HAL open science

MORTALITY DIFFERENTIAL, LABOR TAXATION AND GROWTH: WHAT DO WE LEARN FROM THE BARRO-BECKER MODEL?

Thomas Seegmuller, Stefano Bosi

To cite this version:

Thomas Seegmuller, Stefano Bosi. MORTALITY DIFFERENTIAL, LABOR TAXATION AND GROWTH: WHAT DO WE LEARN FROM THE BARRO-BECKER MODEL?. 2010. halshs-00472732

HAL Id: halshs-00472732

https://shs.hal.science/halshs-00472732

Preprint submitted on 13 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

GREQAM

Groupement de Recherche en Economie
Quantitative d'Aix-Marseille - UMR-CNRS 6579
Ecole des Hautes Etudes en Sciences Sociales
Universités d'Aix-Marseille II et III

Document de Travail
n°2010-05

MORTALITY DIFFERENTIAL, LABOR TAXATION AND GROWTH: WHAT DO WE LEARN FROM THE BARRO-BECKER MODEL?

Stefano BOSI
Thomas SEEGMULLER

February, 2010

WORKING PAPER

Mortality Differential, Labor Taxation and Growth: What Do We Learn from the Barro-Becker Model?*

Stefano BOSI† and Thomas SEEGMULLER‡

February 16, 2010

Abstract

We revisit the seminal paper on endogenous fertility by Barro and Becker (1989) taking into account households' heterogeneity in terms of capital endowments, mortality differential and cost per surviving child. Focusing on an endogenous growth version, we show at first that there exists a unique balanced growth path (BGP) where the population growth rates of all dynasties are identical. Then, we study the long-run effects of shocks on mortality rates (such as epidemics), mortality differential and total factor productivity (TFP) on the economic and demographic growth rates. The main mechanism rests on the adjustment of the average rearing cost of a surviving child. Finally, we extend the model considering the effects of labor taxation. We find that a higher tax rate may, on the one side, enhance growth but, on the other side, raise wealth inequalities.

JEL classification: D20, J13, 040.

Keywords: endogenous fertility, heterogeneous households, mortality differential, labor taxation, endogenous growth.

*We would like to thank Eleni Iliopoulos for her helpful comments to the section on comparative statics. Usual disclaimers apply.

†THEMA, Université de Cergy-Pontoise, 33 bd du Port, 95011 Cergy-Pontoise CEDEX, France. Tel: + 33 1 34 25 22 54. Fax: + 33 1 34 25 62 33. E-mail: stefano.bosi@u-cergy.fr.

‡Corresponding author. CNRS and GREQAM, Centre de la Vieille Charité, 2 rue de la Charité, 13236 Marseille CEDEX 02, France. Tel: + 33 4 91 14 07 27. Fax: + 33 4 91 90 02 27. E-mail: thomas.seegmuller@univmed.fr.

1 Introduction

Barro and Becker (1989) have pioneered the optimal growth literature with endogenous demography and provided a popular framework to represent the interplay between economic and demographic dynamics. As claimed by these authors, they have extended "the literature on optimal economic growth [...] to allow for endogenous and optimizing choices of population growth and intergenerational transfers." In this paper, we want to go one step further, by encompassing households' heterogeneity, specified as heterogeneous capital endowments, mortality rates and costs of rearing children.

We also depart from Barro and Becker (1989) revisiting their contribution in a model of endogenous growth. Indeed, considering dynamics along a balanced growth path (BGP) allows us to highlights the role of shocks on mortality rates (such as epidemics), mortality differential (such as health reforms), total factor productivity (such as industrial revolutions) on economic and population growth. In addition, endogenous growth is also suitable to understand the policy implications, namely the effects of labor taxation, on the growth rates, wealth inequalities and social welfare.

Our model is connected with two streams of literature. On the one hand, some contributions, that introduce heterogeneous households and borrowing constraints in the Ramsey (1928) model, characterize the resulting stationary capital distribution and study how the convergence to the steady state takes place.[1] We depart from these works by considering the fertility choices: endogenous fertility is in fact a way to introduce endogenous discounting. Moreover, we pay attention to the properties of the BGP.

On the other hand, a second stream of literature encompasses endogenous fertility in growth models to study the interaction between economic and demographic growth.[2] In recent years, this topic has been enriched taking into account households' heterogeneity to study the links between demographic transition, inequality and growth.[3] To the best of our knowledge, surprisingly, no effort has been made to include agents' heterogeneity in the seminal work by Barro and Becker (1989) to understand the trade-off between inequality and growth, either economic or demographic. What we can learn about inequality and growth from this benchmark is the focal point and the aim of the paper.[4]

In the spirit of Doepke (2005), we introduce survival probabilities in the Barro and Becker (1989) framework, that can be interpreted as shares of surviving children. Following Becker (1980), we consider also borrowing constraints to avoid negative bequests. This assumption does not matter in presence of a representative household, but becomes crucial when dynasties are heterogeneous, and prevents deviant behaviors of individuals overburdened with debt.

[1] See Becker (1980) for a seminal contribution and Becker (2006) for an extensive survey of this literature.

[2] See Nerlove and Raut (1997) for an introductory survey.

[3] See, among others, de la Croix and Doepke (2003), de la Croix and Sommacal (2009) and Dahan and Tsiddon (1998).

[4] However, following Galor (2005) who argues that the Barro and Becker (1989) model is not suitable to study transitions in the very long-run, we ignore this issue.

For simplicity and tractability, we consider only two heterogeneous dynasties, but heterogeneity is threefold: capital endowments, survival probabilities and time costs of having children. The assumption of heterogeneous survival probabilities implies a mortality differential. Moreover, the time cost of rearing children will be often specified as an affine function of the survival probability (as in Doepke (2005)). In this case, the mortality differential accounts also for the heterogeneity in the costs of children.

To allow for endogenous growth, we assume that production benefits from knowledge externalities that depend on capital intensity. This specification of externalities was pioneered by Frankel (1962), a paper rarely quoted.

We start the BGP characterization, by showing that the borrowing constraints are never binding and so, everybody holds capital along the BGP. Then, we prove the existence and the uniqueness of the BGP. In addition, studying local dynamics, we find the equilibrium determinacy: the jump variables adjust as usual in the basic models of endogenous growth and growth is balanced from the beginning, i.e. there is no room for transitional dynamics. An interesting feature of our model is that, even if agents are heterogeneous, each dynasty experience the same demographic growth rate along the BGP and any fertility differential is ruled out, in contrast with the relative decline of a subpopulation found in de la Croix and Doepke (2003).

Two key mechanisms work along the BGP and account for the behavior of the model. On the one side, we observe a positive relationship between the average cost per surviving child and economic growth; on the other side, a negative link between economic and demographic growth. Both of them can be explained on the ground of the quantity-quality trade-off of rearing children. Following an increase in the cost per surviving children, households would like to have less children, but making higher bequests, which promotes capital accumulation. This slows down the population growth, while enhancing the economic growth. We show that the average cost per surviving child is increasing in the mortality rates of both types of households as well as in the mortality differential. As a consequence, an increase in the mortality rate of one or all the agents, due for instance to epidemics, promotes economic growth, while it pushes down population dynamics. The same happens when the mortality differential becomes larger: economies with more dispersed mortality rates due, for example, to a more unequal access to health services and medical care, experience higher economic growth rates. De la Croix and Sommacal (2009) reach a similar conclusion in a model without endogenous fertility, where the mortality rates depend on medical knowledge. Of course, the mechanism to explain the positive link between mortality differential and economic growth, is different in our paper.

Finally, we extend the model considering labor income taxation. We investigate whether economies experiencing smaller mortality rates or a lower mortality differential, due for instance to a more equal access to medical care, should always cut their tax rates to enhance economic growth and reduce their gap with respect to economies with higher mortality rates or larger mortality differential. We argue that increasing the tax rate on labor income has two opposite effects on economic growth. Indeed, on the one hand, it reduces in-

come, which implies in turn a lower number of children and a higher rate of economic growth, due to the quantity-quality trade-off. On the other hand, it cuts down the time cost of rearing children, which induces households to have more children with a negative impact on economic growth. Therefore, even if public spending is not productive (Barro (1990)) or does not enter households' preferences, we may have a non-monotonic relationship between tax rate and growth. There are also configurations of fundamentals where a higher tax rate always promotes economic growth. In these cases, however, wealth inequalities increase and a positive relation between growth and inequalities emerges. Welfare analysis puts forward another argument which mitigates the benefits of raising growth through the increase of the tax rate. Focusing for simplicity on the model without households' heterogeneity, we prove that welfare is decreasing in the labor tax rate. Indeed, even if a higher level of the tax rate promotes economic growth, it reduces the initial consumption and above all the population growth rate: both these effects have a negative impact on households' welfare.

The paper is organized as follows. In the next section, we present the model. Section 3 is devoted to the definition of equilibrium. In Section 4, we show that everybody holds capital along a BGP, while, in Section 5, we study the existence and uniqueness of the BGP without binding borrowing constraints. Our results on comparative statics are gathered in Section 6. Finally, we extend our framework to take into account labor taxation in Section 7. Many proofs and technical details are relegated in the Appendix.

2 The model

2.1 Dynasties

In the spirit of Becker and Barro (1988) and Barro and Becker (1989), we consider an economy where each person lives for two periods, childhood and adulthood, and has children at the beginning of his adult period. Parents are altruistic towards their children, i.e. utility depends on their own consumption, the number of surviving children and the utility of each child. In addition, we assume that child's utility enters linearly that of his parents. Thus, the utility of an adult of type i belonging to the generation born at $t-1$, is given by:

$$U_{it} = u\left(c_{it}\right) + \left[\alpha\left(\gamma_i n_{it}\right)^{-\varepsilon}\right]\gamma_i n_{it} U_{it+1} \tag{1}$$

c_{it} is the individual consumption giving an instantaneous utility $u\left(c_{it}\right)$, while $\gamma_i n_{it}$ is the number of surviving children. Indeed, n_{it} is the number of children per adult, whereas $\gamma_i \in (0,1]$ can be interpreted as a surviving probability or life expectancy:[5] the larger γ_i, the lower the mortality rate. Finally, we notice that $\alpha\left(\gamma_i n_{it}\right)^{-\varepsilon}$ measures the degree of altruism towards each surviving child,

[5] We notice that there is no uncertainty and γ_i denotes the deterministic share of surviving children. Doepke (2005) observes that introducing uncertainty does not change very much the properties of the Barro-Becker model.

with $\alpha, \varepsilon \in (0,1)$, and U_{it+1} is the utility attained by each surviving child. The utility function (1) augments the preferences in Becker and Barro (1988) and Barro and Becker (1989) to take into account a surviving probability.

Throughout this paper, we consider, as in Becker and Barro (1988), an isoelastic instantaneous utility function, with some parametric restriction to ensure a well-defined maximization program:

Assumption 1 *$u(c_i) \equiv c_i^{\sigma}/\sigma$, with $0 < \sigma < 1 - \varepsilon$.*

To introduce heterogeneity among dynasties, we assume that each generation is formed by two classes of agents $i = 0, 1$ who differ for capital endowments $k_{i0} \geq 0$, child mortality rates γ_i and cost per child. As shown by Becker and Barro (1988), the recursive model (1) can be equivalently written as a discrete-time growth model where the household of type i maximizes a dynastic utility:

$$U_i = \sum_{t=0}^{+\infty} \alpha^t N_{it}^{1-\varepsilon} u(c_{it}) \tag{2}$$

where $U_i \equiv N_{i0}^{1-\varepsilon} U_{i0}$ and N_{it} denotes the size of the ith subpopulation at period t, under the sequence of budget constraints:

$$c_{it} + \gamma_i n_{it} k_{it+1} = R_t k_{it} + (1 - \beta_i n_{it}) w_t \tag{3}$$

borrowing constraints $k_{it+1} \geq 0$,[6] for $t = 0, 1, \ldots$, and given the possibly unequal distribution of initial capital: $k_{00} \neq k_{10}$.

The left-hand side of (3) represents the expenditures. Note that k_{it+1} represents bequest per surviving child. It is materialized by physical as well as human capital used for the production at the next period. The borrowing constraints ensure that bequests can never be negative. Moreover, setting $\gamma_0 \neq \gamma_1$, we take into account a mortality differential.

The right-hand side of (3) represents the disposable income, where $R_t \equiv 1 - \delta + r_t$ and w_t are the gross return on capital and the wage rate, respectively, with $\delta \in (0,1)$ the depreciation rate of capital and r_t the real interest rate.

Each households reduces his working time to educate his children. The opportunity cost of rearing children is given by $\beta_i n_{it} w_t$, where β_i is the constant cost per child in units of time. The cost per child is also heterogeneous and depends on the social class: $\beta_0 \neq \beta_1$. It makes sense to assume that β_i is an increasing function of γ_i: the better off the family, the higher the survival probabilities, the larger the time spend with each child. In this respect, we encompass the affine form considered in Doepke (2005) in a model without heterogeneity: $\beta = a_0 + a_1 \gamma$, with $a_0, a_1 \geq 0$. The child-rearing cost is the sum of an initial cost per child (a_0) and an adding cost when the child survives (a_1). Time devoted to children cannot exceed the endowment of time per period: $\beta_i n_{it} < 1$.

[6] Becker (1980) introduced such borrowing constraints in a Ramsey model with heterogeneous households. In contrast to his paper, discount factor are endogenous in our framework, because of the endogenous fertility.

Let $m_{it} \equiv \gamma_i n_{it}$ be the population growth factor of dynasty i. Hence, the size of a dynasty at time t is given by:

$$N_{it} = m_{it-1} N_{it-1} = N_{i0} \prod_{s=0}^{t-1} m_{is} \tag{4}$$

with $N_{i0} > 0$ given.[7]

Setting $d_t = \prod_{s=0}^{t} R_s^{-1}$, with $d_{-1} = 1$, allows us to write the budget constraint (3) as follows

$$d_t N_{it+1} \left(k_{it+1} + b_i w_t\right) = N_{it} \left[d_{t-1} k_{it} + d_t \left(w_t - c_{it}\right)\right] \tag{5}$$

for $t = 0, 1, \ldots$, where $b_i \equiv \beta_i / \gamma_i$ denotes the cost per surviving child. Instead of choosing an optimal sequence $(c_{it}, k_{it+1}, n_{it})_{t=0}^{+\infty}$, the household i can maximize the utility function (2) with respect to the sequence $(c_{it}, k_{it+1}, N_{it+1})_{t=0}^{+\infty}$ under the budget constraints (5) and borrowing constraints $k_{it+1} \geq 0$.

Deriving the infinite-horizon Lagrangian gives:[8]

$$\begin{aligned} & \alpha \left(1 - \varepsilon\right) u\left(c_{it+1}\right) + \alpha u'\left(c_{it+1}\right) \left(R_{t+1} k_{it+1} + w_{t+1} - c_{it+1}\right) \\ = \; & m_{it}^{\varepsilon} u'\left(c_{it}\right) \left(k_{it+1} + b_i w_t\right) \end{aligned} \tag{6}$$

and

$$m_{it}^{\varepsilon} u'\left(c_{it}\right) \geq \alpha R_{t+1} u'\left(c_{it+1}\right) \tag{7}$$

jointly with the transversality condition[9]

$$\lim_{t \to +\infty} \alpha^t N_{it}^{1-\varepsilon} u'\left(c_{it}\right) m_{it} k_{it+1} = 0 \tag{8}$$

The Euler equation (7) holds with equality when household i owns a positive amount of capital: $k_{it+1} > 0$. In this case, substituting (7) in (6), gives

$$c_{it+1} \left[\left(1 - \varepsilon\right) \frac{u\left(c_{it+1}\right)}{c_{it+1} u'\left(c_{it+1}\right)} - 1\right] = b_i R_{t+1} w_t - w_{t+1}$$

which writes under Assumption 1:

$$c_{it+1} = \frac{\sigma}{1 - \varepsilon - \sigma} \left(b_i R_{t+1} w_t - w_{t+1}\right)$$

As shown in the Appendix, the second-order conditions for utility maximization are also satisfied under Assumption 1.

[7] Heterogeneous initial population sizes are not excluded: $N_{00} \neq N_{10}$.

[8] See the Appendix for more details.

[9] The implicit price of capital k_{it+1} is $\lambda_{it} d_t N_{it+1}$ and the transversality condition writes alternatively $\lim_{t \to +\infty} \lambda_{it} d_t N_{it} m_{it} k_{it+1} = 0$.

2.2 Firms

A continuum, of unit size, of identical firms produce the unique final good using capital K_t and labor L_t. A learning by doing process results in knowledge accumulation. On the one hand, we assume that there are knowledge spillovers and that they play as a non-excludible public good, but rival. On the other hand, we take in account congestion effects, by assuming that the externalities of knowledge depend on the capital intensity instead of on the aggregate capital. More explicitly, the quantity produced by one firm is given by $Y_t = F(K_t, \bar{\kappa}_t L_t)$, where $\bar{\kappa}_t \equiv \overline{K}_t / \overline{L}_t$ denotes the aggregate capital $\overline{K}_t$ per unit of aggregate labor $\overline{L}_t$ and captures the external effect.[10] Technology has standard neoclassical properties:

Assumption 2 *Technology is represented by a constant returns to scale production function $F(X_1, X_2)$ with $X_1 \equiv K_t$ and $X_2 \equiv \bar{\kappa}_t L_t$, defined on $[0, +\infty)^2$. Setting $F_i \equiv \partial F / \partial X_i$ and $F_{ij} \equiv \partial^2 F / (\partial X_j \partial X_i)$, we assume that F is C^2 on $(0, +\infty)^2$, strictly increasing in both arguments ($F_i > 0$), strictly concave ($F_{ii} < 0$, $F_{11}F_{11} - F_{12}F_{21} > 0$). In addition, the boundary (Inada) conditions $\lim_{X_i \to 0} F_i(X_1, X_2) = +\infty$ are satisfied.*

Each firm takes the externalities as constant and maximizes the profit with respect to (K_t, L_t). Demand for inputs fulfills the first-order conditions: $r_t = F_1(K_t, \bar{\kappa}_t L_t)$ and $w_t = F_2(K_t, \bar{\kappa}_t L_t)\bar{\kappa}_t$.[11]

3 Equilibrium

The equilibrium of our economy is a system of three markets, clearing over time. Real prices r_t and w_t adjust to decentralize the equilibrium on the capital market $N_{0t}k_{0t} + N_{1t}k_{1t} = K_t$ and on the labor market $N_{0t}(1 - b_0 m_{0t}) + N_{1t}(1 - b_1 m_{1t}) = L_t$. When these two equalities are fulfilled, the good market also clears, by Walras law.

Because of firms' symmetry and the continuum of unit size, individual and average capital-labor ratios coincide at the equilibrium: $\bar{\kappa}_t = \kappa_t = K_t / L_t$.

We define the total factor productivity $A \equiv F(1,1)$ and the capital share in total income $s \equiv F_1(1,1)/F(1,1)$. Under constant returns to scale, the Euler identity applies, i.e. $F_2(1,1)/F(1,1) = 1 - s$, and the equilibrium prices become:

$$r_t = F_1(1,1) = sA \tag{9}$$

$$w_t = F_2(1,1)\kappa_t = (1-s)A\kappa_t \tag{10}$$

We notice that, as usual in endogenous growth models, the gross interest rate is constant at equilibrium, i.e. $R = 1 - \delta + sA$.

[10] This specification of externalities has been introduced in the seminal but rarely cited paper by Frankel (1962). See also Ljungqvist and Sargent (2004), chapter 14.

[11] Under Assumption 2, the second-order conditions for profit maximization are also satisfied.

Substituting equations (9) and (10) in (3), (6) and (7), and using Assumption 1, we obtain:

$$c_{it} + m_{it}k_{it+1} = Rk_{it} + (1 - b_i m_{it})(1 - s) A\kappa_t \tag{11}$$

$$\left(\frac{c_{it+1}}{c_{it}}\right)^{1-\sigma} = \frac{\alpha}{m_{it}^{\varepsilon}} \frac{Rk_{it+1} + (1 - s) A\kappa_{t+1} + \frac{1-\varepsilon-\sigma}{\sigma} c_{it+1}}{k_{it+1} + (1 - s) Ab_i\kappa_t} \tag{12}$$

$$\left(\frac{c_{it+1}}{c_{it}}\right)^{1-\sigma} \geq \frac{\alpha}{m_{it}^{\varepsilon}} R \tag{13}$$

with

$$\kappa_t = \frac{N_{0t}k_{0t} + N_{1t}k_{1t}}{N_{0t}(1 - b_0 m_{0t}) + N_{1t}(1 - b_1 m_{1t})} \tag{14}$$

and N_{it} is given by (4).

The balanced growth path (BGP) ensues from considering the dynamic system (11)-(13) with a stationary value of the growth factor g.

4 Everybody holds capital along a BGP

In the following, we show that a BGP fails to exist when some borrowing constraints are binding.

Lemma 1 *Under Assumptions 1-2, there exists no BGP with at least one binding borrowing constraint.*

Proof. Obviously, there exists no BGP when $k_{it} = 0$ for $i = 0, 1$. Consider now that only one type of households holds capital. To fix ideas, assume that households of type 1 face binding borrowing constraints, while the others do not. We will prove that $k_{1t} = 0$ along a BGP leads to a contradiction.

For agents of type 0, system (11)-(13) writes:

$$m_{0t}k_{0t+1} = Rk_{0t} + (1 - b_0 m_{0t})(1 - s) A\kappa_t - c_{0t}, \text{ for } t = 0, 1 \ldots \tag{15}$$

$$m_{0t}^{\varepsilon} = \alpha R \left(\frac{Rb_0 - g_{t+1}}{Rb_0 - g_t} g_t\right)^{\sigma-1}, \text{ for } t = 1, 2 \ldots \tag{16}$$

$$c_{0t} = \frac{\sigma}{1 - \varepsilon - \sigma} \left(\frac{Rb_0}{g_t} - 1\right)(1 - s) A\kappa_t, \text{ for } t = 1, 2 \ldots \tag{17}$$

where now $g_t \equiv \kappa_t/\kappa_{t-1}$ and

$$\kappa_t = \frac{N_{0t}k_{0t}}{N_{0t}(1 - b_0 m_{0t}) + N_{1t}(1 - b_1 m_{1t})} \tag{18}$$

At a steady state $g_t = g_{t+1} = g$, equation (16) gives

$$m_0 = \left(\alpha R g^{\sigma-1}\right)^{1/\varepsilon} \tag{19}$$

which is constant, while (15) and (17) entail

$$m_0 g_0 = R + (1-s) A \frac{\kappa_t}{k_{0t}} \left[1 - b_0 m_0 - \frac{\sigma}{1-\varepsilon-\sigma} \left(b_0 \frac{R}{g} - 1 \right) \right] \tag{20}$$

where g_0 is the stationary value of $g_{0t+1} \equiv k_{0t+1}/k_{0t}$. Along the BGP, g_0 is constant and (20) implies that κ_t/k_{0t} is constant as well, which implies in turn $g = g_0$.

Therefore, equation (18) gives:

$$\frac{1 - b_0 m_0 + (1 - b_1 m_{1t}) N_{1t}/N_{0t}}{1 - b_0 m_0 + (1 - b_1 m_{1t+1}) N_{1t+1}/N_{0t+1}} = 1$$

that is

$$m_0 = m_{1t} \frac{1 - b_1 m_{1t+1}}{1 - b_1 m_{1t}}$$

since $m_{it} = N_{it+1}/N_{it}$. We deduce that $m_{1t} = m_1 = m_0$ is also constant.

For agents of type 1, equation (11) becomes $c_{1t} = (1 - b_1 m_{1t})(1-s) A\kappa_t$. Since m_1 constant, we get $c_{1t+1}/c_{1t} = g$. Equation (13) holds with inequality:

$$\left(\frac{c_{1t+1}}{c_{1t}} \right)^{1-\sigma} > \frac{\alpha R}{m_1^{\varepsilon}}$$

i.e. $m_1 > \left(\alpha R g^{\sigma-1} \right)^{1/\varepsilon} = m_0$, which leads to a contradiction with $m_1 = m_0$. ■

5 Growth path when everybody holds capital

Given Lemma 1, we consider that the borrowing constraints are never binding, i.e. $k_{it} > 0$ for $i = 0, 1$ and $t = 0, 1, \ldots$. In this case, the dynamic system (11)-(13) writes:

$$m_{it} k_{it+1} = R k_{it} + (1 - b_i m_{it})(1-s) A\kappa_t - c_{it}, \text{ for } t = 0, 1 \ldots \tag{21}$$

$$m_{it}^{\varepsilon} = \alpha R \left(\frac{R b_i - g_{t+1}}{R b_i - g_t} g_t \right)^{\sigma-1}, \text{ for } t = 1, 2 \ldots \tag{22}$$

$$c_{it} = \frac{\sigma}{1-\varepsilon-\sigma} \left(\frac{R b_i}{g_t} - 1 \right) (1-s) A\kappa_t, \text{ for } t = 1, 2 \ldots \tag{23}$$

where we assume $b_i m_{it} < 1 < b_i R/g_t$.

Now, we study the stationary growth factor which determines a BGP. Along the BGP, we have $g_t = g_{t+1} = g$. Then, equation (22) determines the population growth factor:

$$m_i = \left(\alpha R g^{\sigma-1} \right)^{1/\varepsilon} \equiv m(g) \tag{24}$$

From (4), we have $N_{it} = N_{i0} m^t$. Letting $\mu_i \equiv N_{i0}/(N_{00} + N_{10}) \in (0, 1)$, with $\mu_0 + \mu_1 = 1$, we define b as the average time cost per surviving child, where:

$$b \equiv \mu_0 b_0 + \mu_1 b_1 \tag{25}$$

Using (14), the capital-labor ratio rewrites:

$$\kappa_t = \frac{\mu_0 k_{0t} + \mu_1 k_{1t}}{1 - bm} \tag{26}$$

Along the BGP, we have $(k_{it}, c_{it}) = (k_{i0}, c_{i0})\, g^t$ and $N_{it} = N_{i0} m^t$, with $m = m(g)$ is given by (24). Replacing these paths in the transversality condition (8), we obtain:

$$\lim_{t \to +\infty} \left[\alpha R^{1-\varepsilon} g^{-(1-\varepsilon-\sigma)}\right]^{t/\varepsilon} mgk_{i0} c_{i0}^{\sigma-1} N_{i0}^{1-\varepsilon} = 0$$

Hence, we deduce the parametric restriction $\alpha R^{1-\varepsilon} < g^{1-\varepsilon-\sigma}$, which simplifies to

$$R > gm \tag{27}$$

using (24). We notice that, on a BGP, we have $b_i m_i < 1 < b_i R / g_t$, which ensures that the transversality condition (27) is satisfied.

Now, to study the existence and uniqueness of the BGP, we notice that (21) and (23) imply:

$$mk_{it+1} = Rk_{it} + (1-s) A\kappa_t \left[1 - b_i m - \frac{\sigma}{1-\varepsilon-\sigma}\left(b_i \frac{R}{g} - 1\right)\right] \tag{28}$$

Multiplying both the sides of (28) by μ_i and aggregating over $i = 0, 1$, we get:

$$\omega(g, b) \equiv \varphi(g, b) - \psi(g, b) = 0 \tag{29}$$

where

$$\varphi(g, b) \equiv [1 - bm(g)]\,[R + (1-s)A - gm(g)] \tag{30}$$

$$\psi(g, b) \equiv \frac{(1-s)A\sigma}{1-\varepsilon-\sigma}\left(b\frac{R}{g} - 1\right) \tag{31}$$

and $m(g)$ is given by (24).

A stationary growth path or BGP is a solution g of (29) such that $g < Rb_i$ and $b_i m < 1$. The first inequality requires $g < \overline{g} \equiv R \min_i b_i$, while, from (24), the second one requires $g > \underline{g} \equiv (\alpha R \max_i b_i^{\varepsilon})^{1/(1-\sigma)}$, with $\underline{g} < \overline{g}$ for an appropriate choice of α. In this case, since $\underline{g} < g < \overline{g}$ implies $R > gm$, we have $\varphi(g, b) > 0$ and $\psi(g, b) > 0$ for all $g \in \left(\underline{g}, \overline{g}\right)$. By direct inspection of equations (30) and (31), we immediately see that $\partial\psi(g, b)/\partial g < 0 < \partial\varphi(g, b)/\partial g$. Therefore, there is at most one stationary solution g satisfying (29) provided that $\varphi\left(\underline{g}, b\right) < \psi\left(\underline{g}, b\right)$ and $\psi(\overline{g}, b) < \varphi(\overline{g}, b)$.

Without loss of generality, we rank the time costs per surviving child.

Assumption 3 $b_0 \leq b_1$.

In this case, we have $\underline{g} = (\alpha R b_1^{\varepsilon})^{1/(1-\sigma)}$ and $\overline{g} = Rb_0$, and inequality $\underline{g} < \overline{g}$ is ensured by an appropriate value of α.

Assumption 4 $\alpha < R^{-\sigma} b_0^{1-\sigma} b_1^{-\varepsilon}$.

In addition, we find[12]

$$\varphi\left(\underline{g}, b\right) = \frac{1}{2}\frac{b_1 - b_0}{b_1}\left[R + (1-s)A - \frac{\underline{g}}{b_1}\right] \text{ and } \psi\left(\overline{g}, b\right) = \frac{1}{2}\frac{b_1 - b_0}{b_0}\frac{\sigma(1-s)A}{1-\varepsilon-\sigma}$$

Therefore, when there is no heterogeneity in the time cost per surviving child, i.e. $b_0 = b_1$, $\varphi\left(\underline{g}, b\right) = \psi\left(\overline{g}, b\right) = 0$. Since $\partial\psi(g,b)/\partial g < 0 < \partial\varphi(g,b)/\partial g$, we have, respectively, $\psi\left(\underline{g}, b\right) > 0$ and $\varphi\left(\overline{g}, b\right) > 0$: this ensures the existence of a stationary growth factor g. By continuity, a steady state will exist when the difference between b_1 and b_0 is not too large. More explicitly, the inequalities $\varphi\left(\underline{g}, b\right) < \psi\left(\underline{g}, b\right)$ and $\psi\left(\overline{g}, b\right) < \varphi\left(\overline{g}, b\right)$, jointly sufficient to the existence of a BGP, hold under the following condition:

$$b_1 - b_0 < 2\min_i Z_i(b, b_i) \tag{32}$$

where

$$\begin{aligned} Z_0(b, b_0) &\equiv \frac{1-\varepsilon-\sigma}{\sigma(1-s)A}\left[R + (1-s)A - \left(\alpha R^{\varepsilon+\sigma} b_0^{\varepsilon+\sigma-1}\right)^{\frac{1}{\varepsilon}}\right] \\ &\quad \left[1 - b\left(\alpha R^{\sigma} b_0^{\sigma-1}\right)^{\frac{1}{\varepsilon}}\right] b_0 \\ Z_1(b, b_1) &\equiv \frac{\sigma(1-s)A}{1-\varepsilon-\sigma}\frac{b\left(\alpha R^{\sigma} b_1^{\varepsilon+\sigma-1}\right)^{\frac{1}{\sigma-1}} - b_1}{R + (1-s)A - \left(\alpha R b_1^{\varepsilon+\sigma-1}\right)^{\frac{1}{1-\sigma}}} \end{aligned}$$

Proposition 1 *Under Assumptions 1-4 and inequality (32), there exists a unique BGP g satisfying (29).*

Since $\omega(g, b)$ is increasing in g, the growth rate $g - 1$ of the capital-labor ratio is positive if $\overline{g} > 1$ and $\omega(1, b) < 0$.

Corollary 1 *Under Assumptions 1-4 and inequality (32), the growth rate of the capital-labor ratio $g - 1$ is strictly positive if $Rb_0 > 1$ and*

$$\sigma\left(1 + \frac{(1-s)A(Rb-1)}{\left[1 - b(\alpha R)^{1/\varepsilon}\right]\left[R + (1-s)A - (\alpha R)^{1/\varepsilon}\right]}\right) > 1 - \varepsilon$$

In the Appendix, we analyze the local stability properties of the stationary solution g determined in Proposition 1. We show that the BGP is locally

[12] Under Assumption 3, we have:

$$\begin{aligned} R + (1-s)A - \underline{g}/b_1 &> R + (1-s)A - \overline{g}/b_1 \\ &= R + (1-s)A - Rb_0/b_1 \geq (1-s)A > 0 \end{aligned}$$

unique.[13] In fact, given the initial population sizes N_{i0} and the initial capital stocks k_{i0}, economy experiences the stationary growth rate from period 1 on with suitable choices of c_{i0} and m_{i0}. The lack of transitional dynamics in this model induces us to focus only on the properties of the BGP and especially apply the comparative statics (see Section 6).

Eventually, it is not unworthy to notice that both dynasties experience the same population growth rate along the BGP, equal to $m - 1$. Even if the natality factors $n_i = m/\gamma_i$ differ because of mortality differential, there is no (net) fertility differential. By direct inspection of (28), we remark that, in contrast to de la Croix and Doepke (2003), this P structure is in accordance with individual capital inequality.

6 Comparative statics

In this section, we will focus on the effects on the BGP of demographic or mortality shocks (through different kinds of variations of the survival rates γ_0 and γ_1) and productivity shock (through the total factor productivity (TFP) A). More precisely, we will analyze their role on economic and demographic growth (g and m respectively) and on natality ($n_i - 1$).

As seen above, in a model without heterogeneous households, Doepke (2005) assumes a positive relation between the time cost of rearing a child and the surviving rate of a newborn: $\beta_i = a_0 + a_1\gamma_i$, with $a_0 \geq 0$ the cost per child and $a_1 \geq 0$ the cost associated to the surviving probability. We deduce that the cost per surviving child is given by $b_i = a_0/\gamma_i + a_1$. At this stage, it is interesting to notice that when $a_0 > 0$, it is decreasing and convex in γ_i, meaning that the higher the survival probability, the lower the cost per surviving child.

6.1 Effects of mortality shocks

We will consider the effects on the BGP g, the population growth m and the natality factors n_i of various kinds of shocks resulting from changes in the mortality rates or, equivalently, variations in the survival probabilities γ_i.

By inspection of (29), we see that the key mechanism through which the survival probability affect the BGP goes through the average cost per surviving child b. The rationale rests on the quantity-quality trade-off to have children. Hence, we start by studying how the survival probabilities affect this average cost b.

We notice that, since $b_i = a_0/\gamma_i + a_1$ and $\mu_0 + \mu_1 = 1$, the average cost b can write:

$$b = \left[\mu_0 \frac{1}{\gamma_0} + (1 - \mu_0)\frac{1}{\gamma_1}\right] a_0 + a_1 \tag{33}$$

This expression is appropriate to quantify the effects of γ_i on b. However, since we have mortality differential, we are also interested in determining the

[13] Local uniqueness is analytically proved in the case $b_0 = b_1$ and numerically shown in the case $b_0 < b_1$.

impact of a spread-preserving mean (SPM) and a mean-preserving spread (MPS) of survival probabilities. To do that, we consider the average:

$$\gamma_A \equiv \mu_0\gamma_0 + \mu_1\gamma_1$$

and we define the spread σ_A as usual:

$$\sigma_A^2 \equiv \mu_0 (\gamma_0 - \gamma_A)^2 + \mu_1 (\gamma_1 - \gamma_A)^2 = \mu_0\mu_1 (\gamma_0 - \gamma_1)^2$$

Without loss of generality, we assume that:

Assumption 5 $\gamma_1 < \gamma_0$.

Note that, since b_i is decreasing in γ_i, this assumption is in accordance with $b_0 < b_1$ (see Assumption 3). We deduce that $\sigma_A = \sqrt{\mu_0\mu_1}\,(\gamma_0 - \gamma_1)$ and, therefore,

$$\gamma_1 = \gamma_A - \sigma_A\sqrt{\frac{\mu_1}{\mu_0}} < \gamma_0 = \gamma_A + \sigma_A\sqrt{\frac{\mu_0}{\mu_1}}$$

Substituting these expressions in (33), we get:

$$b = a_0\left(\frac{\mu_0}{\gamma_A + \sigma_A\sqrt{\frac{\mu_0}{\mu_1}}} + \frac{\mu_1}{\gamma_A - \sigma_A\sqrt{\frac{\mu_1}{\mu_0}}}\right) + a_1 \tag{34}$$

A positive shock on SPM corresponds to an increase of γ_A taking σ_A as constant, while a positive shock on MPS to an increase of σ_A taking γ_A as constant.

Using (33) and (34), we obtain a crucial lemma:

Lemma 2 *Under Assumptions 1-5, the average cost per surviving child b is decreasing in γ_0, γ_1 and γ_A, but increasing in σ_A. Under $a_0 = 0$, these effects vanish.*

Proof. See the Appendix. ■

This lemma confirms some results for the average cost b already found for b_0 and b_1. Because of the fixed cost per child a_0, increasing the survival probabilities γ_0 or γ_1, or the average of both will reduce the average cost per surviving child. This also means that every shock which increases the mortality rates, like epidemics, will increase the cost per surviving child.

Lemma 2 and the convexity of the cost per surviving child allow us to affirm that increasing the mortality differential, through a positive shock on σ_A, raises the average cost per surviving child. In other words, societies characterized by more dispersed mortality rates, resulting for instance from a more inegalitarian access to health services and medical care, face a higher average cost per surviving child. Obviously, all these effects disappear if the fixed cost a_0 is zero, because, in this case, the cost per surviving child becomes independent of the mortality rates.

Putting together these preliminary results, we can now derive the effects on the BGP:

Proposition 2 *Under Assumptions 1-5, the BGP g is increasing in b. It follows that g is decreasing in γ_0, γ_1 and γ_A, but increasing in σ_A.*

Proof. See the Appendix. ■

The key mechanism through which the survival probabilities affect the BGP is based on the positive relationship between g and the average cost per surviving child b. It can be explained by the quantity-quality trade-off faced by households when they decide the number of children. Indeed, following an increase in the cost per surviving child, households prefer to improve the quality through the capital accumulation, which promotes growth. Therefore, according to Lemma 2, every increase in the mortality rates, through a decrease of γ_0, γ_1 or γ_A, as in the case of epidemics, fosters economic growth. Furthermore, because a larger mortality differential raises the average cost per surviving child, it also pushes up the growth rate. This suggests that economies facing more unequal mortality rates will experience higher growth rates. De la Croix and Sommacal (2009) reached a similar conclusion in a model which takes into account the progress of medical knowledge, but without endogenous fertility.

We now consider the effects on population growth:

Proposition 3 *Under Assumptions 1-5, the population growth factor m is decreasing in the BGP g. As a consequence, m is increasing in γ_0, γ_1 and γ_A, but decreasing in σ_A.*

Proof. See the Appendix. ■

As stressed in this proposition, the link between population growth and mortality rates goes through the negative relationship between the economic and demographic growth rates. Focus on the Euler equation:

$$\left(\frac{c_{it+1}}{c_{it}}\right)^{1-\sigma} = \frac{\alpha}{m_{it}^{\varepsilon}} R \tag{35}$$

In fact, since the real interest rate R is constant, a higher growth rate c_{it+1}/c_{it} has to compensated by a rise of the endogenous discount factor $\alpha/m_{it}^{\varepsilon}$, which implies in turn a decrease of the number of survival children m_{it}. Therefore, because they increase the average cost per surviving child and economic growth, higher mortality rates also reduce the population growth factor m. According to the quantity-quality trade-off of having children, a higher average cost b reduces the demand for children. This mechanism also explains why increasing mortality differential pushes down population growth.

The results on the natality rates n_0 and n_1 are summarized in the following proposition:

Proposition 4 *Under Assumptions 1-5 and a strictly positive fixed cost a_0, n_i is increasing in γ_j. Moreover, n_i is decreasing in γ_i and γ_A for a_0 weak enough. Finally, n_0 is decreasing in σ_A, whereas n_1 is increasing in σ_A for a sufficiently low a_0.*

Proof. See the Appendix. ■

Let us observe that, for a fixed cost of child-rearing a_0 which is not too large, the two natality rates move in two opposite directions in response to a change in one survival probability. To fix ideas, assume that γ_0 remains constant, whereas γ_1 decreases, that is the mortality rate of households of type 1 increases (suppose, for instance, that the subpopulation 1 has a limited access to vaccination services during an epidemics). In this case, the natality rate n_0 goes down because of the quantity-quality trade-off described above. On the contrary, n_1 goes up. Even if the population growth m decreases, households of type 1 make more children to (partially) compensate the lower survival probability ($n_1 = m/\gamma_1$).

6.2 Effects of the total factor productivity

Focus now on the impact of the TFP A on the BGP. Our findings are summarized in the following proposition:

Proposition 5 *Under Assumptions 1-5, the BGP g is increasing in the TFP A if $gm > \varepsilon R$. However, the impact of A on the population growth m and the natality rates n_0 and n_1 may be positive or negative.*

Proof. See the Appendix. ■

The positive relationship between the TFP A and the growth factor g is far from trivial because, differently from the basic models of endogenous growth without endogenous fertility, the growth factor is not directly determined by the Euler equation. Indeed, the discount factor β of the traditional Euler equation:

$$g = (\beta R)^{1/(1-\sigma)} \tag{36}$$

is replaced by an endogenous discounting α/m^{ε} in our Euler equation:

$$g = \left(\frac{\alpha}{m^{\varepsilon}} R\right)^{1/(1-\sigma)} \tag{37}$$

In (36), A raises g through the gross interest rate R, while in our model A plays a role on g and m not only through (37) but also through the other equilibrium conditions. We derive simple conditions for a positive impact of A on g, while the effects on m and n_i depend on the direct impact of A on R and the indirect impact on g according to $m = \gamma_i n_i = \left(\alpha R g^{\sigma-1}\right)^{1/\varepsilon}$.

7 On the role of labor taxation

In the previous section, we have seen how shocks on the mortality rates and the mortality differential affect the BGP. We have found, for instance, that an economy with a lower mortality differential is characterized by a lower growth rate.

In this section, we introduce a common tax rule and we study its consequences in terms of economic and demographic growth, social inequalities and welfare. This allows us to address whenever policy makers of economies characterized by lower mortality differentials (due, for instance, to a more equal access to medical services) are recommended to raise or lower their tax rates to enhance economic growth and, so, reduce the gap with economies characterized by larger mortality differentials.

The main effects brought out in the comparative statics goes through the time-cost per surviving child, which also determines leisure. In this respect, we consider a pertinent choice to introduce taxation on labor income. Many countries around the world share this fiscal rule. Moreover, for simplicity, we assume a linear tax rate $z \in [0, 1)$. Therefore, the disposable labor income becomes now $(1 - z)(1 - b_i m_{it}) w_t$.

The government spends the fiscal receipts $z(1 - b_i m_{it}) w_t$ to finance public expenditures G_t under the balanced budget rule:

$$G_t = z(1 - b_0 m_{0t}) w_t N_{0t} + z(1 - b_1 m_{1t}) w_t N_{1t}$$

We assume also, for the sake of simplicity, that G_t neither enters the production function, nor the consumers' preferences.

The household of type i still maximizes (2), but faces a new budget constraint at each period t:

$$c_{it} + m_{it} k_{it+1} = R_t k_{it} + (1 - z)(1 - b_i m_{it}) w_t \tag{38}$$

or, equivalently,

$$N_{it} c_{it} + N_{it+1} k_{it+1} = R_t N_{it} k_{it} + (1 - z)(1 - b_i m_{it}) w_t N_{it} \tag{39}$$

Deriving the Lagrangian function, focussing on equilibria where the borrowing constraints are never binding ($k_{it} > 0$) and using (9) and (10), we obtain:

$$m_{it} = (\alpha R)^{\frac{1}{\varepsilon}} \left(\frac{c_{it+1}}{c_{it}} \right)^{\frac{\sigma - 1}{\varepsilon}} \tag{40}$$

$$c_{it+1} = \frac{\sigma (1 - z)(1 - s) A}{1 - \varepsilon - \sigma} (R b_i \kappa_t - \kappa_{t+1}) \tag{41}$$

Along a BGP, we have $g_t = g_{t+1} = g$. From (41), we deduce that $c_{it+1}/c_{it} = g$. Substituting the balanced growth factor into (40), we obtain, as in the model without taxation (see (24)):

$$m_i = m = \left(\alpha R g^{\sigma - 1} \right)^{1/\varepsilon} \equiv m(g) \tag{42}$$

Using (41), equation (38) rewrites:

$$\begin{aligned} m k_{it+1} &= R k_{it} + (1 - z)(1 - b_i m)(1 - s) A \kappa_t \\ &\quad - \frac{\sigma (1 - z)(1 - s) A}{1 - \varepsilon - \sigma} (R b_i \kappa_{t-1} - \kappa_t) \end{aligned} \tag{43}$$

Multiplying by μ_i, aggregating over i and substituting (42), the BGP is given by $\omega(g) = 0$, where now:

$$\begin{aligned} \omega(g) \quad &\equiv \quad [R + (1-z)(1-s)A - gm(g)][1 - bm(g)] \\ &\quad -\frac{\sigma(1-z)(1-s)A}{1-\varepsilon-\sigma}\left(b\frac{R}{g} - 1\right) \end{aligned} \tag{44}$$

The comparative statics on labor taxation starts with the impact of the tax rate on the economic and demographic growth rates (including the natality rates).

7.1 The effects of the tax rate on economic and demographic growth

In order to study the impact of the tax rate z on the economic and demographic growth g and m, respectively, and on the natality factors n_i, let us define:

$$h(g) \equiv \frac{1 - bm(g)}{b\frac{R}{g} - 1}$$

with $h(g) > 0$ and $h'(g) > 0$. Under Assumption 4, we have $\underline{g} = (\alpha R b_1^{\varepsilon})^{1/(1-\sigma)} < \overline{g} = Rb_0$ and so, $h(\underline{g}) < h(\overline{g})$.

Our findings are summarized in the next proposition.

Proposition 6 *Under Assumptions 1-4, the following holds along the BGP. Let $b_0 < b_1$.*

(1) *If $\sigma/(1-\varepsilon-\sigma) < h(\underline{g})$, then $\partial g/\partial z > 0$, $\partial m/\partial z < 0$ and $\partial n_i/\partial z < 0$ $(i = 0, 1)$ for all $g \in (\underline{g}, \overline{g})$.*

(2) *If $h(\underline{g}) < \sigma/(1-\varepsilon-\sigma) < h(\overline{g})$, then there exists $g^* \in (\underline{g}, \overline{g})$ such that:*

- **(2.1)** *$\partial g/\partial z < 0$, $\partial m/\partial z > 0$ and $\partial n_i/\partial z > 0$ $(i = 0, 1)$ for all $g \in (\underline{g}, g^*)$;*
- **(2.2)** *$\partial g/\partial z > 0$, $\partial m/\partial z < 0$ and $\partial n_i/\partial z < 0$ $(i = 0, 1)$ for all $g \in (g^*, \overline{g})$.*

(3) *If $h(\overline{g}) < \sigma/(1-\varepsilon-\sigma)$, then $\partial g/\partial z < 0$, $\partial m/\partial z > 0$ and $\partial n_i/\partial z > 0$ $(i = 0, 1)$ for all $g \in (\underline{g}, \overline{g})$.*

If $b_0 = b_1$, case (2) still applies.

Proof. See the Appendix. ■

Proposition 6 highlights a non-monotonic effect of the tax rate on economic growth. The novelty of this result is that it arises in a model where the externalities of public good neither enter the production function (as in Barro (1990)) nor the utility function. The public spending G_t just works as a component of

the aggregate demand and the economic mechanism is very different from the one found by Barro (1990).

The intuition of Proposition 6 rests on two mechanisms. The first one is the Euler equation and the negative interplay between g and m: population dynamics and natality rates move in response to changes in the tax rate z according to (42). So the endogenous discounting α/m_t^{ε} adjusts under a constant gross rate R.

The second mechanism is the individual budget constraint, which writes at period t:

$$c_{it} + m_{it}\left[k_{it+1} + (1-z)\, b_i w_t\right] = R_t k_{it} + (1-z)\, w_t \tag{45}$$

z has two opposite effects: it lowers the cost of rearing children $(1-z)\, b_i w_t$ on the left-hand side (45) and the disposable income on the right hand-side.

When the first effect dominates, an increase of the tax rate decreases the cost per surviving child. This induces households to have more children, and has a negative impact on economic growth according to the quantity-quality trade-off: we are in presence of a prevailing substitution effect.

When the second effect dominates, households reduce the number of children because of the lower disposable income. This enhances the economic growth because of the negative interaction between g and m stressed above: the economy experiences a dominant income effect.

To conclude, we argue that an increase in labor taxation may raise or lower economic growth depending on the resultant of two classical forces: the substitution effect and the income effect. The mechanism we highlight is very different from Barro (1990) and goes through the endogenous fertility.

Then, a policy makers who faces a low economic growth (due, for instance, to a weak mortality differential), should understand what effect prevails before increasing or decreasing the tax rate on labor income, to avoid any pernicious impact on the BGP.

7.2 The effects of the tax rate on social inequalities

As seen above, a higher tax rate on labor income may promote economic growth. This result can be explained on the ground of the quantity-quality trade-off and, therefore, of capital accumulation. In our model, agents are heterogeneous and labor taxation has an impact on social inequalities. Since the taxation affects the economic growth through the capital accumulation, it is suitable to capture the degree of inequality in terms of wealth, that is of capital distribution.

Governments are not only concerned by a growth performance, but also by social inequalities. In this respect, it is not unworthy to consider also the impact of the tax rate on some measure of social inequality, say the Gini index of wealth distribution, and provide some fiscal policy recommendation to reduce the degree of inequality.

Ranking the agents h in $[0, N_{0t} + N_{1t}]$ according to their increasing wealth

and denoting the cumulative wealth by $W_t(h)$, we define the Gini index as:

$$G_t = 1 - 2\frac{\int_0^{N_{0t}+N_{1t}} W_t(h)\,dh}{(N_{0t}+N_{1t})\,W_t(N_{0t}+N_{1t})} \tag{46}$$

In our economy, there are N_{0t} agents holding k_{0t} and N_{1t} agents holding $k_{1t} \geq k_{0t}$.[14] Therefore, $W_t(h)$ is given by hk_{0t}, if $0 \leq h \leq N_{0t}$, and by $N_{0t}k_{0t} + (h - N_{0t})\,k_{1t}$, if $N_{0t} < h \leq N_{0t} + N_{1t}$. Replacing this function in (46), we obtain a simple formula for the Gini index along the BGP:[15]

$$G = \mu_0\left(1 - \frac{k_{0t}/\kappa_t}{1-bm}\right)$$

The following expression plays a role in the main proposition:

$$H \equiv \frac{1-\varepsilon-\sigma}{\varepsilon}\frac{mg}{R-mg} + \frac{1-\sigma}{\varepsilon}\frac{(b-b_0)\,m}{(1-b_0m)(1-bm)}$$
$$+\frac{\sigma}{1-\varepsilon-\sigma}\frac{(1-z)(1-s)A}{R-gm}\left[b_0\frac{R}{g} + \left(b_0\frac{R}{g}-1\right)\frac{b_0m}{1-b_0m}\frac{1-\sigma}{\varepsilon}\right]\frac{\kappa_t}{k_{0t}}$$

Proposition 7 *Under Assumptions 1-4, H is positive and the tax rate z has two effects on the Gini index evaluated along the BGP:*

(1) *a direct effect, that is $(\mu_0 - G)/(1-z) > 0$, which raises wealth inequalities;*

(2) *an indirect effect through g, that is $(\mu_0 - G)(H/g)\,\partial g/\partial z$, which increases wealth inequalities if and only if $h(g) > \sigma/(1-\varepsilon-\sigma)$.*

Proof. See the Appendix. ■

Proposition 7 shows that a higher labor tax rate, which promotes economic growth, also raises wealth inequalities: so, in the case of a growth-enhancing labor taxation, we find a positive trade-off between economic growth and inequalities. Conversely, when taxation on labor income lowers the economic performance the trade-off between growth and inequalities is missed.[16]

7.3 The effects of the tax rate on welfare

We conclude the section on labor taxation by studying the effect of the tax rate z on social welfare. To keep the analysis as simple as possible, we assume that agents are homogeneous, i.e. $b = b_0 = b_1$ and $k_{0t} = k_{1t} \equiv k_t$, whereas N_{00} and N_{10} may differ.

[14] We obtain this inequality from (43), taking into account that $b_0 \leq b_1$ (see Assumption 3).

[15] k_{0t}/κ_t is constant along the BGP.

[16] The debate about growth and inequalities dates back to the early Nineties, following the seminal papers on endogenous growth, and renews an older debate about the Kuznets curve. This issue has been largely discussed in the literature. See Aghion et al. (1999) for a critical survey.

Substituting in (2) the paths $(k_{it}, c_{it}) = (k_{i0}, c_{i0})\, g^t$ and $N_{it} = N_{i0} m^t$, and using the transversality condition, we can compute the utilities along the BGP:

$$U_i = N_{i0}^{1-\varepsilon} \frac{R}{\sigma} \frac{c_0^\sigma}{R - mg}$$

where $c_0 = k_0 \left[(1-z)(1-s)A + R - gm\right]$ and $m_0 = m$ are the values of the jump variables which ensure that agents are on the BGP from the beginning. Given the initial conditions (k_0, N_{00}, N_{10}), maximizing the welfare function $W(U_0, U_1)$ is equivalent to maximize:[17]

$$V \equiv \frac{R}{\sigma} \frac{\left[(1-z)(1-s)A + R - gm\right]^\sigma}{R - gm} > 0$$

where g and m are given by $\omega(g) = 0$ and $m = \left(\alpha R g^{\sigma-1}\right)^{1/\varepsilon}$. Therefore, the effect of the tax rate z on welfare is summarized by:

$$\frac{\partial V}{\partial z} = \frac{\partial g/\partial z}{\partial g/\partial b} \frac{\partial V}{\partial b} - \frac{\sigma(1-s)AV}{(1-z)(1-s)A + R - gm}$$

with

$$\frac{\partial V}{\partial b} = -\frac{1-\varepsilon-\sigma}{\varepsilon} \frac{(1-z)(1-s)A + (1-\sigma)(R-gm)}{(1-z)(1-s)A + R - gm} \frac{gm}{R-gm} \frac{V}{g} \frac{\partial g}{\partial b} < 0$$

Taking in account that $\partial g/\partial b > 0$, we find, surprisingly, that $\partial g/\partial z > 0$, i.e. $h(g) > \sigma/(1-\varepsilon-\sigma)$, implies $\partial V/\partial z < 0$.[18]

In order to explain this apparent paradox, let us recall that the possibility of $\partial g/\partial z > 0$ is not excluded (see Proposition 6) even if, in contrast with Barro (1990), public spending is not productive. The impact of z on V rests on the interplay of three effects:

(1) a negative direct effect of z on initial consumption,

(2) a positive direct effect of g on future consumption,

(3) a negative indirect effect of g through $m(g)$ on future demography and V, because surviving children matter in the utility function.

We know that $\partial g/\partial z > 0$ entails $\partial m/\partial z < 0$. Thus, if $\partial g/\partial z > 0$, the negative effects on initial consumption (1) and population growth (3) always dominate the positive effect on future consumption (2). This explains why welfare decreases in the tax rate ($\partial V/\partial z < 0$).

The result still holds by continuity if we introduce a degree of agents' heterogeneity. Hence, under a positive impact of taxation on economic growth, a benevolent policy maker is recommended to care also about the negative impact on welfare and the positive impact on inequalities, meaning a larger wealth dispersion (see Proposition 7, point (2)).

[17] Notice that $V = U_0 = U_1$.

[18] However, the case $\partial V/\partial z < 0$ with $\partial g/\partial z < 0$ cannot be excluded.

Of course, if the public spending enters either the production function or the utility function, under a balanced-budget rule, taxation could turn out to be welfare-improving even in the case of a positive impact on g.

8 Appendix

8.1 First-order conditions for utility maximization

To compute a household's optimal behavior, we introduce the infinite-horizon Lagrangian:

$$\begin{aligned} L_i &= \sum_{t=0}^{\infty} \alpha^t N_{it}^{1-\varepsilon} u(c_{it}) \\ &\quad + \sum_{t=0}^{\infty} \lambda_{it} \left[N_{it} \left(d_{t-1} k_{it} + d_t (w_t - c_{it}) \right) - d_t N_{it+1} \left(k_{it+1} + w_t b_i \right) \right] \\ &\quad + \sum_{t=0}^{\infty} \mu_{it} k_{it} \end{aligned} \tag{47}$$

where $\lambda_{it} \geq 0$ and $\mu_{it} \geq 0$ are the multipliers associated to the budget constraint and the borrowing constraint, respectively.

Maximizing this Lagrangian with respect to c_{it}, k_{it+1} and N_{it+1} gives:

$$\frac{\partial L_i}{\partial c_{it}} = \alpha^t N_{it}^{1-\varepsilon} u'(c_{it}) - \lambda_{it} N_{it} d_t = 0 \tag{48}$$

$$\frac{\partial L_i}{\partial k_{it+1}} = -\lambda_{it} d_t N_{it+1} + \lambda_{it+1} d_t N_{it+1} + \mu_{it+1} = 0 \tag{49}$$

$$\begin{aligned} \frac{\partial L_i}{\partial N_{it+1}} &= (1-\varepsilon) \alpha^{t+1} N_{it+1}^{-\varepsilon} u(c_{it+1}) - \lambda_{it} d_t \left(k_{it+1} + w_t b_i \right) \\ &\quad + \lambda_{it+1} \left[d_t k_{it+1} + d_{t+1} (w_{t+1} - c_{it+1}) \right] = 0 \end{aligned} \tag{50}$$

The household's trade-offs (6) and (7) ensue from these three first order conditions after elimination of multipliers. ■

8.2 Second-order conditions for utility maximization

In order to recover the second order-conditions found by Becker and Barro (1986) in our framework with heterogeneous survival probabilities and borrowing constraints, we consider the Lagrangian function (47).

The bordered Hessian H_5 is given by:

$$H_5 \equiv \begin{bmatrix} \frac{\partial^2 L_i}{\partial \lambda_{it}^2} & \frac{\partial^2 L_i}{\partial \lambda_{it} \partial \mu_{it}} & \frac{\partial^2 L_i}{\partial \lambda_{it} \partial k_{it}} & \frac{\partial^2 L_i}{\partial \lambda_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial \lambda_{it} \partial N_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it} \partial \mu_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it}^2} & \frac{\partial^2 L_i}{\partial \mu_{it} \partial k_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it} \partial N_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it} \partial k_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it} \partial k_{it}} & \frac{\partial^2 L_i}{\partial k_{it}^2} & \frac{\partial^2 L_i}{\partial k_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial k_{it} \partial N_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial k_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial c_{it}^2} & \frac{\partial^2 L_i}{\partial c_{it} \partial N_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it} \partial N_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it} \partial N_{it}} & \frac{\partial^2 L_i}{\partial k_{it} \partial N_{it}} & \frac{\partial^2 L_i}{\partial c_{it} \partial N_{it}} & \frac{\partial^2 L_i}{\partial N_{it}^2} \end{bmatrix}$$

with

$$\begin{bmatrix} \frac{\partial^2 L_i}{\partial \lambda_{it}^2} & \frac{\partial^2 L_i}{\partial \lambda_{it} \partial \mu_{it}} & \frac{\partial^2 L_i}{\partial \lambda_{it} \partial k_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it} \partial \mu_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it}^2} & \frac{\partial^2 L_i}{\partial \mu_{it} \partial k_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it} \partial k_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it} \partial k_{it}} & \frac{\partial^2 L_i}{\partial k_{it}^2} \\ \frac{\partial^2 L_i}{\partial \lambda_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial k_{it} \partial c_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it} \partial N_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it} \partial N_{it}} & \frac{\partial^2 L_i}{\partial k_{it} \partial N_{it}} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & N_{it} d_{t-1} \\ 0 & 0 & 1 \\ N_{it} d_{t-1} & 1 & 0 \\ -d_t N_{it} & 0 & 0 \\ d_t (w_t - c_{it}) + d_{t-1} k_{it} & 0 & (\lambda_{it} - \lambda_{it-1}) d_{t-1} \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial^2 L_i}{\partial \lambda_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial \lambda_{it} \partial N_{it}} \\ \frac{\partial^2 L_i}{\partial \mu_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it} \partial N_{it}} \\ \frac{\partial^2 L_i}{\partial k_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial k_{it} \partial N_{it}} \\ \frac{\partial^2 L_i}{\partial c_{it}^2} & \frac{\partial^2 L_i}{\partial c_{it} \partial N_{it}} \\ \frac{\partial^2 L_i}{\partial c_{it} \partial N_{it}} & \frac{\partial^2 L_i}{\partial N_{it}^2} \end{bmatrix}$$

$$= \begin{bmatrix} -d_t N_{it} & d_t (w_t - c_{it}) + d_{t-1} k_{it} \\ 0 & 0 \\ 0 & \lambda_{it} d_{t-1} - \lambda_{it-1} d_{t-1} \\ \alpha^t N_{it}^{1-\varepsilon} u''(c_{it}) & (1-\varepsilon) \alpha^t N_{it}^{-\varepsilon} u'(c_{it}) - \lambda_{it} d_t \\ (1-\varepsilon) \alpha^t N_{it}^{-\varepsilon} u'(c_{it}) - \lambda_{it} d_t & -\varepsilon (1-\varepsilon) \alpha^t N_{it}^{-1-\varepsilon} u(c_{it}) \end{bmatrix}$$

At the point solution of the maximization program, either (1) the individual capital is positive ($k_{it} > 0$) and only one constraint is binding; or (2) the individual capital is zero ($k_{it} = 0$) and two constraints are binding.

In the first case, the Hessian matrix becomes four-dimensional:

$$H_4 \equiv \begin{bmatrix} \frac{\partial^2 L_i}{\partial \lambda_{it}^2} & \frac{\partial^2 L_i}{\partial \lambda_{it} \partial k_{it}} & \frac{\partial^2 L_i}{\partial \lambda_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial \lambda_{it} \partial N_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it} \partial k_{it}} & \frac{\partial^2 L_i}{\partial k_{it}^2} & \frac{\partial^2 L_i}{\partial k_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial k_{it} \partial N_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial k_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial c_{it}^2} & \frac{\partial^2 L_i}{\partial c_{it} \partial N_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it} \partial N_{it}} & \frac{\partial^2 L_i}{\partial k_{it} \partial N_{it}} & \frac{\partial^2 L_i}{\partial c_{it} \partial N_{it}} & \frac{\partial^2 L_i}{\partial N_{it}^2} \end{bmatrix}$$

with

$$\begin{bmatrix} \frac{\partial^2 L_i}{\partial \lambda_{it}^2} & \frac{\partial^2 L_i}{\partial \lambda_{it} \partial k_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it} \partial k_{it}} & \frac{\partial^2 L_i}{\partial k_{it}^2} \\ \frac{\partial^2 L_i}{\partial \lambda_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial k_{it} \partial c_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it} \partial N_{it}} & \frac{\partial^2 L_i}{\partial k_{it} \partial N_{it}} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & N_{it} d_{t-1} \\ N_{it} d_{t-1} & 0 \\ -d_t N_{it} & 0 \\ d_t (w_t - c_{it}) + d_{t-1} k_{it} & (\lambda_{it} - \lambda_{it-1}) d_{t-1} \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial^2 L_i}{\partial \lambda_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial \lambda_{it} \partial N_{it}} \\ \frac{\partial^2 L_i}{\partial k_{it} \partial c_{it}} & \frac{\partial^2 L_i}{\partial k_{it} \partial N_{it}} \\ \frac{\partial^2 L_i}{\partial c_{it}^2} & \frac{\partial^2 L_i}{\partial c_{it} \partial N_{it}} \\ \frac{\partial^2 L_i}{\partial c_{it} \partial N_{it}} & \frac{\partial^2 L_i}{\partial N_{it}^2} \end{bmatrix}$$

$$= \begin{bmatrix} -d_t N_{it} & d_t (w_t - c_{it}) + d_{t-1} k_{it} \\ 0 & \lambda_{it} d_{t-1} - \lambda_{it-1} d_{t-1} \\ \alpha^t N_{it}^{1-\varepsilon} u''(c_{it}) & (1-\varepsilon) \alpha^t N_{it}^{-\varepsilon} u'(c_{it}) - \lambda_{it} d_t \\ (1-\varepsilon) \alpha^t N_{it}^{-\varepsilon} u'(c_{it}) - \lambda_{it} d_t & -\varepsilon (1-\varepsilon) \alpha^t N_{it}^{-1-\varepsilon} u(c_{it}) \end{bmatrix}$$

Since we evaluate the Hessian matrix at a point satisfying the first-order conditions and $\mu_t = 0$, we have $\lambda_{it} = \lambda_{it-1}$ and

$$(1-\varepsilon) \alpha^t N_{it}^{-\varepsilon} u'(c_{it}) - \lambda_{it} d_t = -\varepsilon \alpha^t N_{it}^{-\varepsilon} u'(c_{it})$$

Hence, the bordered Hessian becomes

$$H_4 = \begin{bmatrix} 0 & N_{it} d_{t-1} & -d_t N_{it} & X_t \\ N_{it} d_{t-1} & 0 & 0 & 0 \\ -d_t N_{it} & 0 & \alpha^t N_{it}^{1-\varepsilon} u''(c_{it}) & -\varepsilon \alpha^t N_{it}^{-\varepsilon} u'(c_{it}) \\ X_t & 0 & -\varepsilon \alpha^t N_{it}^{-\varepsilon} u'(c_{it}) & -\varepsilon (1-\varepsilon) \alpha^t N_{it}^{-1-\varepsilon} u(c_{it}) \end{bmatrix}$$

with $X_t \equiv d_t (w_t - c_{it}) + d_{t-1} k_{it}$.

Negative definition requires $(-1)^3 \det H_4 > 0$ and $(-1)^2 \det H_3 > 0$, that is $\det H_4 < 0$ and

$$\det H_3 = \det \begin{bmatrix} 0 & N_{it} d_{t-1} & -d_t N_{it} \\ N_{it} d_{t-1} & 0 & 0 \\ -d_t N_{it} & 0 & \alpha_i^t N_{it}^{1-\varepsilon_i} u''(c_{it}) \end{bmatrix} > 0 \tag{51}$$

Inequality (51) is satisfied, because

$$\det H_3 = -(N_{it} d_{t-1})^2 \alpha^t N_{it}^{1-\varepsilon} u''(c_{it}) > 0$$

Thus, the second-order condition becomes:

$$\det \begin{bmatrix} 0 & N_{it} d_{t-1} & -d_t N_{it} & X_t \\ N_{it} d_{t-1} & 0 & 0 & 0 \\ -d_t N_{it} & 0 & \alpha^t N_{it}^{1-\varepsilon} u''(c_{it}) & -\varepsilon \alpha^t N_{it}^{-\varepsilon} u'(c_{it}) \\ X_t & 0 & -\varepsilon \alpha^t N_{it}^{-\varepsilon} u'(c_{it}) & -\varepsilon (1-\varepsilon) \alpha^t N_{it}^{-1-\varepsilon} u(c_{it}) \end{bmatrix} < 0$$

or, equivalently,

$$-\frac{c_{it}u''(c_{it})}{u'(c_{it})}\frac{u(c_{it})}{c_{it}u'(c_{it})} > \frac{\varepsilon}{1-\varepsilon} \tag{52}$$

Since $u(c_i) \equiv c_i^{\sigma}$, we obtain $\sigma < 1-\varepsilon$, which is satisfied under Assumption 1, as in Becker and Barro (1986).

Case (2) was not considered by Becker and Barro (1986). In this case, H_5 becomes

$$\begin{bmatrix} \frac{\partial^2 L_i}{\partial \lambda_{it}^2} & \frac{\partial^2 L_i}{\partial \lambda_{it}\partial \mu_{it}} & \frac{\partial^2 L_i}{\partial \lambda_{it}\partial k_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it}\partial \mu_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it}^2} & \frac{\partial^2 L_i}{\partial \mu_{it}\partial k_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it}\partial k_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it}\partial k_{it}} & \frac{\partial^2 L_i}{\partial k_{it}^2} \\ \frac{\partial^2 L_i}{\partial \lambda_{it}\partial c_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it}\partial c_{it}} & \frac{\partial^2 L_i}{\partial k_{it}\partial c_{it}} \\ \frac{\partial^2 L_i}{\partial \lambda_{it}\partial N_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it}\partial N_{it}} & \frac{\partial^2 L_i}{\partial k_{it}\partial N_{it}} \end{bmatrix} = \begin{bmatrix} 0 & 0 & N_{it}d_{t-1} \\ 0 & 0 & 1 \\ N_{it}d_{t-1} & 1 & 0 \\ -d_tN_{it} & 0 & 0 \\ X_t & 0 & -\mu_{it}/N_{it} \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial^2 L_i}{\partial \lambda_{it}\partial c_{it}} & \frac{\partial^2 L_i}{\partial \lambda_{it}\partial N_{it}} \\ \frac{\partial^2 L_i}{\partial \mu_{it}\partial c_{it}} & \frac{\partial^2 L_i}{\partial \mu_{it}\partial N_{it}} \\ \frac{\partial^2 L_i}{\partial k_{it}\partial c_{it}} & \frac{\partial^2 L_i}{\partial k_{it}\partial N_{it}} \\ \frac{\partial^2 L_i}{\partial c_{it}^2} & \frac{\partial^2 L_i}{\partial c_{it}\partial N_{it}} \\ \frac{\partial^2 L_i}{\partial c_{it}\partial N_{it}} & \frac{\partial^2 L_i}{\partial N_{it}^2} \end{bmatrix} = \begin{bmatrix} -d_tN_{it} & X_t \\ 0 & 0 \\ 0 & -\mu_{it}/N_{it} \\ \alpha^t N_{it}^{1-\varepsilon}u''(c_{it}) & -\varepsilon\alpha^t N_{it}^{-\varepsilon}u'(c_{it}) \\ -\varepsilon\alpha^t N_{it}^{-\varepsilon}u'(c_{it}) & -\varepsilon(1-\varepsilon)\alpha^t N_{it}^{-1-\varepsilon}u(c_{it}) \end{bmatrix}$$

because now (49) implies $-\lambda_{it-1}d_{t-1} + \lambda_{it}d_{t-1} = -\mu_{it}/N_{it}$. Negative definition requires $(-1)^3 \det H_5 > 0$, that is $\det H_5 < 0$. We notice that

$$\begin{aligned} & \det H_5 \\ = & -\det \begin{bmatrix} 0 & -d_tN_{it} & X_t \\ -d_tN_{it} & \alpha^t N_{it}^{1-\varepsilon}u''(c_{it}) & -\varepsilon\alpha^t N_{it}^{-\varepsilon}u'(c_{it}) \\ X_t & -\varepsilon\alpha^t N_{it}^{-\varepsilon}u'(c_{it}) & -\varepsilon(1-\varepsilon)\alpha^t N_{it}^{-1-\varepsilon}u(c_{it}) \end{bmatrix} \\ = & \alpha^t N_{it}^{1-\varepsilon}u''(c_{it})X_t^2 - 2d_t\alpha^t\varepsilon N_{it}^{1-\varepsilon}u'(c_{it})X_t - d_t^2\alpha^t(1-\varepsilon)\varepsilon N_{it}^{1-\varepsilon}u(c_{it}) \\ \equiv & P(X_t) \end{aligned}$$

Computing the discriminant Δ of this second order polynomial, we obtain:

$$\Delta = \varepsilon\left(2d_t\alpha^t N_{it}^{1-\varepsilon}\right)^2\left[\varepsilon u'(c_{it})^2 + (1-\varepsilon)u(c_{it})u''(c_{it})\right]$$

If $\Delta < 0$, then $\det H_5 < 0$ for all X_t, i.e. the second-order condition for a local maximum is satisfied. The condition $\Delta < 0$ is equivalent to (52). As seen above, since $u(c_i) \equiv c_i^{\sigma}$, (52) becomes $\sigma < 1-\varepsilon$, which is verified under Assumption 1. ■

8.3 Local dynamics when $b_0 = b_1$

Before addressing the local stability issue in the more general case of heterogenous costs per surviving child, we focus on the homogeneous case. We start by computing the reduced dynamic system.

When $b = b_0 = b_1$, equation (22) becomes:

$$\alpha R \left(\frac{Rb - g_{t+1}}{Rb - g_t} g_t \right)^{\sigma - 1} = m_{it}^{\varepsilon} = m_t^{\varepsilon} \tag{53}$$

where $m_t \equiv m_{0t} = m_{1t}$. Moreover, equation (4) simplifies to $N_{it} \equiv N_{i0} \prod_{s=0}^{t-1} m_s$ and capital-labor ratio (14) to:

$$\kappa_t = \frac{\mu_0 k_{0t} + \mu_1 k_{1t}}{1 - b m_t} \tag{54}$$

We deduce that (21) rewrites:

$$m_t k_{it+1} - R k_{it} = \left[1 - b m_t + \frac{\sigma}{1 - \varepsilon - \sigma} \left(1 - b \frac{R}{g_t} \right) \right] (1 - s) A \kappa_t \tag{55}$$

Multiplying both the sides of (55) by μ_i, aggregating over $i = 0, 1$ and using (54), we get:

$$m_t g_{t+1} (1 - b m_{t+1}) = [R + (1 - s) A] (1 - b m_t) + \frac{(1 - s) A \sigma}{1 - \varepsilon - \sigma} \left(1 - b \frac{R}{g_t} \right) \tag{56}$$

Equations (53) and (56) form a two-dimensional system, without predetermined variable, which rules the dynamics of the sequence (g_t, m_t), for $t = 1, \ldots$

Linearizing (53) and (56) around the steady state, given by (24) and (29), we obtain:

$$\frac{dg_{t+1}}{g} = B_1 \frac{dg_t}{g} + (B_1 - 1) B_4 \frac{dm_t}{m_t} \tag{57}$$

$$(B_0 - 1) \frac{dg_{t+1}}{g} + B_0 \frac{dm_{t+1}}{m} = \frac{B_0 - 1}{B_0} \frac{B_1}{B_1 - 1} B_3 \frac{dg_t}{g} + (1 + B_2) \frac{dm_t}{m} \tag{58}$$

where

$$B_0 \equiv bm \in (0, 1) \tag{59}$$

$$B_1 \equiv b \frac{R}{g} > 1 \tag{60}$$

$$B_2 \equiv b \frac{(1 - s) A}{g} > 0 \tag{61}$$

$$B_3 \equiv B_1 - B_0 + B_2 > 0 \tag{62}$$

$$B_4 \equiv \frac{\varepsilon}{1 - \sigma} \in (0, 1) \tag{63}$$

The trace T and the determinant D of the associated Jacobian matrix are given by:

$$D = \frac{B_1}{B_0} \left[1 + B_3 \left(1 + B_4 \frac{1 - B_0}{B_0} \right) \right] > 1$$

$$T = 1 + \frac{1 + B_3}{B_0} + (B_1 - 1) \left(1 + B_4 \frac{1 - B_0}{B_0} \right) > 2$$

In addition, we have:

$$D = T - 1 + (B_1 - 1)\frac{B_3}{B_0} + \left[(B_1 - 1)(1 - B_4) + \frac{B_1 B_3 B_4}{B_0}\right]\frac{1 - B_0}{B_0} > T - 1$$

Inequalities $D > T - 1 > 1$ entail that the steady state is a source. Since both the dynamic variables g_t and m_t are non-predetermined, the equilibrium is locally unique under rational expectations, i.e. one jumps on the steady state (24) and (29). We summarize this proof in a proposition.

Proposition 8 *Under Assumptions 1-4 and $b_0 = b_1$, the equilibrium is locally unique under rational expectations.*

8.4 Local dynamics when $b_0 < b_1$

Focus now on the case of heterogenous costs per surviving child. As in the previous section, we first derive the reduced dynamic system which drives the dynamics.

We introduce four normalized variables:

$$\eta_{it} \equiv N_{it}/m^t \tag{64}$$

$$\pi_{it} \equiv k_{it}/\kappa_t \tag{65}$$

for $i = 0, 1$, where $m = \left(\alpha R g^{\sigma-1}\right)^{1/\varepsilon}$ is the number of surviving children along the BGP.

Substituting (23) in (21), we find:

$$m_{it} g_{t+1} \pi_{it+1} = R\pi_{it} + (1 - s) A \left[1 - b_i m_{it} + \frac{\sigma}{1 - \varepsilon - \sigma}\left(1 - b_i \frac{R}{g_t}\right)\right] \tag{66}$$

Moreover, using (64) and (65), (14) rewrites:

$$1 = \frac{\eta_{0t}\pi_{0t} + \eta_{1t}\pi_{1t}}{\eta_{0t}(1 - b_0 m_{0t}) + \eta_{1t}(1 - b_1 m_{1t})} \tag{67}$$

while (4) gives:

$$m_{it} = m\eta_{it+1}/\eta_{it} \tag{68}$$

Substituting (68) in (22), (66) and (67), we get a five-dimensional system driving the sequence $(g_t, \pi_{0t}, \pi_{1t}, \eta_{0t}, \eta_{1t})$, for $t = 1, \ldots$:

$$\left(m\frac{\eta_{it+1}}{\eta_{it}}\right)^{\varepsilon} = \alpha R\left(\frac{Rb_i - g_{t+1}}{Rb_i - g_t} g_t\right)^{\sigma-1}, \text{ for } i = 0, 1 \tag{69}$$

$$m g_{t+1}\pi_{it+1}\frac{\eta_{it+1}}{\eta_{it}} = R\pi_{it} + (1 - s) A \left[1 - m b_i \frac{\eta_{it+1}}{\eta_{it}} + \frac{\sigma}{1 - \varepsilon - \sigma}\left(1 - b_i \frac{R}{g_t}\right)\right], \text{ for } i = 0, 1 \tag{70}$$

$$1 = \frac{\eta_{0t}\pi_{0t} + \eta_{1t}\pi_{1t}}{\eta_{0t} - m b_0 \eta_{0t+1} + \eta_{1t} - m b_1 \eta_{1t+1}} \tag{71}$$

We notice that η_{0t} and η_{1t} inherit the property of predetermined variable from N_{0t} and N_{1t}, while g_t, π_{0t}, π_{1t} are independently non-predetermined variables.

Before analyzing local dynamics, we reexamine the steady state using the dynamic system (69)-(71). Equation (69) defines $m(g) \equiv \left(\alpha R g^{\sigma-1}\right)^{1/\varepsilon}$, as in (24), while (70) determines π_i as a function of g:

$$\pi_i(g) = (1-s)\,A \frac{1 - b_i m(g) + \frac{\sigma}{1-\varepsilon-\sigma}\left(1 - b_i \frac{R}{g}\right)}{g m(g) - R}, \text{ for } i = 0, 1$$

Finally, along the BGP, (71) becomes:

$$1 = \frac{\mu_0 \pi_0(g) + \mu_1 \pi_1(g)}{1 - b m(g)} \tag{72}$$

and gives the stationary growth factor g as a function of the initial population shares (μ_0, μ_1).

Linearizing (69)-(71) around the steady state gives:

$$\frac{dg_{t+1}}{g} + (1 - B_{1i}) B_4 \frac{d\eta_{it+1}}{\eta_i} = B_{1i} \frac{dg_t}{g} + (1 - B_{1i}) B_4 \frac{d\eta_{it}}{\eta_i} \tag{73}$$

$$\pi_i \frac{dg_{t+1}}{g} + \pi_i \frac{d\pi_{it+1}}{\pi_i} + (\pi_i + B_{2i}) \frac{d\eta_{it+1}}{\eta_i} = \frac{B_1 B_{2i} B_5}{B_0} \frac{dg_t}{g} + \frac{\pi_i B_1}{B_0} \frac{d\pi_{it}}{\pi_i} + (\pi_i + B_{2i}) \frac{d\eta_{it}}{\eta_i} \tag{74}$$

for $i = 0, 1$, and

$$B_{60} \frac{d\eta_{0t+1}}{\eta_0} + B_{61} \frac{d\eta_{1t+1}}{\eta_1} = -\frac{\pi_0 B_{70}}{1-\pi_0} \frac{d\pi_{0t}}{\pi_0} - \frac{\pi_1 B_{71}}{1-\pi_1} \frac{d\pi_{1t}}{\pi_1} + B_{70} \frac{d\eta_{0t}}{\eta_0} + B_{71} \frac{d\eta_{1t}}{\eta_1} \tag{75}$$

with

$$B_{1i} \equiv b_i \frac{R}{g}, \; B_{2i} \equiv b_i \frac{(1-s)\,A}{g}, \; B_5 \equiv \frac{\sigma}{1-\varepsilon-\sigma},$$

$$B_{6i} \equiv \frac{b_i \eta_i}{b_0 \eta_0 + b_1 \eta_1}, \; B_{7i} \equiv \frac{(1-\pi_i)\,\eta_i}{(1-\pi_0)\,\eta_0 + (1-\pi_1)\,\eta_1}$$

According to the our numerical computations presented below, the associated Jacobian matrix presents two unit roots and three eigenvalues outside the unit circle.[19] The existence of two unit roots ensues from the persistence of changes in η_{0t} and η_{1t} and rests on the nature of the BGP which now depends on the initial conditions, namely those on the shares of population (μ_0, μ_1) (see (72)). Initial distributions of population and wealth, jointly with (69)-(71), uniquely determine the BGP. So, the shape of the BGP is informative about these initial distributions. This kind of informational persistence results in the existence of two unit roots.

[19] Notice that the system is non-hyperbolic and the Hartman-Großman Theorem no longer applies.

8.5 Numerical computations for local dynamics when $b_0 < b_1$

Let us set the parameters of the model to fit the empirical features of western economies during the last decades. We consider 40 years as length of a period (the adulthood of a generation). We choose for α[20] a standard value for discounting according to the RBC literature: $\alpha = 0.2 \approx 0.99^{160}$ (0.99 per quarter).[21] In order to take be compatible with the restriction $\sigma < 1 - \varepsilon$, we fix $\varepsilon = 0.6$ and $\sigma = 0.3$. The survival probability for a child is set equal to $\gamma_0 = 0.9$ for the privileged dynasty and to $\gamma_1 = 0.8$ for the disadvantaged one. We set common values $a_0 = 0.1$ and $a_1 = 0.2$ for both the classes: in this case, the affine form $\beta_i = a_0 + a_1\gamma_i$ introduced by Doepke (2005) fits the evidence on the opportunity cost of a child in terms of parents' time endowment found by Haveman and Wolfe (1995). In the spirit of Mankiw, Romer and Weil (1992), we consider that K_t represents a mix of physical and human capital. So, we accept a value of $s = 2/3$ for the capital share in total income, while, taking into account that the human capital depreciates less than the physical one, we set the average depreciation rate to $\delta = 0.96$, that is about 8% per year (Aghion and Howitt (1998)). Finally, we calibrate the TFP $A = 15.3$ to obtain a yearly growth rate of about 2% which is in line with the value experienced by most of western countries in the last decades.

According to this parametrization, we obtain the following stationary values:[22] $g = 2.2089$, $m = 1.3102$, $n_0 = 1.4559$, $n_1 = 1.6378$, $\pi_0 = 0.5099$ and $\pi_1 = 0.6567$ with $\eta_0 = 0.9$ and $\eta_1 = 1.1$. We check also the positivity of labor supplies: $1 - \beta_0 n_0 = 0.5924$ and $1 - \beta_1 n_1 = 0.5742$.

Computing the eigenvalues of the Jacobian matrix of system (73)-(75), we find: $x_1 = x_2 = 1$, $x_3 = 3.1065$, $x_4 = 3.5382$, $x_5 = 5.7407$. Thus, there are two unit roots and three roots outside the unit circle, as seen in the previous section.

8.6 Proof of Lemma 2

On the one hand, using equation (33), we get:

$$\begin{aligned} \frac{\partial b}{\partial \gamma_0} &= -a_0 \frac{\mu_0}{\gamma_0^2} \leq 0 \\ \frac{\partial b}{\partial \gamma_1} &= -a_0 \frac{\mu_1}{\gamma_1^2} \leq 0 \end{aligned}$$

[20] Note that α represents the discount factor for one surviving child.

[21] That is an interest rate per year of about 4%.

[22] For notational parsimony, we show only 4 digits in results, but computations accuracy is of 16 digits.

On the other hand, using (34), we obtain:

$$\begin{aligned}\frac{\partial b}{\partial \gamma_A} &= -a_0\left(\frac{\mu_0}{\gamma_0^2}+\frac{\mu_1}{\gamma_1^2}\right)\leq 0\\ \frac{\partial b}{\partial \sigma_A} &= a_0\sqrt{\mu_0\mu_1}\left(\frac{1}{\gamma_1^2}-\frac{1}{\gamma_0^2}\right)\geq 0\end{aligned}$$

because $\gamma_1 < \gamma_0$. ■

8.7 Proof of Proposition 2

The BGP g is determined by equation (29) and we obtain, with some notational misuse:

$$\frac{\partial g}{\partial b}=\frac{g}{b}\frac{\frac{bR/g-bm}{1-bm}}{b\frac{R}{g}+\left(b\frac{R}{g}-1\right)\frac{bm}{1-bm}\frac{1-\sigma}{\varepsilon}+m\frac{1-bm}{(1-s)A/g}\frac{(1-\varepsilon-\sigma)^2}{\varepsilon\sigma}}>0$$

Then, using Lemma 2, we find:

$$\frac{\partial g}{\partial \gamma_0}=\frac{\partial g}{\partial b}\frac{\partial b}{\partial \gamma_0}<0 \text{ and } \frac{\partial g}{\partial \gamma_1}=\frac{\partial g}{\partial b}\frac{\partial b}{\partial \gamma_1}<0$$

In addition, we have

$$\frac{\partial g}{\partial \gamma_A}=-a_0\left(\frac{\mu_0}{\gamma_0^2}+\frac{\mu_1}{\gamma_1^2}\right)\frac{\partial g}{\partial b}<0 \text{ and } \frac{\partial g}{\partial \sigma_A}=a_0\sqrt{\mu_0\mu_1}\left(\frac{1}{\gamma_1^2}-\frac{1}{\gamma_0^2}\right)\frac{\partial g}{\partial b}>0$$

■

8.8 Proof of Proposition 3

Differentiating $m=\left(\alpha R g^{\sigma-1}\right)^{1/\varepsilon}$ (see (24)), we obtain:

$$\frac{\partial m}{\partial b}=-\frac{1-\sigma}{\varepsilon}\frac{m}{g}\frac{\partial g}{\partial b}$$

Applying Lemma 2, we find the following results:

$$\begin{aligned}\frac{\partial m}{\partial \gamma_i} &= \frac{\partial m}{\partial b}\frac{\partial b}{\partial \gamma_i}=-\frac{1-\sigma}{\varepsilon}\frac{m}{g}\frac{\partial g}{\partial b}\frac{\partial b}{\partial \gamma_i}=a_0\frac{1-\sigma}{\varepsilon}\frac{\mu_i}{\gamma_i^2}\frac{m}{g}\frac{\partial g}{\partial b}>0\\ \frac{\partial m}{\partial \gamma_A} &= \frac{\partial m}{\partial b}\frac{\partial b}{\partial \gamma_A}=a_0\frac{1-\sigma}{\varepsilon}\left(\frac{\mu_0}{\gamma_0^2}+\frac{\mu_1}{\gamma_1^2}\right)\frac{m}{g}\frac{\partial g}{\partial b}>0\\ \frac{\partial m}{\partial \sigma_A} &= \frac{\partial m}{\partial b}\frac{\partial b}{\partial \sigma_A}=a_0\sqrt{\mu_0\mu_1}\frac{1-\sigma}{\varepsilon}\left(\frac{1}{\gamma_0^2}-\frac{1}{\gamma_1^2}\right)\frac{m}{g}\frac{\partial g}{\partial b}<0\end{aligned}$$

■

8.9 Proof of Proposition 4

Using $n_i = m/\gamma_i$, we obtain:

$$\begin{aligned}
\frac{\partial n_i}{\partial \gamma_i} &= \frac{1}{\gamma_i}\frac{\partial m}{\partial \gamma_i} - \frac{m}{\gamma_i^2} = \frac{m}{\gamma_i^2}\left(a_0 \frac{1-\sigma}{\varepsilon}\frac{\mu_i}{\gamma_i}\frac{1}{g}\frac{\partial g}{\partial b} - 1\right) \\
\frac{\partial n_i}{\partial \gamma_j} &= \frac{1}{\gamma_i}\frac{\partial m}{\partial \gamma_j} = a_0 \frac{1-\sigma}{\varepsilon}\frac{1}{\gamma_i}\frac{\mu_j}{\gamma_j^2}\frac{m}{g}\frac{\partial g}{\partial b}
\end{aligned}$$

Since $db/dg > 0$, we deduce that $\partial n_i/\partial \gamma_j > 0$ for a strictly positive a_0 and $\partial n_i/\partial \gamma_i < 0$ for a sufficiently weak a_0.

Using

$$\gamma_1 = \gamma_A - \sigma_A \sqrt{\frac{\mu_1}{\mu_0}} < \gamma_0 = \gamma_A + \sigma_A \sqrt{\frac{\mu_0}{\mu_1}}$$

we get:

$$\frac{\partial \gamma_i}{\partial \gamma_A} = 1,\ \frac{\partial \gamma_0}{\partial \sigma_A} = \sqrt{\frac{\mu_0}{\mu_1}},\ \frac{\partial \gamma_1}{\partial \sigma_A} = -\sqrt{\frac{\mu_1}{\mu_0}}$$

and, eventually,

$$\begin{aligned}
\frac{\partial n_i}{\partial \gamma_A} &= \frac{\partial n_i}{\partial m}\frac{\partial m}{\partial \gamma_A} + \frac{\partial n_i}{\partial \gamma_i}\frac{\partial \gamma_i}{\partial \gamma_A} = \frac{1}{\gamma_i}\frac{\partial m}{\partial \gamma_A} - \frac{m}{\gamma_i^2}\frac{\partial \gamma_i}{\partial \gamma_A} \\
&= \frac{1-\sigma}{\varepsilon}\frac{a_0}{\gamma_i}\left(\frac{\mu_0}{\gamma_0^2} + \frac{\mu_1}{\gamma_1^2}\right)\frac{m}{g}\frac{\partial g}{\partial b} - \frac{m}{\gamma_i^2}
\end{aligned}$$

which implies that $\partial n_i/\partial \gamma_A < 0$ for a sufficiently weak a_0.

Finally, we compute:

$$\frac{\partial n_i}{\partial \sigma_A} = \frac{\partial n_i}{\partial m}\frac{\partial m}{\partial \sigma_A} + \frac{\partial n_i}{\partial \gamma_i}\frac{\partial \gamma_i}{\partial \sigma_A} = \frac{1}{\gamma_i}\frac{\partial m}{\partial \sigma_A} - \frac{m}{\gamma_i^2}\frac{\partial \gamma_i}{\partial \sigma_A}$$

which gives, more explicitly,

$$\begin{aligned}
\frac{\partial n_0}{\partial \sigma_A} &= \frac{1}{\gamma_0} a_0 \sqrt{\mu_0 \mu_1}\frac{1-\sigma}{\varepsilon}\left(\frac{1}{\gamma_0^2} - \frac{1}{\gamma_1^2}\right)\frac{m}{g}\frac{\partial g}{\partial b} - \frac{m}{\gamma_0^2}\sqrt{\frac{\mu_0}{\mu_1}} \\
\frac{\partial n_1}{\partial \sigma_A} &= \frac{1}{\gamma_1} a_0 \sqrt{\mu_0 \mu_1}\frac{1-\sigma}{\varepsilon}\left(\frac{1}{\gamma_0^2} - \frac{1}{\gamma_1^2}\right)\frac{m}{g}\frac{\partial g}{\partial b} + \frac{m}{\gamma_1^2}\sqrt{\frac{\mu_1}{\mu_0}}
\end{aligned}$$

It ensues that $\partial n_0/\partial \sigma_A < 0$, whereas $\partial n_1/\partial \sigma_A > 0$ for a sufficiently low a_0. ■

8.10 Proof of Proposition 5

We notice that A affects $R = 1 - \delta + sA$ and therefore m (see equation (24)), but has no impact on the average cost b. With some notational misuse, the BGP satisfies:

$$R - gm(g, R) = (1-s)A\left[\frac{\sigma}{1-\varepsilon-\sigma}\frac{b\frac{R}{g} - 1}{1 - bm(g, R)} - 1\right] \tag{76}$$

with $m \equiv m(g,R) = \left(\alpha R g^{\sigma-1}\right)^{1/\varepsilon}$. Differentiating (76) with respect to A and g, we obtain:

$$\frac{A}{g}\frac{dg}{dA} = \frac{\left[\frac{sA}{R}\frac{b\frac{R}{g}+\left(b\frac{R}{g}-1\right)\frac{bm}{1-bm}\frac{1}{\varepsilon}}{1-bm} + \frac{b\frac{R}{g}-1}{1-bm} - \frac{1-\varepsilon-\sigma}{\sigma}\right] + \frac{1-\varepsilon-\sigma}{\varepsilon\sigma}\frac{s}{1-s}\frac{gm-\varepsilon R}{R}}{\frac{b\frac{R}{g}+\left(b\frac{R}{g}-1\right)\frac{bm}{1-bm}\frac{1-\sigma}{\varepsilon}}{1-bm} + \frac{(1-\varepsilon-\sigma)^2}{\varepsilon\sigma}\frac{gm}{(1-s)A}}$$

Transversality condition along the BGP writes $R > gm(g,R)$. Hence, equation (76) implies:

$$\frac{b\frac{R}{g}-1}{1-bm} > \frac{1-\varepsilon-\sigma}{\sigma}$$

Therefore, a sufficient condition to have $dg/dA > 0$ is $gm > \varepsilon R$.

Still using $m = \left(\alpha R g^{\sigma-1}\right)^{1/\varepsilon}$, we find:

$$\begin{aligned}
&\frac{A}{m}\frac{\partial m}{\partial A} = \frac{1}{\varepsilon}\frac{sA}{R} - \frac{1-\sigma}{\varepsilon}\frac{A}{g}\frac{dg}{dA} \\
= \quad & \frac{1}{\varepsilon}\frac{sA}{R} - \frac{1-\sigma}{\varepsilon}\frac{\left[\frac{sA}{R}\frac{b\frac{R}{g}+\left(b\frac{R}{g}-1\right)\frac{bm}{1-bm}\frac{1}{\varepsilon}}{1-bm} + \frac{b\frac{R}{g}-1}{1-bm} - \frac{1-\varepsilon-\sigma}{\sigma}\right] + \frac{1-\varepsilon-\sigma}{\varepsilon\sigma}\frac{s}{1-s}\frac{gm-\varepsilon R}{R}}{\frac{b\frac{R}{g}+\left(b\frac{R}{g}-1\right)\frac{bm}{1-bm}\frac{1-\sigma}{\varepsilon}}{1-bm} + \frac{(1-\varepsilon-\sigma)^2}{\varepsilon\sigma}\frac{gm}{(1-s)A}}
\end{aligned}$$

The impact of A on m can be positive or negative. Since

$$\frac{A}{n_i}\frac{\partial n_i}{\partial A} = \frac{A}{m}\frac{\partial m}{\partial A}$$

the same conclusion holds for n_0 and n_1. ■

8.11 Proof of Proposition 6

Deriving (44) with respect to z and g, we obtain:

$$\begin{aligned}
\frac{\partial \omega}{\partial z} &= (1-s)A\left[\left(b\frac{R}{g}-1\right)\frac{\sigma}{1-\varepsilon-\sigma} + bm - 1\right] \\
\frac{\partial \omega}{\partial g} &= m\left((1-bm)\frac{1-\varepsilon-\sigma}{\varepsilon} + \left[b\frac{R+(1-z)(1-s)A}{g} - bm\right]\frac{1-\sigma}{\varepsilon}\right) \\
&\quad + b\frac{R}{g}\frac{(1-z)(1-s)A}{g}\frac{\sigma}{1-\varepsilon-\sigma}
\end{aligned}$$

We notice that $\partial\omega/\partial g > 0$, while $\partial\omega/\partial z > 0$ if and only if $h(g) < \sigma/(1-\varepsilon-\sigma)$. Hence,

$$\frac{\partial g}{\partial z} = -\frac{\partial\omega/\partial z}{\partial\omega/\partial g} < 0 \text{ if and only if } h(g) < \frac{\sigma}{1-\varepsilon-\sigma}$$

Under $b_0 < b_1$ and Assumption 4, we have $0 < h(\underline{g}) < h(\overline{g})$. Therefore, (1) if $\sigma/(1-\varepsilon-\sigma) < h(\underline{g})$, then $\sigma/(1-\varepsilon-\sigma) < h(g)$ for all $g \in (\underline{g}, \overline{g})$,

which implies $\partial g/\partial z > 0$. (2) If $h\left(\underline{g}\right) < \sigma/\left(1-\varepsilon-\sigma\right) < h\left(\overline{g}\right)$, then there exists $g^* \in \left(\underline{g}, \overline{g}\right)$ such that (2.1) $h\left(g\right) < \sigma/\left(1-\varepsilon-\sigma\right)$, that is $\partial g/\partial z < 0$, for all $g \in \left(\underline{g}, g^*\right)$ and (2.2) $\sigma/\left(1-\varepsilon-\sigma\right) < h\left(g\right)$, that is $\partial g/\partial z > 0$, for all $g \in \left(g^*, \overline{g}\right)$. (3) If $h\left(\overline{g}\right) < \sigma/\left(1-\varepsilon-\sigma\right)$, then $h\left(g\right) < \sigma/\left(1-\varepsilon-\sigma\right)$, that is $\partial g/\partial z < 0$, for all $g \in \left(\underline{g}, \overline{g}\right)$.

Since

$$\frac{\partial m}{\partial z} = -\frac{1-\sigma}{\varepsilon}\frac{m}{g}\frac{\partial g}{\partial z} \text{ and } \frac{\partial n_i}{\partial z} = -\frac{1}{\gamma_i}\frac{1-\sigma}{\varepsilon}\frac{m}{g}\frac{\partial g}{\partial z}$$

we have that $\partial m/\partial z > 0$ or $\partial n_i/\partial z > 0$ if and only if $\partial g/\partial z < 0$.

Without heterogeneity $b_0 = b_1 = b$ and

$$0 = h\left(\underline{g}\right) < \frac{\sigma}{1-\sigma-\varepsilon} < h\left(\overline{g}\right) = +\infty$$

Therefore, case (2) still applies. ■

8.12 Proof of Proposition 7

From (43), we have, along the BGP,

$$\frac{k_{0t}/\kappa_t}{1-bm} = \frac{\left(1-z\right)\left(1-s\right)A}{R-gm}\frac{1-b_0 m}{1-bm}\left(\frac{\sigma}{1-\varepsilon-\sigma}\frac{b_0\frac{R}{g}-1}{1-b_0 m}-1\right)$$

Using this expression, we are able to compute the derivative of G with respect to z as follows:

$$\frac{\partial G}{\partial z} = \left(\mu_0 - G\right)\left(\frac{1}{1-z} + \frac{H}{g}\frac{\partial g}{\partial z}\right)$$

where $\mu_0 > G$ and $H > 0$. The proposition immediately follows. ■

References

[1] Aghion, P., E. Caroli and C. Garcia-Penalosa (1999), Inequality and Economic Growth: The Perspective of the New Growth Theories, *Journal of Economic Literature* **37**, 1615-1660.

[2] Aghion, P. and P. Howitt (1998), *Endogenous Growth Theory*, MIT Press.

[3] Barro, R. J. (1990), Government Spending in a Simple Model of Endogenous Growth, *Journal of Political Economy* **98**, s103-s125.

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

[5] Becker, G. S. and R. J. Barro (1986), A Reformulation of the Economic Theory of Fertility, NBER Working Paper 1793.

[6] Becker, G. S. and R. J. Barro (1988), A Reformulation of the Economic Theory of Fertility, *The Quarterly Journal of Economics* **103**, 1-25.

[7] Becker, R. A. (1980), On the Long-Run Steady State in a Simple Dynamic Model of Equilibrium with Heterogeneous Households, *Quarterly Journal of Economics* **95**, 375-382.

[8] Becker, R. A. (2006), Equilibrium Dynamics with Many Agents, in *Handbook on Optimal Growth* (R.- A. Dana, C. Le Van, T. Mitra and K. Nishimura, eds.), Springer.

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

[10] de la Croix, D. and A. Sommacal (2009), A Theory of Medical Effectiveness, Differential Mortality, Income Inequality and Growth for Pre-Industrial England, *Mathematical Population Studies* **16**, 2-35.

[11] Dahan, M. and D. Tsiddon (1998), Demographic Transition, Income Distribution, and Economic Growth, *Journal of Economic Growth* **3**, 29-52.

[12] Doepke, M. (2005), Child Mortality and Fertility Decline: Does the Barro-Becker Model Fit the Fact?, *Journal of Population Economics* **18**, 337-366.

[13] Frankel, M. (1962), The Production Function in Allocation and Growth: A Synthesis, *American Economic Review* **52**, 996-1022.

[14] Galor, O. (2005), From Stagnation to Growth: Unified Growth Theory, in *Handbook of Economic Growth* (P. Aghion and S. N. Durlauf, eds.), Elsevier.

[15] Haveman, R. and B. Wolfe (1995), The Determinants of Children's Attainments: A Review of Methods and Findings, *Journal of Economic Literature* **33**, 1829-1878.

[16] Ljungqvist, L. and T. J. Sargent (2004), *Recursive macroeconomic theory*, Cambridge Ma: MIT Press.

[17] Mankiw, N. G., D. Romer and D. N. Weil (1992), A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics* **107**, 407-437.

[18] Nerlove, M. and L. K. Kraut (1997), Growth Models with Endogenous Population: A General Framework, in *Handbook of Population and Family Economics* (M. R. Rosenzweig and O. Stark, eds.), Elsevier.

[19] Ramsey, F. P. (1928), A Mathematical Theory of Saving, *Economic Journal* **38**, 543-559.