *International Journal of Economic Theory*, 4, 503-518.

[9] Boucekkine, R., B. Martínez and J. R. Ruiz-Tamarit (2014). Optimal Sustainable Policies under Pollution Ceiling: the Demographic side. *Mathematical Modelling of Natural Phenomena*, 9 (4), 38-64.

[10] Caballé, J. and M.S. Santos (1993). On Endogenous Growth with Physical and Human Capital. *Journal of Political Economy*, 101, 1042-1067.

[11] Canton, E. and L. Meijdam (1997). Altruism and the Macroeconomic Effects of Demographic Changes. *Journal of Population Economics*, 10 (3), 317-334.

[12] Chiang, A. C. (1992). *Elements of Dynamic Optimization*. McGraw-Hill, Singapore.

[13] Constantinides, M. A. (1988). Optimal Population Growth and the Social Welfare Function. *Eastern Economic Journal* 14 (2), 229-238.

[14] de la Croix, D. and M. Doepke (2003). Inequality and Growth: Why Differential Fertility Matters. *American Economic Review*, 93 (4), 1091-1113.

[15] Dasgupta, P. (1969) On the concept of Optimum Population. *Review of Economic Studies* 36, 295-318.

[16] Dasgupta, P. (2005). Regarding Optimum Population. *Journal of Political Philosophy*, 13, 414-442.

[17] Edgeworth, F. (1925). *Papers Relating to Political Economy*. Vol. III, Macmillan.

[18] Marin-Solano, J. and J. Navas (2009). Non-Constant Discounting in Finite Horizon: The Free Terminal Time Case. *Journal of Economic Dynamics and Control* 33, 666-675.

[19] Nairay, A. (1984). Asymptotic Behavior and Optimal Properties of a Consumption-Investment Model with Variable Time Preference. *Journal of Economic Dynamics and Control*, 7 (3), 283-313.

[20] Nerlove, M., A. Razin and E. Sadka (1985). Population Size: Individual Choice and Social Optima. *The Quarterly Journal of Economics*, 100 (2), 321-334.

[21] Obstfeld, M. (1990). Intertemporal Dependence, Impatience, and Dynamics. *Journal of Monetary Economics*, 26, 45–75.

[22] Palivos, Th., P. Wang and J. Zhang (1997). On the Existence of Balanced Growth Equilibrium. *International Economic Review*, 38 (1), 205-224.

[23] Palivos, T., and C. K. Yip (1993). Optimal Population Size and Endogenous Growth. *Economics Letters*, 41 (1), 107-110.

[24] Parfit, D. (1984). *Reasons and persons*. Oxford University Press.

[25] Pittel, K. (2002). *Sustainability and Endogenous Growth*. Edward Elgar, Massachusetts.

[26] Razin, A., and C. Yuen (1995). Utilitarian Trade-off between Population Growth and Income Growth. *Journal of Population Economics*, 8, 81-87.

[27] Schumacher, I. (2011). Endogenous Discounting and the domain of the Felicity Function. *Economic Modelling*, 28, 574-581.

[28] Seierstad, A. and K. Sydsaeter (1987). *Optimal Control Theory with Economic Applications*. Elsevier Science, Amsterdam.

[29] Sethi. S. P. and G. L. Thompson (2000). *Optimal Control Theory. Applications to Management Science and Economics*. Springer, New York.