



Munich Personal RePEc Archive

Heterogeneous human capital, inequality and growth: the role of patience and skills

Borissov, Kirill and Bosi, Stefano and Ha-Huy, Thai and Modesto, Leonor

European University of St. Petersburg, EPEE, Univ Evry, Université Paris-Saclay, EPEE, Univ Evry, Université Paris-Saclay, UCP, Catolica Lisbon School of Business and Economics

April 2018

Online at <https://mpra.ub.uni-muenchen.de/86314/>
MPRA Paper No. 86314, posted 24 Apr 2018 08:44 UTC

Heterogeneous human capital, inequality and growth: the role of patience and skills

Kirill Borissov^a, Stefano Bosi^b,
Thai Ha-Huy^c and Leonor Modesto^d

^aEuropean University of St. Petersburg

^bEPEE, Univ Evry, Université Paris-Saclay

^cEPEE, Univ Evry, Université Paris-Saclay; TIMAS, Thang Long University

^dUCP, Católica Lisbon School of Business and Economics

April 22, 2018

Abstract

We extend the Lucas' 1988 model introducing two classes of agents with heterogeneous skills, discount factors and initial human capital endowments. We consider two regimes according to the planner's political constraints. In the meritocratic regime, the planner faces individual constraints. In the redistributive regime, the planner faces an aggregate constraint. We find that heterogeneity matters, particularly with redistribution. In the meritocratic regime, the optimal solution coincides with the BGP found by Lucas (1988) for the representative agent's case. In contrast, in the redistribution case, the solution for time devoted to capital accumulation is never interior for both agents. Either the less talented agents do not accumulate human capital or the more skilled agents do not work. Moreover, social welfare under the redistribution regime is always higher than under meritocracy and it is optimal to exploit existing differences. Finally, we find that inequality in human capital distribution increases in time and that, in the long run, inequality always promotes growth.

Keywords: human capital, heterogenous patience and skills, inequality and growth. **JEL codes:** J24, O15, O40.

1 Introduction

Though the consideration of human capital as a source of economic growth and development goes back at least to Adam Smith (1776),¹ for a long time the concept of human capital was dormant. Interest in the economic role of human capital was awakened in the late 1950s and early 1960s through the writings of Mincer (1958, 1962), Schultz (1961, 1962) and Becker (1962, 1964), who rekindled this concept by emphasizing its importance in explaining earnings differentials and its links with economic growth.

Human capital as an engine of growth was incorporated into growth theory by Uzawa (1965). The emergence of a new endogenous growth literature stimulated the interest of economists in the role of human capital. In his (1988) paper, Lucas shows that the growth rate of per capita income depends on the growth rate of human capital, which in turn depends on the time individuals use for acquiring skills. In a more recent work Manuelli and Seshadri (2014) “suggest that human capital has a central role in determining the wealth of nations and that the quality of human capital varies systematically with the level of development”.² Lucas (2015) proposes a simple model where “all growth is driven by schooling and on-the-job learning”.

Most of endogenous growth models with human capital accumulation assume a representative agent, which is only a fair approximation if income and wealth inequality play a negligible role in the process of economic development.³ However, it is widely recognized that inequality has a strong impact on economic growth, although it is not clear if this impact is positive or negative. The accumulation of savings (Stiglitz, 1969; Bourguignon, 1981), unobservable effort (Rebelo, 1991) and the investment project size (Barro, 2000) represent the main channels through which inequality may enhance growth. On the other hand, in the presence of credit market imperfections, inequality has a negative impact on investment (Bénabou, 1996 and 2000) and entrepreneurial activity (Banerjee and Newman, 1993), and hence on

¹Adam Smith included in the capital stock of a nation the inhabitants’ acquired and useful talents, and noticed that human skills increase wealth for society.

²Illustrative accounting estimates by Jones (2014) “suggest that human capital variation can be substantially amplified, including to the point where capital variation could possibly fully account for cross-country income differences”.

³Glomm and Ravikumar (1992) also study the evolution of a (lognormal) distribution of human capital over time in an endogenous growth model, but, unlike us, they consider overlapping generations.

economic growth. Empirical studies are also generally inconclusive. While Alesina and Rodrik (1994), Persson and Tabellini (1994) and Perotti (1996) show a negative relationship between inequality and growth, more recent works by Partridge (1997), Forbes (2000) and Frank (2009) find a positive relationship.⁴

Economists have also paid attention to the relationship between the accumulation of human capital and inequality. Becker and Tomes (1979), Viaene and Zilcha (2003) and Galor and Moav (2004) emphasize educational attainment as one of the causes of greater income inequality. The latter authors find that when human capital replaces physical capital as the driving force of economic growth, a greater equality improves the growth performance. Galor and Zeira (1993) and Banerjee and Newman (1993) identify credit market constraints as the channel relating the accumulation of human capital and inequality. Galor and Ozak (2016) establishes significant positive effect of the geographical origins of time preference on human capital. Bénabou (2002) compares educational and fiscal redistributions and finds that the former is better for growth. Eicher and García-Peñalosa (2001) predict a non-monotonic relationship between educational attainment and inequality, which drives the lack of a clear relation between inequality and growth. More recently, Turnovsky (2011) and Turnovsky and Mitra (2013) propose two-sector endogenous growth models linking human capital accumulation and income inequality. They find that an increase in the growth rate resulting from productivity enhancement in the human capital sector will be accompanied by an increase in inequality, whereas a productivity boost in the final output sector results in a reduction in inequality.

In this paper we introduce a model of human capital accumulation and economic growth with heterogeneous agents. Not only we assume that agents are heterogeneous in terms of (1) human capital endowments, but also in term of (2) patience and (3) cognitive skills. Moreover, we assume that agents with higher cognitive skills are more patient.

There is an empirical literature about the heterogeneity in patience and cognitive skills, and the relationship between patience and skills.

In almost all growth models with infinitely-lived agents, patience plays a key role. In exogenous growth models with physical capital, higher patience implies a higher propensity to save and hence a higher steady state

⁴Barro (2000) finds that the effect of income inequality on economic growth can differ between poor and rich economies.

stock of physical capital. In the context of endogenous growth models with human capital accumulation, higher patience implies greater incentives to devote time to the acquisition of skills and hence leads to higher rates of growth. Recent results of Hübner and Vannoorenberghe (2015) suggest that increasing patience by one standard deviation raises per-capita income by between 34% and 78%. Dohmen et al. (2015) show that average patience explains a considerable fraction of the variation in growth rates both in the medium run and in the long run and about 40% of the between-country variation in income. Their results establish that, within countries, average patience in geographical regions predicts average years of education and, in individual-level analyses, that individual patience predicts educational attainment within countries and regions. Patience varies not only between countries, but also within countries. Falk et al. (2015), using the first global survey on measuring preferences, which includes 80000 individuals from 76 countries, that represent 90 percent of world population and income, find that between-country variance accounts for about 13.5 percent of total variation in patience. These authors also find a positive relation between patience and cognitive skills which "is statistically significant in almost all countries. This reveals that the relationship is relatively universal, and arguably not the product of institutions or specific educational and value systems".

There is a literature on models with infinitely-lived agents heterogeneous in their discount factor (see a very good survey by Becker (2006)). At the same time, to the best of our knowledge, all models with human capital accumulation assume either a representative agent or agents with an identical discount factor. In fact, the only type of heterogeneity considered in these models was heterogeneity in initial human capital.⁵

Suen (2014) also addresses the issue of the relationship between heterogeneity in time preference and economic inequality in a market economy with human and physical capital accumulation. Preference for wealth prevents the economy from collapsing into a degenerate distribution of capital, while human capital formation generates a positive correlation between earnings and capital income. Differently from Suen (2014), we consider the planner's solution and we leave aside the physical capital accumulation.

A recent empirical literature emphasizes the increasing role of cognitive skills in promoting economic well-being and wage determination (see Burks et al. (2009) and Hanushek and Woessmann (2008) among others).

⁵See Turnovsky (2011) and Turnovsky and Mitra (2013) cited above.

Hanushek and Woessmann (2012) introduce direct measures of cognitive skills from international tests of mathematics and science and show that the achievement of the population is closely related to cognitive skills and, importantly, that a causal interpretation is likely warranted. In particular, they regress growth on initial levels of GDP and international test scores over the period 1960-2000 for the set of 50 countries and obtain a very good results: their simple model can explain three-quarters of the variance in growth rates.

Cognitive skills play an important role in individual economic performance and wage determination. A large literature has documented a significant increase in wage inequality, measured as the difference between the wage of skilled and unskilled workers, over the past decades. The existing literature has proposed the following explanations of the rise in skill premium: an increase in the relative demand for skills caused by skill biased technical change; product demand shifts across sectors with different skill intensities; a slowdown in the growth of the relative supply of skilled workers (comprehensive discussions of this literature are found in Lemieux (2010) and Autor et al. (2008)).

Lemieux (2006) notes that "explanations for the growth in wage inequality linked to standard human capital variables like experience and education are limited by the fact that these variables explain only about a third of the variance of wages" (p. 461). It seems reasonable to explain the residual wage inequality by heterogeneity on cognitive skills particularly against the backdrop of increasing role of human capital in economic growth. It should be noted that often the terms education and skill are used interchangeably. What evidence do we have that cognitive skills (numeracy and problem-solving) are rewarded more than educational credentials? Autor (2014) considers representative samples of adults in 22 OECD countries and finds that "cognitive skills are substantially rewarded in the labor market across all twenty-two economies. The average wage premium to a one "unit" (i.e., one standard deviation) increase in measured cognitive skills is 18 percent. In addition, cognitive earnings premia differ substantially across countries. The premium is below 13 percent in the Sweden, the Czech Republic, and Norway. It is above 20 percent in six countries. Third, the U.S. stands out as having the highest measured return to skill, with a premium of 28 percent per unit increment to cognitive ability. Concretely, comparing two U.S. workers who are one standard deviation above and one standard deviation below the population average of cognitive ability, respectively, we

would expect their full-time weekly earnings to differ by 50 to 60 percent."

It also should be noted that though the wage inequality literature has typically focused on wage differentials between education groups, wages also vary within education categories. It was noted that much of the overall increase in wage inequality has been due to an increase in inequality within education groups (see e.g. Katz (1999), Gosling et al. (2000) and Lemieux (2006)). Lindley and McIntosh (2015) investigate why such wage inequality has grown. They show that in U.K. most of the increase in graduate wage inequality has occurred within degree subjects, rather than between. Also they investigate the increase in the variance of childhood cognitive test scores amongst graduates in the same subject and the widening variety of jobs performed by graduates with degrees in the same subject as potential explanations of the increase in wage inequality. Lindley and McIntosh (2015, p. 101) find that "both of these factors have played a role in explaining growing graduate wage inequality within subjects, though the largest is by far from the increased variance of test scores."

We know less about the relationships between cognitive skills and patience, though the notion that more intelligent people are more patient (they discount future rewards less) was advanced by Rae (1834) and supported by recent studies.

The relationship between cognitive ability and patience has been tested in several (rather unsystematic) psychological studies surveyed in Frederick (2005). Considering short-term choices between monetary rewards, Frederick (2005) finds that those who scored higher on the Cognitive Reflection Test were generally more patient, but for longer horizons, temporal preferences seem weakly related or unrelated to the Cognitive Reflection Test scores. A study by Burks et al. (2009) shows that individuals with better cognitive skills are more patient, in both short and long run. Falk et al. (2015), using a new data set that covers 80000 individuals from 76 countries, find that patience is more pronounced among individuals with higher cognitive ability.

Our model focuses more on the basic mechanism of saving/investment in human capital in a model with central planner and heterogeneous agents that differ in their discount factors, their skills in accumulating capital and initial human capital endowments. There is no room for (credit) market imperfection as in Galor and Zeira (1993) and Bénabou (1996 and 2000) and, thus, conclusions are quite different: inequalities have a positive impact on growth as was the case in the seminal papers by Stiglitz (1969) and Bourguignon

(1981). More precisely, in order to better understand the trade-off between inequalities and growth, we extend the Lucas' 1988 model introducing two classes of agents with heterogenous skills, patience and initial human wealth,⁶ and we consider two regimes according to the planner's political constraints. In the first regime, that we call meritocracy, the planner faces individual constraints. In the second regime the planner redistributes the total product and, then, formally, faces an aggregate constraint. We find that heterogeneity matters, specially with redistribution. Indeed, although in the meritocratic regime the solution coincides with the representative agent's Lucas BGP, in the redistribution case the optimal solution for non-leisure time devoted to capital accumulation is never interior for both agents. Either the less talented agents do not accumulate human capital or the more skilled agents do not work. Obviously, the social welfare under the redistribution regime is always higher than under meritocracy because the planner's program is less constrained. Third, the redistribution of consumption is not affected by the distribution of skills, depending only on differences in patience. In contrast, the allocation of tasks takes into account skills differences. Finally we find that inequality in the distribution of human capital increases in time, and that inequality is associated with higher rates of growth in both regimes.

The rest of the paper is organized as follows. In the next section we present the fundamentals of the models considered. The meritocratic and the redistributive cases are analyzed in Sections 3 and 4 respectively. In Section 5 we compare both cases and further discuss our results. Finally, in Section 6, we provide some concluding remarks. Proofs and computations are relegated to the Appendix.

2 Fundamentals

The models considered follow closely the Lucas (1988) framework, extending it to account for agents heterogeneity. Since we want to focus on the role of heterogeneity in human capital accumulation on inequality and growth, we ignore technological change and physical capital accumulation.

We consider two classes of agents: patient and impatient households, labeled respectively by 0 and 1. They have different discount rates, ρ_i , $i = 0, 1$

⁶Note that differently from the Ramsey model with physical capital accumulation (Becker, 1980), our extensions exhibit long-run distributions of capital (heterogenous human capital) that are non-degenerated.

with $\rho_0 < \rho_1$.

We denote the individual labor supply by l_{it} and the size of each class by π_i . We normalize the size of the entire population to one so that $\pi_0 + \pi_1 = 1$. Then, l_t , aggregate labor supply is given by

$$l_t = \sum_{i=0}^1 \pi_i l_{it} \quad (1)$$

We consider a linear constant returns technology.

Assumption 1 *Technology is represented by a production function*

$$y_t = Al_t \quad (2)$$

where y_t denotes aggregate production and $A > 0$ is a scaling parameter.

Leisure time is exogenous. Non-leisure time is normalized to one and spent either working or accumulating human capital (education and health). Individual labor supply is given by the product of human capital, h_{it} , and working time, u_{it} , i.e.,

$$l_{it} \equiv h_{it} u_{it} \quad (3)$$

This means that both factors are necessary in order to supply labor. The remaining non-leisure time, $1 - u_{it}$, is devoted to human capital accumulation.

The specification chosen for human capital accumulation of each class of agents is identical to the one considered in Lucas (1988). However we assume that the two classes have different skills in accumulating human capital.⁷

Assumption 2 *The law of human capital accumulation is given by*

$$\dot{h}_{it}/h_{it} = B_i (1 - u_{it}) \quad (4)$$

If an individual does not devote any non-leisure time to human capital accumulation then there is no accumulation. If an individual devotes all his non-leisure time to human capital accumulation then his human capital grows at its maximal rate, B_i , which is specific to each class, denoting different skills. As in Lucas (1988) we assume that $B_i > \rho_i$.

We denote by c_{it} *real consumption* of an individual of class i , and by c_t aggregate consumption in the economy. Production is entirely consumed so that $c_t \equiv \sum_{i=0}^1 \pi_i c_{it} = y_t$.

⁷We do not consider the human capital externalities resulting from learning spillovers, that can account for a positive effect of equality on growth.

For simplicity we assume no labor or capital accumulation disutility.

Assumption 3 *Preferences are rationalized by a logarithmic felicity*

$$\int_0^{\infty} e^{-\rho_i t} \ln c_{it} dt \quad (5)$$

with $\rho_0 < \rho_1$.

As in Lucas (1988) we consider a planner who maximizes the weighted sum of utilities

$$\sum_{i=0}^1 \pi_i \int_0^{\infty} e^{-\rho_i t} \ln c_{it} dt = \int_0^{\infty} \sum_{i=0}^1 \pi_i e^{-\rho_i t} \ln c_{it} dt \quad (6)$$

an intertemporal welfare functional, under (4).⁸

In the following, we consider two regimes according to the planner's political constraints. In the first regime, that we call "meritocracy", the planner's faces individual constraints:

$$c_{it} \leq A l_{it} \quad (7)$$

Remark that when instead of two classes of individuals the subscript i denotes two countries, the meritocratic regime corresponds to the autarky situation.

In the second regime, that we call "redistribution", the planner faces an aggregate constraint:

$$\sum_{i=0}^1 \pi_i c_{it} \leq A l_t \quad (8)$$

When i denotes countries instead of classes of individuals, this regime corresponds to the case where the two countries are integrated in one union.

3 Meritocracy

In this case the planner maximizes (6) subject to (4) and (7). The Hamiltonian is written as follows:

$$\sum_{i=0}^1 \pi_i e^{-\rho_i t} \ln (A h_{it} u_{it}) + \sum_{i=0}^1 \lambda_{it} h_{it} B_i (1 - u_{it})$$

⁸The question of time-consistency can be raised in the case of a welfare function maximization with heterogenous agents. As noted by Zuber (2011) and Heal and Millner (2015), stationarity and time-consistency of aggregate preferences hold together only if agents have the same discount factor.

We have two state variables (h_{it}) and two controls (u_{it}). The strict concavity of the objective function and the concavity of the law of motion with respect to (h_{it}, u_{it}) satisfy the Arrow-Mangasarian sufficient condition for Hamiltonian maximization and imply the uniqueness of the planner's solution.

Proposition 1. *The optimal solution is the Balanced Growth Path (BGP):*

$$u_{it}^* = \frac{\rho_i}{B_i} \text{ and } h_{it}^* = h_{i0}e^{(B_i - \rho_i)t} \quad (9)$$

for $t \geq 0$. This solution is unique.

Proof. See the Appendix. ■

This solution coincides with that forwarded by Lucas (1988) in the representative agent's case. In fact, what the planner is doing in the meritocratic case is simply solving two independent problems, each of them featuring identical agents. Note that with the functional forms chosen we are able to prove analytically Lucas (1988) conjecture that the BGP is the optimal solution.

We can see that the equilibrium growth rate of human capital of class i increases with the class ability in investing in human capital, B_i , and declines with increases in the class discount rate, ρ_i . However, what really matters for human capital growth is the difference between these two parameters. This means that more skilled but very impatient agents may accumulate human capital at a lower rate than less skilled but more patient ones. However, we will rule out this possibility since it is not supported by empirical studies. Indeed, Falk et al. (2015), using a new data set that covers 80000 individuals from 76 countries, find that patience is more pronounced among individuals with higher cognitive ability. Importantly, this relationship holds when controlling for income, education, and country fixed effects. Earlier studies for the US, Germany, and Chile, have also found that higher cognitive ability is associated with greater patience.⁹ Therefore, we assume that the growth rate of human capital accumulation of the patient consumer, $B_0 - \rho_0$, always exceeds the growth rate of the impatient one, $B_1 - \rho_1$, i.e. we consider that the patient consumer, $i = 0$, is the dominant one. Let $\delta \equiv (B_0 - \rho_0) - (B_1 - \rho_1)$ denote the difference between the growth rates of patient and impatient consumers.¹⁰

Assumption 4 $\delta > 0$.

⁹See Frederick (2005), Dohmen et al. (2010) and Benjamin et al. (2013), respectively.

¹⁰The stock of human capital/wealth of those households who discount future outcomes

3.1 Inequality and growth in the meritocratic case

Let us now discuss the trade-off between social inequalities and growth in the meritocratic case. The first choice that we have to make is how to measure these two concepts. Most studies use income inequality data as a proxy for social inequality and the growth rate of income to measure economic growth. In this work we chose to measure social inequalities using the Gini index of consumption, while economic growth is measured by the growth rate of aggregate consumption. In the meritocratic case, as consumption and income (output) coincide, both at the individual and aggregate level, this choice is irrelevant. In contrast, in the redistribution case, although aggregate consumption is still identical to aggregate output, the real welfare of each individual is given by consumption and not by output. Therefore, social inequalities are better proxied using a measure of consumption inequality. Also, since in this work we focus on the effects of human capital on inequality and growth we start by presenting the Gini index of human wealth, which measures inequality in the distribution of human capital.

Proposition 2. *If, without loss of generality, $h_{0t} > h_{1t}$, the period t Gini index of human wealth is given by*

$$g_{ht} = \pi_1 - \frac{\pi_1 h_{1t}}{\pi_0 h_{0t} + \pi_1 h_{1t}} \quad (10)$$

Proof. See the Appendix. ■

The stock of human capital that a class has accumulated until period t , h_{it} , depends not only on the class rate of growth of human capital accumulation, $B_i - \rho_i$, but also on its initial human capital endowment, h_{i0} . However, it is easy to see that, in the long run, the growth rate effect will dominate. We have therefore the following proposition.

Proposition 3 (human wealth inequality). *Let Assumption 4 hold.*

1. *If $h_{00} > h_{10}$, then $h_{0t} > h_{1t}$ for every t and*

$$g_{ht} = \pi_1 - \frac{\pi_1 h_{10} e^{-\delta t}}{\pi_0 h_{00} + \pi_1 h_{10} e^{-\delta t}} \quad (11)$$

more heavily, will either lack permanently behind or be eventually exceeded by the human capital/wealth of more patient households. A similar story of overtaking is presented by Doepke and Zilibotti (2008) with regard to the Industrial Revolution. They attributes the overtaking of the wealthy aristocracy by middle-class industrialists, as the dominant (in economic terms) group, to heterogeneous degrees of patience that are actually endogenous and diffused endogenously through a process of intergenerational cultural transmission.

Thus, g_t increases monotonically from $g_{h0} = \pi_1 - \pi_1 h_{10} / (\pi_0 h_{00} + \pi_1 h_{10})$ to $g_\infty = \pi_1$.

2. If $h_{00} < h_{10}$, then there is a critical date

$$T_h = \frac{\ln(h_{10}/h_{00})}{\delta} \quad (12)$$

beyond which the human capital stock of the dominant (patient) class exceeds the human capital stock of the impatient class. In this case, the Gini index of human wealth is given by

$$g_{ht} = \pi_0 - \frac{\pi_0 h_{00}}{\pi_0 h_{00} + \pi_1 h_{10} e^{-\delta t}} \text{ for } t \leq T_h \quad (13)$$

$$g_{ht} = \pi_1 - \frac{\pi_1 h_{10} e^{-\delta t}}{\pi_0 h_{00} + \pi_1 h_{10} e^{-\delta t}} \text{ for } t > T_h \quad (14)$$

Proof. See the Appendix. ■

Since human capital and consumption of an individual of type i grow at the same rate along the BGP, the observed differences in the evolution of the Gini index of human wealth and the Gini index of consumption reflect mainly differences in initial endowments, being therefore relevant only in the short run.

Definition 4. *If, without loss of generality, $c_{0t} > c_{1t}$, the Gini index of consumption is given by*

$$g_{ct} = \pi_1 - \frac{\pi_1 c_{1t}}{\pi_0 c_{0t} + \pi_1 c_{1t}} \quad (15)$$

Proposition 5 (consumption inequality). *Let Assumption 4 hold.*

1. If $c_{00} > c_{10}$, then $c_{0t} > c_{1t}$ for every t and

$$g_{ct} = \pi_1 - \frac{\pi_1 c_{10} e^{-\delta t}}{\pi_0 c_{00} + \pi_1 c_{10} e^{-\delta t}} \quad (16)$$

Thus, g_t increases monotonically from $g_0 = \pi_1 - \pi_1 c_{10} / (\pi_0 c_{00} + \pi_1 c_{10})$ to $g_\infty = \pi_1$.

2. If $c_{00} < c_{10}$, then there is a critical date

$$T_c = \frac{\ln(c_{10}/c_{00})}{\delta} \quad (17)$$

beyond which consumption of the dominant (patient) class exceeds consumption of the impatient class. In this case, the Gini index of consumption is given by

$$g_{ct} = \pi_0 - \frac{\pi_0 c_{00}}{\pi_0 c_{00} + \pi_1 c_{10} e^{-\delta t}} \text{ for } t \leq T_c \quad (18)$$

$$g_{ct} = \pi_1 - \frac{\pi_1 c_{10} e^{-\delta t}}{\pi_0 c_{00} + \pi_1 c_{10} e^{-\delta t}} \text{ for } t > T_c \quad (19)$$

Proof. See the Appendix. ■

Remark that $h_{00} > h_{10}$ does not imply $c_{00} > c_{10}$. Indeed, using (3), (7) and (9), we can rewrite this last inequality as

$$h_{00} > h_{10} \frac{\rho_1 B_0}{\rho_0 B_1}$$

where $\rho_1/\rho_0 > 1$, so that we have $c_{10} > c_{00}$ if B_1 is sufficiently low.

We now describe the evolution of the growth rate of aggregate consumption.

Proposition 6 (consumption growth rate). *The dynamics of the aggregate consumption growth rate are given by*

$$\gamma_{ct} \equiv \frac{\dot{c}_t}{c_t} = \frac{\pi_0 c_{00} (B_0 - \rho_0) e^{\delta t} + \pi_1 c_{10} (B_1 - \rho_1)}{\pi_0 c_{00} e^{\delta t} + \pi_1 c_{10}} \quad (20)$$

Proof. See the Appendix. ■

We can now characterize the trade-off between inequality and growth in the meritocratic case.

Proposition 7 (trade-off inequality-growth). *Let Assumption 4 hold.*

In case (1) of Proposition 5 ($c_{00} > c_{10}$), the higher the consumption inequality, the higher the growth rate:

$$\gamma_c = \pi_0 (B_0 - \rho_0) + \pi_1 (B_1 - \rho_1) + \delta g_c \quad (21)$$

In case (2) of Proposition 5 ($c_{00} < c_{10}$), this trade-off holds after T_c and is reversed before:

$$\gamma_c = \pi_0 (B_0 - \rho_0) + \pi_1 (B_1 - \rho_1) - \delta g_c \text{ for } t \leq T_c \quad (22)$$

$$\gamma_c = \pi_0 (B_0 - \rho_0) + \pi_1 (B_1 - \rho_1) + \delta g_c \text{ for } t > T_c \quad (23)$$

Proof. See the Appendix. ■

We conclude that, in the long run, higher inequality is always associated with higher growth in the meritocratic case. However, in the short run, depending on the initial distribution of skills and endowments, this result may be reversed. Indeed, in the subcase $t \leq T_c$ of case (2) where $c_{00} < c_{10}$, the lower the social inequality, the higher the growth rate.

4 Redistribution

In this case the planner maximizes the same social welfare functional:

$$\sum_{i=0}^1 \pi_i \int_0^\infty e^{-\rho_i t} \ln c_{it} dt = \int_0^\infty \sum_{i=0}^1 \pi_i e^{-\rho_i t} \ln c_{it} dt$$

under the resource constraint

$$\sum_{i=0}^1 \pi_i c_{it} \leq A \sum_{i=0}^1 \pi_i h_{it} u_{it} \quad (24)$$

and the law of motion of human capital

$$\dot{h}_{it} = B_i (1 - u_{it}) h_{it}$$

with $0 \leq u_{it} \leq 1$.

There are two state variables (h_{it}) and four controls (c_{it} and u_{it}). In this case, since there is redistribution, the planner can solve its problem in two stages. In the first stage, given h_{it} and u_{it} , the planner solves a static problem:

$$\max_{c_{0t}, c_{1t}} \sum_{i=0}^1 \pi_i e^{-\rho_i t} \ln c_{it}$$

subject to (24). The solution of this static program is given by:

$$c_{it} = A e^{-\rho_i t} \frac{\sum_{j=0}^1 \pi_j h_{jt} u_{jt}}{\sum_{j=0}^1 \pi_j e^{-\rho_j t}} \quad (25)$$

Note that the division of consumption among individuals is only determined by the degree of impatience, ρ_i , not being influenced by the distribution of skills, B_i . Since $c_t = y_t$, using (24) and (25) we obtain the consumption share

$$s_{it} \equiv \frac{\pi_i c_{it}}{c_t} = \frac{\pi_i e^{-\rho_i t}}{\sum_{j=0}^1 \pi_j e^{-\rho_j t}}$$

Substituting now (25) in the original problem it becomes

$$\max \int_0^\infty \left[\sum_{i=0}^1 \pi_i e^{-\rho_i t} \ln \frac{A e^{-\rho_i t}}{\sum_{j=0}^1 \pi_j e^{-\rho_j t}} + \left(\sum_{i=0}^1 \pi_i e^{-\rho_i t} \right) \ln \sum_{j=0}^1 \pi_j h_{jt} u_{jt} \right] dt$$

or, equivalently,

$$\max \int_0^\infty \left(\sum_{i=0}^1 \pi_i e^{-\rho_i t} \right) \ln \sum_{j=0}^1 \pi_j h_{jt} u_{jt} dt$$

subject to $\dot{h}_{it} = B_i (1 - u_{it}) h_{it}$ with $0 \leq u_{it} \leq 1$.

We now have to obtain the optimal solution for h_{it} and u_{it} . To better understand the mechanisms involved we will start by analyzing the case where $B_0 = B_1$, the only difference between the two types of agents being their degree of patience.

4.1 Redistribution without heterogeneity in skills

In this case $B_0 = B_1 = B$, and $0 < \rho_0 < \rho_1$.

Proposition 8. *If $B_0 = B_1 = B$, and $0 < \rho_0 < \rho_1$ the optimal trajectory $(h_{it}, u_{it})_{i=0,1}$ is given by:*

$$h_{it} = \phi \zeta_i \sum_{j=0}^1 \frac{\pi_j}{\rho_j} e^{(B-\rho_j)t} \quad \text{and} \quad u_{it} = \frac{1}{B} \frac{\sum_{j=0}^1 \pi_j e^{(B-\rho_j)t}}{\sum_{j=0}^1 \frac{\pi_j}{\rho_j} e^{(B-\rho_j)t}} \in (0, 1) \quad (26)$$

for $t \geq 0$, where

$$\phi = \frac{\sum_{i=0}^1 \pi_i h_{i0}}{\sum_{i=0}^1 \frac{\pi_i}{\rho_i}} \text{ and } \zeta_i = \frac{h_{i0}}{\sum_{j=0}^1 \pi_j h_{j0}} \quad (27)$$

Proof. See the Appendix. ■

Observe that in this case ($B_0 = B_1 = B$), we have $u_{0t} = u_{1t} = u_t$, which converges to ρ_0/B . The total human capital stock $H_t = \sum_{i=0}^1 \pi_i h_{it}$, satisfies $\dot{H}_t = BH_t(1 - u_t)$.

4.2 Redistribution with heterogeneity in skills

In this subsection we obtain the optimal trajectory $(h_{it}, u_{it})_{i=0,1}$ for $B_0 \neq B_1$, and $0 < \rho_0 < \rho_1$.

Our first important result is given in the proposition below.

Proposition 9. *Let $B_0 \neq B_1$. For every t , there exists an i such that $u_{it} \in \{0, 1\}$.*

Proposition 9 tells us that, in contrast to the meritocratic case and to the redistribution case without heterogeneity in skills, we never have an interior solution for working time, u_{it} , of both classes of agents.

Using this proposition, we can find the explicit planner's solution, which is the optimal trajectory. To this purpose we introduce the following assumption.

Assumption 5 $B_0 > B_1 > \rho_1 > \rho_0$.

According to Assumption 5, and in line with empirical evidence,¹¹ the dominant class is not only more patient ($\rho_0 < \rho_1$), but is also more talented in accumulating human capital ($B_0 > B_1$). Indeed, as discussed above, Falk et al. (2015) find that the relationship between cognitive ability and patience is unequivocally positive in almost all countries. Assumption 5 is of course more restrictive than Assumption 4. Indeed, Assumption 5 implies $\delta > 0$, while the reverse is not true.

Let us define the average discount rate, $\rho(t)$, as a time-dependent harmonic mean of discount rates

$$\rho(t) \equiv \left(\sum_{i=0}^1 \frac{1}{\rho_i} \frac{\pi_i e^{-\rho_i t}}{\sum_{j=0}^1 \pi_j e^{-\rho_j t}} \right)^{-1} \in (\rho_0, \rho_1) \quad (28)$$

¹¹See again Falk et al. (2015), Dohmen et al. (2010), Benjamin et al. (2013) and Frederick (2005).

Of course, under Assumption 5, $B_0 > B_1 > \rho(t)$. Moreover, $\rho(t)$ decreases from $\rho(0) \in (\rho_0, \rho_1)$ to $\rho(\infty) = \rho_0$. If $t = 0$, (28) simplifies to:

$$\rho(0) = \left(\sum_{i=0}^1 \frac{\pi_i}{\rho_i} \right)^{-1}$$

which is a harmonic mean where the weights are the size of each class, π_i .

In the following,

$$H_{it} \equiv \pi_i h_{it} \text{ and } \sigma_{it} \equiv \frac{H_{it}}{H_{0t} + H_{1t}} \quad (29)$$

will denote respectively the human capital of class i at time t and its aggregate share. Of course, $\sum_{i=0}^1 \sigma_{it} = 1$.

To simplify the presentation we will consider separately the following two cases:

1. $\sigma_{10} \leq \sigma_{10}^* \equiv \rho(0) / B_0$,
2. $\sigma_{10} > \sigma_{10}^*$.

Note that in the first case, since σ_{10} is sufficiently small, the more patient and talented class is also relatively well endowed in initial human capital.

4.2.1 The case where $\sigma_{10} \leq \sigma_{10}^*$

The optimal trajectory in this case is given in Proposition 10 below.

Proposition 10. *Let Assumption 5 hold and assume that $\sigma_{10} \leq \sigma_{10}^* \equiv \rho(0) / B_0$. Then, the optimal trajectory $(h_{it}, u_{it})_{i=0,1}$ is given by*

$$h_{0t} = \frac{H_{10} \rho(0)}{\sigma_{10} \pi_0} \sum_{i=0}^1 \pi_i \left[\frac{e^{(B_0 - \rho_i)t}}{\rho_i} - \frac{\sigma_{10}}{\rho_i} \right] \quad (30)$$

$$u_{0t} = \frac{\sum_{i=0}^1 \pi_i \left[\frac{e^{(B_0 - \rho_i)t}}{B_0} - \frac{\sigma_{10}}{\rho_i} \right]}{\sum_{i=0}^1 \pi_i \left[\frac{e^{(B_0 - \rho_i)t}}{\rho_i} - \frac{\sigma_{10}}{\rho_i} \right]} \in (0, 1) \quad (31)$$

and

$$h_{1t} = h_{10} \quad (32)$$

$$u_{1t} = 1 \quad (33)$$

for every $t \geq 0$.

Proof. See the Appendix. ■

We notice that $0 < u_{0t} < 1$ and $u_{1t} = 1$ for every $t \geq 0$. This means that when the initial human capital endowment of the patient and talented class is relatively important, the less patient and less talented class *never* invests in human capital, devoting all its non-leisure time to work. This result can be seen as an application of the comparative advantage principle, according to which agents will specialize in the activity where they are relatively better. However, the more patient and talented agent also devotes some non-leisure to work every period. Remember that in order to supply labor, contributing therefore to production, an agent must devote some time to work. See (3). In this case, as the initial capital endowment of the skilled class is sufficiently important, these agents divide each period their non-leisure time between work and capital accumulation, contributing to the production effort since the beginning. Also, since, under Assumption 5, $B_0 > \rho_1$, we get $\lim_{t \rightarrow \infty} u_{0t} = \rho_0/B_0 < 1$.

It is easy to conclude that since only the more talented and patient class accumulates human capital, its human capital share, σ_{0t} , strictly increases in time. Replacing (30) and (32) in (29) we obtain the dynamics of the human capital shares.

Lemma 11 (shares of human capital). *Under Assumption 5 and $\sigma_{10} \leq \sigma_{10}^* \equiv \rho(0)/B_0$, the shares of human capital at time t are given by*

$$\begin{aligned}\sigma_{0t} &= 1 - \sigma_{1t} \\ \sigma_{1t} &= \frac{\sigma_{10}}{\sum_{i=0}^1 w_i e^{(B_0 - \rho_i)t}}\end{aligned}\tag{34}$$

for any t , where the denominator is an average with weights

$$w_i \equiv \frac{\pi_i}{\rho_i} / \sum_{j=0}^1 \frac{\pi_j}{\rho_j}$$

4.2.2 The case where $\sigma_{10} > \sigma_{10}^*$

We start by introducing four important critical values: T_0, T_1, T_2 , with $T_1 < T_2$, and λ_{10} . We define T_0 and T_2 (λ_{10}) respectively as the solutions of

the two following equalities below

$$\begin{aligned}\sigma_{1T_0} &= \sigma_{1T_0}^* \equiv \frac{\rho(T_0)}{B_0} \\ \sigma_{1T_2} &= \sigma_{1T_2}^* \equiv \frac{\rho(T_2)}{B_0}\end{aligned}$$

that we can rewrite respectively as:

$$\frac{B_0}{\rho(T_0)} = 1 + \frac{H_{00}}{H_{10}} e^{B_0 T_0} \quad (35)$$

$$\frac{B_0}{\rho(T_2)} = 1 + \frac{H_{00}}{H_{10}} \frac{e^{B_0 T_2}}{e^{B_1 T_1}} \left[1 + \frac{\sum_{i=0}^1 \frac{\pi_i}{\rho_i} (1 - e^{-\rho_i T_1})}{\sum_{i=0}^1 \frac{\pi_i}{B_1} e^{-\rho_i T_1}} \right] \quad (36)$$

where $T_1(\lambda_{10})$ is the smallest solution of

$$\lambda_{10} h_{10} = \sum_{i=0}^1 \frac{\pi_i}{B_1} e^{-\rho_i T_1} + \sum_{i=0}^1 \frac{\pi_i}{\rho_i} (1 - e^{-\rho_i T_1}) \quad (37)$$

Note that at T_1 we have $u_{1t} = 1$. Notice also that T_2 equals T_0 when $T_1 = 0$. Finally, let λ_{10} be the solution of

$$f(\lambda_{10}) \equiv \frac{\frac{1}{\rho(T_2(\lambda_{10}))} - \frac{1}{B_0}}{\frac{1}{\rho(T_1(\lambda_{10}))} - \frac{1}{B_1}} - \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i T_1(\lambda_{10})}}{\sum_{i=0}^1 \pi_i e^{-\rho_i T_2(\lambda_{10})}} = 0 \quad (38)$$

λ_{10} determines in turn $T_1(\lambda_{10})$ and $T_2(\lambda_{10})$.

The optimal trajectory in this case, where $\sigma_{10} > \sigma_{10}^*$, i.e., the initial human capital share of the less patient and less talented class exceeds the critical threshold $\sigma_{10}^* \equiv \rho(0)/B_0$, is given in Proposition 12 below.

Proposition 12. *Let Assumption 5 hold and assume that $\sigma_{10} > \sigma_{10}^* \equiv \rho(0)/B_0$.*

1. If

$$B_1 \leq B_1^* \equiv \frac{1}{\frac{1}{\rho(0)} - \left[\frac{1}{\rho(T_0)} - \frac{1}{B_0} \right] \sum_{i=0}^1 \pi_i e^{-\rho_i T_0}} \quad (39)$$

then the optimal trajectory $(h_{it}, u_{it})_{i=0,1}$ is given by

$$h_{0t} = h_{00}e^{B_0t} \text{ for } 0 \leq t \leq T_0 \quad (40)$$

$$h_{0t} = \frac{H_{10}}{\pi_0} \left[\frac{\sum_{i=0}^1 \frac{\pi_i e^{(B_0-\rho_i)t}}{\rho_i}}{\sum_{i=0}^1 \frac{\pi_i e^{(B_0-\rho_i)T_0}}{B_0}} - 1 \right] \text{ for } t > T_0 \quad (41)$$

$$u_{0t} = 0 \text{ for } 0 \leq t \leq T_0 \quad (42)$$

$$u_{0t} = \frac{\sum_{i=0}^1 \frac{\pi_i e^{(B_0-\rho_i)t}}{B_0} - \sum_{i=0}^1 \frac{\pi_i e^{(B_0-\rho_i)T_0}}{B_0}}{\sum_{i=0}^1 \frac{\pi_i e^{(B_0-\rho_i)t}}{\rho_i} - \sum_{i=0}^1 \frac{\pi_i e^{(B_0-\rho_i)T_0}}{B_0}} \in (0, 1) \text{ for } t > T_0 \quad (43)$$

and

$$h_{1t} = h_{10} \text{ for any } t \quad (44)$$

$$u_{1t} = 1 \text{ for any } t \quad (45)$$

2. If

$$B_1 > B_1^* \equiv \frac{1}{\frac{1}{\rho(0)} - \left[\frac{1}{\rho(T_0)} - \frac{1}{B_0} \right] \sum_{i=0}^1 \pi_i e^{-\rho_i T_0}}$$

then the optimal trajectory $(h_{it}, u_{it})_{i=0,1}$ is given by

$$h_{0t} = h_{00}e^{B_0t} \text{ for } 0 \leq t \leq T_2 \quad (46)$$

$$h_{0t} = \frac{H_{1T_2}}{\pi_0} \left[\frac{\sum_{i=0}^1 \frac{\pi_i e^{(B_0-\rho_i)t}}{\rho_i}}{\lambda_{1T_2} h_{1T_2} e^{B_0 T_2}} - 1 \right] \text{ for } t \geq T_2 \quad (47)$$

$$u_{0t} = 0 \text{ for } 0 \leq t \leq T_2$$

$$u_{0t} = \frac{\sum_{i=0}^1 \frac{\pi_i e^{(B_0-\rho_i)t}}{B_0} - \lambda_{1T_2} h_{1T_2} e^{B_0 T_2}}{\sum_{i=0}^1 \frac{\pi_i e^{(B_0-\rho_i)t}}{\rho_i} - \lambda_{1T_2} h_{1T_2} e^{B_0 T_2}} \in (0, 1) \text{ for } t \geq T_2 \quad (48)$$

and

$$h_{1t} = h_{10}e^{B_1 t} \left[1 - \frac{1}{\lambda_{10} h_{10}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} (1 - e^{-\rho_i t}) \right] \text{ for } 0 \leq t \leq T_1 \quad (49)$$

$$h_{1t} = h_{1T_1} \text{ for } t \geq T_1$$

$$u_{1t} = \frac{1}{B_1 \lambda_{10} h_{10} + \sum_{i=0}^1 \frac{\pi_i}{\rho_i} (e^{-\rho_i t} - 1)} \in (0, 1) \text{ for } 0 \leq t < T_1 \quad (50)$$

$$u_{1t} = 1 \text{ for } t \geq T_1$$

where $\lambda_{10}h_{10}$ is given by (37),

$$\lambda_{1T_2}h_{1T_2} = \sum_{i=0}^1 \frac{\pi_i}{B_1} e^{-\rho_i T_1} + \sum_{i=0}^1 \frac{\pi_i}{\rho_i} (e^{-\rho_i T_2} - e^{-\rho_i T_1}) \quad (51)$$

and $H_{1T_2} = \pi_1 h_{1T_1}$.

Proof. See the Appendix. ■

In case 1 of Proposition 12, although the initial capital share of the less talented and less patient individuals is now above the critical threshold $\rho(0)/B_0$, as they are not sufficiently skilled, $B_1 \leq B_1^*$, these individuals still never accumulate human capital, devoting all their non-leisure time to work, i.e. we have, as in the previous case, $h_{1t} = h_{10}$ and $u_{1t} = 1$ for any t . However, now the dominant class does not work initially, devoting all their non-leisure time to capital accumulation, that, until period T_0 increases at the maximal accumulation rate, B_0 . See (42). Therefore, in this case for $t < T_0$, specialization is more intense: the less talented agents just work, while the others only accumulate capital. Nevertheless, once the share of human capital of the more talented class reaches the critical level $1 - \rho(t)/B_0$, i.e. when $\sigma_{1T_0} = \sigma_{1T_0}^*$, the dominant class starts working and, since less time is devoted to capital accumulation, the rate of growth of human capital decreases over time.

As the less talented agents never accumulate capital $H_{1t} = H_{10}$ for any t . Then, we get immediately the following result.

Lemma 13 (shares of human capital). *In case 1 of Proposition 12, the shares of human capital at time t are given by*

$$\begin{aligned} \sigma_{0t} &= 1 = \sigma_{1t} \text{ for any } t \geq 0 \\ \sigma_{1t} &= \frac{\sigma_{10}}{\sigma_{10} + (1 - \sigma_{10}) e^{B_0 t}} \text{ for } 0 \leq t \leq T_0 \end{aligned} \quad (52)$$

$$\sigma_{1t} = \frac{\sigma_{1T_0}}{\sum_{i=0}^1 w_i e^{(B_0 - \rho_i)(t - T_0)}} \text{ for } t > T_0 \quad (53)$$

where the denominator in (52) is an average with weights

$$w_i \equiv \frac{\frac{\pi_i}{\rho_i} e^{(B_0 - \rho_i)T_0}}{\sum_{j=0}^1 \frac{\pi_j}{\rho_j} e^{(B_0 - \rho_j)T_0}}$$

and

$$\sigma_{1T_0} = \frac{\sum_{i=0}^1 \frac{\pi_i}{B_0} e^{(B_0 - \rho_i)T_0}}{\sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B_0 - \rho_i)T_0}}$$

We observe that (34) is a particular case of (53) with $T_0 = 0$.

In case 2 of Proposition 12, since the starting value of the capital share of the more talented class is below $1 - \sigma_{10}^*$, again the dominant class does not work initially, only accumulating capital, which until period T_2 , where σ_{0T_2} reaches the critical value $1 - \rho(T_2)/B_0$, increases at the maximal accumulation rate, B_0 . See (46). As in the previous case, after date T_2 the more talented agents start working and accumulate capital at a lower pace. However, as, in this case, the more impatient agents are more skilled in accumulating capital, they now accumulate capital until period T_1 , dividing their non-leisure time between capital accumulation and work. Note however that the time they devote to capital accumulation decreases continuously from period 0 to T_1 , where $u_{1T_1} = 1$. After period T_1 they stop accumulating capital, devoting all their non-leisure time to work. As $T_1 \leq T_2$ and until T_2 the other class is accumulating capital at the maximal rate, their capital share declines steadily until period T_1 . The evolution in time of the capital shares of both classes is given below in Lemma 14.

Lemma 14. *In case 2 of Proposition 12, the shares of human capital at time t are given by*

$$\begin{aligned} \sigma_{0t} &= 1 - \sigma_{1t} \text{ for any } t \geq 0 \\ \sigma_{1t} &= \frac{\sigma_{10}}{\sigma_{10} + (1 - \sigma_{10}) \frac{\lambda_{10} h_{10} e^{(B_0 - B_1)t}}{\lambda_{10} h_{10} - \sum_{i=0}^1 \frac{\pi_i}{\rho_i} (1 - e^{-\rho_i t})}} \text{ for } 0 \leq t \leq T_1 \\ \sigma_{1t} &= \frac{\sigma_{1T_1}}{\sigma_{1T_1} + (1 - \sigma_{1T_1}) e^{B_0(t - T_1)}} \text{ for } T_1 \leq t \leq T_2 \\ \sigma_{1t} &= \frac{\sigma_{1T_2}}{\sum_{i=0}^1 w_i e^{(B_0 - \rho_i)(t - T_2)}} \text{ for } t \geq T_2 \end{aligned} \tag{54}$$

where $\lambda_{10} h_{10}$ is given by (37) and the denominator in (54) is an average with weights

$$w_i \equiv \frac{\frac{\pi_i}{\rho_i} e^{(B_0 - \rho_i)T_2}}{\sum_{j=0}^1 \frac{\pi_j}{\rho_j} e^{(B_0 - \rho_j)T_2}}$$

and

$$\sigma_{1T_2} = \frac{\lambda_{1T_2} h_{1T_2} e^{B_0 T_2}}{\sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B_0 - \rho_i) T_2}}$$

4.2.3 Summarizing interpretation

In this section we considered two classes with different skills, different degrees of patience and different initial capital stocks. We assumed, in line with empirical plausibility, that the more talented class is also more patient. We showed that in this case we never have an interior solution for working time of both classes of agents. Moreover, several optimal solutions exist, depending on the initial distribution of human capital and on the level of skills of the less talented class. If the initial human capital share of the less skilled individuals is not high enough, they will never accumulate human capital, devoting all their non-leisure time to work. If their initial human capital share is sufficiently high two cases are possible. Either they are not sufficiently skilled and again they only work, never accumulating human capital, or they are sufficiently skilled and devote initially some of their non-leisure time to human capital accumulation. However, the time they devote to human capital accumulation decreases steadily in time, so that in the long run we find that, as before, the less talented agents never accumulate human capital. In contrast, the more talented individuals always accumulate human capital. When their initial human capital share exceeds a critical value, these individuals will also work every period. However, if their initial share of human capital is not high enough, they will not work initially, devoting all their non-leisure time to capital accumulation, that grows at the maximal possible rate. This behavior continues until this critical level of their capital share is reached. After, they will start working and accumulating capital as in the previous solution. It is interesting to note two things. First, what matters for the choice of the optimal solution is the distribution of human capital and not its level.¹² Second, in the long run, independently of the initial conditions and of the level of skills, we always get the same result: less talented individuals never accumulate capital and the more skilled agents work and accumulate capital.

¹²Remark that the critical share of one class is always equal to one minus the critical share of the other class.

4.3 Inequality and growth in the redistribution case

Let us now discuss the trade-off between social inequalities and growth in the redistribution case with heterogenous skills and discounting. From Lemmas 11, 13, and 14, we obtain immediately the dynamics of the Gini index of wealth in this case.

Proposition 15 (human wealth inequality). *If, without loss of generality, $h_{0t} > h_{1t}$, the Gini index of human wealth is given by*

$$g_{ht} = \pi_1 - \sigma_{1t} \quad (55)$$

The dynamics of wealth inequality are the following.

When $\sigma_{10} \leq \sigma_{10}^$:*

$$g_{ht} = \pi_1 - \frac{\sigma_{10}}{\sum_{i=0}^1 w_i e^{(B_0 - \rho_i)t}} \text{ for any } t \geq 0$$

When $\sigma_{10} > \sigma_{10}^$ and $B_1 \leq B_1^*$:*

$$g_{ht} = \pi_1 - \frac{\sigma_{10}}{\sigma_{10} + (1 - \sigma_{10}) e^{B_0 t}} \text{ for } 0 \leq t \leq T_0$$

$$g_{ht} = \pi_1 - \frac{\sigma_{1T_0}}{\sum_{i=0}^1 w_i e^{(B_0 - \rho_i)(t - T_0)}} \text{ for } t > T_0$$

When $\sigma_{10} > \sigma_{10}^$ and $B_1 > B_1^*$:*

$$g_{ht} = \pi_1 - \frac{\sigma_{10}}{\sigma_{10} + (1 - \sigma_{10}) \frac{\lambda_{10} h_{10} e^{(B_0 - B_1)t}}{\lambda_{10} h_{10} - \sum_{i=0}^1 \frac{\pi_i}{\rho_i} (1 - e^{-\rho_i t})}} \text{ for } 0 \leq t \leq T_1$$

$$g_{ht} = \pi_1 - \frac{\sigma_{1T_1}}{\sigma_{1T_1} + (1 - \sigma_{1T_1}) e^{B_0(t - T_1)}} \text{ for } T_1 \leq t \leq T_2$$

$$g_{ht} = \pi_1 - \frac{\sigma_{1T_2}}{\sum_{i=0}^1 w_i e^{(B_0 - \rho_i)(t - T_2)}} \text{ for } t \geq T_2$$

In any case, the optimal Gini index of human wealth inequality increases over time.

Proposition 16 (consumption inequality). *The Gini index of consumption is given by*

$$g_{ct} = \pi_1 - \frac{\pi_1 c_{1t}}{\pi_0 c_{0t} + \pi_1 c_{1t}} = \pi_1 - \frac{\pi_1}{\pi_1 + \pi_0 e^{(\rho_1 - \rho_0)t}} \quad (56)$$

The Gini index of consumption increases over time from 0 to π_1 .

The dynamics of the aggregate consumption growth rate are complicated and depend on the regime considered. For simplicity, we focus on the more plausible situation. i.e. the case considered in subsection 4.2.1, whose dynamics are given Proposition 10. In this case, $\sigma_{10} \leq \sigma_{10}^*$ meaning that the more patient and talented class is relatively well endowed in human capital at the beginning.

Proposition 17 (consumption growth rate). *In the case of Proposition 10, the dynamics of aggregate consumption and its growth rate are given by*

$$c_t = AH_{10} \frac{\rho(0)}{\sigma_{10}} \sum_{i=0}^1 \frac{\pi_i}{B_0} e^{(B_0 - \rho_i)t} \quad (57)$$

$$\gamma_{ct} \equiv \frac{\dot{c}_t}{c_t} = \frac{\sum_{i=0}^1 \pi_i (B_0 - \rho_i) e^{(B_0 - \rho_i)t}}{\sum_{i=0}^1 \pi_i e^{(B_0 - \rho_i)t}} \quad (58)$$

Proposition 16 shows that, in the case of redistribution, the Gini index of consumption converges to π_1 or, equivalently, the consumption share of the impatient class converges to 0. The same happens in the case of meritocracy (Proposition 5). However, in the case of meritocracy, the growth rate of consumption of the impatient class is $B_1 - \rho_1$, while, in the case of redistribution considered in Proposition 17, it is equal to $B_0 - \rho_1$. Thus, in this case of redistribution, the consumption of the impatient class grows at a higher rate.

We can now obtain the trade-off between inequality and growth.

Proposition 18 (trade-off inequality-growth). *The (aggregate) consumption growth rate γ_c is a function of the Gini index of consumption g_c . More precisely, the growth rate is an average whose weights depend on the Gini index:*

$$\gamma_c(g_c) = \sum_{i=0}^1 (B_0 - \rho_i) w_i(g_c) \quad (59)$$

with

$$w_i(g_c) \equiv \frac{\pi_i \left(\frac{\pi_1}{\pi_0} \frac{\pi_0 + g_c}{\pi_1 - g_c} \right)^{\frac{B_0 - \rho_i}{\rho_1 - \rho_0}}}{\sum_{j=0}^1 \pi_j \left(\frac{\pi_1}{\pi_0} \frac{\pi_0 + g_c}{\pi_1 - g_c} \right)^{\frac{B_0 - \rho_j}{\rho_1 - \rho_0}}}$$

$\gamma_c(g_c)$ is an increasing function, increasing from $\gamma_c(0) = B_0 - (\pi_0 \rho_0 + \pi_1 \rho_1)$ to $\gamma_c(\pi_1) = B_0 - \rho_0$. Thus, the social inequality and the consumption growth rate are positively correlated.

This proposition means that higher inequality is associated with higher growth, even in the case of redistribution. The planner raises the consumption share of low discount rate agents and reduces that of higher discount rate agents as time passes. This directly increases the inequality. On the other hand, the consumption growth rate of low discount rate agents is higher than that of high growth rate agents and, thus, the rise of the consumption share of low discount rate agents raises the growth rate of aggregate consumption.

These results are consistent with empirical evidence. Many advanced economies have experienced increasing (income) inequality since the 1980s, see Atkinson (1999) and Goldin and Katz (2008). Moreover, according to Turnovsky and Mitra (2013), this recent increase in inequality is explained by the increasing role of human capital as an engine of growth.¹³

5 Discussion of the results and further comments

Our first important result is that heterogeneity matters. Without heterogeneity, redistribution is not an issue and, therefore, the two policy regimes considered are indistinguishable. Also the effect of heterogeneity is much more pronounced in the case of redistribution. Indeed, in the meritocratic case the optimal solution for each agent is identical to the Lucas (1988) representative agent BGP, whereas with redistribution the optimal solution is never the BGP. The type of heterogeneity also plays an important role.

Although we did not discuss it previously, it is easy to see that, if we only had considered heterogeneity in the initial capital endowments, the optimal solution would converge to the Lucas BGP, being identical in the long run under both the meritocratic and redistributive regimes. When we introduce also different degrees of patience, the role of heterogeneity becomes more important and the optimal solutions of the two regimes are no longer asymptotically identical. However, they are both interior in what concerns working time. Finally, when we also consider heterogeneity in skills, in the redistribution case, the optimal solution for non-leisure time devoted to capital accumulation is never interior for both agents, being therefore dramatically

¹³Goldin and Katz (1999, 2001) and Abramovitz and David (2000) find that the contribution of human capital to growth almost doubled during the 20th century in the United States, while the contribution of physical capital decreased.

different both from the meritocracy and from the representative agent solutions. A second important, although trivial result is that social welfare under the redistribution regime is always higher than under meritocracy. Indeed, the ability to redistribute cannot reduce overall welfare since one option available to the planner is to choose to not redistribute. Since we have seen that this is not the optimal path, the planner must do strictly better redistributing.

Together, these two results imply that it is optimal to exploit existing differences. We conjecture that, provided there is redistribution, welfare is higher (within mean preserving transformations) when we move from the representative agent case to an unequal distribution of skills. The proof is left for further works.

It is also interesting to note that, in the redistribution regime, the distribution of consumption is not affected by skills heterogeneity. Indeed only differences in patience, i.e. in preferences, are taken into account by the planner when she allocates consumption to each agent. However, skills differences are determinant for the allocation of tasks between agents. We found that, in the long run, the less talented class never invests in human capital, regardless of the initial distribution of human capital and of the efficiency of the less skilled agents in accumulating human capital. Also, whenever the human capital share of the more skilled class is below a certain critical threshold, these agents do not supply labor, devoting all their non-leisure time to capital accumulation at the beginning. We conclude that selection in the access to education and specialization are optimal.

We also find that, in both regimes, the inequality in the distribution of human capital increases in time. In the meritocratic regime, this is due to differences in skills and in patience, which have therefore permanent effects, while with redistribution specialization is also an important driver of this result. Discussing now the relation between inequality and growth we find that in both regimes, in the long run, inequality is associated with higher rates of growth. It is easy to see that, within our framework, human capital accumulation is the channel through which inequality is associated with higher growth. This happens, even with redistribution, suggesting that although redistribution increases utility and growth, it does not prevent a deterioration in the distribution of consumption during the growth process.

6 Concluding remarks

In this paper we extended the Lucas (1988) framework, introducing simultaneously heterogeneity in patience, in skills and in initial capital endowments. We considered two policy regimes: "meritocracy" and "redistribution". Our main conclusions are the following. First, heterogeneity changes significantly the optimal solution, specially in the presence of redistribution. Second, cooperation is always better, i.e. utility under the redistribution regime is always higher than under meritocracy. We conjecture that welfare is higher (within mean preserving transformations) when we move from a representative agent economy to an economy with an unequal distribution of skills. This means that it is optimal to exploit existing differences. Third, the redistribution of consumption only depends on preferences. In contrast, the distribution of tasks takes into account skills differences. Finally, we find that inequality is associated with higher rates of growth.

These results are novel, showing that heterogeneity really matters, which implies of course that the representative agent approach may be misleading. Another point of our contribution we want to stress, is that heterogeneity, instead of being considered a problem, should be optimally exploited in order to increase welfare. However, for this outcome redistribution is essential. In particular, we found that with redistribution, differences in skills should be translated into specialization in tasks and in a differential access to education. Note however that these results were obtained using a very stylized model. Therefore they should not be seen as policy recommendations. Nevertheless, in spite of the simplicity of the framework considered, the message that heterogeneity implies choices that are not only radically different from the ones associated with a representative agent economy, but which also dominate them in terms of welfare, should still be true in more general settings. Finally, in this work we have only considered the first best (planner's solution). Future work on the decentralized market solution is therefore welcome.

7 Appendix

Proof of Proposition 1. The Hamiltonian is written as follows:

$$\sum_{i=0}^1 \pi_i e^{-\rho_i t} \ln(Ah_{it}u_{it}) + \sum_{i=0}^1 \lambda_{it} h_{it} B_i (1 - u_{it})$$

We have two state variables (h_{it}) and two controls (u_{it}).

We derive the first-order conditions: $\partial H_t / \partial \lambda_{it} = \dot{h}_{it}$, $\partial H_t / \partial h_{it} = -\dot{\lambda}_{it}$, $\partial H_t / \partial u_{it} = 0$ and the transversality condition $\lim_{t \rightarrow \infty} (\lambda_{it} h_{it}) = 0$. The first-order conditions write

$$\begin{aligned} \dot{u}_{it} &= (B_i u_{it} - \rho_i) u_{it} \\ \dot{h}_{it} &= B_i (1 - u_{it}) h_{it} \end{aligned} \quad (60)$$

jointly with the transversality condition $\lim_{t \rightarrow \infty} (e^{-\rho_i t} / u_{it}) = 0$.

Focus on the solution of ODE (60).

There are three cases.

(1) If $u_{i0} < \rho_i / B_i$,

$$u_{it} = \frac{1}{B_i} \frac{\rho_i}{1 + e^{\rho_i(c_i+t)}}$$

with

$$c_i = \frac{1}{\rho_i} \ln \left(\frac{\rho_i / B_i}{u_{i0}} - 1 \right)$$

(2) If $u_{i0} = \rho_i / B_i$,

$$u_{it} = \frac{\rho_i}{B_i}$$

(3) If $u_{i0} > \rho_i / B_i$,

$$\begin{aligned} u_{it} &= \frac{1}{B_i} \frac{\rho_i}{1 - e^{\rho_i(c_i+t)}} \leq 1 \text{ for } t \leq t^* \\ u_{it} &= 1 \text{ for } t > t^* \end{aligned}$$

with

$$\begin{aligned} c_i &= \frac{1}{\rho_i} \ln \frac{B_i u_{i0} - \rho_i}{B_i u_{i0}} < 0 \\ t^* &\equiv \frac{1}{\rho_i} \ln \left(1 - \frac{\rho_i}{B_i} \right) - c_i > 0 \end{aligned}$$

(1) Focus on the case $u_{i0} < \rho_i / B_i$.

We evaluate the utility along this trajectory:

$$\int_0^\infty e^{-\rho_i t} \ln (A h_{it} u_{it}) dt = \frac{1}{\rho_i} \left[\ln (A h_{i0} u_{i0}) + \frac{B_i}{\rho_i} - 1 \right]$$

(2) Consider the second case: $u_{i0} = \rho_i / B_i$.

The economy is at the steady state from the beginning.
We evaluate the utility at the steady state.

$$\int_0^{\infty} e^{-\rho_i t} \ln(Ah_{it}u_{it}) dt = \frac{1}{\rho_i} \left[\ln(Ah_{i0}u_{i0}) + \frac{B_i}{\rho_i} - 1 \right]$$

Therefore, in the cases (1) and (2), we find:

$$\begin{aligned} & \arg \max_{u_{i0} \leq \rho_i/B_i} \int_0^{\infty} e^{-\rho_i t} \ln(Ah_{it}u_{it}) dt \\ &= \arg \max_{u_{i0} \leq \rho_i/B_i} \frac{1}{\rho_i} \left[\ln(Ah_{i0}u_{i0}) + \frac{B_i}{\rho_i} - 1 \right] = \frac{\rho_i}{B_i} \end{aligned}$$

and

$$\max_{u_{i0} \leq \rho_i/B_i} \int_0^{\infty} e^{-\rho_i t} \ln(Ah_{it}u_{it}) dt = \frac{1}{\rho_i} \left[\ln \left(Ah_{i0} \frac{\rho_i}{B_i} \right) + \frac{B_i}{\rho_i} - 1 \right]$$

(3) Focus on the case $u_{i0} > \rho_i/B_i$.

The solution becomes

$$\begin{aligned} u_{it} &= \frac{1}{B_i} \frac{\rho_i}{1 - e^{\rho_i(c_i+t)}} \leq 1 \text{ for } t \leq t^* \\ u_{it} &= 1 \text{ for } t > t^* \end{aligned}$$

where

$$\begin{aligned} c_i &= \frac{1}{\rho_i} \ln \frac{B_i u_{i0} - \rho_i}{B u_{i0}} < 0 \\ t^* &\equiv \frac{1}{\rho_i} \ln \left(1 - \frac{\rho_i}{B_i} \right) - c_i > 0 \end{aligned}$$

We evaluate the utility along this trajectory.

$$\begin{aligned} & \int_0^{\infty} e^{-\rho_i t} \ln(Ah_{it}u_{it}) dt = \int_0^{t^*} e^{-\rho_i t} \ln(Ah_{it}u_{it}) dt + \int_{t^*}^{\infty} e^{-\rho_i t} \ln(Ah_{it^*}) dt \\ &= \frac{1}{\rho_i} \left[\frac{1 - u_{i0}}{u_{i0}} + \ln(Ah_{i0}u_{i0}) \right] \end{aligned}$$

Let us maximize

$$\int_0^{\infty} e^{-\rho_i t} \ln(Ah_{it}u_{it}) dt = \frac{1}{\rho_i} \left[\frac{1 - u_{i0}}{u_{i0}} + \ln(Ah_{i0}u_{i0}) \right]$$

with respect to u_{i0} . Notice that

$$\frac{d}{du_{i0}} \left[\ln(Ah_{i0}u_{i0}) + \frac{1-u_{i0}}{u_{i0}} \right] = \frac{1}{u_{i0}} - \frac{1}{u_{i0}^2} < 0$$

for $u_{i0} \in [\rho_i/B_i, 1)$. Then, $\int_0^\infty e^{-\rho_i t} \ln(Ah_{it}u_{it}) dt$ decreases in $[\rho_i/B_i, 1)$ and attains its maximum at $u_{i0}^* = \rho_i/B_i$.

The value at $u_{i0}^* = \rho_i/B_i$ is

$$\int_0^\infty e^{-\rho_i t} \ln(Ah_{it}u_{it}) dt = \frac{1}{\rho_i} \left(\ln \frac{Ah_{i0}\rho_i}{B_i} + \frac{B_i}{\rho_i} - 1 \right)$$

We conclude that $u_{i0}^* = \rho_i/B_i$ maximizes the utility whatever the case we consider. Since $u_{i0}^* = \rho_i/B_i$ is also the steady state, we find that the BGP (9) is the planner's solution. ■

Proof of Proposition 2. Under the assumption $h_{0t} > h_{1t}$, the Gini index of wealth is given by:

$$g_{ht} = \frac{\frac{\pi_0 h_{0t} + \pi_1 h_{1t}}{2} - \left(\frac{\pi_1 \pi_1 h_{1t}}{2} + \pi_0 \pi_1 h_{1t} + \frac{\pi_0 \pi_0 h_{0t}}{2} \right)}{\frac{\pi_0 h_{0t} + \pi_1 h_{1t}}{2}}$$

that is by (10). ■

Proof of Proposition 3. In order to compute the dynamics of the wealth Gini index, we consider the optimal individual wealth dynamics (9): $h_{it} = h_{i0}e^{(B_i - \rho_i)t}$.

Under Assumption 4, $B_0 - \rho_0 > B_1 - \rho_1$. There are two cases: (1) $h_{00} > h_{10}$ and (2) $h_{00} < h_{10}$.

(1) In the first case, $h_{0t} > h_{1t}$ for every t and, then,

$$g_{ht} = \pi_1 - \frac{\pi_1 h_{10} e^{(B_1 - \rho_1)t}}{\pi_0 h_{00} e^{(B_0 - \rho_0)t} + \pi_1 h_{10} e^{(B_1 - \rho_1)t}}$$

that is (11). It is easy to check that g_{ht} increases in time.

(2) In the second case, $h_{00} < h_{10}$. Beyond the critical date T_h , the dominant consumer earns the higher revenue. Hence, T_h is solution of $h_{0T} = h_{1T}$, that is of $h_{00}e^{(B_0 - \rho_0)T} = h_{10}e^{(B_1 - \rho_1)T}$. We get (12).

Therefore, the Gini index of wealth is given by

$$\begin{aligned} g_{ht} &= \pi_0 - \frac{\pi_0 h_{0t}}{\pi_0 h_{0t} + \pi_1 h_{1t}} \text{ for } t \leq T_h \\ g_{ht} &= \pi_1 - \frac{\pi_1 h_{1t}}{\pi_0 h_{0t} + \pi_1 h_{1t}} \text{ for } t > T_h \end{aligned}$$

Replacing the expressions $h_{it} = h_{i0}e^{(B_i-\rho_i)t}$, we obtain (13) and (14). ■

Proof of Proposition 5. Under the assumption $c_{0t} > c_{1t}$, the Gini index of consumption is given by:

$$g_{ct} = \frac{\frac{\pi_0 c_{0t} + \pi_1 c_{1t}}{2} - \left(\frac{\pi_1 \pi_1 c_{1t}}{2} + \pi_0 \pi_1 c_{1t} + \frac{\pi_0 \pi_0 c_{0t}}{2} \right)}{\frac{\pi_0 c_{0t} + \pi_1 c_{1t}}{2}}$$

that is by (15).

In order to compute the dynamics of the consumption Gini index, we consider the individual income dynamics: $c_{it} = Ah_{it}u_{it} = Ah_{i0}\rho_i e^{(B_i-\rho_i)t}/B_i = c_{i0}e^{(B_i-\rho_i)t}$.

Under Assumption 4, $B_0 - \rho_0 > B_1 - \rho_1$. There are two cases: (1) $c_{00} > c_{10}$ and (2) $c_{00} < c_{10}$.

(1) In the first case, $c_{0t} > c_{1t}$ for every t and, then,

$$g_{ct} = \pi_1 - \frac{\pi_1 c_{10} e^{(B_1-\rho_1)t}}{\pi_0 c_{00} e^{(B_0-\rho_0)t} + \pi_1 c_{10} e^{(B_1-\rho_1)t}}$$

that is (16). It is easy to check that g_{ct} increases in time.

(2) In the second case, $c_{00} < c_{10}$. Beyond the critical date T_c , the dominant consumer earns the higher revenue. Hence, T_c is solution of $c_{0T} = c_{1T}$, that is of $c_{00}e^{(B_0-\rho_0)T} = c_{10}e^{(B_1-\rho_1)T}$. We get (17).

Therefore, the Gini index is given by

$$\begin{aligned} g_{ct} &= \pi_0 - \frac{\pi_0 c_{0t}}{\pi_0 c_{0t} + \pi_1 c_{1t}} \text{ for } t \leq T_c \\ g_{ct} &= \pi_1 - \frac{\pi_1 c_{1t}}{\pi_0 c_{0t} + \pi_1 c_{1t}} \text{ for } t > T_c \end{aligned}$$

Replacing the expressions $c_{it} = c_{i0}e^{(B_i-\rho_i)t}$, we obtain (18) and (19). ■

Proof of Proposition 6. The aggregate consumption growth rate is given by

$$\gamma_t \equiv \frac{\dot{c}_t}{c_t} = \frac{\left[\sum_{i=0}^1 \pi_i c_{i0} e^{(B_i-\rho_i)t} \right]'}{\sum_{i=0}^1 \pi_i c_{i0} e^{(B_i-\rho_i)t}} = \frac{\sum_{i=0}^1 \pi_i c_{i0} (B_i - \rho_i) e^{(B_i-\rho_i)t}}{\sum_{i=0}^1 \pi_i c_{i0} e^{(B_i-\rho_i)t}}$$

that is by (20). ■

Proof of Proposition 7. Focus on case (1): $c_{00} > c_{10}$. From (16), we find

$$e^{-\delta t} = \frac{\pi_0 c_{00} \pi_1 - g t}{\pi_1 c_{10} \pi_0 + g t}$$

Replacing this in (20), we obtain the trade-off (21) between inequality and growth.

Focus on case (2): $c_{00} < c_{10}$. In the case $t \leq T$, we have

$$e^{-\delta t} = \frac{\pi_0 c_{00} \pi_1 + g t}{\pi_1 c_{10} \pi_0 - g t}$$

Replacing this in (20), we find the trade-off between the inequality and growth (22). Similarly, we obtain (23). ■

Proof of Proposition 8. In this case $B_0 = B_1 = B$, and $0 < \rho_0 < \rho_1$. By supposing that we have interior solution, the Hamiltonian is written as follows:

$$H_t = \left(\sum_{i=0}^1 \pi_i e^{-\rho_i t} \right) \ln \sum_{i=0}^1 \pi_i h_{it} u_{it} + B \sum_{i=0}^1 \lambda_{it} (1 - u_{it}) h_{it}$$

Since h_{it} maximizes H_t , $\partial H_t / \partial u_{it} = 0$, which implies:

$$B \lambda_{it} = \pi_i \frac{\sum_{j=0}^1 \pi_j e^{-\rho_j t}}{\sum_{j=0}^1 \pi_j h_{jt} u_{jt}} \quad (61)$$

Replacing (61) in $\partial H_t / \partial h_{it} = -\dot{\lambda}_{it}$, we find

$$-\dot{\lambda}_{it} = \lambda_{it} B u_{it} + \lambda_{it} B (1 - u_{it}) = B \lambda_{it}$$

Which implies $\lambda_{it} = \lambda_{i0} e^{-Bt}$. We have also for any t ,

$$\frac{\lambda_{0t}}{\lambda_{1t}} = \frac{B \lambda_{0t}}{B \lambda_{1t}} = \frac{\pi_0}{\pi_1}$$

Define

$$\phi = \frac{\pi_0}{\lambda_{00}} = \frac{\pi_1}{\lambda_{10}}$$

From (61), we get

$$\sum_{i=0}^1 \pi_i h_{it} u_{it} = \frac{\phi}{B} \sum_{i=0}^1 \pi_i e^{(B-\rho_i)t}$$

Define $H_t \equiv \sum_{i=0}^1 H_{it} = \sum_{i=0}^1 \pi_i h_{it}$. We have

$$\dot{H}_t = \sum_{i=0}^1 \pi_t \dot{h}_{it} = \sum_{i=0}^1 \pi_i B h_{it} (1 - u_{it}) = B H_t - \phi \sum_{i=0}^1 \pi_i e^{(B-\rho_i)t}$$

Define

$$z_t = H_t - \phi \sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B-\rho_i)t}$$

We get

$$\begin{aligned} \dot{z}_t &= \dot{H}_t - \phi \sum_{i=0}^1 \frac{\pi_i}{\rho_i} (B - \rho_i) e^{(B-\rho_i)t} \\ &= B H_t - \phi \sum_{i=0}^1 \pi_i e^{(B-\rho_i)t} - \phi \sum_{i=0}^1 \frac{\pi_i}{\rho_i} (B - \rho_i) e^{(B-\rho_i)t} \\ &= B z_t \end{aligned}$$

This implies $z_t = z_0 e^{Bt}$ and

$$H_t = z_0 e^{Bt} + \phi \sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B-\rho_i)t}$$

From the transversality condition, we have

$$\lim_{t \rightarrow \infty} (\lambda_{it} h_{it}) = \frac{\lambda_{i0}}{\pi_i} \lim_{t \rightarrow \infty} (e^{-Bt} \pi_i h_{it}) = 0$$

Then, $\lim_{t \rightarrow \infty} (e^{-Bt} \pi_i h_{it}) = 0$ and

$$0 = \lim_{t \rightarrow \infty} \left(e^{-Bt} \sum_{i=0}^1 \pi_i h_{it} \right) = \lim_{t \rightarrow \infty} (e^{-Bt} H_t) = \lim_{t \rightarrow \infty} \left(z_0 + \phi \sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{-\rho_i t} \right) = z_0$$

which implies $z_t = 0$ for any t . Thus,

$$\sum_{i=0}^1 \pi_i h_{it} = H_t = \phi \sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B-\rho_i)t} \quad (62)$$

We can guess that the solution must satisfies, for $i = 0, 1$, equation (26) where ζ_0 and ζ_1 satisfy $\sum_{i=0}^1 \pi_i \zeta_i = 1$. We observe that ζ_i in (27) satisfies this condition.

We can also guess that

$$h_{it}u_{it} = \frac{\phi}{B}\zeta_i \sum_{j=0}^1 \pi_j e^{(B-\rho_j)t} \quad (63)$$

We observe that (63) satisfies (26).

This is equivalent to

$$u_{it} = \frac{\frac{\phi}{B}\zeta_i \sum_{j=0}^1 \pi_j e^{(B-\rho_j)t}}{\phi\zeta_i \sum_{j=0}^1 \frac{\pi_j}{\rho_j} e^{(B-\rho_j)t}}$$

that is (26).

We can now construct the solution. Considering (62) with $t = 0$, we get ϕ in (26).

Considering (63) with $t = 0$, we get

$$\zeta_i = \frac{h_{i0}}{\phi \sum_{j=0}^1 \pi_j / \rho_j}$$

that is ζ_i in (26). We observe that $\sum_{i=0}^1 \pi_i \zeta_i = 1$.

Define $\lambda_{i0} = \pi_i / \phi$ yielding $\lambda_{it} = e^{-Bt} \pi_i / \phi$.

Observe that $0 < u_{it} < 1$ and $\dot{h}_{it} = Bh_{it} - Bh_{it}u_{it} = Bh_{it}(1 - u_{it})$.

We can verify easily that $(h_{it}, u_{it}, \lambda_{it})$ satisfies Pontryagin's conditions.

Observe that in this case ($B_0 = B_1 = B$), we have $u_{0t} = u_{1t} = u_t$, which converges to ρ_0/B . The total human capital stock $H_t = \sum_{i=0}^1 \pi_i h_{it}$, satisfies $\dot{H}_t = BH_t(1 - u_t)$. ■

Proof of Proposition 9. Suppose the contrary: there exists t such that, for every i , $0 < u_{it} < 1$. This implies the existence of $t_1 < t_2$ such that $0 < u_{it} < 1$ for every $t_1 < t < t_2$. On the interval (t_1, t_2) the Hamiltonian is written as follows:

$$H_t = \left(\sum_{i=0}^1 \pi_i e^{-\rho_i t} \right) \ln \sum_{i=0}^1 \pi_i h_{it} u_{it} + \sum_{i=0}^1 \lambda_{it} B_i (1 - u_{it}) h_{it}$$

For every $t_1 < t < t_2$, we have $\partial H_t / \partial u_{it} = 0$, $\partial H_t / \partial \lambda_{it} = \dot{h}_{it}$ and $\partial H_t / \partial h_{it} = -\dot{\lambda}_{it}$. The first equation implies that, for every i :

$$B_i \lambda_{it} = \pi_i \frac{\sum_{j=0}^1 \pi_j e^{-\rho_j t}}{\sum_{j=0}^1 \pi_j h_{jt} u_{jt}} \quad (64)$$

Replacing (64) in $\partial H_t / \partial h_{it} = -\dot{\lambda}_{it}$, we find

$$-\dot{\lambda}_{it} = \lambda_{it} B_i u_{it} + \lambda_{it} B_i (1 - u_{it}) = \lambda_{it} B_i$$

Integrating from t_1 to t , we obtain $\lambda_{it} = \lambda_{it_1} e^{-B_i(t-t_1)}$ for any $t_1 < t < t_2$. From (64), we have also $B_0 \lambda_{0t} / (B_1 \lambda_{1t}) = \pi_0 / \pi_1$. This implies that $e^{-(B_0 - B_1)(t-t_1)} = \pi_0 B_1 \lambda_{1t_1} / (\pi_1 B_0 \lambda_{0t_1})$ for any $t_1 < t < t_2$. This leads to a contradiction with $B_0 \neq B_1$ since, in this case, the left-hand side changes over time. ■

Proof of Proposition 10. The Hamiltonian is written as follows:

$$H_t = \left(\sum_{i=0}^1 \pi_i e^{-\rho_i t} \right) \ln \sum_{i=0}^1 \pi_i h_{it} u_{it} + \sum_{i=0}^1 \lambda_{it} B_i (1 - u_{it}) h_{it}$$

Proposition 9 allows us to focus on a candidate solution which is a corner solution for one agent: there exists $T \geq 0$ such that

$$0 < u_{0t} < 1 \text{ and } u_{1t} = 1 \quad (65)$$

for every $t \geq T$. We will show that $T = 0$ and the Pontryagin's optimal conditions are satisfied.

If (65) is solution, we have

$$\left. \frac{\partial H_t}{\partial u_{0t}} \right|_{0 < u_{0t} < 1} = \pi_0 h_{0t} \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} - B_0 \lambda_{0t} h_{0t} = 0 \quad (66)$$

$$\left. \frac{\partial H_t}{\partial u_{1t}} \right|_{u_{1t}=1} = \pi_1 h_{1t} \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} - B_1 \lambda_{1t} h_{1t} \geq 0 \quad (67)$$

Therefore,

$$B_0 \lambda_{0t} = \pi_0 \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} \quad (68)$$

$0 < u_{0t} < 1$ implies

$$\dot{\lambda}_{0t} = -\frac{\partial H_t}{\partial h_{0t}} = -\pi_0 u_{0t} \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} - B_0 \lambda_{0t} (1 - u_{0t}) = -B_0 \lambda_{0t} \quad (69)$$

(69) entails

$$\lambda_{0t} = \lambda_{0T} e^{-B_0(t-T)} \quad (70)$$

for any $t \geq T$.

$u_{1t} = 1$ implies

$$\dot{\lambda}_{1t} = -\frac{\partial H_t}{\partial h_{1t}} = -\pi_1 u_{1t} \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} - B_1 \lambda_{1t} (1 - u_{1t}) = -\frac{\pi_1}{\pi_0} B_0 \lambda_{0t}$$

Moreover,

$$\dot{\lambda}_{1t} = -\frac{\pi_1}{\pi_0} B_0 \lambda_{0t} = \frac{\pi_1}{\pi_0} \dot{\lambda}_{0t}$$

that is

$$\lambda_{1t} = \frac{\pi_1}{\pi_0} (\lambda_{0t} - \lambda_{0T}) + \lambda_{1T} = \frac{\pi_1}{\pi_0} \lambda_{0T} e^{-B_0(t-T)} + \lambda_{1T} - \frac{\pi_1}{\pi_0} \lambda_{0T} \quad (71)$$

A candidate solution satisfies the transversality condition, that is

$$\lim_{t \rightarrow \infty} (\lambda_{1t} h_{1t}) = \lim_{t \rightarrow \infty} \left[\frac{\pi_1}{\pi_0} \lambda_{0T} e^{-B_0(t-T)} h_{1t} \right] + \lim_{t \rightarrow \infty} \left[\left(\lambda_{1T} - \frac{\pi_1}{\pi_0} \lambda_{0T} \right) h_{1t} \right] = 0$$

which holds only if $\lambda_{1T} - \lambda_{0T} \pi_1 / \pi_0 = 0$. According to (71), we find

$$\lambda_{1t} = \frac{\pi_1}{\pi_0} \lambda_{0t}$$

Replacing (70) in equation (68), we get

$$\sum_{i=0}^1 \pi_i h_{it} u_{it} = \frac{\pi_0}{B_0 \lambda_{0T} e^{B_0 T}} \sum_{i=0}^1 \pi_i e^{(B_0 - \rho_i)t}$$

and, noticing that $u_{1t} = 1$ and $h_{1t} = h_{1T}$ for every $t \geq T$, we obtain

$$h_{0t} u_{0t} = \frac{1}{B_0 \lambda_{0T} e^{B_0 T}} \sum_{i=0}^1 \pi_i e^{(B_0 - \rho_i)t} - \frac{\pi_1}{\pi_0} h_{1T} \quad (72)$$

Therefore,

$$\dot{h}_{0t} = B_0 h_{0t} - B_0 u_{0t} h_{0t} = B_0 h_{0t} + B_0 \frac{\pi_1}{\pi_0} h_{1T} - \frac{\sum_{i=0}^1 \pi_i e^{(B_0 - \rho_i)t}}{\lambda_{0T} e^{B_0 T}}$$

This differential equation is equivalent to $\dot{z}_t = B_0 z_t$, where

$$z_t \equiv h_{0t} + \frac{\pi_1}{\pi_0} h_{1T} - \frac{\sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B_0 - \rho_i)t}}{\lambda_{0T} e^{B_0 T}}$$

The solution of $\dot{z}_t = B_0 z_t$ is $z_t = z_T e^{B_0(t-T)}$ for $t \geq T$. Thus

$$h_{0t} = z_T e^{B_0(t-T)} - \frac{\pi_1}{\pi_0} h_{1T} + \frac{\sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B_0 - \rho_i)t}}{\lambda_{0T} e^{B_0 T}} \quad (73)$$

A candidate solution satisfies the transversality condition, that is, according to (70),

$$\lim_{t \rightarrow \infty} (\lambda_{0t} h_{0t}) = \lim_{t \rightarrow \infty} (\lambda_{0T} z_T) + \lim_{t \rightarrow \infty} \left(\lambda_{0t} \left[\frac{\sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B_0 - \rho_i)t}}{\lambda_{0T} e^{B_0 T}} - \frac{\pi_1}{\pi_0} h_{1T} \right] \right) = 0$$

which holds only if $z_T = 0$, that is only if

$$\pi_1 h_{1T} + \pi_0 h_{0T} = \pi_0 \frac{\sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B_0 - \rho_i)T}}{\lambda_{0T} e^{B_0 T}} = \frac{\pi_0}{\lambda_{0T}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{-\rho_i T} \quad (74)$$

(68) and (74) imply

$$\frac{B_0}{\rho(T)} = \frac{\pi_0 h_{0T} + \pi_1 h_{1T}}{\pi_1 h_{1T} + \pi_0 h_{0T} u_{0T}} \quad (75)$$

where $\rho(T)$ is given by (28).

We compute u_{0T} from (75) and we impose the economic restriction $0 \leq u_{0T} \leq 1$ or, equivalently,

$$1 \leq \frac{B_0}{\rho(T)} \leq \frac{1}{\sigma_{1T}} \quad (76)$$

($0 \leq u_{0T}$ is equivalent to the RHS, while $u_{0T} \leq 1$ to the LHS).

Inequality in the LHS of (76) is satisfied by Assumption 5: $B_0 > \rho_1 > \rho(T)$.

Thus, the candidate solution holds for every $t \geq 0$ only if the RHS holds at $T = 0$. It is the case because (76) with $T = 0$ is equivalent to the initial condition $\sigma_{10} \leq \sigma_{10}^*$.

Let us now provide the explicit trajectory for h_{it} and u_{it} and show that the optimal conditions are verified.

From (75) with $T = 0$, we have

$$u_{00} = \frac{\rho(0)}{B_0} - \left[1 - \frac{\rho(0)}{B_0} \right] \frac{H_{10}}{H_{00}}$$

As seen above, $0 \leq u_{00} \leq 1$ under Assumption 5 and inequality $\sigma_{10} \leq \rho(0)/B_0$.

From (68) and (74) with $T = 0$, we find λ_{00} .

Moreover, (73) with $z_T = 0$ and $T = 0$ implies

$$h_{0t} = \frac{1}{\lambda_{00}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B_0 - \rho_i)t} - \frac{\pi_1}{\pi_0} h_{10} \quad (77)$$

Replacing (77) in (72) with $T = 0$ and solving for u_{0t} , we get (31).

Let us show that

$$\lambda_{0t} = \frac{\pi_0 e^{-B_0 t}}{H_{00} + H_{10}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} \quad (78)$$

$$\lambda_{1t} = \frac{\pi_1}{\pi_0} \lambda_{0t} \quad (79)$$

and (32) to (31) satisfy the Maximum Principle by Pontryagin.

First, replacing (78), (30) and (31) in (66), we find

$$\left. \frac{\partial H_t}{\partial u_{0t}} \right|_{0 < u_{0t} < 1} = 0$$

Assumption 5 ($B_0 > B_1$) implies

$$B_1 \lambda_{1t} < B_0 \lambda_{1t} = B_0 \lambda_{0t} \frac{\pi_1}{\pi_0} = \frac{\pi_1 \sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\pi_0 h_{0t} u_{0t} + \pi_1 h_{10}}$$

and, so, according to (67), also

$$\left. \frac{\partial H_t}{\partial u_{1t}} \right|_{u_{1t}=1} \geq 0$$

Moreover,

$$\begin{aligned} \dot{\lambda}_{0t} &= -\lambda_{0t} B_0 = -\frac{\partial H_t}{\partial h_{0t}} \\ \dot{\lambda}_{1t} &= -\frac{\pi_1 \sum_i \pi_i e^{-\rho_i t}}{\pi_1 h_{10} + \pi_0 h_{0t} u_{0t}} = -\frac{\pi_1}{\pi_0} \lambda_{0t} B_0 = -\frac{\partial H_t}{\partial h_{1t}} \end{aligned}$$

Finally, we check the transversality conditions. Since

$$\begin{aligned}\lim_{t \rightarrow \infty} (\lambda_{it} h_{it}) &= \lim_{t \rightarrow \infty} (\lambda_{i0} e^{-B_i t} h_{it}) = \lambda_{i0} \lim_{t \rightarrow \infty} (e^{-B_i t} h_{it}) \\ &= \lambda_{i0} \lim_{t \rightarrow \infty} \left[e^{-B_i t} h_{i0} e^{\int_0^t B_i (1-u_{is}) ds} \right] = \lambda_{i0} e^{-B_i \int_0^\infty u_{is} ds}\end{aligned}$$

the transversality conditions are equivalent to $\int_0^\infty u_{is} ds = \infty$. These equalities are satisfied because $\lim_{t \rightarrow \infty} u_{0t} = \rho_0/B_0 > 0$ and $u_{1t} = 1$ for every t .

■

Proof of Proposition 12. (1) Focus first on the case $B_1 \leq B_1^*$.

Proposition 9 allows us to focus on a candidate solution which is a corner solution for one agent. We prove the first subcase by considering a potential solution such that

$$\begin{aligned}u_{0t} &= 0 \text{ for } 0 \leq t \leq T \\ 0 < u_{0t} &< 1 \text{ for } t > T \\ u_{1t} &= 1 \text{ for any } t\end{aligned}\tag{80}$$

by deriving its explicit trajectory and checking that this trajectory satisfies the Pontryagin's Maximum Principle.

According to the Maximum Principle, we need

$$\left. \frac{\partial H_t}{\partial u_{0t}} \right|_{u_{0t}=0} = \pi_0 h_{0t} \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} - B_0 \lambda_{0t} h_{0t} \leq 0\tag{81}$$

$$\left. \frac{\partial H_t}{\partial u_{1t}} \right|_{u_{1t}=1} = \pi_1 h_{1t} \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} - B_1 \lambda_{1t} h_{1t} \geq 0\tag{82}$$

for any $t \in [0, T]$, that is

$$B_0 \lambda_{0t} \geq \frac{\pi_0 \sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\pi_1 h_{10}} \text{ and } B_1 \lambda_{1t} \leq \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{h_{10}}$$

Moreover, for any $t \in [0, T]$,

$$\begin{aligned}\dot{\lambda}_{0t} &= -\frac{\partial H_t}{\partial h_{0t}} = -\pi_0 u_{0t} \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} - B_0 \lambda_{0t} (1 - u_{0t}) \\ \dot{\lambda}_{1t} &= -\frac{\partial H_t}{\partial h_{1t}} = -\pi_1 u_{1t} \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} - B_1 \lambda_{1t} (1 - u_{1t})\end{aligned}$$

that is

$$\dot{\lambda}_{0t} = -B_0\lambda_{0t} \text{ and } \dot{\lambda}_{1t} = -\frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{h_{10}}$$

with solutions

$$\lambda_{0t} = \lambda_{00}e^{-B_0t} \text{ and } \lambda_{1t} = \lambda_{10} - \frac{1}{h_{10}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} (1 - e^{-t\rho_i})$$

If $t > T$, according to the Maximum Principle, we require

$$\left. \frac{\partial H_t}{\partial u_{0t}} \right|_{0 < u_{0t} < 1} = \pi_0 h_{0t} \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} - B_0 \lambda_{0t} h_{0t} = 0$$

that is

$$B_0 \lambda_{0t} = \pi_0 \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} \quad (83)$$

which implies in turn

$$\dot{\lambda}_{0t} = -u_{0t} \left(\pi_0 \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} \right) - B_0 \lambda_{0t} (1 - u_{0t}) = -B_0 \lambda_{0t}$$

with explicit solution $\lambda_{0t} = \lambda_{00}e^{-B_0t}$. Thus, this trajectory holds for every t .

If $0 \leq t \leq T$, we obtain $\dot{h}_{0t} = B_0 h_{0t}$ from $\dot{h}_{0t} = B_0 (1 - u_{0t}) h_{0t}$, that is $h_{0t} = h_{00}e^{B_0t}$.

If $t \geq T$, (83) and $u_{1t} = 1$ entail

$$\dot{\lambda}_{1t} = -\pi_1 u_{1t} \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} - B_1 \lambda_{1t} (1 - u_{1t}) = -\frac{\pi_1}{\pi_0} B_0 \lambda_{0t} = \frac{\pi_1}{\pi_0} \dot{\lambda}_{0t}$$

and

$$\lambda_{1t} = \frac{\pi_1}{\pi_0} (\lambda_{0t} - \lambda_{0T}) + \lambda_{1T} = \frac{\pi_1}{\pi_0} \lambda_{0T} e^{-B_0(t-T)} + \lambda_{1T} - \frac{\pi_1}{\pi_0} \lambda_{0T} \quad (84)$$

According to the Maximum Principle, a candidate trajectory satisfies also the transversality condition:

$$\lim_{t \rightarrow \infty} (\lambda_{1t} h_{1t}) = \lim_{t \rightarrow \infty} \left[\frac{\pi_1}{\pi_0} \lambda_{0T} e^{-B_0(t-T)} h_{1t} \right] + \lim_{t \rightarrow \infty} \left[\left(\lambda_{1T} - \frac{\pi_1}{\pi_0} \lambda_{0T} \right) h_{1t} \right] = 0$$

which holds only if $\lambda_{1T} - \lambda_{0T}\pi_1/\pi_0 = 0$. Therefore, (84) implies $\lambda_{1t} = \lambda_{0t}\pi_1/\pi_0$, for any $t \geq T$. In particular, we get $\lambda_{1T} = \lambda_{0T}\pi_1/\pi_0$. Replacing $\lambda_{0t} = \lambda_{00}e^{-B_0t}$ in equation (83), we get

$$\sum_{i=0}^1 \pi_i h_{it} u_{it} = \frac{\pi_0 \sum_{i=0}^1 \pi_i e^{-\rho_i t}}{B_0 \lambda_{00} e^{-B_0 t}} = \frac{\pi_0 \sum_{i=0}^1 \pi_i e^{(B_0 - \rho_i)t}}{B_0 \lambda_{0T} e^{B_0 T}}$$

and, noticing that $u_{1t} = 1$ and $h_{1t} = h_{1T}$ for every $t \geq T$, we obtain

$$h_{0t} u_{0t} = \frac{\sum_{i=0}^1 \pi_i e^{(B_0 - \rho_i)t}}{B_0 \lambda_{0T} e^{B_0 T}} - \frac{\pi_1}{\pi_0} h_{1T}$$

Therefore, for $t \geq T$,

$$\dot{h}_{0t} = B_0 h_{0t} - B_0 h_{0t} u_{0t} = B_0 h_{0t} + B_0 \frac{\pi_1}{\pi_0} h_{1T} - \frac{\sum_{i=0}^1 \pi_i e^{(B_0 - \rho_i)t}}{\lambda_{0T} e^{B_0 T}}$$

or, equivalently, $\dot{z}_t = B_0 z_t$ with

$$z_t \equiv h_{0t} + \frac{\pi_1}{\pi_0} h_{1T} - \frac{\sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B_0 - \rho_i)t}}{\lambda_{0T} e^{B_0 T}}$$

whose solution is $z_t = z_T e^{B_0(t-T)}$. Thus,

$$h_{0t} = z_T e^{B_0(t-T)} - \frac{\pi_1}{\pi_0} h_{1T} + \frac{\sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B_0 - \rho_i)t}}{\lambda_{0T} e^{B_0 T}}$$

for $t \geq T$.

As above, a candidate solution satisfies also the transversality condition:

$$\lim_{t \rightarrow \infty} (\lambda_{0t} h_{0t}) = \lim_{t \rightarrow \infty} (\lambda_{0T} z_T) + \lim_{t \rightarrow \infty} \left(\lambda_{0t} \left[\frac{\sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B_0 - \rho_i)t}}{\lambda_{0T} e^{B_0 T}} - \frac{\pi_1}{\pi_0} h_{1T} \right] \right) = 0$$

which holds only if $z_T = 0$ or, equivalently,

$$\pi_1 h_{1T} + \pi_0 h_{0T} = \pi_0 \frac{\sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B_0 - \rho_i)T}}{\lambda_{0T} e^{B_0 T}} = \frac{\pi_0}{\lambda_{0T}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{-\rho_i T} \quad (85)$$

(83) and (85) imply

$$\frac{B_0}{\rho(T)} = \frac{\pi_0 h_{0T} + \pi_1 h_{1T}}{\pi_0 h_{0T} u_{0T} + \pi_1 h_{1T}} \quad (86)$$

Since $u_{0T} = 0$, $h_{1T} = h_{10}$, $h_{0T} = h_{00}e^{B_0T}$ and $H_{i0} \equiv \pi_i h_{i0}$, (86) is equivalent to (35).

The critical time T of the candidate trajectory (80) is precisely the solution T_0 of equation (35). This solution exists and is positive. Indeed, we know that $\rho(T) \in [\rho_0, \rho(0)]$, a bounded interval. Moreover, (35) is equivalent to $e^{B_0T} H_{00}/H_{10} = B_0/\rho(T) - 1$, whose RHS is positive according to Assumption 5 ($B_0 > \rho_1 > \rho(T)$) and bounded from above. The LHS goes from H_{00}/H_{10} to ∞ . If $H_{00}/H_{10} < B_0/\max \rho(T) - 1 = B_0/\rho(0) - 1$, then a solution T_0 exists. But this inequality is precisely equivalent to inequality $\sigma_{10} > \sigma_{10}^*$, precisely the case we are considering in Proposition 12.

From (85), we get

$$\lambda_{0T_0} = \frac{\pi_0 \sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{-\rho_i T_0}}{H_{00} e^{B_0 T_0} + H_{10}}$$

and, from (35),

$$\lambda_{0T_0} = \frac{\pi_0 \sum_{i=0}^1 \pi_i e^{-\rho_i T_0}}{B_0 H_{10}} \quad (87)$$

We want to prove that the trajectory defined by equations

$$\lambda_{0t} = \lambda_{0T_0} e^{B_0(T_0-t)} \text{ for any } t \quad (88)$$

$$\lambda_{1t} = \frac{\pi_1}{\pi_0} \lambda_{0T_0} + \frac{1}{h_{10}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} (e^{-\rho_i t} - e^{-\rho_i T_0}) \text{ for } 0 \leq t \leq T_0 \quad (89)$$

$$\lambda_{1t} = \frac{\pi_1}{\pi_0} \lambda_{0T_0} e^{B_0(T_0-t)} \text{ for } t > T_0 \quad (90)$$

and (44) to (43) satisfies the Pontryagin's Maximum Principle.

It is easy to check that λ_{0t} and λ_{1t} are differentiable functions with respect to t in $[0, +\infty)$ with $\dot{\lambda}_{it} = -\partial H_t / \partial h_{it}$ and that, for any t , we have $\dot{h}_{0t} = B_0(1 - u_{0t})h_{0t}$.

Let us show that u_{0t} and u_{1t} maximize the Hamiltonian. For $t > T_0$, apply the proof of Proposition 10. Focus now on $0 \leq t \leq T_0$.

We want to prove that $\partial H_t / \partial u_{0t} |_{u_{0t}=0} \leq 0$.

Since $\sum_{i=0}^1 \pi_i e^{(B_0 - \rho_i)t}$ is strictly increasing and

$$\frac{B_0 \sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B_0 - \rho_i)T_0}}{1 + e^{B_0 T_0} H_{00}/H_{10}} = \sum_{i=0}^1 \pi_i e^{(B_0 - \rho_i)T_0}$$

then, for any $t \in [0, T_0]$, we have

$$\frac{B_0 \sum_{i=0}^1 \frac{\pi_i e^{(B_0 - \rho_i)T_0}}{\rho_i}}{1 + e^{B_0 T_0} H_{00}/H_{10}} \geq \sum_{i=0}^1 \pi_i e^{(B_0 - \rho_i)T_0}$$

or, equivalently,

$$B_0 \lambda_{0t} = B_0 \lambda_{0T_0} e^{B_0(T_0 - t)} \geq \pi_0 \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}}$$

because $u_{0t} = 0$ for $0 \leq t \leq T_0$ and $u_{1t} = 1$ for any t . Then, (81) is verified. This means that $u_{0t} = 0$ maximizes the Hamiltonian.

We want to prove that $\partial H_t / \partial u_{1t} |_{u_{1t}=1} \geq 0$.

(39) is equivalent to

$$\sum_{i=0}^1 \left(\frac{1}{\rho_i} - \frac{1}{B_1} \right) \pi_i e^{-\rho_i 0} \leq \sum_{i=0}^1 \left(\frac{1}{\rho_i} - \frac{1}{B_0} \right) \pi_i e^{-\rho_i T_0}$$

which implies

$$\sum_{i=0}^1 \left(\frac{1}{\rho_i} - \frac{1}{B_1} \right) \pi_i e^{-\rho_i t} \leq \sum_{i=0}^1 \left(\frac{1}{\rho_i} - \frac{1}{B_0} \right) \pi_i e^{-\rho_i T_0} \quad (91)$$

for any $t \in [0, T_0]$, because the LHS of (91), under Assumption 5, is a decreasing function of t . By definition of T_0 ,

$$\sum_{i=0}^1 \left(\frac{1}{\rho_i} - \frac{1}{B_0} \right) \pi_i e^{-\rho_i T_0} = \frac{1}{B_0} \frac{H_{00}}{H_{10}} \sum_{i=0}^1 \pi_i e^{(B_0 - \rho_i)T_0}$$

Inequality (91) becomes

$$\sum_{i=0}^1 \left(\frac{1}{\rho_i} - \frac{1}{B_1} \right) \pi_i e^{-\rho_i t} \leq \frac{1}{B_0} \frac{H_{00}}{H_{10}} \sum_{i=0}^1 \pi_i e^{(B_0 - \rho_i)T_0} = \sum_{i=0}^1 \left(\frac{1}{\rho_i} - \frac{1}{B_0} \right) \pi_i e^{-\rho_i T_0} \quad (92)$$

for any $t \in [0, T_0]$. (89) and (87) imply

$$\sum_{i=0}^1 \left(\frac{1}{\rho_i} - \frac{1}{B_0} \right) \pi_i e^{-\rho_i T_0} = \sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{-\rho_i t} - \lambda_{1t} h_{10}$$

Therefore, (92) becomes

$$B_1 \lambda_{1t} \leq \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{h_{10}} = \pi_1 \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}}$$

(because $u_{0t} = 0$ for $0 \leq t \leq T_0$ and $u_{1t} = 1$ for any t), that is, according to (82), $\partial H_t / \partial u_{1t} |_{u_{1t}=1} \geq 0$. This means that $u_{1t} = 1$ maximizes the Hamiltonian.

(2) Focus now on the case $B_1 > B_1^*$.

Proposition 9 allows us to focus on a candidate solution which is corner solution for one agent. We prove the second subcase by considering a potential solution such that

$$\begin{aligned} u_{0t} &= 0 \text{ for } 0 \leq t \leq T_2 \\ 0 &< u_{0t} < 1 \text{ for } t > T_2 \\ 0 &< u_{1t} < 1 \text{ for } 0 \leq t < T_1 \\ u_{1t} &= 1 \text{ for } t \geq T_1 \end{aligned}$$

with $T_1 < T_2$, by deriving its explicit trajectory and checking that this trajectory satisfies the Pontryagin's Maximum Principle.

Applying the arguments of point (2.1) to determine T_0 (namely, the transversality condition), we find T_2 as solution of

$$\begin{aligned} \frac{B_0}{\rho(T_2)} &= 1 + \frac{H_{0T_2}}{H_{1T_2}} \\ &= 1 + \frac{H_{00}}{H_{10}} \frac{e^{B_0 T_2}}{e^{B_1 T_1} \left[1 + \frac{1}{\lambda_{10} h_{10}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} (e^{-\rho_i T_1} - 1) \right]} \end{aligned}$$

that is of (36).

For $0 \leq t \leq T_1$ we have

$$\left. \frac{\partial H_t}{\partial u_{1t}} \right|_{0 < u_{1t} < 1} = \pi_1 h_{1t} \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} - B_1 \lambda_{1t} h_{1t} = 0$$

that is

$$B_1 \lambda_{1t} = \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{h_{1t} u_{1t}} \tag{93}$$

since $u_{0t} = 0$ and $0 < u_{1t} < 1$ (because $T_1 < T_2$).

Moreover, using (93), we find

$$\dot{\lambda}_{1t} = -\frac{\partial H_t}{\partial h_{1t}} = -\pi_1 u_{1t} \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} - B_1 \lambda_{1t} (1 - u_{1t}) = -B_1 \lambda_{1t}$$

Therefore, for $0 \leq t \leq T_1$, $\lambda_{1t} = \lambda_{10} e^{-B_1 t}$ and, from (93),

$$h_{1t} u_{1t} = \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{B_1 \lambda_{1t}} = \frac{1}{B_1 \lambda_{10}} \sum_{i=0}^1 \pi_i e^{(B_1 - \rho_i)t} \quad (94)$$

The law of motion $\dot{h}_{1t} = B_1 (1 - u_{1t}) h_{1t}$ and (94) imply

$$\dot{h}_{1t} = B_1 h_{1t} - \frac{1}{\lambda_{10}} \sum_{i=0}^1 \pi_i e^{(B_1 - \rho_i)t} \quad (95)$$

We define

$$z_t \equiv h_{1t} - \frac{1}{\lambda_{10}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B_1 - \rho_i)t} \quad (96)$$

Using (95), it is easy to verify that $\dot{z}_t = B_1 z_t$. Thus, $z_t = z_1 e^{B_1 t}$. From (96), we obtain

$$h_{1t} = z_1 e^{B_1 t} + \frac{1}{\lambda_{10}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} e^{(B_1 - \rho_i)t} \text{ with } z_1 = h_{10} - \frac{1}{\lambda_{10}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} \quad (97)$$

(95) and (97) imply $\dot{h}_{1t} = e^{B_1 t} \phi(t)$ where

$$\phi(t) \equiv B_1 h_{10} - \frac{1}{\lambda_{10}} \sum_{i=0}^1 \pi_i \left[\frac{B_1}{\rho_i} + \left(1 - \frac{B_1}{\rho_i}\right) e^{-\rho_i t} \right]$$

Assume that

$$\frac{1}{B_1 h_{10}} < \lambda_{10} < \frac{1}{h_{10}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i}$$

Therefore,

$$\phi(0) = B_1 h_{10} - \frac{1}{\lambda_{10}} > 0 > B_1 h_{10} - \frac{B_1}{\lambda_{10}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} = \phi(\infty)$$

Thus, $\phi(t) = 0$ (that is $\dot{h}_{1t} = 0$) has solution. Let $T_1 > 0$ be the smallest solution. This solution depends on λ_{10} . $\phi(T_1) = 0$ is equivalent to (37). Function $T_1(\lambda_{10})$ behaves as follows.

$$\begin{aligned}\lambda_{10} &\rightarrow \frac{1}{B_1 h_{10}} \Leftrightarrow T_1(\lambda_{10}) \rightarrow 0 \\ \lambda_{10} &\rightarrow \frac{1}{h_{10}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} \Leftrightarrow T_1(\lambda_{10}) \rightarrow \infty\end{aligned}$$

There exists a constant C such that $h_{1t} < C e^{B_1 t}$ for any $t \geq 0$. Hence for T_1 big enough we have

$$\frac{h_{0T_1}}{h_{1T_1}} > \frac{e^{B_0 T_1}}{C e^{B_1 T_1}} = \frac{e^{(B_0 - B_1)T_1}}{C} > \frac{\pi_1}{\pi_0} \left[\frac{B_0}{\rho(T_1)} - 1 \right]$$

or, equivalently, $B_0/\rho(T_1) < 1 + H_{0T_1}/H_{1T_1}$. Since $B_0/\rho(0) > 1 + H_{00}/H_{10}$, there exists $\hat{\lambda}$ such that for $T_1 = T_1(\hat{\lambda})$:

$$\frac{B_0}{\rho(T_1)} = 1 + \frac{H_{0T_1}}{H_{1T_1}}$$

We observe that $T_1(\lambda_{10}) < T_2(\lambda_{10})$ for $1/(B_1 h_{10}) < \lambda_{10} < \hat{\lambda}$. Differentiating $\phi(t) = 0$ with respect to t and λ_{10} , we find at $t = T_2$:

$$T_1'(\lambda_{10}) = \frac{B_1 h_{10} / \sum_{j=0}^1 \pi_j e^{-\rho_j T_1}}{B_1 - \sum_{i=0}^1 \rho_i \frac{\pi_i e^{-\rho_i T_1}}{\sum_{j=0}^1 \pi_j e^{-\rho_j T_1}}} > 0$$

Moreover, $T_1(\hat{\lambda}) = T_2(\hat{\lambda})$.

Therefore,

$$\dot{\lambda}_{1t} = -\frac{\partial H_t}{\partial h_{1t}} = -\pi_1 u_{1t} \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} - B_1 \lambda_{1t} (1 - u_{1t}) = -\frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{h_{1T_1}} \quad (98)$$

At $t = T_1$, we have $\partial H_t / \partial u_{1t} = 0$, that is

$$\lambda_{1T_1} = \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i T_1}}{B_1 h_{1T_1}} \quad (99)$$

Integrating (98) from T_1 to $t \leq T_2$, we get

$$\lambda_{1t} = \lambda_{1T_1} + \frac{1}{h_{1T_1}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} (e^{-\rho_i t} - e^{-\rho_i T_1})$$

Replacing (99), we obtain

$$\lambda_{1t} = \frac{1}{h_{1T_1}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} \left[e^{-\rho_i t} + \left(\frac{\rho_i}{B_1} - 1 \right) e^{-\rho_i T_1} \right] \text{ for } T_1 \leq t \leq T_2 \quad (100)$$

and, in particular, (51).

At time T_2 , the optimal solution satisfies $\partial H_t / \partial u_{0t} = 0$ or, equivalently,

$$B_0 \lambda_{0t} = \pi_0 \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} \quad (101)$$

that is

$$\lambda_{0T_2} = \frac{\pi_0}{B_0 H_{1T_1}} \sum_{i=0}^1 \pi_i e^{-\rho_i T_2} \quad (102)$$

Using (101), we find

$$\dot{\lambda}_{1t} = -\frac{\partial H_t}{\partial h_{1t}} = -\pi_1 u_{1t} \frac{\sum_{i=0}^1 \pi_i e^{-\rho_i t}}{\sum_{i=0}^1 \pi_i h_{it} u_{it}} - B_1 \lambda_{1t} (1 - u_{1t}) = -\frac{\pi_1}{\pi_0} B_0 \lambda_{0t} = \frac{\pi_1}{\pi_0} \dot{\lambda}_{0t}$$

Integrating (98) from T_1 to $t \leq T_2$, we get

$$\lambda_{1t} = \frac{\pi_1}{\pi_0} (\lambda_{0t} - \lambda_{0T_1}) + \lambda_{1T_1} = \frac{\pi_1}{\pi_0} \lambda_{0T_1} e^{-B_0(t-T_1)} + \lambda_{1T_1} - \frac{\pi_1}{\pi_0} \lambda_{0T_1}$$

A candidate solution satisfies the transversality condition, that is

$$\lim_{t \rightarrow \infty} (\lambda_{1t} h_{1t}) = \lim_{t \rightarrow \infty} \left[\frac{\pi_1}{\pi_0} \lambda_{0T_1} e^{-B_0(t-T_1)} h_{1t} \right] + \lim_{t \rightarrow \infty} \left[\left(\lambda_{1T_1} - \frac{\pi_1}{\pi_0} \lambda_{0T_1} \right) h_{1t} \right] = 0$$

which holds only if $\lambda_{1T_1} - \lambda_{0T_1} \pi_1 / \pi_0 = 0$. This implies $\lambda_{1t} = \lambda_{0t} \pi_1 / \pi_0$. In particular, at the point $T_2 > T_1$, we have $\lambda_{1T_2} = \lambda_{0T_2} \pi_1 / \pi_0$ and, according to (102),

$$\lambda_{1T_2} = \frac{\pi_1}{\pi_0} \lambda_{0T_2} = \frac{1}{B_0 h_{1T_1}} \sum_{i=0}^1 \pi_i e^{-\rho_i T_2}$$

Using (51), we obtain (38). To conclude, we must prove that there exists λ_{10} with

$$\frac{1}{h_{10}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i} < \lambda_{10} < \hat{\lambda}$$

such that the equation (38) is verified. Indeed, for

$$\lambda_{10} = \frac{1}{h_{10}} \sum_{i=0}^1 \frac{\pi_i}{\rho_i}$$

we have $T_1(\lambda_{10}) = 0$ and $T_2(\lambda_{10}) = T_0$ and, so,

$$f(\lambda_{10}) \equiv \frac{\frac{1}{\rho(T_0)} - \frac{1}{B_0}}{\frac{1}{\rho(0)} - \frac{1}{B_1}} - \frac{1}{\sum_{i=0}^1 \pi_i e^{-\rho_i T_0}} < 0$$

because of (??). For $\lambda = \hat{\lambda}$, we have $T_1(\lambda_{10}) = T_2(\lambda_{10})$ and, so,

$$f(\lambda_{10}) \equiv \frac{\frac{1}{\rho(T_1(\lambda_{10}))} - \frac{1}{B_0}}{\frac{1}{\rho(T_1(\lambda_{10}))} - \frac{1}{B_1}} - 1 > 0$$

since $B_0 > B_1$. f is a continuous function. Then, there exists λ_{10} such that $f(\lambda_{10}) = 0$, that is (38) is satisfied.

Finally, we can construct the explicit solution with λ_{10} solution of (38).

We have $\lambda_{0t} = \lambda_{00} e^{-B_0 t}$ for any t with $\lambda_{00} = \lambda_{1T_2} e^{B_0 T_2} \pi_0 / \pi_1$. Moreover, $h_{0t} = h_{00} e^{B_0 t}$ for $0 \leq t \leq T_2$, while for $t \geq T_2$, we obtain (47) similarly to (41). We have also $u_{0t} = 0$ for $0 \leq t \leq T_2$, while, for $t \geq T_2$, we obtain (48) similarly to (43).

For $0 \leq t \leq T_1$, from (94) and (97), we obtain (50). We observe that $h_{1t} = h_{1T_1}$ for $t \geq T_1$.

For $0 \leq t \leq T_1$, it is easy to see that $\lambda_{1t} = \lambda_{10} e^{-B_0 t}$, while, for $T_1 \leq t \leq T_2$, λ_{1t} is given by (100). For $t \geq T_2$, we find $\lambda_{1t} = \lambda_{1T_2} e^{-B_0(t-T_2)}$. ■

Proof of Proposition 15. Under the assumption $h_0 > h_1$, the Gini index of wealth is given by:

$$g_{ht} = \frac{\frac{\pi_0 h_{0t} + \pi_1 h_{1t}}{2} - \left(\frac{\pi_1 \pi_1 h_{1t}}{2} + \pi_0 \pi_1 h_{1t} + \frac{\pi_0 \pi_0 h_{0t}}{2} \right)}{\frac{\pi_0 h_{0t} + \pi_1 h_{1t}}{2}}$$

that is by

$$g_{ht} = \pi_1 - \frac{\pi_1 h_{1t}}{\pi_0 h_{0t} + \pi_1 h_{1t}} = \pi_1 - \frac{H_{1t}}{H_{0t} + H_{1t}}$$

or, equivalently, by (55). ■

Proof of Proposition 16. According to (25), $\rho_0 < \rho_1$ implies $c_{0t} > c_{1t}$. Thus, the Gini index of consumption is given by

$$g_{ct} = \pi_1 - \frac{\pi_1 c_{1t}}{\pi_0 c_{0t} + \pi_1 c_{1t}} \quad (103)$$

Replacing (25) in (103) we get (56). ■

Proof of Proposition 17. Focus on Proposition 10. According to (25), we obtain the aggregate consumption:

$$c_t = \sum_{i=0}^1 \pi_i A e^{-\rho_i t} \frac{\sum_{j=0}^1 \pi_j h_{jt} u_{jt}}{\sum_{j=0}^1 \pi_j e^{-\rho_j t}} = A (\pi_0 h_{0t} u_{0t} + \pi_1 h_{1t} u_{1t})$$

and, so, replacing expressions (30) to (33), we find (57). Computing the time derivative of (57), we obtain the aggregate consumption growth rate (58). ■

Proof of Proposition 18. From equation (56), we find

$$e^{(B_0 - \rho_i)t} = \left(\frac{\pi_1 \pi_0 + g_{ct}}{\pi_0 \pi_1 - g_{ct}} \right)^{\frac{B_0 - \rho_i}{\rho_1 - \rho_0}} \quad (104)$$

Replacing (104) in (58) we find (59). Finally, it is easy to show that γ_c is an increasing function of g_c . ■

References

- [1] Abramovitz M. and P.A. David (2000). American macroeconomic growth in the era of knowledge-based progress: the long-run perspective. In Engerman S.L. and R.E. Gallman editors, *The Cambridge Economic History of the United States*, Cambridge University Press, New York.
- [2] Alesina A. and D. Rodrik (1994). Distributive politics and economic growth. *Quarterly Journal of Economics* 109, 465-490.

- [3] Atkinson A.B. (1999). The distribution of income in the UK and OECD countries in the twentieth century. *Oxford Review of Economic Policy*, 15, 56-75.
- [4] Autor D.H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science* 344, 843-851.
- [5] Autor D.H., L.F. Katz and M.S. Kearney (2008). Trends in U.S. wage inequality: revising the revisionists. *Review of Economics and Statistics* 90, 300-323.
- [6] Banerjee A. and A.F. Newman (1993). Occupational choice and the process of development. *Journal of Political Economy* 101, 274-298.
- [7] Barro R.J. (2000). Inequality and growth in a panel of countries. *Journal of Economic Growth* 5, 5-32.
- [8] Becker G.S. (1962). Investment in human capital: a theoretical analysis. *Journal of Political Economy* 70, 9-49.
- [9] Becker G.S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. National Bureau of Economic Research, New York.
- [10] Becker G.S. and N. Tomes (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy* 87, 1153-1189.
- [11] Becker R.A. (2006). Equilibrium dynamics with many agents. In Dana R.-A., C. Le Van, T. Mitra, and K. Nishimura editors, *Handbook of Optimal Growth 1*, Springer-Verlag.
- [12] Bénabou R. (1996). Inequality and growth. In Bernanke B.S. and J.J. Rotemberg editors, *National Bureau of Economic Research Macro Annual* vol. 11, MIT Press, Cambridge.
- [13] Bénabou R. (2000) Unequal societies: income distribution and the social contract. *American Economic Review*, 90, 96-129.
- [14] Bénabou R. (2002). Tax and education policy in a heterogeneous-agent economy: what levels of redistribution maximize growth and efficiency? *Econometrica* 70, 481-517.

- [15] Benjamin D.J., S.A. Brown and J.M. Shapiro (2013). Who is "behavioral"? Cognitive ability and anomalous preferences. *Journal of the European Economic Association* 11, 1231-1255.
- [16] Bourguignon F. (1981). Pareto-superiority of unegalitarian equilibria in Stiglitz' model of wealth distribution with convex savings function. *Econometrica* 49, 1469-1475.
- [17] Burks S.V., J.P. Carpenter, L. Goette and A. Rustichini (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences* 106, 7745-7750.
- [18] Campbell J. (1999). Asset prices, consumption and the business cycle. In J.B. Taylor and M. Woodford editors, *Handbook of Macroeconomics*, North-Holland, Amsterdam, 1231-1303.
- [19] Doepke M. and F. Zilibotti (2008). Occupational choice and the spirit of capitalism. *Quarterly Journal of Economics* 123, 747-793.
- [20] Dohmen T., A. Falk, D. Huffman and U. Sunde (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review* 100, 1238-1260.
- [21] Dohmen T., B. Enke, A. Falk, D. Huffman and U. Sunde (2015). Patience and the wealth of nations. Mimeo.
- [22] Eckstein Z. and I. Zilcha (1994). The effects of compulsory schooling on growth, income distribution and welfare. *Journal of Public Economics* 54, 339-359.
- [23] Eicher T.S. and C. García-Peñalosa (2001). Inequality and growth: the dual role of human capital in development. *Journal of Development Economics*, 66, 173-197.
- [24] Falk A., A. Becker, T. Dohmen, B. Enke, U. Sunde and D. Huffman (2015). The nature of human preferences: global evidence. *CESIFO Area Conference on Behavioral Economics*, 23-24 October.
- [25] Forbes K. (2000). A reassessment of the relationship between inequality and growth. *American Economic Review* 90, 869-887.

- [26] Frank M. (2009). Inequality and growth in the United States: evidence from a new state-panel of income inequality measures. *Economic Inquiry* 47, 55-68.
- [27] Frederick S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives* 19, 25-42.
- [28] Galor O. and O. Moav (2004). From physical to human capital accumulation: inequality and the process of development. *Review of Economic Studies* 71, 1001-1126.
- [29] Galor O. and O. Ozak (2016). The agricultural origins of time preference. *American Economic Review* 106, 3064-3103.
- [30] Galor O. and J. Zeira (1993). Income distribution and macroeconomics. *Review of Economic Studies* 60, 35-52.
- [31] Glomm G. and B. Ravikumar (1992). Public versus private investment in human capital: endogenous growth and income inequality. *Journal of Political Economy* 100, 818-834.
- [32] Goldin C. and L.F. Katz (1999). The returns to skill across the twentieth century United States. Working Paper NBER 7126, Cambridge, MA.
- [33] Goldin C. and L.F. Katz (2001). The legacy of U.S. educational leadership: notes on distribution and economic growth in the 20th century. *American Economic Review, Papers and Proceedings* 91, 18-23.
- [34] Goldin C. and L.F. Katz (2008). *The Race between Education and Technology*. Harvard University Press, Cambridge, MA.
- [35] Gosling A., S. Machin, C. Meghir (2000). The changing distribution of male wages in the U.K. *Review of Economic Studies* 67, 635-666.
- [36] Hanushek E.A. and L. Woessmann (2008). The role of cognitive skills in economic development. *Journal of Economic Literature* 46, 607-668.
- [37] Hanushek E.A. and L. Woessmann (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth* 17, 267-321.

- [38] Heal, G. and Millner, A. (2015). Collective intertemporal choice: time consistency vs time invariance. Working Paper 247, Centre for Climate Change Economics and Policy.
- [39] Hübner M. and G. Vannoorenberghe (2015). Patience and long-run growth, *Economics Letters* 137, 163-167.
- [40] Jones B.F. (2014). The human capital stock: a generalized approach. *American Economic Review* 104, 3752-3777.
- [41] Katz L.F. (1999). Changes in the wage structure and earnings inequality. In Ashenfelter O. and D. Card editors, *Handbook of Labor Economics* vol. 3A, Elsevier Science, North-Holland, Amsterdam.
- [42] Lemieux T. (2006). Increasing residual wage inequality: composition effects, noisy data, or rising demand for skill? *American Economic Review* 96, 461-498.
- [43] Lemieux T. (2010). What do we really know about changes in wage inequality? In Abraham K.G., J.R. Spletzer and M. Harper editors, *Labor in the New Economy*, University of Chicago Press, p. 17-59.
- [44] Lindley J. and S. McIntosh (2015). Growth in within graduate wage inequality: the role of subjects, cognitive skill dispersion and occupational concentration. *Labour Economics* 37, 101-111.
- [45] Lucas R.E. (1988). On the mechanics of economic development. *Journal of Monetary Economics* 22, 3-42.
- [46] Lucas R.E. (1993). Making a miracle. *Econometrica* 61, 251-172.
- [47] Lucas R.E. (2015). Reflections of new growth theory, human capital and growth. *American Economic Review, Papers and Proceedings* 105, 85-88.
- [48] Manuelli R.E. and A. Seshadri (2014). Human capital and the wealth of nations. *American Economic Review* 104, 2736-2762.
- [49] Mincer J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy* 66, 281-302.
- [50] Mincer J. (1962). On-the-job training: costs, returns, and some implications. *Journal of Political Economy* 70, 50-79.

- [51] Partridge M.D. (1997). Is inequality harmful for growth? Comment. *American Economic Review* 87, 1019-1032.
- [52] Perotti R. (1996). Growth, income distribution and democracy: what the data say. *Journal of Economic Growth* 1, 149-187.
- [53] Persson T. and G. Tabellini (1994). Is inequality harmful for growth? Theory and evidence. *American Economic Review* 84, 600-621.
- [54] Rae J. (1834). *The New Principles of Political Economy*. Reprinted in 1905 as *The Sociological Theory of Capital*, Macmillan, New York.
- [55] Rebelo S. (1991). Long-run policy analysis and long-run growth. *Quarterly Journal of Economics* 107, 255-284.
- [56] Saint-Paul G. and T. Verdier (1993). Education, democracy and growth. *Journal of Development Economics* 42, 399-407.
- [57] Schultz T.W. (1961). Investment in human capital. *American Economic Review* 51, 1-17.
- [58] Schultz T.W. (1962). Reflections on investment in man. *Journal of Political Economy* 70, 1-8.
- [59] Shane F. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives* 19, 25-42.
- [60] Stiglitz J.E. (1969). Distribution of income and wealth among individuals. *Econometrica* 37, 382-397.
- [61] Suen R. (2014). Time preference and the distributions of wealth and income. *Economic Inquiry* 52, 364-381.
- [62] Turnovsky S.J. (2011). The accumulation of human capital and income inequality in a two-sector economy. *Journal of Human Capital* 5, 418-452.
- [63] Turnovsky S.J. and A. Mitra (2013). The interaction between human and physical capital accumulation and the growth-inequality trade-off. *Journal of Human Capital* 7, 26-75.

- [64] Uzawa H. (1965). Optimum technical change in an aggregative model of economic growth. *International Economic Review* 6, 18-31.
- [65] Viaene J.-M. and I. Zilcha (2003). Human capital formation, income inequality, and growth. In Eicher T.S. and S.J. Turnovsky editors. *Inequality and Growth: Theory and Policy Implications*. MIT Press, Cambridge, MA.
- [66] Zuber S. (2011). The aggregation of preferences: can we ignore the past? *Theory and Decision* 70, 367-384.