Working Paper Series

Inequality and development: the role of opportunities and free-will

Gustavo A. Marrero
Juan G. Rodríguez

ECINEQ WP 2014 - 327

ECINEQ 2014 - 327
March 2014

www.ecineq.org

Inequality and development: the role of opportunities and free-will*

Gustavo A. Marrero†
Universidad de La Laguna, IUDR and CAERP, Spain

Juan G. Rodríguez
Universidad Complutense de Madrid, Spain

Abstract

Individual income is determined by free-will actions related to the level of effort exerted and by opportunities determined by aspects beyond the individual's control, such as family background, race, place of birth or health endowments. Taking human capital as the main engine of development, we build a general equilibrium overlapping generations model to study the effect of inequality of free-will and inequality of opportunity on real per capita income. When the process of human capital accumulation is convex, we show that inequality of opportunity has a negative and monotone impact on this degree of development, while the effect of inequality of free-will is positive. However, when human capital is accumulated according to a non-convex process, multiplicity of equilibria and poverty traps arise. For economies with a low percentage of trapped dynasties (rich), the same result as for the convex model is obtained. However, for poor economies, while the relationship between free-will inequality and development generally remains positive, that between inequality of opportunity and human capital takes on an inverted U-shape.

Keywords: Inequality of opportunity, inequality of free-will, human capital accumulation.

JEL Classification: D63, E24, O15, O40.

*We acknowledge helpful comments of F. Bourguignon and the audiences at the IARIW Conference (Boston), ECINEQ (Bari) and XX Encuentro de Economía Pública (Sevilla). The authors acknowledge the financial support of the Spanish Ministry through project ECO2010-17590. The usual disclaimer applies.

†Corresponding author. **Contacts details**: G.A.Marrero, Dpto. Análisis Económico, U. of La Laguna, Campus de Guajara, 38071, Canary Islands, Spain, email: gmarrero@ull.es. J.G.Rodriguez, Dpto. de Fundamentos del Análisis Econémico I, U. Complutense de Madrid, Campus de Somosaguas, Pozuelo de Alarcón, 28223 Madrid, Spain, email: juangabr@pdi.ucm.es.

1 Introduction

Since the pioneering work by Kuznetz (1955), economists have traditionally questioned whether inequality is good or bad for long-term growth. In a world of perfect equality, there would be no rewards for hard work and entrepreneurial activity, while in a world of vast inequality, those with silver spoons might coast along on their parents' wealth rather than create more, and those with wooden spoons might not develop their talent. Both extreme situations would cripple growth and, for this reason, the literature has extensively searched for the optimal amount of inequality, although no general consensus has been reached.[1] In this vein, the recent debate between those who think that top income share increases in the U.S. since the 1970s have not translated into higher economic growth (Piketty et al., 2011) and those who believe that incomes at the top have grown much faster than average because they have made significant economic contributions (Mankiw, 2013), could be explained by the different perception of the position reached by the economy, whether or not it is beyond the optimal level of inequality.

Instead, we argue in this paper that the critical issue in the inequality-growth debate is to make a distinction between the different types of inequality. Following the World Bank (2006), Bourguignon et al. (2007b) and Marrero and Rodríguez (2013), as with cholesterol, there are two types of inequality, inequality that enhances growth (good) and inequality that deters growth (bad). For example, if rent-seeking is the primary driving force behind the growing incomes of the rich (Stiglitz, 2012), opportunity for the rest of population would be reduced causing an increase in income inequality and a reduction in growth; on the contrary, if the root cause is the change in technology (Goldin and Katz, 2008; Mankiw, 2013), hard-working and talented individuals could command superstar incomes, prompting not only inequality but also growth. Unfortunately, both types of inequality are likely to be present at the same time, so they will offset each other, causing the final impact of overall inequality on growth to be ambiguous, as emphasized by the empirical literature.[2] Despite the fact that this distinction between bad and good inequality is at the center of the current debate on inequality, there is a lack of studies analyzing their simultaneous functioning. The main task of this paper is precisely to analyze the simultaneous impact of different sources of inequality on the economy under an integrated macro-framework.

Our framework departs from the idea that individual income and implied inequality is mainly determined by two factors (Roemer, 1993): first, free-will actions related to the

[1] See the surveys on this issue in Bénabou (1996), Aghion et al. (1999) and Bertola et al. (2005), among many others.

[2] See, for example, Banerjee and Duflo (2003) on the inconclusiveness of the cross-country empirical literature on economic inequality and growth. Previously, Barro (2000) and Voitchovsky (2005) pointed out that inequality might affect growth through distinct avenues that may offset each other.

level of exerted effort; second, opportunities, which are beyond the individual's control because they depend on circumstances like family background, race, place of birth or health endowments. Taking human capital as the main engine of development, we build a general equilibrium overlapping generations model to study the effect of inequality of free-will (or pure effort) and inequality of opportunity on growth.[3] The model combines the basic principles of the wage determination and human capital accumulation literature (Glomm and Ravikumar, 1992; Boldrin and Montes, 2005) with that of inequality-of-opportunity (Roemer, 1998; Fleurbaey, 2008). The economy is populated by a continuum of dynasties with warm-glove preferences, where effort is considered as a non-monetary factor which generates disutility, but is needed for the individual to accumulate human capital. Following Roemer (1998), the level of effort is considered to be an endogenous decision that depends not only on individual free-will, but also on the set of circumstances.[4] For example, high parental human capital creates a better environment for the accumulation of human capital (Galor and Tsidon, 1997), and favors a bequest to the offspring in the form of the quality of schooling (Glomm and Ravikumar, 1992). With all these ingredients, the ultimate sources of heterogeneity in the model come from individual free-will, initial parental human capital (an endogenous circumstance) and exogenous circumstances like race, place of birth or health endowments.

In the first part of the paper (Section 2), human capital is accumulated following a convex process, whose solution is characterized by a globally stable, dynasty-specific steady-state equilibrium. In this context, we are able to reproduce the classical decomposition of total income inequality into inequality of opportunity and inequality of pure effort (Ruiz-Castillo, 2003; Checchi and Peragine, 2010). More importantly, for a widely used human capital accumulation process, we show that human capital accumulation is concave with respect to the set of circumstances but convex with respect to individual free-will. As a result, a more equal distribution of circumstances increases the average human capital, while the opposite happens for free-will. Accordingly, the impact of total inequality on average human capital is ambiguous and the sign of the effect depends on

[3] In a first empirical attempt, Marrero and Rodríguez (2013), for panel data for the U.S. states between 1970 and 2000, found robust evidence that inequality of effort is growth enhancing, while inequality due to differences in opportunities is growth deterring. On the theoretical side, as far as we are aware, only the static general equilibrium model in Mejia and St-Pierre (2008) has proposed, for a given distribution function of exogenous circumstances, that there is no trade-off between the average level of human capital and equality of opportunity.

[4] The important role of circumstances has largely been emphasized in the literature on inequality of opportunity. For example, Roemer (2000) and Bowles et al. (2005) have shown that even if individuals have high inborn talent, the likelihood of their being able to realize the benefits of that talent (for example, in terms of admission to university or access to employment) will be affected by social conditions. This reasoning also applies to individuals with strong preferences for effort. See Arrow et al. (2000), Hertz et al. (2008), Blume and Durlauf (2001), Durlauf (2003) and Loury (1989).

which type of inequality dominates the aggregate inequality. In addition, we show that the impact of each source of inequality on the aggregate, and hence the sign of the relationship, strongly depends on key characteristics, such as the degree of meritocracy in the economy.

Based on the extensive literature that emphasizes the relevance of considering poverty traps and multiplicity of equilibria in development models, the second part of the paper (Section 3) extends the initial framework by adopting a non-convex process for the accumulation of individual human capital.[5] In this case, each dynasty faces, a priory, two potential equilibria, a low one and a high one. The low equilibrium can be interpreted as a human-capital trap from which the amount of effort required to escape is too high. This situation could arise even for a dynasty with large free-will to exert effort if it possesses strongly unfavorable circumstances and/or lives in a poor economy where the return to effort is very small. After characterizing multiplicity, we analyze the average human capital of the aggregate economy, its dynamics and its relationship with the different sources of inequality in the initial period and over the long-run. We find that in general the effect of the different sources of inequality on human capital depends on the initial levels of development and inequality in the economy. Because poor economies are characterized by large poverty rates (so a large fraction of their population is trapped), increasing any type of inequality (opportunity or pure effort) moves dynasties out of the trap, which benefits the average human capital. But, when the level of development increases and, therefore, a higher fraction of dynasties are not trapped, predictions converge to those of the convex model: the negative impact of inequality of opportunity and the positive impact of inequality of pure effort on per capita income.

The results in Sections 2 and 3 are complemented with some simulations and an empirical illustration (Section 4). Based on existing estimations of inequality of opportunity for 40 countries, we contrast the main results of the paper with a simple correlation analysis. Finally, the main conclusions of the paper are set out in Section 5.

2 A basic model of human capital and inequality

Our framework is a small open overlapping-generations economy with heterogeneous agents, populated by a continuum of dynasties, each one indexed by $i \equiv [0,1]$, and with perfectly competitive markets. Time t is discrete and a single homogenous good, y, is produced every period according to the following Cobb-Douglas technology:

$$y_t = A \cdot k_t^{\lambda} \widetilde{l_t}^{1-\lambda}, \; A > 0, \, \lambda \in (0,1). \tag{1}$$

[5] We assume a standard non-convex function for human capital without specifying the mechanism by which the multiple equilibria emerge. See, among others, Azariadis and Drazen (1990), Banerjee and Duflo (2003) or Azariadis and Stachurski (2005).

The Arrow-neutral technological term A is assumed to be constant, k_t is physical capital and $\widetilde{l_t} = l_t \cdot h_t$ is the efficiency units of labor, with l_t as raw labor and h_t as the average level of human capital,

$$h_t = \int_0^1 h_t(i) \cdot dG\left[h(i)\right], \tag{2}$$

where $G\left[h(i)\right]$ is the distribution function of individual human capital.

Raw labor is perfectly inelastically supplied and, without loss of generality, normalized to 1. The small open economy has unrestricted international borrowing and lending, thus the real interest rate is exogenous and equal to the stationary world interest rate $\overline{r}$.[6] Since producers operate in a perfectly competitive environment, $\overline{r}$ determines the k_t/h_t constant ratio,

$$\overline{r} = y'_k = A \cdot \lambda \cdot \left(\frac{k_t}{h_t}\right)^{\lambda - 1} \Rightarrow \left(\frac{k_t}{h_t}\right) = \left(\frac{A \cdot \lambda}{\overline{r}}\right)^{1/(1-\lambda)}, \tag{3}$$

and the wage per unit of human capital (or effective labor) is given by,

$$w = y'_{\widetilde{l}} = A \cdot (1 - \lambda) \cdot \left(\frac{k_t}{h_t}\right)^{\lambda} = A^{1/(1-\lambda)} \cdot (1 - \lambda) \cdot \left(\frac{\lambda}{\overline{r}}\right)^{\lambda/(1-\lambda)}, \tag{4}$$

which increases with A, decreases with $\overline{r}$ and is constant as long as A and $\overline{r}$ are also constant.[7] Plugging (3) into (1),

$$y_t = A \cdot \left(\frac{k_t}{h_t}\right)^{\lambda} h_t = A^{1/(1-\lambda)} \left(\frac{\lambda}{\overline{r}}\right)^{\lambda/(1-\lambda)} h_t. \tag{5}$$

Moreover, given h_t from (2), k_t is obtained from (3). Hence, given A, λ and $\overline{r}$, real output y_t and k_t are fully determined by h_t, which is used accordingly as a proxy of development.

Each dynasty i born at t consists of a common individual who lives two periods, childhood and adulthood. During adulthood, the individual gives birth to another individual so the overall population remains constant over time. Individuals are born with a set of factors, called circumstances and denoted by $\theta_t(i)$ (Roemer, 1993; Fleurbaey, 2008), which are beyond their own control but affect their actions. Circumstances are generally associated with factors related to parental socio-economic background (family status, social connections, child nourishment, etc.), represented in our model by parental human capital, $h_{t-1}(i)$ and the bequest devoted to the quality of the offpring's education, $x_{t-1}(i)$

[6] The choice of a small open economy is to simplify the notation and based on the fact that interest rates do not change significantly in the course of economic growth (Galor and Tsidon, 1997).

[7] We are not assuming the existence of a global technological externality in the economy, which, for example, would require A to be a function of h_t (Benabou, 1996; Galor and Tsidon, 1997). This assumption is not needed to obtain our main results and it just complicates the solution of the model.

(Card and Krueger, 1992; Glomm and Ravikumar, 1992), and by other factors such as race, nationality or health endowments, which, for simplicity, are all grouped in $a(i)$ (Bourguignon et al., 2007a; Rodríguez, 2008; Li Donni et al., 2012).[8] Following Roemer (1998) and Ferreira and Gignoux (2011), among others, $\theta_t(i)$ can be seen as a composite index of $h_{t-1}(i)$, $x_{t-1}(i)$ and $a(i)$,[9]

$$\theta_t(i) = a(i)^{1-\alpha-\varphi} \cdot x_{t-1}(i)^{\alpha} \cdot h_{t-1}(i)^{\varphi};\ \alpha, \varphi \in (0,1),\ \alpha + \varphi < 1. \tag{6}$$

Individuals show warm-glove preferences, which depend on consumption, $c_t(i)$, the bequest to their offspring, $x_t(i)$, and total effort, $e_t(i)$, during adulthood

$$u_t(i) = \pi(\eta) \cdot c_t(i)^{\eta} \cdot x_t(i)^{1-\eta} - \gamma(i) \cdot e_t(i)^{1+\beta}, \tag{7}$$

where, without loss of generality, we assume that consumption during childhood is included in the consumption of the parents (Benabou, 2000), $\eta \in (0,1)$ is a parameter of relative preferences between $c(i)$ and $x(i)$, and $\pi(\eta) = \eta^{-\eta}(1-\eta)^{-(1-\eta)}$ is a normalization factor (Acemoglu, 2010). Effort is a non-monetary factor that generates disutility, but is needed to accumulate human capital (Agion and Bolton, 1997; Roemer, 1998). We assume $\beta > 0$ so that the marginal disutility of effort is increasing.[10] Moreover, individual preferences for bundles of effort and consumption (and bequest) are determined by a *free-will* parameter, $\gamma(i) > 0$, related to the individual's preferences for effort and which society views as within the jurisdiction of personal responsibility because it is independent of $\theta_t(i)$ (Roemer, 1998).

We start by assuming a simple individual human capital accumulation rule and leave the treatment of non-convexities for the next section. Thus, $h_t(i)$ is determined by two non-purchasable but complementary factors (Sen, 1980; Roemer, 1993), effort and circumstances, showing constant returns to scale,

$$h_t(i) = \theta_t(i)^{\psi} \cdot e_t(i)^{1-\psi},\ \psi \in (0,1), \tag{8}$$

[8] Macroeconomists have extensively shown that parental education and resources devoted to the offspring's education have significant effects on the individual's human capital, while, for example, school characteristics have relatively little importance in determining individual achievement (Coleman et al., 1966; Becker and Tomes, 1986; Hanushek, 1996; Ginther et al., 2000). In our framework, it can be interpreted that h_{t-1} creates a better environment for the accumulation of human capital (i.e., a local home environmental externality, Galor and Tsidon, 1997), while x_{t-1} favors the bequest to the offspring in the form of quality of schooling (Glomm and Ravikumar, 1992). See also Boldrin and Montes (2005) for a general discussion of this issue.

[9] Because inborn ability or talent is less than perfectly correlated between generations, a model that explicitly models how it evolves in the dynasty over time would be required (Hasler and Rodriguez-Mora, 2000). Another source of inequality beyond the scope of this paper is luck (Lefranc et al., 2009). Mejía and St-Pierre (2008) consider the whole set of $\theta(i)$ as exogenous.

[10] The effort function $\gamma(i) \cdot e_t(i)^{1+\beta}$ is convex, as always assumed in the literature on the economics of information (Macho-Stadler and Pérez-Castrillo, 1996).

where ψ denotes the relative importance of circumstances with respect to effort in determining $h(i)$.[11] Given $h(i)$, individuals work during their adulthood (supplying one unit of labor inelastically) and earn labor income,

$$w_t(i) = w \cdot h_t(i). \tag{9}$$

The source of heterogeneity comes from differences in $\Gamma(i) = \{a(i), h_{-1}(i), \gamma(i)\}$, which characterizes each dynasty. Following Benabou (1996), we assume that a, h_{-1} and γ follow mean invariant log normal independent distributions:[12]

$$\ln a \sim N\left(\ln \hat{a} - \frac{\Delta_a^2}{2}, \Delta_a^2\right), \tag{10}$$

$$\ln h_{-1} \sim N\left(\ln \widehat{h} - \frac{\Delta_{-1}^2}{2}, \Delta_{-1}^2\right), \tag{11}$$

$$\ln \gamma \sim N\left(\ln \widehat{\gamma} - \frac{\Delta_\gamma^2}{2}, \Delta_\gamma^2\right). \tag{12}$$

In this manner, a, γ and h_{-1} have constant means equal to $\hat{a}$, $\widehat{\gamma}$ and $\widehat{h}$, that are independent of the corresponding variances. Moreover, the variance term is closely related to the class of relative inequality indices consistent with the Lorenz curve (Cowell, 2009), such as the Gini or Mean Logarithmic Deviation (MLD). In fact, the MLD index (T_0) is half the variance under log-normality.[13] For simplicity, a and γ do not affect initial parental human capital h_{-1}, although they will strongly affect human capital accumulation along the transition (more on this later).

2.1 Solving the model

Each individual takes $\theta_t(i)$ and $\Gamma(i)$ as given and maximizes (7) subject to

$$c_t(i) + x_t(i) \leq w_t(i), \tag{13}$$

which is satisfied with equality because utility is strictly monotonic in c and x. For simplicity, time subscript is omitted from now on whenever it is not strictly necessary.

[11] For $\psi = 0$ the role of circumstances is null, hence the economy is purely meritocratic (Lucas, 1995); meanwhile, the higher ψ, the greater the degree of nepotism.

[12] Two reasons justify the general use of the lognormal distribution. First, lognormal distributions capture reasonably well the negative skewness that characterizes income distributions in practice. Second, the product of independent normal distributions converges asymptotically to a lognormal distribution (Gibrat, 1957). Accordingly, we can view income generation as the product of multiple factors over time.

[13] The MLD index has a path-independent additive decomposition (Foster and Shneyerov, 2000) and, for this reason, it is the inequality index used most in the empirical literature on inequality of opportunity (Ferreira and Gignoux, 2011; Marrero and Rodríguez, 2011).

The problem is solved in two steps. First, taking $h(i)$ as given, (7) is maximized subject to (13) and (9), obtaining

$$e^1(i) = \left[\frac{h(i)}{\theta(i)^\psi}\right]^{\frac{1}{1-\psi}}, \tag{14}$$

$$c(i) = \eta{\cdot}w{\cdot}h(i), \tag{15}$$

$$x(i) = (1-\eta){\cdot}w{\cdot}h(i). \tag{16}$$

Substituting these expressions into (7), the following indirect utility function V, in terms of $h(i)$, is obtained:

$$V = w{\cdot}h(i) - \gamma(i){\cdot}\left[\frac{h(i)}{\theta(i)^\psi}\right]^{\frac{1+\beta}{1-\psi}}. \tag{17}$$

In the second step, (17) is maximized with respect to $h(i)$, and the solution is substituted into (9) and (14),

$$e(i) = \left[\frac{(1-\psi){\cdot}w}{\gamma(i){\cdot}(1+\beta)}\right]^{\frac{1}{\beta+\psi}} \theta(i)^{\frac{\psi}{\beta+\psi}}, \tag{18}$$

$$h(i) = \left[\frac{(1-\psi){\cdot}w}{\gamma(i){\cdot}(1+\beta)}\right]^{\frac{1-\psi}{\beta+\psi}} \theta(i)^{\frac{(1+\beta){\cdot}\psi}{\beta+\psi}}, \tag{19}$$

$$w(i) = \left[\frac{(1-\psi){\cdot}w^{\frac{1+\beta}{1-\psi}}}{\gamma(i){\cdot}(1+\beta)}\right]^{\frac{1-\psi}{\beta+\psi}} \theta(i)^{\frac{(1+\beta){\cdot}\psi}{\beta+\psi}}. \tag{20}$$

Expressions for $c(i)$ and $x(i)$ are obtained by plugging (19) into (15) and (16), respectively. Three comments are in order. First, absolute effort $e(i)$ depends on the characteristics of the aggregate economy, w, and on dynasty-specific characteristics, $\theta(i)$ and $\gamma(i)$. Second, the free-will parameter $\gamma(i)$, while affecting effort, is by definition independent of circumstances. In this manner, free will represents what in the inequality-of-opportunity literature is referred to as *pure effort* (Roemer, 1998), the part of total effort not influenced by individual circumstances. Third, the set of circumstances affects individual human capital and wages through two different channels: a *direct* channel related to the returns-to-effort and represented by $\theta(i)^\psi$ in (8), and an *indirect* channel related to the positive impact of circumstances on absolute effort and represented by $\theta(i)^{\frac{\psi}{\beta+\psi}}$ in (18).[14]

[14] For example, Roemer observes that "Asian children generally work hard in school and thereby do well because parents press them to do so." (Roemer, 1998, p.22). Family pressure is clearly a circumstance that affects children´s effort but is outside their control.

2.2 Human capital dynamics and inequality

Using (16), we can rewrite $\theta(i)$ in terms of $h_{t-1}(i)$,

$$\theta(i) = (1-\eta)^{\alpha} \cdot w^{\alpha} \cdot a(i)^{1-\alpha-\varphi} h_{t-1}(i)^{\alpha+\varphi} \tag{21}$$

and using (21) and (18) we can derive the following dynamic equation for $h_t(i)$:

$$\begin{aligned} h_t(i) &= \zeta[h_{t-1}(i)] = \exp^{\frac{G}{\beta+\psi}} \left[\frac{a(i)^{[(1+\beta)\cdot\psi-\vartheta]} h_{t-1}(i)^{\vartheta}}{\gamma(i)^{1-\psi}} \right]^{\frac{1}{\beta+\psi}}, \\ G &= (1-\psi)\ln\left(\frac{(1-\psi)\cdot w}{1+\beta}\right) + (1+\beta)\psi \ln\left[(1-\eta)^{\alpha} w^{\alpha}\right], \\ \vartheta &= (1+\beta)\psi\,(\alpha+\varphi), \end{aligned} \tag{22}$$

where $\zeta[0] = 0$, $\zeta[\cdot]$ is C^2 on $(0, +\infty)$ and, because $\vartheta < \beta+\psi$, $\zeta[h_{t-1}(i)]$ is strictly increasing and strictly concave in $h_{t-1}(i)$. Under these conditions, there exists a unique steady-state that solves the fixed point $h_\infty(i) = \zeta[h_\infty(i)]$,

$$h_\infty(i) = \exp^{\frac{G}{\beta+\psi-\vartheta}} \left[\frac{a(i)^{[(1+\beta)\psi-\vartheta]}}{\gamma(i)^{1-\psi}} \right]^{\frac{1}{\beta+\psi-\vartheta}}. \tag{23}$$

Moreover, the steady-state is globally stable, i.e., all dynasties reach their long-run equilibrium regardless of their initial level of parental education.

Let's denote by $\mu_t = E[\ln h_t]$ our proxy of average per capita real income, in logs (recall (5)), and by $\Delta_t^2 = Var(\ln h_t)$ our proxy of wage inequality. In fact, under log-normality, Δ_t^2 is exactly twice the MLD index of wage inequality. However, we do not need the log-normality assumption to obtain the main results of this section, it just simplifies the algebra and permits their interpretation in terms of inequality.[15] Then, taking logs in (22), it is easy to show that $\ln h_t \sim N[\mu_t, \Delta_t^2]$ for all t, where the dynamics of the first and second moments are characterized in the following proposition.

Proposition 1 *For any generation $t \geq 0$, the dynamics of μ_t and Δ_t^2 are:*

[15] The variance of the logarithms, the index of inequality most broadly used in macroeconomics, is an appropriate index of inequality only if the distribution under consideration is lognormal, otherwise it does not fulfill a normative minimum, the principle of progressive transfers (Foster and Ok, 2003). This principle says that any transfer of income from a richer to a poorer individual, which does not reverse which of the two is richer, reduces inequality. Accordingly, we assume log-normality to guarantee that the variance of the logarithms is consistent with the Lorenz curve.

$$\mu_t = \xi^{t+1}\mu_{-1} + \left[1-\xi^{t+1}\right]\mu_\infty, \quad (24)$$

$$\Delta_t^2 = \xi^{2(t+1)}\Delta_{-1}^2 + \left[1-\xi^{t+1}\right]^2\Delta_\infty^2, \quad (25)$$

where $\xi = \frac{\vartheta}{\beta+\psi}$, $\mu_{-1} = \ln\widehat{h} - \Delta_{-1}^2/2$ and Δ_{-1}^2 are the initial conditions, and

$$\mu_\infty = \frac{G}{\beta+\psi-\vartheta} + \frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}E\ln a - \frac{1-\psi}{\beta+\psi-\vartheta}E\ln\gamma, \quad (26)$$

$$\Delta_\infty^2 = \left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}\right)^2\Delta_a^2 + \left(\frac{1-\psi}{\beta+\psi-\vartheta}\right)^2\Delta_\gamma^2, \quad (27)$$

are the steady-states of μ_t and Δ_t^2, respectively, which are globally stable.
Proof. *See Appendix A1* ■

For illustrative purposes, (24) and (25) are expressed for $t=0$,

$$E[\ln h_0] = \mu_0 = \frac{G}{\beta+\psi} + \frac{(1+\beta)\psi-\vartheta}{\beta+\psi}E\ln a + \xi\cdot E\ln h_{-1} - \frac{1-\psi}{\beta+\psi}E\ln\gamma, \quad (28)$$

$$Var(\ln h_0) = \Delta_0^2 = \left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi}\right)^2\Delta_a^2 + \xi^2\cdot\Delta_{-1}^2 + \left(\frac{1-\psi}{\beta+\psi}\right)^2\Delta_\gamma^2. \quad (29)$$

where μ_0 and $\Delta_0^2/2$ are the degree of development and inequality in the very short-run.

Taking μ_t and μ_{t-1} as the average child and parent outcomes, respectively, the parameter ξ in (24) and (28) can be interpreted as the elasticity of intergenerational mobility in human capital for an extended version of the canonical Galton model (Galton, 1869).[16] In fact, given $\alpha+\varphi>0$, a society with perfect intergenerational mobility, $\xi=0$, would be equivalent to a pure meritocracy economy, $\psi=0$.[17] Moreover, strict concavity of the function $h(i)$ is equivalent to having $\xi<1$, which is fully consistent with the empirical evidence found in the related literature (Corak, 2013). This fact provides confidence with the results set out in the following three corollaries, since they are based on the concavity of the $h(i)$ function.

Three results arise from the analysis of (24)-(27). First, the next corollary reproduces the classical decomposition of total income inequality into inequality of opportunity (rep-

[16] The canonical Galton model is the regression model $\ln Y_{i,t} = \upsilon + \xi\ln Y_{i,t-1} + \varepsilon_i$, where the coefficient ξ is the so-called elasticity of intergenerational mobility. In our case, we have $\mu_t = \frac{G}{\beta+\psi} + \frac{(1+\beta)\psi-\vartheta}{\beta+\psi}E\ln a - \frac{1-\psi}{\beta+\psi}E\ln\gamma + \frac{\vartheta}{\beta+\psi}\mu_{t-1}$, after taking logs in (22) and then expected values.

[17] In general, we have that $\frac{\partial\xi}{\partial\psi}>0$, $\frac{\partial\xi}{\partial(\alpha+\varphi)}>0$ and $\frac{\partial\xi}{\partial\beta}<0$.

resented by Δ_a^2 and Δ_{-1}^2) and inequality of (pure) effort (related to Δ_γ^2) (Ruíz-Castillo, 2003; Checchi and Peragine, 2010).[18]

Corollary 1. *For any generation $t \geq 0$, Δ_t^2 is a linear function of the three primitive sources of inequality Δ_a^2, Δ_γ^2 and Δ_{-1}^2:*

$$\begin{aligned}\Delta_t^2 &= \Delta_a^2\left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}\right)^2\left[1-\xi^{t+1}\right]^2 \\ &+\Delta_\gamma^2\left(\frac{1-\psi}{\beta+\psi-\vartheta}\right)^2\left[1-\xi^{t+1}\right]^2 \\ &+\Delta_{-1}^2\xi^{2(t+1)}\end{aligned} \quad (30)$$

Proof. *This result is straightforward from* (25) and (27) ■

Second, the next corollary shows that depending on the type of inequality (free-will or opportunity), its relationship with average human capital is positive (inequality of pure effort) or negative (inequality of opportunity).[19]

Corollary 2. *The human capital average μ_t and real wage average w_t are:*
i) negatively affected by Δ_a^2 for all $t \geq 0$.
ii) negatively affected by Δ_{-1}^2 for $t \neq \infty$, otherwise no affected by Δ_{-1}^2.
iii) positively affected by Δ_γ^2 for all $t \geq 0$.

Proof. *From (10)-(12), we have $E\ln a = \ln\widehat{a} - \Delta_a^2/2$, $E\ln\gamma = \ln\widehat{\gamma} - \Delta_\gamma^2/2$ and $E\ln h_{-1} = \ln\widehat{h} - \Delta_{-1}^2/2$. Then, from (24), it is clear that $\partial\mu_t/\partial\Delta_a^2 = -\frac{1}{2}\frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}\left[1-\xi^{t+1}\right] < 0$, $\partial\mu_t/\partial\Delta_{-1}^2 = -\frac{1}{2}\xi^{t+1} < 0$ and $\partial\mu_t/\partial\Delta_\gamma^2 = \frac{1}{2}\frac{1-\psi}{\beta+\psi-\vartheta}\left[1-\xi^{t+1}\right] > 0$ for $t \neq \infty$. However, at steady-state (26), the signs of $\partial\mu_\infty/\partial\Delta_a^2$ and $\partial\mu_\infty/\partial\Delta_\gamma^2$ remain unchanged, while $\partial\mu_\infty/\partial\Delta_{-1}^2 = 0$. From (9), obtaining the corresponding results for w_t is straightforward* ■

This result comes from the properties of (22) and (23).[20] While $h_t(i)$ is strictly increasing and strictly concave with respect to $h_{t-1}(i)$ and $a(i)$, $h_t(i)$ is decreasing and strictly convex with respect to $\gamma(i)$. Accordingly, keeping means $\widehat{a}$, $\widehat{h}$ and $\widehat{\gamma}$ unchanged, equalizing

[18] An alternative way of interpreting this decomposition is the following. Assume that people are grouped by their circumstances. Then, total inequality could be decomposed into a between-groups component (the inequality-of-opportunity part) and a within-groups component (the inequality-of-pure-effort part).

[19] This result has already been discussed in World Bank (2006) and Bourguignon et al. (2007b), and tested empirically in Marrero and Rodríguez (2013), but a theoretical model was missing.

[20] Bearing in mind that $E[h(t)] = e^{\mu_t + \Delta_t^2/2}$, it can be shown that the effects of Δ_a^2, Δ_{-1}^2 and Δ_γ^2 on $E(h_t)$ have the same sign as the corresponding effects on $\mu_t = E[\ln h_t]$. Again, the concavity (or intergenerational elasticity) condition, $\frac{\vartheta}{\beta+\psi} < 1$, is necessary for the result.

circumstances and/or increasing the dispersion of free-will (i.e., of pure effort) raises the average level of human capital and thus the real income in the economy. A relevant aspect of this result is that it has been obtained through a very simple framework where the concavity of the human capital accumulation function has not been imposed, but rather endogenously obtained.

The empirical literature that studies the relationship between total inequality and growth has mainly focused on the estimation of the equation $g_t = a + bT_0(w_t) + cZ$, where g_t is per capita income (in logs) or income growth, $T_0(w_t)$ is a measure of total income inequality and Z is an array of other controls. Although no agreement has been reached about the sign of b (recall the Introduction and references therein), the corollary below attempts to shed some light on this issue. Taking μ_t as a proxy of per capita income (or growth), for expository purposes, the corollary is divided into two parts: Part A assumes that $T_0(w_t)$ and Z are independent, while Part B considers that Z and $T_0(w_t)$ are strongly related through some of the inequality components, $T_0(a)$, $T_0(h_{-1})$ or $T_0(\gamma)$.

Corollary 3. Part A) *Given the equation* $\mu_t = a + bT_0(w_t) + cZ$, *with* $T_0(w_t) \perp Z$, *the impact of overall income inequality,* $T_0(w_t)$, *on* μ_t *depends on which component of inequality (pure effort or opportunity) dominates in* $T_0(w_t)$*:*

i) if the change in $T_0(w_t)$ *comes entirely from* $T_0(\gamma)$, *then* μ_t *is positively affected by* $T_0(w_t)$ *.*

ii) if the change in $T_0(w_t)$ *comes entirely from* $T_0(a)$, *or* $T_0(h_{-1})$, *or both, then* μ_t *is negatively affected by* $T_0(w_t)$.

iii) if the change in $T_0(w_t)$ *comes from simultaneous changes in* $T_0(a)$ *and* $T_0(\gamma)$, *or in* $T_0(h_{-1})$ *and* $T_0(\gamma)$, *or in* $T_0(a)$, $T_0(\gamma)$ *and* $T_0(h_{-1})$, μ_t *is ambiguously affected by* $T_0(w_t)$, *although the effect is positive if* $\psi \to 0$ *and negative if* $\psi \to 1$.

Part B) *Given the equation* $\mu_t = a + bT_0(w_t) + cZ$, *the sign of* b *depends on the relationship between* Z *and the different components of inequality* $T_0(a)$, $T_0(h_{-1})$ *and* $T_0(\gamma)$. *Two extreme, but illustrative cases are:*

iv) if $Z = (T_0(a), T_0(h_{-1}))$, *then* μ_t *is positively affected by* $T_0(w_t)$.

v) if $Z = (T_0(a), T_0(\gamma))$ *or* $Z = (T_0(h_{-1}), T_0(\gamma))$, *then* μ_t *is negatively affected by* $T_0(w_t)$.

Proof. *See Appendix A2* ■

When total inequality $T_0(w_t)$ and the set of controls Z are independent (Part A), the effect of $T_0(w_t)$ on μ_t is determined by the component of inequality, $T_0(a)$, $T_0(h_{-1})$ or $T_0(\gamma)$ that dominates in $T_0(w_t)$. Setting reasonable values for the parameters in the model, we simulate $\mu_t = E[\ln h_t]$ under alternative situations and illustrate the results of this Corollary (Part A) in Figures 1a-1f. First, when the change in $T_0(w_t)$ comes exclusively from changes in $T_0(\gamma)$, the effect of $T_0(w_t)$ on μ_t is positive in the short- and long-run

(Figure 1c), while it is negative when the change comes exclusively from changes in $T_0(a)$ and/or $T_0(h_{-1})$ (Figures 1a and 1b). Nevertheless, note that the relationship between μ_t and $T_0(h_{-1})$ is null in the very long-run because the steady-state is globally stable (the dashed line in Figure 1b). On the contrary, when the change in $T_0(w_t)$ is driven by opportunity and free will at the same time, its final impact on μ_t will strongly depend on the meritocracy parameter ψ (alternatively, on the elasticity of intergenerational mobility ξ). If, for example, we consider an equal variation for the three components of inequality, the effect is null in the short and long-run when $\psi = 0.5$, our benchmark calibration (Figure 1d). This effect turns out to be negative when a low meritocracy economy is assumed ($\psi = 0.75$, Figure 1e), while it becomes positive when the level of meritocracy in the economy is high ($\psi = 0.25$, Figure 1f). Hence, the lower the level of meritocracy in society, the more important are individual circumstances for the process of human capital accumulation.

[FIGURES 1a-1f ABOUT HERE]

The intuition of Part B in the Corollary is the following. If the set of controls Z is more positively related to inequality of opportunity than with inequality of pure effort (for example, race, land property, etc.), $T_0(w_t)$ would then be capturing the inequality-of-pure-effort component more closely and, therefore, the coefficient of $T_0(w_t)$ would be positive. The opposite happens when Z is more positively related to aspects of inequality of pure effort (for example, hours of working, occupational choice, etc.). In this case, b would turn negative because $T_0(w_t)$ would behave as a proxy of inequality of opportunity. It is interesting to note here that some empirical studies have found that the effect of income inequality on growth is sensitive to the inclusion of alternative explanatory variables (Birdsall et al., 1995), while the relationship between initial land inequality, which captures more closely opportunity than income, and growth is negative and robust to the introduction of different explicative variables (Deininger and Squire, 1998).

3 A non-convex model of human capital and inequality

An extensive literature has emphasized the relevance of considering non-convexities and poverty traps in growth models.[21] In keeping with this literature, we follow Acemoglu (2010) and Roemer (1998), and assume the following non-convex process for individual

[21] Among many others, see Azariadis and Drazen (1990), Banerjee and Duflo (2003) and Azariadis and Stachurski (2005).

human capital accumulation:

$$h(i) = \Pi[e(i), \theta(i)] = \begin{cases} \overline{h} & e(i) \leq \widetilde{e}(i) \\ \theta(i)^{\psi} \cdot e(i)^{1-\psi} & e(i) > \widetilde{e}(i) \end{cases}, \; \widetilde{e}(i) = \left(\frac{\overline{h}}{\theta(i)^{\psi}} \right)^{1/(1-\psi)}, \tag{31}$$

where $\overline{h} \geq 0$, common to all dynasties and economies, represents a human capital trap; the threshold $\widetilde{e}(i)$, which depends inversely on the set of circumstances (Roemer, 1998), denotes the minimum level of effort that an individual needs to make to accumulate $h(i)$ above $\overline{h}$.[22] The function $\Pi[\cdot]$ is continuous and increasing in both $e(i)$ and $\theta(i)$.

In this context, the individual must first decide whether or not to exert positive effort. On the one hand, the solution is trivial for the non-effort case (allocations denoted with a 0 superscript): $e^0(i) = 0$, $h^0(i) = \overline{h}$, $w^0(i) = w \cdot \overline{h}$, $c^0(i) = \eta \cdot w \cdot \overline{h}$, $x^0(i) = (1-\eta) \cdot w \cdot \overline{h}$ and $V^0(i) = w \cdot \overline{h}$.[23] On the other hand, the solution is given by (18)-(20) if the individual decides to exert positive effort. Thus, the solution of the dynasty problem is characterized by an incentive-to-effort condition. The following Lemma establishes this condition in terms of the minimum level of circumstances required, $\widetilde{\theta}(i)$, and also in terms of the minimum level of parental human capital, $\widetilde{h}(i)$.

Lemma 1. *For any dynasty $i \in [0,1]$ and generation $t \geq 0$, the incentive-to-effort condition $V^1(i) - V^0(i) \geq 0$ is fulfilled if and only if,*

$$\begin{aligned} \theta(i) &\geq \widetilde{\theta}(i), \\ \widetilde{\theta}(i) &= \left(\frac{\overline{h}\,(1+\beta)}{\beta+\psi} \right)^{\frac{\beta+\psi}{(1+\beta)\cdot\psi}} \cdot \left[\frac{\gamma(i) \cdot (1+\beta)}{(1-\psi) \cdot w} \right]^{\frac{1-\psi}{(1+\beta)\cdot\psi}}, \end{aligned} \tag{32}$$

or alternatively,

$$\begin{aligned} h_{t-1}(i) &\geq \widetilde{h}(i), \\ \widetilde{h}(i) &= \exp^{-\frac{G}{\vartheta}} \left(\frac{\overline{h}\,(1+\beta)}{\beta+\psi} \right)^{\frac{\beta+\psi}{\vartheta}} \left[\frac{\gamma(i)^{1-\psi}}{a(i)^{(1+\beta)\psi-\vartheta}} \right]^{\frac{1}{\vartheta}}. \end{aligned} \tag{33}$$

Moreover, both conditions imply that $h^1(i) \geq \overline{h}$ and $e^1(i) \geq \widetilde{e}(i)$, which makes the incentive-to-effort condition consistent with the non-convex process for individual human capital accumulation in (31).

[22] For the sake of simplicity, we assume that $\overline{h}$ is exogenous and common to all dynasties. A more sophisticated model, beyond the scope of this paper, would consider $\overline{h}$ to be related, for example, to publicly provided funds and be dynasty and/or country specific.

[23] Since the individual has decided not to exert effort above $\widetilde{e}(i)$, it is obvious that the optimal decision is $e(i) = 0$.

Proof. *See Appendix A3* ■

According to (33), parental human capital needs to be high enough to overcome the requirements imposed by $\widetilde{h}(i)$, which depends on $a(i)$ and $\gamma(i)$ as well as on the aggregate economy (w and $\overline{h}$); otherwise, the descendants will be trapped at $\overline{h}$. From (33), (31) can be rewritten in terms of $h_{t-1}(i)$ as follows:

$$h_t(i) = \Omega\left[h_{t-1}(i)\right] = \begin{cases} \overline{h} & h_{t-1}(i) \leq \widetilde{h}(i) \\ \zeta[h_{t-1}(i)] & h_{t-1}(i) > \widetilde{h}(i) \end{cases}, \tag{34}$$

where $\zeta[h_{t-1}(i)]$ is given by (22), with $\zeta' > 0$ and $\zeta'' < 0$; $\Omega\left[0\right] = \overline{h}$ and, since $\psi < 1$, $\Omega\left[\widetilde{h}(i)\right] = \frac{\overline{h}\cdot(1+\beta)}{\beta+\psi} > \overline{h}$, then $\Omega\left[\cdot\right]$ is increasing and concave in $h_{t-1}(i)$. A new implication of non-convexities is the multiplicity of steady-states in (34): one *low* and common to all dynasties, given by $\overline{h}$, and another *high*, dynasty specific, given by the solution of $h_\infty(i) = \zeta[h_\infty(i)]$ for $h_\infty(i) \geq \widetilde{h}(i)$.[24] Given the properties of $\Omega\left[\cdot\right]$, it is easy to show that $\widetilde{h}(i) > \overline{h}$ is a necessary and sufficient condition for the existence of the *low* steady-state, which would be at least locally stable. Likewise, there exists a *high* steady-state if and only if $h_\infty(i) \geq \widetilde{h}(i)$, while its local stability is guaranteed by the strict concavity of $\zeta[\cdot]$. The Lemma 2 below will illustrate these conditions and the steady-state multiplicity.

From the previous section, we know that individual circumstances can affect $h_t(i)$ through a direct and an indirect channel. In a non-convex setting, $\theta(i)$ may also affect $h_t(i)$ through the threshold $\widetilde{e}(i)$ and, consequently, through the probability p_t of being trapped,

$$p_t = \Pr\left[h_{t-1}(i) \leq \widetilde{h}(i)\right]. \tag{35}$$

Since this probability can also be seen as the percentage of individuals earning $w{\cdot}\overline{h}$, if we consider $w{\cdot}\overline{h}$ to be a relative poverty line, p_t would be a measure of relative poverty (Ravallion et al., 1991).[25] Moreover, the average human capital (in logarithms) in this framework is given by:

$$E\left[\ln h_t\right] = p_t \ln \overline{h} + (1-p_t){\cdot}E\left[\ln\zeta(h_{t-1}) \Big/ \ln h_{t-1} > \ln\widetilde{h}\right]. \tag{36}$$

[24]Non-convexities and multiple steady-state equilibria have traditionally been justified in the context of imperfect credit markets (Galor and Zeira, 1993; Banerjee and Newman, 1993), although multiple equilibria are also possible when there are no convexities if credit markets are imperfect and the marginal propensity to save is higher for richer dynasties (Galor and Moav, 2004, Galor and Tsidon, 1997). Instead, we assume the process in (31), which makes explicit the role of circumstances in the accumulation of human capital.

[25]While we are assuming that $\overline{h}$ is common to all individuals and countries, w is common to the individuals within a particular economy. Accordingly, the difference between a rich and a poor country would be the level of w.

The first term is the average human capital accumulated by the individuals in the trap, while the second term is associated with *non-trapped* individuals. Thus, the model presented in Section 2 can be interpreted as a particular case of this more general setup when $p_t = 0$ or $\widetilde{h}$ is small enough. From the definition of $\widetilde{h}$ in (33), it is easy to show that this particular case is more likely to be found in rich economies, characterized by large values of w. Accordingly, the results in Section 2 are more feasible in those economies with a sufficiently high income level. We will illustrate this idea using simulations in Section 4 and real data in Section 5.

In Section 2 we found one single type of dynasty. Now, depending on the ranking of $\widetilde{h}(i)$, $\overline{h}$ and $h_\infty(i)$, we have up to six different types of dynasties, though two of them are unfeasible and another two are redundant. Thus, dynasties can be classified in three different types as characterized in the following Lemma and illustrated in Figures 2a-2c (Figure 2d shows an unfeasible case).

Lemma 2. *A dynasty $i \in [0, 1]$ belongs to any of the following cases:*

Case 1. $\widetilde{h}(i) < \overline{h}$: $h(i)$ always converges to $h_\infty(i)$ regardless the value of $h_{-1}(i)$. In this case, $h_\infty(i)$ is globally stable.

Case 2. $\overline{h} \leq \widetilde{h}(i) \leq h_\infty(i)$ (at least one with inequality): depending on whether $h_{-1}(i)$ is below or above $\widetilde{h}(i)$, the dynasty converges to either $\overline{h}$ or $h_\infty(i)$, respectively. In this case, both equilibria are locally stable.

Case 3. $h_\infty(i) < \widetilde{h}(i)$: the dynasty always converges to $\overline{h}$ regardless of the value of $h_{-1}(i)$. In this case, $\overline{h}$ is globally stable.

Proof. *The trap $\overline{h}$ cannot be higher than $\widetilde{h}(i)$ and $h_\infty(i)$ at the same time because it would imply $\Omega\left[\widetilde{h}(i)\right] < \overline{h}$, which is unfeasible. For this reason, Case 1, $\widetilde{h}(i) < \overline{h} < h_\infty(i)$, and Case 3, $\overline{h} < h_\infty(i) < \widetilde{h}(i)$ or $h_\infty(i) < \overline{h} < \widetilde{h}(i)$, are just fully characterized by $\widetilde{h}(i) < \overline{h}$ (Figure 2a) and $h_\infty(i) < \widetilde{h}(i)$ (Figure 2c), respectively. The condition for Case 2 follows immediately from Figure 2b. Figure 2d represents $h_\infty(i) < \widetilde{h}(i) < \overline{h}$, one of the two unfeasible cases.* ■

Dynasties belonging to Case 1 (Figure 2a) always escape from the trap because they have good exogenous circumstances and/or low $\gamma(i)$ (i.e., high free-will to exert effort) so $\widetilde{h}(i)$ is sufficiently small. Also note that living in a rich economy with high w facilitates belonging to this group. In this case, when an individual is initially accumulating $\overline{h}$ (i.e., $h_{-1}(i) \leq \widetilde{h}(i)$), the next generation is able to move *upward* and to converge to their high steady-state.[26] In contrast, dynasties belonging to Case 3 (Figure 2c) will always end up trapped because they show high $\gamma(i)$ and/or low $a(i)$, so $\widetilde{h}(i)$ is high. Unlike

[26] Note that the concept of intergenerational mobility, ξ, presented in Section 2 referred to the society as a whole.

in the previous case, living in a poor economy with low w facilitates belonging to this group. When a dynasty of this type initially accumulates human capital above $\overline{h}$, their descendants will experience, sooner or later, *downward* mobility. Finally, the steady-state achieved by those dynasties belonging to Case 2 is fully characterized by whether $h_{-1}(i)$ is initially lower or higher than $\widetilde{h}(i)$ so their descendants do not experience any mobility (Figure 2b).[27]

[FIGURES 2a-2d ABOUT HERE]

According to Lemma 2, after $t=0$, upward mobility will occur for those individuals that belong to $B^u = \left\{\ln\widetilde{h} < \ln\overline{h}\right\} \cap \left\{\ln h_{-1} \leq \ln\widetilde{h}\right\}$ while downward mobility will occur for those in $B^d = \left\{\ln\widetilde{h} > \ln h_\infty\right\} \cap \left\{\ln h_{-1} > \ln\widetilde{h}\right\}$. The corresponding probabilities, $p^u = \Pr\left[B^u\right]$ and $p^d = \Pr\left[B^d\right]$, can be written as conditional probabilities:

$$p^u = p_{C1} \cdot \Pr\left[\ln h_{t-1} \leq \ln\widetilde{h} / \ln\widetilde{h} < \ln\overline{h}\right], \quad (37)$$

$$p^d = p_{C3} \cdot \Pr\left[\ln h_{t-1} > \ln\widetilde{h} / \ln\widetilde{h} > \ln h_\infty\right], \quad (38)$$

where $p_{C1} = \Pr\left[\ln\widetilde{h} < \ln\overline{h}\right]$ and $p_{C3} = \Pr\left[\ln\widetilde{h} > \ln h_\infty\right]$ are the probabilities of belonging to Case 1 and Case 3, respectively.

The existence of upward and downward mobility makes the distinction between $t=0$ and $t>0$ necessary for the characterization of $E\left[\ln h_t\right]$ and its relationship with the different sources of inequality. As we show next, obtaining theoretical results for $t=0$ is relatively easy because h_{-1} and $\widetilde{h}$ are independent log-normal variables, p_0 can be characterized and the term $\left[\ln\zeta(h_{-1}) \middle/ \ln h_{-1} > \ln\widetilde{h}\right]$ is also log-normal.[28] However, for

[27] For simplicity, we have assumed that $a(i)$ and $\gamma(i)$ are given by the initial generation. Otherwise, a dynasty might move randomly from one type of dynasty to another at any period t, depending on the realizations of $a(i)$ and $\gamma(i)$.

[28] Note that taking logs in (33), it can be shown that $\ln\widetilde{h} \sim N\left(\widetilde{\mu}, \widetilde{\Delta}^2\right)$, where

$$\widetilde{\mu} = J - \frac{1-\alpha-\varphi}{\alpha+\varphi}\left(\ln\widehat{a} - \frac{\Delta_a^2}{2}\right) + \frac{1-\psi}{\vartheta}\left(\ln\widehat{\gamma} - \frac{\Delta_\gamma^2}{2}\right), \quad (39)$$

$$\widetilde{\Delta}^2 = \left(\frac{1-\alpha-\varphi}{\alpha+\varphi}\right)^2 \Delta_a^2 + \left(\frac{1-\psi}{\vartheta}\right)^2 \Delta_\gamma^2, \quad (40)$$

$$J = \frac{1}{\vartheta}\left[\ln\left(\frac{\overline{h}\,(1+\beta)}{\beta+\psi}\right)^{\beta+\psi} - G\right].$$

$t > 0$ (and $p_t > 0$), the analytical expression for $E[\ln h_t]$ will be obscure since p_t may change over time and h_t will not follow a log-normal distribution. For this reason, the illustration of results for $t > 0$ will require simulations, which will show that, in general, the theoretical results extracted for $t = 0$ apply also for $t > 0$ and steady-state.

3.1 Short-run human capital and inequality

From (36), it is straightforward to see that $E[\ln h_0]$ is:

$$E[\ln h_0] = p_0 \ln \overline{h} + (1-p_0) E\left[\ln \zeta(h_{-1}) \big/ \ln h_{-1} > \ln \widetilde{h}\right], \tag{41}$$

where $p_0 = \Pr\left[\ln h_{-1} \leq \ln \widetilde{h}\right]$ and $\left(\ln h_{-1} - \ln \widetilde{h}\right) \sim N(\mu_x, \sigma_x^2)$, with μ_x and σ_x^2 obtained from (11), (39) and (40). Then, using (28) and (29), p_0, μ_x and σ_x^2 can be expressed in terms of μ_0 and Δ_0^2 as follows:

$$p_0 = 1 - \Phi\left(\frac{\mu_x}{\sigma_x}\right), \tag{42}$$

$$\mu_x = \frac{1}{\xi}\mu_0 - \frac{1}{\xi}\ln\left(\frac{\overline{h}(1+\beta)}{\beta+\psi}\right), \tag{43}$$

$$\sigma_x^2 = \frac{1}{\xi^2}\Delta_0^2, \tag{44}$$

where Φ is the $N(0,1)$ cumulative distribution function. We observe that σ_x^2 is directly related to the initial inequality of wages for the group of dynasties with $h_{-1} > \widetilde{h}$, while, for a given level of inequality, μ_x reflects the initial degree of development of the economy. Consequently, from (42), it is easy to understand why p_0 (the initial relative poverty rate) falls with μ_x and rises with σ_x^2. Note also that $1 - p_0$ and $E\left[\ln \zeta(h_{-1}) \big/ \ln h_{-1} > \ln \widetilde{h}\right]$ affect $E[\ln h_0]$ positively because $E\left[\ln \zeta(h_{-1}) \big/ \ln h_{-1} > \ln \widetilde{h}\right] > \ln \overline{h}$ by the definition of $\Omega[h_{t-1}(i)]$. In fact, $E\left[\ln \zeta(h_{-1}) \big/ \ln h_{-1} > \ln \widetilde{h}\right] \geq E \ln \zeta(h_{-1})$ because $\ln \zeta(h_{-1})$ and $\left(\ln h_{-1} - \ln \widetilde{h}\right)$ are positively correlated so the truncated mean is pushed to the right. The following Lemma gives the detailed expression of $E\left[\ln \zeta(h_{-1}) \big/ \ln h_{-1} > \ln \widetilde{h}\right]$.

Lemma 3. *Assuming that* $\left[\ln \zeta(h_{-1}), \left(\ln h_{-1} - \ln \widetilde{h}\right)\right]$ *follows a bivariate normal distribution, the average human capital (in logs) at t=0 for those dynasties with* $h_{-1}(i) > \widetilde{h}(i)$ *is:*

$$E\left[\ln \zeta(h_{-1}) \big/ \ln h_{-1} > \ln \widetilde{h}\right] = \mu_0 + \xi \frac{\sigma_x}{1-p_0} \phi\left(\frac{-\mu_x}{\sigma_x}\right), \tag{45}$$

where $\phi(\cdot)$ is the standard normal density function, and $\mu_0 = E[\ln \zeta(h_{-1})]$ is the unconditional human capital average in (28).

Proof. *See Appendix A4* ■

Several theoretical results for $E[\ln h_0]$ regarding its relationship with the different sources of inequality can be characterized. Using (41) and (45), $E[\ln h_0]$ can be rewritten as follows:

$$E[\ln h_0] = p_0 \cdot \ln \overline{h} + (1 - p_0) \cdot \mu_0 + M_0(\mu_x, \sigma_x), \quad (46)$$

where μ_0 is given by (28) and $M_0(\mu_x, \sigma_x) = \xi \cdot \sigma_x \cdot \phi\left(-\frac{\mu_x}{\sigma_x}\right)$. This expression shows the three routes through which the alternative sources of inequality can affect $E[\ln h_0]$ at $t = 0$: i) the unconditional mean μ_0 (considering all dynasties); ii) the probability of being trapped, p_0; iii) the term $M_0(\mu_x, \sigma_x)$, which represents the *extra* human capital accumulated by those dynasties *initially* non-trapped, i.e., with $h_{-1}(i) > \widetilde{h}(i)$. We already showed in Section 2 that $\partial\mu_0/\partial\Delta^2_\gamma > 0$, $\partial\mu_0/\partial\Delta^2_a < 0$ and $\partial\mu_0/\partial\Delta^2_{-1} < 0$. Unfortunately, the effects through the other two channels, p_0 and $M_0(\mu_x, \sigma_x)$, are ambiguous because they depend on the relative magnitude of μ_x and σ^2_x (initial degree of development and initial inequality). Therefore, the overall impact of the different sources of inequality on $E[\ln h_0]$ will depend, as shown in the Proposition below, on the initial degrees of development and inequality and, for this reason, may be ambiguous.[29]

Proposition 2. *The effect of the different sources of inequality, Δ^2_j, $j = a,h,\gamma$, on $E[\ln h_0]$ is characterized by the following condition:*

$$\frac{\partial E[\ln h_0]}{\partial \Delta^2_j} > 0 \quad \text{iff} \quad \Pi_j(\mu_x, \sigma_x) < 0, \quad (47)$$

$$\Pi_j(\mu_x, \sigma_x) = \frac{\mu_x}{\sigma^2_x} + \frac{1}{\varepsilon_j} - \xi \cdot \frac{1 - \frac{\sigma_x}{\varepsilon_j}\lambda\left(-\frac{\mu_x}{\sigma_x}\right)}{\ln\left(\frac{1+\beta}{\beta+\psi}\right)}, \quad (48)$$

where $\varepsilon_a = \frac{1-\alpha-\varphi}{\alpha+\varphi}$, $\varepsilon_h = 1$ and $\varepsilon_\gamma = -\frac{1-\psi}{\vartheta}$; $\lambda\left(-\frac{\mu_x}{\sigma_x}\right) = \Phi\left(\frac{\mu_x}{\sigma_x}\right)/\phi\left(-\frac{\mu_x}{\sigma_x}\right)$ is the Mill's ratio at $\left(-\frac{\mu_x}{\sigma_x}\right)$; and $\Pi_j(\mu_x, \sigma_x)$ is an implicit function on (μ_x, σ_x) with the following properties:

i) $\Pi_j(\mu_x, \sigma_x)$ is C^2 on the $(-\infty, +\infty) \times [0, +\infty)$ space for $j = a, h, \gamma$;

ii) $\lim_{\mu_x \to -\infty} \Pi_j(\mu_x, \sigma_x) = -\infty$ for $j = a, h, \gamma$;

iii) $\lim_{\mu_x \to +\infty} \Pi_j(\mu_x; \sigma_x) = +\infty$ for $j = a, h$; $\lim_{\mu_x \to +\infty} \Pi_j(\mu_x; \sigma_x) = -\infty$ for $j = \gamma$;

[29] From (43) and (44), it is clear that the following Proposition and Corollary can be alternatively expressed in terms of p_0 and Δ^2_0.

iv) $\Pi_j(\mu_x;\sigma_x)$ *is strictly monotone (increasing) in* μ_x *for* $j = a, h$, *and no monotone for* $j = \gamma$;

v) $\Pi_j(\mu_x;\sigma_x)$ *is convex in* μ_x *for* $j = a, h$, *and concave for* $j = \gamma$ *with a global maximum at* $\mu_x^{\gamma\max}$.

Proof. *See Appendix A5* ■

In general, condition (47) implies that the impact of any source of inequality on $E[\ln h_0]$ is ambiguous. The following corollary elaborates on the characteristics of these ambiguities, distinguishing between the case of inequality of opportunity (Part A) and inequality of pure effort (Part B).

Corollary 4. Part A. *Inequality of opportunity:* Δ_a^2, Δ_{-1}^2.

For $j = a, h$, *the function* $\Pi_j(\mu_x;\sigma_x)$ *shows, for a given* σ_x, *a unique root,* $\widehat{\mu}_x^j$, *such that* $\partial E[\ln h_0]/\partial\Delta_j^2 < 0$ *iff* $\mu_x > \widehat{\mu}_x^j$ *and* $\partial E[\ln h_0]/\partial\Delta_j^2 \geq 0$ *otherwise. The roots* $\widehat{\mu}_x^a$ *and* $\widehat{\mu}_x^h$ *verify the following conditions:*

i) $\widehat{\mu}_x^h \leq \widehat{\mu}_x^a$ *iff* $\alpha + \varphi \leq \frac{1}{2}$, *so that the range of* μ_x *under which* Δ_a^2 *harmfully affects* $E[\ln h_0]$ *can be larger or smaller than for* Δ_{-1}^2, *depending on the value of* $\alpha + \varphi$.

ii) $\widehat{\mu}_x^j < 0$ *iff* $\sigma_x > \sqrt{2/\pi}\left[\varepsilon_j - \frac{(\beta+\psi)\ln\left(\frac{1+\beta}{\beta+\psi}\right)}{\vartheta}\right]$, *so that* $\widehat{\mu}_x^a$ *and* $\widehat{\mu}_x^h$ *are negative for a sufficient level of initial inequality.*

Part B. *Inequality of pure effort:* Δ_γ^2

For $j = \gamma$, *the effect of* Δ_γ^2 *on* $E[\ln h_0]$ *depends on the sign of* $\Pi_\gamma^{\max} = \Pi_\gamma(\mu_x^{\gamma\max})$:

i) if $\Pi_\gamma^{\max} < 0$, *it is true that* $\partial E[\ln h_0]/\partial\Delta_\gamma^2 > 0$. *A sufficient condition for this is* $\mu_x^{\gamma\max} < 0$, *or equivalently* $\sigma_x^2 > \frac{(1-\psi)}{\xi\cdot\vartheta}\ln\left(\frac{1+\beta}{\beta+\psi}\right)$.

ii) if $\Pi_\gamma^{\max} \geq 0$, *there exist two positive roots,* $\underline{\mu}_x^\gamma > \overline{\mu}_x^\gamma > 0$, *that divide the real line into three zones:* $\mu_x < \underline{\mu}_x^\gamma$; $\mu_x \in \left[\underline{\mu}_x^\gamma, \overline{\mu}_x^\gamma\right]$ *and* $\mu_x > \overline{\mu}_x^\gamma$. *For the 1st and 3rd zones,* $\partial E[\ln h_0]/\partial\Delta_\gamma^2 > 0$, *while* $\partial E[\ln h_0]/\partial\Delta_\gamma^2 < 0$ *for the 2nd zone.*

Proof. *See Appendix A6* ■

The results from Proposition 2 and Corollary 4 are illustrated in Figures 3a and 3b, where the functions $\Pi_j(\cdot)$, $j = a, h, \gamma$, are shown for two alternative situations. On the one hand, Figure 3a illustrates an economy with sufficiently high initial inequality, i.e., with $\sigma_x^2 > \frac{1-\psi}{\xi\cdot\vartheta}\ln\left(\frac{1+\beta}{\beta+\psi}\right)$.[30] Under this situation, $\Pi_\gamma(\cdot)$ is always negative so increases in the inequality of pure effort always benefit $E[\ln h_0]$. On the other hand, Figure 3b

[30] This condition on σ_x^2 is equivalent to $\Delta_0^2 > \frac{1-\psi}{\beta+\psi}\ln\left(\frac{1+\beta}{\beta+\psi}\right)$.

considers that $\sigma_x^2 \leq \frac{1-\psi}{\xi \cdot \vartheta} \ln\left(\frac{1+\beta}{\beta+\psi}\right)$ and illustrates the case of multiple roots in $\Pi_\gamma(\mu_x, \sigma_x)$. In this case, inequality of pure effort is harmful for $E\left[\ln h_0\right]$ when $\mu_x \in \left[\underline{\mu}_x^\gamma, \overline{\mu}_x^\gamma\right]$. Using simulations (see below), we can check that this special situation is highly unrealistic and, in fact, the impact of Δ_γ^2 on $E\left[\ln h_0\right]$ is positive in most of the cases.

With respect to inequality of opportunity (functions $\Pi_a(\cdot)$ and $\Pi_h(\cdot)$), two main possibilities are found. First, inequality of opportunity is always harmful for $E\left[\ln h_0\right]$ when the economy is sufficiently rich (or p_0 is sufficiently small), i.e., when $\mu_x > \max(\widehat{\mu}_x^a, \widehat{\mu}_x^h)$. In this situation, the impact of inequality of opportunity through μ_0 prevails over the other two channels, p_0 and M_0, because most dynasties are not trapped and, therefore, accumulate human capital above $\overline{h}$ (this situation is similar to the result shown in Section 2). Second, only very poor economies, where μ_x is very low (or p_0 is very high), benefit in terms of $E\left[\ln h_0\right]$ from an increase in inequality of opportunity. Because a high percentage of dynasties are initially trapped at $\overline{h}$, raising any source of inequality (Δ_γ^2 but also Δ_a^2 and Δ_{-1}^2) helps those dynasties with better circumstances and free-will parameters to escape from the trap and, therefore, to start accumulating human capital above $\overline{h}$. This situation is more likely when, in addition, the levels of inequality are low.

INSERT FIGURE 3a and 3b ABOUT HERE

Summing up, the results obtained in Section 2 (Corollaries 1-3) apply to this more general setting for rich enough economies (high μ_x or low p_0): inequality of opportunity is harmful for human capital accumulation, while inequality of pure effort is beneficial. However, for poor enough economies, and even more so for low initial inequality, an increase of any source of inequality might have a positive impact on the average level of human capital. Using simulations, these results are illustrated in subsection 3.3 (Figures 4a-4f).

3.2 Steady-state human capital and inequality

After $t = 0$, dynasties can move upward or downward, which means that $p_t \neq p_0$. In fact, at steady-state, p_∞ is the proportion of dynasties initially with $\ln h_{-1} \leq \ln \widetilde{h}$, plus the proportion of dynasties moving downward, minus the proportion of dynasties moving upward,

$$p_\infty = p_0 + p^d - p^u. \tag{49}$$

Hence, p_∞ differs from p_0 by the *net* mobility term, $p^d - p^u$. Likewise, the expected steady-state level of human capital can be expressed as follows:

$$\begin{aligned} E\left[\ln h_\infty\right] &= p_\infty \cdot \ln \overline{h} + (1-p_0) E\left[\ln h_\infty \middle/ \ln h_{-1} > \ln \widetilde{h}\right] \\ &\quad + p^u E\left[\ln h_\infty / B^u\right] - p^d E\left[\ln h_\infty / B^d\right], \end{aligned} \tag{50}$$

and using (49), it can be rewritten as,

$$E\left[\ln h_{\infty}\right] = E\left[\ln h_{\infty}^{NM}\right] + p^{u}\left[E\left(\ln h_{\infty}/B^{u}\right) - \ln \overline{h}\right] \quad (51)$$

$$-p^{d}\left[E\left(\ln h_{\infty}/B^{d}\right) - \ln \overline{h}\right], \quad (52)$$

$$E\left[\ln h_{\infty}^{NM}\right] = p_{0}\ln \overline{h} + (1-p_{0})E\left[\ln h_{\infty} \Big/ \ln h_{-1} > \ln \widetilde{h}\right], \quad (53)$$

where $E\left[\ln h_{\infty}^{NM}\right]$ denotes the expected steady-state level of $\ln h(t)$ if dynasties neither move upward or downward, i.e., $p^{u} = p^{d} = 0$.

Following the proof of Lemma 3, we can obtain the expression of $E\left[\ln h_{\infty} \Big/ \ln h_{-1} > \ln \widetilde{h}\right]$ (see Appendix A7), but solving the terms p^{u} and p^{d} requires the calculation of the cumulative density function of a truncated distribution. Worse still, solving $E\left(\ln h_{\infty}/B^{u}\right)$ and $E\left(\ln h_{\infty}/B^{d}\right)$ requires the calculation of the first-order moments of two trivariate truncated log-normal distributions. The detailed expressions of these moments are shown in Appendices A8 and A9, but they are too complex to obtain any clear-cut conclusions. For this reason, we will use these expressions in subsection 3.3 to simulate the model and reach its main intuitions.

Nevertheless, a detailed examination of (51) and (53) allows us to highlight an important result. Since all terms in brackets in (51) are positive, it is clear that increasing upward mobility and/or reducing downward mobility must be good for the economy. Better still, because h_{∞} and $\widetilde{h}$ are negatively correlated (see Appendix A7), it can be seen that $E\left(\ln h_{\infty}/B^{u}\right) > E\left(\ln h_{\infty}/B^{d}\right)$. Consequently, if, for example, p^{u} and p^{d} decrease by the same amount (i.e., $\nabla p^{u} = \nabla p^{d}$), the change in $E\left[\ln h_{\infty}\right]$ caused by ∇p^{u} will be higher than the variation coming from ∇p^{d}. In this situation, the reduction of p^{u} harms the achievement of those individuals with the potential to accumulate high levels of $h_{\infty}(i)$, while the reduction of p^{d} alleviates the achievement of those individuals that accumulate low levels of $h_{\infty}(i)$. We conclude, therefore, that reducing (increasing) mobility in a symmetric way is harmful (beneficial) for $E\left[\ln h_{\infty}\right]$. Upward mobility would benefit those individuals with the potential to accumulate higher levels of human capital, while downward mobility would push those individuals already accumulating low levels of human capital into the trap.

3.3 Numerical simulations of the non-convex model

To untangle the relationship between the average human capital and the sources of initial inequality in the non-convex model, we simulate conditions (46), (51) and (53). According to the previous discussion, we want to show this relationship in the short- and long-run for a rich (high w, hence high μ_{x} and low p_{0}) and a poor economy (low w, hence low μ_{x}

and high p_0).[31]

First, we need to assume reasonable values for the parameters and initial distributions in the model. Some parameters are taken from the relevant literature, while others are calibrated in order to replicate certain characteristics of an economy in steady-state. The value of parameter λ in the Cobb-Douglas technology is set to 0.36 as commonly assumed in the literature, while η in the utility function is set to replicate a c/y ratio of 55% if all agents were homogenous, i.e., assuming $h(i) = h$ for all i. Combining (4), (5) and (15), it is easy to show that $c/y = \eta\cdot(1-\lambda)$, hence $\eta = 0.86$ in this case. More difficult is to pin down the precise values of α and φ in the composite circumstance index, $\theta(i)$. Nevertheless, empirical studies suggest that they, and their sum, are clearly less than one, thereby we set $\alpha = 0.22$ and $\varphi = 0.26$. In addition, as benchmark values we set the parameter ψ representing the level of meritocracy in the economy at 0.50, and the convexity β of the effort function at 1. With these values, we obtain an elasticity of intergenerational mobility ξ equal to 0.32, which is close to the average of existing estimations (Corak, 2013). Nevertheless, the robustness analysis carried out around these values does not change the main conclusions.

Given these parameters, we initially assume $\Delta^2_{-1} = \Delta^2_a = \Delta^2_\gamma = 1$ in (10)-(12), although later on we modify these variances to reproduce the empirical range for $\Delta^2_0/2$ (i.e, the MLD index of wages for dynasties with $h_{-1} > \widetilde{h}$) between 0.10, approximately the average value for Nordic European countries, and 0.55, approximately the average value for Latin America, the region with largest levels of inequality according to recent World Bank estimates. Lastly, to simplify the numerical exercise, we set $\hat{a} = \widehat{h} = \widehat{\gamma} = 1$, the real interest rate $\overline{r}$ to 0.05, normalize $\overline{h}$ to 1, and adjust A and w, which are one-to-one related from (4), to replicate initial levels of poverty p_0 to 10% (low) and 95% (high).[32]

The strategy of the simulations is the following: first, we evaluate the expressions for the means $\widetilde{\mu}$, μ_x and μ_∞, and their corresponding second moments $\widetilde{\sigma}^2$, σ^2_x and Δ^2_∞; second, the set of probabilities p_0, p^u, p^d, the resultant p_∞ and the expressions for $M_0(\cdot)$ and $M_\infty(\cdot)$ are obtained; finally, we simulate $E\left[\ln h_\infty / B^u\right]$ and $E\left[\ln h_\infty / B^d\right]$ using the expressions developed in the Appendix A9. For all cases, we check whether functions $\Pi_j(\cdot)$ and $\Pi^\infty_j(\cdot)$ show the desired properties. Figures 4a-4f show the relationship between the average human capital and $T_0(w_t)$, when the change in $T_0(w_t)$ comes exclusively from one of the alternative sources of inequality, Δ^2_a, Δ^2_{-1} or Δ^2_γ.[33] As commented above, the selected range for Δ^2_a, Δ^2_{-1} and Δ^2_γ is such that $T_0(w_t)$ ranges from 0.1 to 0.55. The solid-line curves and the dashed-line curves represent the expected level of human capital

[31] The results for intermediate economies, although not shown, are similar to those for rich economies.

[32] To replicate $p_0 = 0.10$, we set $w = 57$ and $A = 8.693$, while for $p_0 = 0.95$, we set $w = 1.5$ and $A = 0.935$.

[33] Recall that in our framework, any change in Δ^2_{-1}, Δ^2_a or Δ^2_γ is equivalent to modifying the associated level of inequality, while keeping the mean of the distribution and the rest of parameters unchanged.

(in logs) for $t = 0$ and steady-state, respectively. Any curve simulated for any period along the transition (not shown for simplicity) falls between these two.

We first give the results for rich economies (Figures 4a, 4c and 4e). For this kind of economy, most dynasties converge to their high steady-state so our results are, at least locally, equivalent to those obtained in Section 2 (Corollary 2 and Figures 1a-1c). For rich (enough) economies, inequality of opportunity generally harms the accumulation of human capital and real income, while inequality of pure effort benefits the economy. Note that, as in Section 2 (Figures 1a-1c), the effect of Δ^2_{-1} on the average human capital at steady-state is null. However, this clear-cut result blurs when a very poor economy is considered (Figures 4b, 4d and 4f). In this situation, most dynasties are, sooner or later, trapped at $\overline{h}$. The key (and simple) mechanism to improving human capital is generating upward mobility, pushing more advantaged dynasties out of the trap. For this result, given that the means remain unchanged by construction, we only need an increase in any kind of inequality. For very poor economies and low levels of inequality, the relationship between the average human capital and inequality is, in general, positive regardless of the type of inequality considered. However, as inequality increases and more dynasties move out of the trap, the results come closer to those obtained for rich economies. As a result, when the increase in total inequality is mainly explained by the inequality-of-opportunity components, the relationship shows -except for Δ^2_{-1} at steady-state- an inverted U shape (Figure 4b and 4d). On the contrary, the relationship is positive and convex when variations in total inequality are caused by changes in the pure-effort-inequality component (Figure 4f).[34]

Summing up, this simulation exercise shows that increasing any type of inequality benefits average human capital accumulation (and thus average wages and income) when the economy has a large initial poverty rate and low initial inequality. The reduction of the percentage of individuals in the trap caused by the upward-mobility effect prevails over the other mechanisms or channels. The important implication for empirical applications is that mixing economies with large differences in levels of development, poverty and inequality, as well as degrees of meritocracy and intergenerational mobility, might generate non-robust and perhaps misleading conclusions regarding the relationship between growth and the different sources of inequality.

[34] For intermediate economies a strong increase in any source of inequality promotes downward mobility, so that the results for inequality of pure effort in Figure 4e and 4f might be reversed, while the negative impact of inequality of opportunity is always present.

4 Empirical evidence

To test empirically the results depicted in the previous sections, we need a sufficiently large sample of country-year information on income inequality (I) and inequality of opportunity (IO). Unfortunately, there are still few micro databases with such information because estimations of IO require not only comparable measures of individual disposable income but also information on individual circumstances. For this reason, our limited sample of countries will prevent us from carrying out a comprehensive quantitative analysis of the causal relationship between IO and human capital, although it will allow us to check whether the empirical correlations are consistent with the main theoretical results in Section 2 and 3.[35]

Although the final sample is not very large, fortunately for our analysis estimates of I and IO exist for countries with different degrees of development (see Table 1). For developed countries with low poverty rates (EU-18) and transitory economies (East European countries), we use the estimates in Marrero and Rodríguez (2012) computed from the EU Survey on Income, Social Inclusion and Living Conditions (EU-SILC) for 2005.[36] For developing countries with intermediate poverty rates, we consider the sample of Latin American countries (Brazil, Colombia, Ecuador, Guatemala, Panama and Peru) in Ferreira and Gignoux (2011), and the IO estimates for Chile in Baéz et al. (2011) and Turkey in Ferreira et al. (2011). Finally, for less developed countries with high poverty rates, we use the IO estimates of African countries (Ghana, Guinea, Ivory Coast, Madagascar and Uganda) in Cogneau and Mesplé-Somps (2009), South Africa in Piraino (2012), Egypt in Belhaj-Hassine (2012) and India in Singh (2011).[37] For these countries, Table 1 shows also the information on per capita real income (in US dollars, 2005), Ypc, obtained from

[35]Because it is implausible to observe all the relevant circumstances (in most cases, only parental education, race and the place of birth are observed), the resultant IO estimates are a lower bound. As a result, since the standard way to calculate the inequality of effort component (IE) is IO from overall inequality, IE estimates could be significantly biased by unobserved circumstances. For this reason, our analysis has not considered IE estimates and, instead, has focused on total inequality and IO. See Ferreira and Gignoux (2011) for a detailed discussion on this issue.

[36]Originally, Marrero and Rodríguez (2012) estimated overall inequality and IO for 23 European countries. We enlarge their original sample by also including Iceland, Cyprus and Luxembourg.

[37]Although the databases from which these estimates are obtained present differences in quality, the applied methodology is not always the same and the number of circumstances differs, we use all of them in order to have as many observations as possible. All the studies applied the parametric technique proposed in Ferreira and Gignoux (2011) except for Baéz et al. (2011) and Cogneau and Mesplé-Somps (2009). With respect to the inequality index, all the studies considered the MLD, except Belhaj-Hassine (2012) and Cogneau and Mesplé-Somps (2009) who used the Theil index and Ferreira et al. (2011) who used the variance. Finally, Ferreira et al. (2011) and Cogneau and Mesplé-Somps (2009) used consumption instead of income.

The indices of inequality and IO used for India and Egypt are those referring to the first three cohorts (21-50 years old for India and 20-49 years old for Egypt) as is usual in the IO literature.

the UN Statistics Division, and the average of human capital, h, measured as the average percentage of individuals among those aged 25-64 years with at least a second level of education in 2010, according to the Barro and Lee database (v. 1.3, 04/13).[38]

Looking at the values in Table 1, we first make some illustrative comparisons. Focusing on the Western EU (developed) countries, it can be seen that the Nordic countries (Denmark, Norway, Sweden, Finland and Iceland), Germany, Austria and the Netherlands are not only the richest countries but also have very low IO; on the contrary, Portugal, Greece, Cyprus and Spain present lower levels of per capita income and higher IO in comparison with the rest of the Western EU countries. This clear-cut division disappears when looking at middle income regions, i.e., Eastern Europe and Latin America. Meanwhile, when looking at the poorest region (African countries and India), it can be observed that South Africa and Egypt present both high values of per capita real income and IO, which contradicts the evidence for developed countries and, therefore, is consistent with the non-convexities found in Section 3.

In Table 2 some simple but appealing regressions are shown. In particular, we estimate by OLS five different models, all of them including regional dummies (see the division in Table 1), and focus on the partial correlation between I, IO and average human capital (in logs). In the first column, it is observed that overall inequality has a positive but insignificant correlation with human capital. Moreover, this correlation remains insignificant when controlling for per capita income. Since we are controlling for per capita GDP, these two results accord with Corollary 3 (Section 2), which said that the impact of overall income inequality on μ_t is undefined a priori and depends on which component of inequality (opportunity or effort) dominates. Following the arguments in that Corollary, we introduce IO into the regression as an additional control (column 3). Now, while the partial correlation between IO and μ_t is negative and significant, the coefficient of I is positive but becomes significant - now, I is capturing more the pure-effort component of total inequality.

In addition, we try to evaluate some theoretical consequences of the non-convex model (columns 4 and 5). In particular, we cross the IO term with the regional dummies and introduce them into the regression. We observe that the coefficient estimated for IO is strongly negative and significant for EU-18, while for regions with intermediate levels of development (i.e., East European countries and Latin America), the estimated coefficients are positive and negative, respectively, but insignificant. However, for the set poorest region, the estimated coefficient turns positive and significant. However, when controlling for the degree of inequality and development, the log of per capita GDP, the correlation of all the IO-crossed variables becomes negative although it is significant only for the richest countries. Again, the estimated coefficient of total inequality is positive and significant

[38] Guinea and Madagascar are not included in Table 1 because the Barro-Lee database does not contain these two countries.

(as in column 3).

In conclusion, applying some caution, we can say that the theoretical results highlighted in Sections 2 and 3 are generally consistent with the existing observations. In particular, we observe that the impact of overall inequality on μ_t is positive, negative, or null depending on the control set considered, that the equalization of opportunity exerts a positive effect on μ_t, and that the higher a country's degree of development, the larger the negative effect of IO on the average level of human capital.

5 Conclusions

The way overall income inequality affects growth is more complex than what the literature has commonly assumed since total inequality is actually a composite measure of different inequalities which might have opposite effects on growth. Taking human capital accumulation as the main engine of growth, we consider two complementary sources of heterogeneity, namely differences in free-will to exert effort and differences in factors (circumstances) beyond the individual's control such as parental background, race, place of birth or health endowments. In this setting, the functioning of credit markets is irrelevant because free-will and individual circumstances are non-purchasable and, therefore, there are no markets for them.

Under a stylized convex process of human capital accumulation, we find that the returns to human capital are decreasing. Hence, compensating for bad circumstances is growth enhancing since marginal returns to human capital are higher for those individuals who have less favorable circumstances. In the same manner, rewarding effort would enhance growth because the marginal returns to human capital are larger for those individuals with a higher aversion to effort. The lack of robustness of the literature on the impact of total inequality on growth is therefore explained, at least partially, on the grounds of which of the different components, opportunity or free-will (pure effort), dominates the change in total inequality. When the change in overall inequality comes from equal changes in inequality of free-will and inequality of opportunity, the impact of total inequality on average human capital becomes ambiguous and depends on certain key characteristics of the economy, such as the degree of meritocracy.

This is not, however, the whole story, since traps in the accumulation of human capital may exist. After considering the existence of non-convexities in the process of human capital accumulation, the impact of inequality of opportunity and inequality of free-will on growth becomes less evident because the multiplicity of steady-states implies a complex non-linear relationship between the sources of total inequality (opportunities and free-will) and human capital. On the one hand, inequality of opportunity may affect growth positively if the economy presents a high poverty rate and a relatively low level of inequality, but negatively otherwise. On the other hand, inequality of pure effort will still benefit human capital accumulation in general and hence ongoing growth.

The following conclusion emerges. In the early stages of development, when the economy is poor and inequality is low, increases in any kind of inequality are good because they allow dynasties to get out of the trap. Later on, as the economy becomes richer and fewer individuals are trapped, an increase in inequality of opportunity will be a deterrent to growth, while an increase in inequality of free-will will enhance growth. As a result, in order to avoid misleading conclusions in future empirical work on the relationship between inequality and the different sources of inequality and growth, researchers should be careful when mixing economies or regions with large differences in poverty rates, total inequality and other crucial characteristics like the degree of meritocracy.

References

[1] Acemoglu, D. (2010), Introduction to Modern Economic Growth. Princeton University Press.

[2] Aghion, P. and Bolton P. (1997), “A trickle-down theory of growth and development with debt overhang”, Review of Economic Studies, 64, 151–162.

[3] Aghion, P. Caroli, E. and García-Peñalosa, C. (1999), “Inequality and economic growth: the perspective of the new growth theories”, Journal of Economic Literature, 37, 1615–1660.

[4] Arrow, K., Bowles, S. and Durlauf, S. (2000), Meritocracy and economic inequality. Princeton University Press.

[5] Azariadis, C. and Drazen, A. (1990), "Threshold Externalities in Economic Development", The Quarterly Journal of Economics, 105, 501–26.

[6] Azariadis, C. and Stachurski, J. (2005), "Poverty traps", in Handbook of Economic Growth, P. Aghion and S. Durlauf (Eds.), vol. 1 (Ch. 5), Elsevier.

[7] Baez-Montenegro, A., Rodríguez-Prieto, J. and Salas R. (2011), “Changes over time in parental education and inequality of opportunities in Chile”, in Inequality of Opportunity: Theory and Measurement (Research on Economic Inequality, Volume 19), Juan Gabriel Rodríguez (ed.), Emerald Group Publishing Limited, 157–177.

[8] Banerjee, A. and Newman, A. (1993), "Occupational choice and the process of development", Journal of Political Economy 101, 274–298.

[9] Banerjee, A. and Duflo, E. (2003), “Inequality and growth: what can the data say?”, Journal of Economic Growth, 8, 267–299.

[10] Baricz, A. (2008), "Mills’ ratio: Monotonicity patterns and functional inequalities", Journal of Mathematical Analysis and Applications, 340, 1362–1370.

[11] Barro, R. J. (2000), "Inequality and Growth in a Panel of Countries", Journal of Economic Growth, 5, 5–32.

[12] Becker, G. and Tomes, N. (1986), "Human Capital and the Rise and Fall of Families", Journal of Labor Economics, 4, 1–39.

[13] Belhaj-Hassine, N. (2012), “Inequality of Opportunity in Egypt”, World Bank Economic Review, 26, 265–295.

[14] Bénabou, R. (1996), "Equity and efficiency in human capital investment: the local connection", Review of Economic Studies, 63, 237–264.

[15] Bénabou, R. (2000), "Unequal Societies: Income Distribution and the Social Contract", American Economic Review, 90, 96–129.

[16] Bertola, G., Foellmi, R. and Zweimüller, J. (2005), Income distribution in macroeconomic models. Princeton: Princeton University Press.

[17] Birdsall, N., Ross, D. and Sabot, R. (1995), "Inequality and growth reconsidered: lessons from East Asia", World Bank Economic Review, 9, 477–508.

[18] Blume, L. E. and Durlauf, S. N. (2001), Social Dynamics. Brookings Institution Press, Washington D.C.

[19] Boldrin, M. and Montes A. (2005), "The Intergenerational State Education and Pensions", Review of Economic Studies, 72, 651–664.

[20] Bourguignon, F., Ferreira, F. and Menéndez, M. (2007a), "Inequality of opportunity in Brazil", Review of Income and Wealth, 53, 585–618.

[21] Bourguignon, F., Ferreira, F. and Walton, M. (2007b), "Equity, efficiency and inequality traps: a research agenda", Journal of Economic Inequality, 5, 235–256.

[22] Bowles, S., Gintis, H. and Osborne-Groves, M. (2005), Unequal chance: family background and economic success. Russell Sage Foundation and Princeton University Press, New York.

[23] Card, D. and Krueger A. (1992), "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States", Journal of Political Economy, 100, 1–40.

[24] Checchi, D. and Peragine, V. (2010), "Inequality of opportunity in Italy", Journal of Economic Inequality, 8, 429–450.

[25] Cogneau D. and S. Mesple-Somps (2009), "Inequality of Opportunity for Income in Five Countries of Africa", in Inequality of Opportunity: Papers from the Second ECINEQ Society Meeting (Research on Economic Inequality, Volume 16), John A. Bishop and Buhong Zheng (ed.), Emerald Group Publishing Limited, 99–128.

[26] Coleman, J., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F. and York, R. (1966), Equality of educational opportunity. Department of health, education, and welfare, Washington D.C.

[27] Corak, M. (2013), "Income Inequality, Equality of Opportunity, and Intergenerational Mobility", Journal of Economic Perspectives, 27, 79–102.

[28] Cowell, F. A. (2009), Measuring Inequality (2nd Edition). Hemel Hempstead: Prentice Hall / Harvester Wheatsheaf.

[29] Deininger, K. and Squire, L. (1998), "New ways of looking at old issues: inequality and growth", Journal of Development Economics, 57, 259–287.

[30] Durlauf, S. N. (2003), "Groups, social influences and inequality: a memberships theory on poverty traps", Santa Fe Institute Working Paper N^o 03-03020.

[31] Ferreira, F. and Gignoux, J. (2011), "The measurement of inequality of opportunity: theory and an application to Latin America", Review of Income and Wealth, 57, 622–657.

[32] Ferreira, F., Gignoux J. and Aran M. (2011), "Measuring Inequality of Opportunity with Imperfect Data: The case of Turkey", Journal of Economic Inequality, 9, 651–680.

[33] Fleurbaey, M. (2008), Fairness, responsibility, and welfare. Oxford: Oxford University Press.

[34] Foster, J. and Shneyerov, A. (2000), "Path independent inequality measures", Journal of Economic Theory, 91, 199–222.

[35] Foster, J. and Ok, E. (1999), "Lorenz dominance and the variance of logarithms", Econometrica, 67, 901–907.

[36] Galor, O. and Moav, O. (2004), "From Physical to Human Capital Accumulation: Inequality and the Process of Development", Review of Economic Studies, 71, 1001–1026.

[37] Galor, O. and Tsiddon, D. (1997), "The Distribution of Human Capital and Economic Growth", Journal of Economic Growth, 2, 93–124.

[38] Galor, O. and Zeira, J. (1993), "Income Distribution and Macroeconomics", Review of Economic Studies, 60, 35–52.

[39] Galton, F. (1869), Hereditary Genius: an Inquiry into its Laws and Consequences. London, UK: Macmillan and Co.

[40] Gibrat, R. (1957), "On economic inequalities", International Economic Papers, 7, 53–70.

[41] Ginther, D., Haveman, R. and Wolfe, B. (2000), "Neighborhood attributes as determinants of children's outcomes: how robust are the relationships, Journal of Human Resources, 35, 603–642.

[42] Goldin, C. and Katz, L.F. (2008), The Race between Education and Technology. Cambridge, MA: Harvard University Press.

[43] Glomm, G. and B. Ravikumar (1992), "Public versus private investment in human capital endogenous growth and income inequality", Journal of Political Economy, 100, 818–34.

[44] Green, W.H. (2008), Econometric Analysis. Pearson, 6th Edition.

[45] Hanushek, E. (1996), "Measuring investment in education", Journal of Economic Perspectives, 10, 9–30.

[46] Hassler, J. and Rodríguez-Mora, J. V. (2000), "Intelligence, Social Mobility, and Growth", American Economic Review, 90, 888–908.

[47] Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N. and Verashchagina, A. (2008), "The inheritance of education inequality: international comparisons and fifty-year trends", The Berkeley Electronic Journal of Economic Analysis and Policy, 7.

[48] Kuznets, S. (1955), "Economic growth and income inequality", American Economic Review, 45, 1–28

[49] Lefranc, A., Pistolesi, N. and Trannoy A. (2009), "Equality of opportunity and luck: Definitions and testable conditions, with an application to income in France", Journal of Public Economics, 93, 1189–1207.

[50] Li Donni, P., Rodríguez, J. G. and Rosas-Dias, P. (2012), "Empirical definition of social types in the analysis of inequality of opportunity: a latent classes approach", mimeo.

[51] Loury, G. C. (1989), "Why should we care about group inequality?", in The Question of Discrimination: Racial Inequality in the U.S. Labor Market, S. Shulman and W. Darity Jr. (eds.), Wesleyan, 1989.

[52] Lucas, J. R. (1995), Responsibility. Oxford: Clarendon Press.

[53] Macho-Stadler, I. and Pérez-Castrillo, J. D. (1996), An introduction to the economics of information. Oxford University Press. Oxford, England.

[54] Mankiw, N.G. (2013), "Defending the one percent", Journal of Economic Perspectives, 27, 21–34.

[55] Manjunath, B. G. and Wilhelm, S. (2009), "Moments Calculation for the Double Truncated Multivariate Normal Density", Social Science Research Network, working papers series (http://ssrn.com/abstract=1472153).

[56] Marrero, A. G. and Rodríguez, J. G. (2011), "Inequality of opportunity in the U.S.: Trends and decomposition", in Juan Gabriel Rodríguez (ed.) Inequality of Opportunity: Theory and Measurement (Research on Economic Inequality, Volume 19), Emerald Group Publishing Limited, 217–246.

[57] Marrero, A. G. and Rodríguez, J. G. (2012): "Inequality of opportunity in Europe", The Review of Income and Wealth, 58, 597–621.

[58] Marrero, A. G. and Rodríguez, J. G. (2013), "Inequality of opportunity and growth", Journal of Development Economics, 104, 107–122.

[59] Mejía, D. and St-Pierre, M. (2008), "Unequal Opportunities and Human Capital Formation", Journal of Development Economics, 86, 395–413.

[60] Piraino P. (2012), "Inequality of opportunity and intergenerational mobility in South Africa", paper presented at the 2nd World Bank Conference on Equity. June 27th, 2012; Washington DC, USA.

[61] Piketty, T., Saez, E. and Stantcheva, S. (2011), "Optimal taxation of top labor incomes: a tale of three elasticities", NBER Working Paper N. 17616.

[62] Ravallion, M., Datt, G. and van de Walle, D. (1991), "Quantifying absolute poverty in the developing world", Review of Income and Wealth, 37, 345–361.

[63] Rodríguez, J. G. (2008), "Partial equality-of-opportunity orderings", Social Choice and Welfare, 31, 435–456.

[64] Roemer, J. E. (1993), "A pragmatic approach to responsibility for the egalitarian planner", Philosophy & Public Affairs, 20, 146–166.

[65] Roemer, J.E. (1998), Equality of opportunity. Harvard University Press, Cambridge, M.A.

[66] Roemer, J.E. (2000), "Equality of opportunity", in Meritocracy and economic inequality, Arrow, K., Bowles, S. and Durlauf, S. (eds.), Princeton University Press.

[67] Ruiz-Castillo, J. (2003), “The measurement of the inequality of opportunities”, in Inequality, Welfare and Poverty: Theory and Measurement (Research on Economic Inequality, Volume 9), Yoram Amiel and John A. Bishop (ed.), Elsevier Science Limited, 1–34.

[68] Tallis, G.M. (1961), "The moment generating function of the truncated multi-normal distribution", Journal of the Royal Statistical Society, Series B (Methodological), 23, 223–229.

[69] Sen, A. (1980), “Equality of what?”, in Tanner Lectures on Human Values, Vol. I, Crambridge University Press.

[70] Singh A. (2011), “Inequality of opportunity in earnings and consumption expenditure: The case of Indian men”, Review of Income and Wealth, 58, 79–106.

[71] Stiglitz, J. E. (2012), The Price of Inequality: How Today’s Divided Society Endangers Our Future. W. W. Norton and Company.

[72] Voitchovsky, S. (2005), “Does the profile of income inequality matter for economic growth?: distinguishing between the effects of inequality in different parts of the income distribution”, Journal of Economic Growth, 10, 273–296.

[73] World Bank (2006), "World development report 2006: equity and development", Washington, DC: The World Bank and Oxford University Press.

6 Appendix

A1. Proof of Proposition 1

To characterize the distribution of h_t, we take logs in (22),

$$\ln h_t(i) = \frac{G}{\beta+\psi} + \frac{(1+\beta)\psi-\vartheta}{\beta+\psi}\ln a(i) - \frac{1-\psi}{\beta+\psi}\ln\gamma(i) + \frac{\vartheta}{\beta+\psi}\ln h_{t-1}(i) \tag{54}$$

where $G = (1-\psi)\ln\left(\frac{(1-\psi)\cdot w}{1+\beta}\right) + (1+\beta)\psi\ln\left[(1-\eta)^{\alpha}w^{\alpha}\right]$. Given the assumptions of the model, it is clear that $\ln h_0$ is normally distributed, $\ln h_0 \sim N(\mu_0, \Delta_0^2)$, with

$$\mu_0 = \frac{G}{\beta+\psi} + \frac{(1+\beta)\psi-\vartheta}{\beta+\psi}E\ln a - \frac{1-\psi}{\beta+\psi}E\ln\gamma + \xi\cdot\mu_{-1}, \tag{55}$$

$$\Delta_0^2 = \left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi}\right)^2\Delta_a^2 + \left(\frac{1-\psi}{\beta+\psi}\right)^2\Delta_\gamma^2 + \xi^2\Delta_{-1}^2. \tag{56}$$

Hence, $\ln h_t$ is also normally distributed for all t, $\ln h_t \sim N(\mu_t, \Delta_t^2)$. To obtain μ_t we need to take expected values in (54), and solve the resulting linear equation in first differences,

$$\mu_t = \frac{G}{\beta+\psi} + \frac{(1+\beta)\psi-\vartheta}{\beta+\psi}E\ln a - \frac{1-\psi}{\beta+\psi}E\ln\gamma + \xi\cdot\mu_{t-1}. \tag{57}$$

Solving this equation for the steady-state, $\mu_t = \mu_{t-1} = \mu_\infty$, we have

$$\mu_\infty = \frac{G}{\beta+\psi-\vartheta} + \frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}E\ln a - \frac{1-\psi}{\beta+\psi-\vartheta}E\ln\gamma, \tag{58}$$

which is globally stable because $\xi = \frac{\vartheta}{\beta+\psi} < 1$. The solution to (57) is

$$\mu_t = \mu_\infty + (\mu_0 - \mu_\infty)\,\xi^t. \tag{59}$$

Using (55) and (58), it is straighforward to obtain

$$\mu_0 - \mu_\infty = \frac{-\xi\cdot\left[G + ((1+\beta)\psi-\vartheta)\,E\,(\ln a(i)) - (1-\psi)\,E\ln\gamma\right]}{\beta+\psi-\vartheta} + \xi\cdot E\ln h_{-1}, \tag{60}$$

and plugging this term into (59) we have,

$$\mu_t = \mu_\infty\left[1-\xi^{t+1}\right] + \mu_{-1}\xi^{t+1}. \tag{61}$$

The process for the variance Δ_t^2 is computed following similar steps. Taking variance in (54),

$$\begin{aligned}\Delta_t^2 &= \left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi}\right)^2\Delta_a^2+\left(\frac{1-\psi}{\beta+\psi}\right)^2\Delta_\gamma^2+\xi^2\Delta_{t-1}^2 \\ &\quad +2\left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi}\right)\xi\cdot Cov\left[\ln a,\ln h_{t-1}\right] \\ &\quad -2\left(\frac{1-\psi}{\beta+\psi}\right)\xi\cdot Cov\left[\ln\gamma,\ln h_{t-1}\right].\end{aligned} \tag{62}$$

Computing the covariances is straighforward, hence we can rewritte (62) as

$$\Delta_t^2=\left[\left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi}\right)^2\Delta_a^2+\left(\frac{1-\psi}{\beta+\psi}\right)^2\Delta_\gamma^2\right]\left[1+2\xi\sum_{j=0}^{t-1}\xi^j\right]+\xi^2\Delta_{t-1}^2. \tag{63}$$

To simplify $\sum_{j=0}^{t-1}\xi^j$, we use the general expression $\sum_{j=0}^{t-1}a^j=\frac{1-a^t}{1-a}$, given that $a=\xi<1$ in our case, it leads to

$$\Delta_t^2=\left[\left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi}\right)^2\Delta_a^2+\left(\frac{1-\psi}{\beta+\psi}\right)^2\Delta_\gamma^2\right]\left[\frac{1+\xi-2\xi^{t+1}}{1-\xi}\right]+\xi^2\Delta_{t-1}^2. \tag{64}$$

Its steady-state equilibrium - i.e., setting $\Delta_t^2=\Delta_{t-1}^2=\Delta_\infty^2$ and taking limits when $t\to\infty$ - is

$$\Delta_\infty^2=\left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}\right)^2\Delta_a^2+\left(\frac{1-\psi}{\beta+\psi-\vartheta}\right)^2\Delta_\gamma^2, \tag{65}$$

which is globally stable because $\xi=\frac{\vartheta}{\beta+\psi}<1$. The solution to (64) is

$$\Delta_t^2=\Delta_\infty^2+\left(\Delta_0^2-\Delta_\infty^2\right)\xi^{2t} \tag{66}$$

and using (56) and (65), after tedious calculations, we obtain

$$\Delta_t^2=\Delta_\infty^2\left[1-\xi^{t+1}\right]^2+\Delta_{-1}^2\xi^{2(t+1)}. \tag{67}$$

A2. Proof of Corollary 3

We are interested in finding out the impact of overall income inequality, $T_0(w_t)$, on μ_t. Assuming that the linear relationship

$$\mu_t=a+bT_0(w_t)+cZ \tag{68}$$

exists, we look for the sign of b. We find two possible cases:

Part A) $T_0(w_t) \perp Z$

From (24) we know that

$$\mu_t = \pi_0 + \left[\frac{1-\psi}{\beta+\psi-\vartheta}T_0(\gamma) - \frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}T_0(a)\right]\left[1-\xi^{t+1}\right] - T_0(h_{-1})\xi^{t+1} \tag{69}$$

where $\pi_0 = \left[\frac{G}{\beta+\psi-\vartheta} + \frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}\ln\hat{a} - \frac{1-\psi}{\beta+\psi-\vartheta}\ln\widehat{\gamma}\right]\left[1-\xi^{t+1}\right] + \ln\widehat{h}\cdot\xi^{t+1}$, $T_0(\gamma) = \frac{\Delta_\gamma^2}{2}$, $T_0(a) = \frac{\Delta_a^2}{2}$ and $T_0(h_{-1}) = \frac{\Delta_{-1}^2}{2}$. Noting that $T_0(w_t) = T_0(h_t)$ and differentiating the expression in (25) we have

$$dT_0(w_t) = \left[\left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}\right)^2 dT_0(a) + \left(\frac{1-\psi}{\beta+\psi-\vartheta}\right)^2 dT_0(\gamma)\right]\left[1-\xi^{t+1}\right]^2 + dT_0(h_{-1})\xi^{2(t+1)}. \tag{70}$$

Furthermore, $dT_0(a)$ and $dT_0(h_{-1})$ can be written as linear functions of $dT_0(\gamma)$ as follows:

$$dT_0(a) = \varsigma dT_0(\gamma) \tag{71}$$

$$dT_0(h_{-1}) = \varkappa dT_0(\gamma) \tag{72}$$

where ς, $\varkappa \in \mathbb{R}$. Inserting the last two expressions in (70) we have, $\frac{dT_0(\gamma)}{dT_0(w_t)} = \frac{1}{\left[\left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}\right)^2\varsigma + \left(\frac{1-\psi}{\beta+\psi-\vartheta}\right)^2\right]\left[1-\xi^{t+1}\right]^2 + \varkappa\xi^{2(t+1)}}$, $\frac{dT_0(a)}{dT_0(w_t)} = \varsigma\frac{dT_0(\gamma)}{dT_0(w_t)}$ and $\frac{dT_0(h_{-1})}{dT_0(w_t)} = \varkappa\frac{dT_0(\gamma)}{dT_0(w_t)}$. Now, consider (68) and (69) and derive μ_t with respect to $T_0(w_t)$ in both expressions. Then, the result is:

$$b = \frac{d\mu_t}{dT_0(w_t)} = \frac{\left[\frac{1-\psi}{\beta+\psi-\vartheta} - \frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}\varsigma\right]\left[1-\xi^{t+1}\right] - \varkappa\xi^{t+1}}{\left[\left(\frac{1-\psi}{\beta+\psi-\vartheta}\right)^2 + \left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}\right)^2\varsigma\right]\left[1-\xi^{t+1}\right]^2 + \varkappa\xi^{2(t+1)}} \tag{73}$$

Although the slope b can be positive, negative or null in general, assuming particular values for ς and $\varkappa$ allows us to predict a particular sign for b. Let's study the most interesting cases:

i) If the change of $T_0(w_t)$ comes entirely from $T_0(\gamma)$, the slope b will be positive. In this case, the parameters ς and $\varkappa$ converge to zero so we find: $\lim_{\varsigma,\varkappa\to 0} b = \frac{1}{\left[1-\xi^{t+1}\right]\left(\frac{1-\psi}{\beta+\psi-\vartheta}\right)} > 0$.

ii) If the change of $T_0(w_t)$ comes entirely from $T_0(a)$, or $T_0(h_{-1})$, or both, the slope b will be negative. In the first case, we have $\lim_{\varsigma\to\infty,\varkappa\to 0} b = -\frac{1}{\left[1-\xi^{t+1}\right]\left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}\right)} < 0$. For the second case we obtain $\lim_{\varsigma\to 0,\varkappa\to\infty} b = -\frac{1}{\xi^{t+1}} < 0$, while for the last case we find $\lim_{\varsigma,\varkappa\to\infty} b = -\frac{\left[1-\xi^{t+1}\right]\frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta} + \xi^{t+1}}{\left[1-\xi^{t+1}\right]^2\left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}\right)^2 + \xi^{2(t+1)}} < 0$.

iii) If the change of $T_0(w_t)$ is due to the same change in $T_0(a)$ and $T_0(\gamma)$, or in $T_0(h_{-1})$ and $T_0(\gamma)$, or in $T_0(a)$, $T_0(\gamma)$ and $T_0(h_{-1})$ the slope b will have an unknown sign. For these three cases we have, respectively: $\lim\limits_{\varsigma\to 1,\varkappa\to 0} b = \frac{\frac{(1-\psi)-(1+\beta)\psi(1-(\alpha+\varphi))}{\beta+\psi-\vartheta}}{\left[\left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}\right)^2+\left(\frac{1-\psi}{\beta+\psi-\vartheta}\right)^2\right]\left[1-\xi^{t+1}\right]}$, $\lim\limits_{\varsigma\to 0,\varkappa\to 1} b = \frac{\frac{1-\psi}{\beta+\psi-\vartheta}\left[1-\xi^{t+1}\right]-\xi^{t+1}}{\left(\frac{1-\psi}{\beta+\psi-\vartheta}\right)^2\left[1-\xi^{t+1}\right]^2+\xi^{2(t+1)}}$ and $\lim\limits_{\varsigma\to 1,\varkappa\to 1} b = \frac{\left[\frac{(1-\psi)-(1+\beta)\psi(1-(\alpha+\varphi))}{\beta+\psi-\vartheta}\right]\left[1-\xi^{t+1}\right]-\xi^{t+1}}{\left[\left(\frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}\right)^2+\left(\frac{1-\psi}{\beta+\psi-\vartheta}\right)^2\right]\left[1-\xi^{t+1}\right]^2+\xi^{2(t+1)}}$. However, these limits have a well defined sign when, in addition, $\psi \to 0$: $\lim\limits_{\varsigma\to 1,\varkappa\to 0,\psi\to 0} b = \lim\limits_{\varsigma\to 0,\varkappa\to 1,\psi\to 0} b = \lim\limits_{\varsigma\to 1,\varkappa\to 1,\psi\to 0} b = \beta > 0$, or $\psi \to 1$: $\lim\limits_{\varsigma\to 1,\varkappa\to 0,\psi\to 1} b = -\frac{1}{1-\left(\frac{\vartheta}{\beta+1}\right)^{t+1}} < 0$. $\lim\limits_{\varsigma\to 0,\varkappa\to 1,\psi\to 1} b = -\frac{1}{\left(\frac{\vartheta}{\beta+1}\right)^{t+1}} < 0$ and $\lim\limits_{\varsigma\to 1,\varkappa\to 1,\psi\to 1} b = -\frac{1}{\left[1-\left(\frac{\vartheta}{\beta+1}\right)^{t+1}\right]^2+\left(\frac{\vartheta}{\beta+1}\right)^{2(t+1)}} < 0$.

Part B) $T_0(w_t)$ and Z are not independent. In this case, we find two interesting results.

iv) $Z = (T_0(a), T_0(h_{-1}))$

We first rewrite (25) in terms of $T_0(w_t)$, $T_0(a)$, $T_0(h_{-1})$ and $T_0(\gamma)$, and solve it for $T_0(\gamma)$,

$$T_0(\gamma) = \left[\frac{\left(\frac{\beta+\psi-\vartheta}{1-\psi}\right)}{1-\xi^{t+1}}\right]^2 T_0(w_t) - \left(\frac{(1+\beta)\psi-\vartheta}{1-\psi}\right)^2 T_0(a) - \left[\frac{\left(\frac{\beta+\psi-\vartheta}{1-\psi}\right)\xi^{t+1}}{1-\xi^{t+1}}\right]^2 T_0(h_{-1}). \quad (74)$$

Then, plugging the last expression into equation (24), we obtain:

$$\mu_t = \pi_0 - \pi_1 \cdot T_0(a) - \pi_2 \cdot T_0(h_{-1}) + \pi_3 \cdot T_0(w_t), \quad (75)$$

where $\pi_0 = \left[\frac{G}{\beta+\psi-\vartheta} + \frac{(1+\beta)\psi-\vartheta}{\beta+\psi-\vartheta}\ln\hat{a} - \frac{1-\psi}{\beta+\psi-\vartheta}\ln\widehat{\gamma}\right]\left[1-\left(\frac{\vartheta}{\beta+\psi}\right)^{t+1}\right] + \ln\widehat{h}\cdot\xi^{t+1}$, $\pi_1 = \frac{((1+\beta)\psi-\vartheta)(1+\beta\psi-\vartheta)\left[1-\xi^{t+1}\right]}{(\beta+\psi-\vartheta)(1-\psi)} > 0$, $\pi_2 = \xi^{t+1}\left[1+\frac{\left(\frac{\beta+\psi-\vartheta}{1-\psi}\right)\xi^{t+1}}{1-\xi^{t+1}}\right] > 0$ and $\pi_3 = \frac{\left(\frac{\beta+\psi-\vartheta}{1-\psi}\right)}{1-\xi^{t+1}} > 0$.
In this case, it is clear that $T_0(w_t)$ has a positive impact on the average human capital, while the impacts of $T_0(h_{-1})$ and $T_0(a)$ are negative.

v) $Z = (T_0(a), T_0(\gamma))$

Let us rewrite again (25) in terms of $T_0(w_t)$, $T_0(a)$, $T_0(h_{-1})$ and $T_0(\gamma)$, but solve it for $T_0(h_{-1})$. Then, substituting the resultant expression into (24), we have:

$$\mu_t = \pi_0 - \delta_1 \cdot T_0(a) + \delta_2 \cdot T_0(\gamma) - \delta_3 \cdot T_0(w_t), \quad (76)$$

where $\delta_1 = \frac{((1+\beta)\psi-\vartheta)\left[1-\xi^{t+1}\right]}{\beta+\psi-\vartheta}\left[1-\frac{((1+\beta)\psi-\vartheta)\left[1-\xi^{t+1}\right]}{(\beta+\psi-\vartheta)\xi^{t+1}}\right]$, $\delta_2 = \frac{(1-\psi)\left[1-\xi^{t+1}\right]}{\beta+\psi-\vartheta}\left[1+\frac{(1-\psi)\left[1-\xi^{t+1}\right]}{(\beta+\psi-\vartheta)\xi^{t+1}}\right]$ > 0 and $\delta_3 = \frac{1}{\xi^{t+1}} > 0$. Now, $T_0(w_t)$ shows a negative impact on μ_t, while the impact of $T_0(\gamma)$ is positive. In this case, the impact of $T_0(a)$ remains uncertain.

A similar reasoning applies for $Z = (T_0(h_{-1}), T_0(\gamma))$.

A3. Proof of Lemma 1

The proof of the first part of the Lemma comes directly from plugging *(18) into* $V^1(i)$ *in (17), using* $V^0(i) = w{\cdot}\overline{h}$ *and solving the inequality* $V^1(i) - V^0(i) \geq 0$ *for* $\theta(i)$. The second part of the Lemma comes from using $V^0(i) = w{\cdot}\overline{h}$ and rewriting (17) in terms of $e^1(i)$. Hence, the incentive-to-effort condition can be rewritten as follows:

$$V^1(i) - V^0(i) \geq 0 \Leftrightarrow w{\cdot} \left[h^1(i) - \overline{h}\right] - \gamma(i){\cdot}e^1(i)^{1+\beta} \geq 0. \tag{77}$$

From this expression, it is clear that the incentive-to-effort condition implies that $h^1(i) \geq \overline{h}$. Plugging (18) and (20) into (77), condition (77) can be rewritten as

$$V^1(i) - V^0(i) \geq 0 \Leftrightarrow \Phi\left[\theta(i), \gamma(i), w\right] \geq \overline{h}, \tag{78}$$

where

$$\Phi\left[\theta(i), \gamma(i), w\right] = \left(\frac{\beta+\psi}{1+\beta}\right)\cdot\left[\frac{(1+\psi)\cdot w}{\gamma(i){\cdot}(1+\beta)}\right]^{\frac{1-\psi}{\beta+\psi}}{\cdot}\theta(i)^{\frac{(1+\beta)\cdot\psi}{\beta+\psi}}. \tag{79}$$

Using (79), it is easy to solve (78) for $\theta(i)$ and obtain the incentive-to-effort condition $\theta(i) \geq \widetilde{\theta}(i)$ given by (32). By plugging (21) into (32) and solving the resultant inequality for $h_{t-1}(i)$, the incentive-to-effort condition can be rewritten in terms of parental human capital

Finally, using the expressions for $\widetilde{e}(i)$ and $e^1(i)$ in (31) and (20), respectively, and the definition of $\Phi\left[\cdot\right]$ in (79), we have

$$\Phi\left[\theta(i), \gamma(i), w\right] \geq \overline{h} \Leftrightarrow e^1(i) \geq \left[\left(\frac{1+\beta}{\psi+\beta}\right)\cdot\left(\frac{(1+\psi)\cdot w{\cdot}\theta(i)^{\psi}}{\gamma(i){\cdot}(1+\beta)}\right)^{\frac{\psi}{\beta+\psi}}\right]\widetilde{e}(i). \tag{80}$$

Because the term in brackets is always positive, (80) proves that the incentive-to-effort condition is consistent with $e^1(i) \geq \widetilde{e}(i)$.

A4. Proof of Lemma 3

Let's define the rnadom variable $X_0 = h_{-1} - \widetilde{h}$. Next, let's assuming that $(\ln\zeta\left(h_{-1}\right), X_0)$ follows a bivariate normal distribution, we can apply the following result for truncated bivariate normal distributions (see Green, 2008, pp. 883):

$$E\left[\ln\zeta\left(h_{-1}\right)/X_0 > 0\right] = E\left[\ln\zeta\left(h_{-1}\right)\right] + \frac{Cov(\ln\zeta\left(h_{-1}\right), X_0)}{\sigma_x}\kappa(\alpha_x), \tag{81}$$

where

$$Cov(\ln\zeta(h_{-1}), X_0) = \frac{\vartheta\cdot\Delta_{-1}^2}{\beta+\psi} + \frac{(1-\alpha-\varphi)^2(1+\beta)\psi\cdot\Delta_a^2}{(\beta+\psi)(\alpha+\varphi)} + \frac{(1-\psi)^2\cdot\Delta_\gamma^2}{\vartheta(\beta+\psi)} \quad (82)$$
$$= \xi\cdot\sigma_x^2,$$
$$\kappa(\alpha_x) = \frac{\phi\left(\frac{-\mu_x}{\sigma_x}\right)}{1-\Phi\left(\frac{-\mu_x}{\sigma_x}\right)} = \frac{\phi\left(\frac{-\mu_x}{\sigma_x}\right)}{1-p_0}. \quad (83)$$

Thus, the result in (45) is straighforward.

A5. Proof of Proposition 2

Deriving the expression in (46), we have $\frac{\partial E[\ln h_0]}{\partial\Delta_j^2} = \frac{\partial p_0}{\partial\Delta_j^2}\ln\overline{h} + \frac{\partial(1-p_0)}{\partial\Delta_j^2}\mu_0 + (1-p_0)\frac{\partial\mu_0}{\partial\Delta_j^2} + \frac{\partial M_0(\mu_x,\sigma_x)}{\partial\Delta_j^2}$ for $j = a, h$ and γ. From the condition in (28) and the definition of G, we have $\mu_0 = \ln\overline{h} + \ln\left(\frac{1+\beta}{\beta+\psi}\right) + \xi\mu_x$. Then, substituting the latter expression into the former and considering that $p_0 = 1-\Phi\left(\frac{\mu_x}{\sigma_x}\right)$, we obtain:

$$\frac{\partial E[\ln h_0]}{\partial\Delta_j^2} = \Phi'\left(\frac{\mu_x}{\sigma_x}\right)\frac{\partial\left(\frac{\mu_x}{\sigma_x}\right)}{\partial\Delta_j^2}\left[\ln\left(\frac{1+\beta}{\beta+\psi}\right) + \xi\cdot\mu_x\right] + \frac{\vartheta\cdot\Phi\left(\frac{\mu_x}{\sigma_x}\right)}{\beta+\psi}\frac{\partial\mu_x}{\partial\Delta_j^2} + \frac{\partial M_0(\mu_x,\sigma_x)}{\partial\Delta_j^2} \quad (84)$$

for $j = a, h$ and γ.

From the definitions of μ_x and σ_x^2 it can be shown that $\frac{\partial\mu_x}{\partial\Delta_j^2} = -\frac{1}{2}\varepsilon_j$, $\frac{\partial\sigma_x}{\partial\Delta_j^2} = \frac{1}{2\sigma_x}\varepsilon_j^2$ and $\frac{\partial\left(\frac{\mu_x}{\sigma_x}\right)}{\partial\Delta_j^2} = -\frac{\varepsilon_j}{2\sigma_x}\left(1+\varepsilon_j\frac{\mu_x}{\sigma_x^2}\right)$ where $\varepsilon_a = \frac{1-\alpha-\varphi}{\alpha+\varphi}$, $\varepsilon_h = 1$ and $\varepsilon_\gamma = -\frac{1-\psi}{\vartheta}$. Moreover, it is true that $\Phi'\left(\frac{\mu_x}{\sigma_x}\right) = \phi\left(\frac{\mu_x}{\sigma_x}\right) = \phi\left(-\frac{\mu_x}{\sigma_x}\right)$, $\phi'\left(-\frac{\mu_x}{\sigma_x}\right) = \phi\left(-\frac{\mu_x}{\sigma_x}\right)\frac{\mu_x}{\sigma_x}$ and $\frac{\partial M_0}{\partial\Delta_j^2} = \xi\frac{\varepsilon_j^2}{2\sigma_x}\left(1+\frac{1}{\varepsilon_j}\mu_x+\frac{\mu_x^2}{\sigma_x^2}\right)\phi\left(-\frac{\mu_x}{\sigma_x}\right)$. As a result, we can divide both sides of (84) by $\phi\left(-\frac{\mu_x}{\sigma_x}\right)$ and obtain $\frac{1}{\phi\left(-\frac{\mu_x}{\sigma_x}\right)}\frac{\partial E[\ln h_0]}{\partial\Delta_j^2} = -\frac{\varepsilon_j^2}{2\sigma_x}\left(\frac{1}{\varepsilon_j}+\frac{\mu_x}{\sigma_x^2}\right)\left[\ln\left(\frac{1+\beta}{\beta+\psi}\right)+\xi\mu_x\right] - \frac{\xi\lambda\left(-\frac{\mu_x}{\sigma_x}\right)\varepsilon_j}{2} + \frac{\xi\varepsilon_j^2}{2\sigma_x}\left(1+\frac{1}{\varepsilon_j}\mu_x+\frac{\mu_x^2}{\sigma_x^2}\right)$ where $\lambda(\cdot)$ is the Mill's ratio (Greene, 2008). Hence, $\frac{\partial E[\ln h_0]}{\partial\Delta_j^2} > 0$ iff $-\left(\frac{1}{\varepsilon_j}+\frac{\mu_x}{\sigma_x^2}\right)\left[\ln\left(\frac{1+\beta}{\beta+\psi}\right)+\xi\mu_x\right] - \frac{\xi\lambda\left(-\frac{\mu_x}{\sigma_x}\right)\sigma_x}{\varepsilon_j} + \xi\left(1+\frac{1}{\varepsilon_j}\mu_x+\frac{\mu_x^2}{\sigma_x^2}\right) > 0$. Rearranging and simplifying terms, this condition can be rewritten as $\frac{1}{\xi}\left(\frac{1}{\varepsilon_j}+\frac{\mu_x}{\sigma_x^2}\right)\ln\left(\frac{1+\beta}{\beta+\psi}\right) + \frac{\sigma_x}{\varepsilon_j}\lambda\left(-\frac{\mu_x}{\sigma_x}\right) - 1 < 0$.

Thus, $\frac{\partial E[\ln h_0]}{\partial\Delta_j^2} > 0$ iff $\Pi_j(\mu_x,\sigma_x) = \frac{\mu_x}{\sigma_x^2} + \frac{1}{\varepsilon_j} - \xi\frac{1-\frac{\sigma_x}{\varepsilon_j}\lambda\left(-\frac{\mu_x}{\sigma_x}\right)}{\ln\left(\frac{1+\beta}{\beta+\psi}\right)} < 0$. The intuition behind this ambiguity is the following. From above, it can be shown that $\partial\mu_0/\partial\Delta_j^2 = -\frac{\xi\varepsilon_j}{2}$, hence the effect of Δ_j^2 on $E\left[\ln\zeta\left(h_{-1}\right)\right]$ is strictly positive for Δ_γ^2 and strictly negative for Δ_h^2 and Δ_a^2. However, we find that $\partial(1-p_0)/\partial\Delta_j^2 > 0$ iff $\varepsilon_j\left(1+\varepsilon_j\frac{\mu_x}{\sigma_x^2}\right) < 0$ and $\partial M_0(\mu_x,\sigma_x)/\partial\Delta_j^2 > 0$ iff $\frac{\mu_x}{\sigma_x^2}+\frac{1}{\varepsilon_j}+\frac{1}{\mu_x} > 0$ for $j = a$, h and γ. Therefore, we see that the last two effects are ambiguous (they depend on the relative magnitude of μ_x and σ_x^2) and, consequently, the overall impact of Δ_j^2 ($j = a$, h and γ) on $E\left[\ln h_0\right]$ is ambiguous in general.

Regarding the properties of $\Pi_j(\mu_x,\sigma_x)$, we use the following characteristics of the Mill's ratio (Baricz, 2008): the Mill's ratio $\lambda(z)$ is C^2 on the $(-\infty,+\infty)$ space; it is strictly monotone (decreasing), $\lambda'(z) < 0$, and convex, $\lambda''(z) \geq 0$; and its limits are $\lim_{z\to-\infty}\lambda(z) = +\infty$ and $\lim_{z\to+\infty}\lambda(z) = 0$. From the first property the proof of *i)* is trivial. Moreover, we have that $\lim_{\mu_x\to-\infty}\Pi_j(\mu_x,\sigma_x) = -\infty$ for all $j = a$, h and γ, because $\lim_{\mu_x\to-\infty}\lambda(-\frac{\mu_x}{\sigma_x}) = 0$. On the other hand, $\lim_{\mu_x\to+\infty}\Pi_j(\mu_x,\sigma_x) = +\infty$ for $j = a$ and h, because in these cases $\varepsilon_j > 0$; however, $\varepsilon_j < 0$ for $j = \gamma$ so $\lim_{\mu_x\to+\infty}\Pi_j(\mu_x,\sigma_x) = \infty - \infty$. After applying the L'Hopital rule, we find that $\lim_{\mu_x\to+\infty}\Pi_j(\mu_x,\sigma_x) = -\infty$ for $j = \gamma$ (note that the Mill's ratio is convex while $\frac{\mu_x}{\sigma_x^2}$ is linear in μ_x, therefore, the former converges faster to $-\infty$ than the latter to $+\infty$). To prove condition *iv)*, we calculate the derivative $\frac{\partial\Pi(\mu_x,\sigma_x)}{\partial\mu_x} = \frac{1}{\sigma_x^2} - \frac{\vartheta}{\beta+\psi}\frac{\lambda'\left(-\frac{\mu_x}{\sigma_x}\right)}{\varepsilon_j\ln\left(\frac{1+\beta}{\beta+\psi}\right)}$, which is strictly positive for $j = a$ and h because their associated ε_a and ε_h are positive and $\lambda'\left(-\frac{\mu_x}{\sigma_x}\right) < 0$. On the contrary, $\frac{\partial\Pi(\mu_x,\sigma_x)}{\partial\mu_x}$ can be positive or negative for $j = \gamma$, because $\varepsilon_\gamma = -\frac{1-\psi}{\vartheta}$ so that the final result depends on the levels of μ_x and σ_x. Finally, for condition *v)*, the second derivative of $\Pi_j(\cdot)$ is equal to $\frac{\vartheta}{\beta+\psi}\frac{1}{\varepsilon_j\sigma_x}\frac{\lambda''\left(-\frac{\mu_x}{\sigma_x}\right)}{\ln\left(\frac{1+\beta}{\beta+\psi}\right)}$ for all j, which is positive for $j = a$ and h, but negative for $j = \gamma$. Hence, for $j = \gamma$, the solution of $\frac{\partial\Pi(\mu_x,\sigma_x)}{\partial\mu_x} = 0$ defines a global maximum.

A6. Proof of Corollary 4

Part A.

From the definition of $\Pi_j\left(\mu_x,\sigma_x\right)$ ($j = a$ and h) in Proposition 3, it is straightforward to show that $\widehat{\mu}_x^h \leq \widehat{\mu}_x^a$ iff $\varepsilon_h \leq \varepsilon_a$. Then, considering the values $\varepsilon_a = \frac{1-\alpha-\varphi}{\alpha+\varphi}$ and $\varepsilon_h = 1$, we obtain condition *i)*. To prove condition *ii)*, notice that $\Pi_j(0,\sigma_x) = \frac{1}{\varepsilon_j}\left[1+\frac{\xi}{\ln\left(\frac{1+\beta}{\beta+\psi}\right)}\left(\sqrt{\frac{\pi}{2}}\sigma_x-\varepsilon_j\right)\right]$ because $\lambda(0) = \sqrt{\pi/2}$ (Baricz, 2008). Moreover, there is a negative root, $\widehat{\mu}_x^j < 0$, if and only if $\Pi_j\left(0,\sigma_x\right) > 0$ because $\Pi_j(\cdot)$ is monotone increasing for $j = a$ and h. Hence, $\widehat{\mu}_x^j < 0$ iff $\sigma_x > \sqrt{2/\pi}\left[\varepsilon_j - \frac{\ln\left(\frac{1+\beta}{\beta+\psi}\right)}{\xi}\right]$.

Part B.

The existance of cases *i)* (no ambiguity) and *ii)* (ambiguity) is trivial from the properties of $\Pi_\gamma(\mu_x, \sigma_x)$ (see Proposition 3). With respect to the first case, the sufficient condition is obtained as follows. First, we set $\varepsilon_\gamma = -\frac{1-\psi}{\vartheta}$ and compute $\Pi_\gamma(\mu_x^{\gamma\max}, \sigma_x)$ $= \frac{\mu_x^{\gamma\max}}{\sigma_x^2} - \frac{\vartheta}{1-\psi} - \xi\frac{1+\frac{\vartheta\cdot\sigma_x}{1-\psi}\lambda\left(-\frac{\mu_x^{\gamma\max}}{\sigma_x}\right)}{\ln\left(\frac{1+\beta}{\beta+\psi}\right)}$. From this result it is clear that $\mu_x^{\gamma\max} < 0$ guarantees that $\Pi_\gamma(\mu_x^{\gamma\max}, \sigma_x) < 0$. Next, we ellaborate on $\frac{\partial\Pi_\gamma(\mu_x,\sigma_x)}{\partial\mu_x} = \frac{1}{\sigma_x^2} + \frac{\xi\vartheta}{(1-\psi)}\frac{\lambda'\left(-\frac{\mu_x}{\sigma_x}\right)}{\ln\left(\frac{1+\beta}{\beta+\psi}\right)} = 0$ to obtain the corresponding condition for $\mu_x^{\gamma\max}$. By using $\lambda'\left(-\frac{\mu_x}{\sigma_x}\right) = -\frac{\mu_x}{\sigma_x}\lambda\left(-\frac{\mu_x}{\sigma_x}\right) - 1$, we have $\mu_x^{\gamma\max} = \frac{\sigma_x}{\lambda\left(-\frac{\mu_x^{\gamma\max}}{\sigma_x}\right)}\left[\frac{1}{\sigma_x^2}\frac{(1-\psi)}{\xi\vartheta}\ln\left(\frac{1+\beta}{\beta+\psi}\right) - 1\right]$. Hence, $\mu_x^{\gamma\max} < 0$ is equivalent to $\sigma_x^2 > \frac{1-\psi}{\xi\vartheta}\ln\left(\frac{1+\beta}{\beta+\psi}\right)$. With respect to the second case, given the previous result, it is easy to see that $\mu_x^{\gamma\max} \geq 0$ is a necessary (but not sufficient) condition for this case. In addition, it is true that $\Pi_\gamma(0, \sigma_x) < 0$ for $\varepsilon_\gamma = -\frac{1-\psi}{\vartheta}$ so that whenever the roots of $\Pi_\gamma(\mu_x, \sigma_x)$ exist, they are always positive.

A7. Detailed expression of $E\left[\ln\zeta(h_\infty)/X_0 > 0\right]$

Following the same strategy as in Lemma 3 for $t = 0$, we obtain

$$E\left[\ln\zeta(h_\infty)/X_0 > 0\right] = \mu_\infty + \frac{\vartheta}{\beta+\psi-\vartheta}\frac{\sigma_x^2 - \Delta_{-1}^2}{(1-p_0)\sigma_x}\phi\left(-\frac{\mu_x}{\sigma_x}\right). \tag{85}$$

We know that $Cov(\ln h_\infty, X_0) = -Cov(\ln h_\infty, \ln\widetilde{h})$ because h_∞ does not depend on h_{-1}. Taking logs in the definitions of h_∞ and $\widetilde{h}$ and considering the fact that $cov(a,\gamma) = 0$, we obtain $Cov(\ln h_\infty, X_0) = \frac{(1+\beta)\cdot\psi\cdot(1-\alpha-\varphi)^2}{(\beta+\psi-\vartheta)(\alpha+\varphi)}\Delta_a^2 + \frac{(1-\psi)^2}{(\beta+\psi-\vartheta)\cdot\vartheta}\Delta_\gamma^2$. Then, recalling from (44) and (29) that $\sigma_x^2 = \Delta_{-1}^2 + \left(\frac{1-\alpha-\varphi}{\alpha+\varphi}\right)^2\Delta_a^2 + \left(\frac{1-\psi}{\vartheta}\right)^2\Delta_\gamma^2$ the result in (85) is straightforward.

A8. Detailed expressions of p^u and p^d

Recalling from the main text that $X_0 = \left(\ln h_{-1} - \ln\widetilde{h}\right) \sim N\left(\mu_x, \sigma_x^2\right)$ and $\ln\widetilde{h} \sim N\left(\widetilde{\mu}, \widetilde{\Delta}^2\right)$, we have that $p^u = p_{C1}\cdot\Pr\left[X_0 \leq 0/Y > 0\right]$ where $Y = \ln\overline{h} - \ln\widetilde{h}$. That is, $p^u = p_{C1}\cdot\Phi\left(\frac{-\mu_{X_0/Y>0}}{\sigma_{X_0/Y>0}}\right)$, where

$$\mu_{X_0/Y>0} = E\left[X_0/Y>0\right] = \mu_x - \frac{Cov(X_0, \ln \widetilde{h})}{\widetilde{\Delta}}\kappa(\alpha_Y), \tag{86}$$

$$Cov(X_0, \ln \widetilde{h}) = -\widetilde{\Delta}^2 \tag{87}$$

$$\kappa(\alpha_Y) = \frac{\phi\left(\alpha_Y\right)}{1-\Phi\left(\alpha_Y\right)}, \alpha_Y = \frac{\widetilde{\mu} - \ln \overline{h}}{\widetilde{\Delta}} \tag{88}$$

$$\sigma^2_{X_0/Y>0} = Var\left[X_0/Y>0\right] = \sigma^2_x\left[1 - \frac{Cov(X_0, \ln \widetilde{h})^2 \cdot \delta\left(\alpha_Y\right)}{\widetilde{\Delta}^2 \sigma^2_x}\right], \tag{89}$$

$$\delta\left(\alpha_Y\right) = \kappa(\alpha_Y)\left[\kappa(\alpha_Y) - \alpha_Y\right]. \tag{90}$$

Rearranging terms, we have:

$$\mu_{X_0/Y>0} = \mu_x + \widetilde{\Delta} \cdot \kappa(\alpha_Y), \tag{91}$$

$$\kappa(\alpha_Y) = \frac{\phi\left(\alpha_Y\right)}{1-\Phi\left(\alpha_Y\right)}, \alpha_Y = \frac{\widetilde{\mu} - \ln \overline{h}}{\widetilde{\Delta}} \tag{92}$$

$$\sigma^2_{X_0/Y>0} = \sigma^2_x\left[1 - \frac{\widetilde{\Delta}^2}{\sigma^2_x} \cdot \delta\left(\alpha_Y\right)\right], \tag{93}$$

$$\delta\left(\alpha_Y\right) = \kappa(\alpha_Y)\left[\kappa(\alpha_Y) - \alpha_Y\right]. \tag{94}$$

Likewise, we have that $p^d = p_{C3} \cdot \Pr\left[X_0 > 0/R > 0\right]$ where $R = \left(\ln \widetilde{h} - \ln h_\infty\right) \sim N(\mu_R, \Delta^2_R)$ with $\mu_R = \widetilde{\mu} - \mu_\infty$ and $\Delta^2_R = \widetilde{\Delta}^2 + \Delta^2_\infty - 2 \cdot cov(\ln h_\infty, \ln \widetilde{h})$. We know from Appendix A7 that $Cov(\ln h_\infty, \ln \widetilde{h}) = -\frac{(1+\beta)\cdot\psi\cdot(1-\alpha-\varphi)^2}{(\beta+\psi-\vartheta)(\alpha+\varphi)}\Delta^2_a - \frac{(1-\psi)^2}{(\beta+\psi-\vartheta)\cdot\vartheta}\Delta^2_\gamma$, and using the corresponding expressions of $\widetilde{\Delta}^2$ and Δ^2_∞, we find that $\Delta^2_R = \frac{(1-\alpha-\varphi)^2(\beta+\psi)^2}{(\beta+\psi-\vartheta)^2(\alpha+\varphi)^2}\Delta^2_a + \frac{(1-\psi)^2(\beta+\psi)^2}{(\beta+\psi-\vartheta)^2\vartheta^2}\Delta^2_\gamma$, which is indeed equal to $\frac{(\beta+\psi)^2}{(\beta+\psi-\vartheta)^2}\widetilde{\Delta}^2$. As a result, $p^d = p_{C3} \cdot \left(1 - \Phi\left(\frac{-\mu_{X_0/R>0}}{\sigma_{X_0/R>0}}\right)\right)$, where

$$\mu_{X_0/R>0} = E\left[X_0/R>0\right] = \mu_x + \frac{\beta+\psi-\vartheta}{\beta+\psi}\frac{Cov(X_0,R)}{\widetilde{\Delta}}\kappa(\alpha_R), \tag{95}$$

$$Cov(X_0,R) = -\frac{(\beta+\psi)}{(\beta+\psi-\vartheta)}\widetilde{\Delta}^2 \tag{96}$$

$$\kappa(\alpha_R) = \frac{\phi\left(-\alpha_R\right)}{1-\Phi\left(-\alpha_R\right)}, \alpha_R = \frac{\beta+\psi-\vartheta}{\beta+\psi}\frac{\mu_\infty-\widetilde{\mu}}{\widetilde{\Delta}} \tag{97}$$

$$\sigma^2_{X_0/R>0} = Var\left[X_0/R>0\right] = \tag{98}$$

$$= \sigma_x^2\left[1-\frac{(\beta+\psi-\vartheta)^2}{(\beta+\psi)^2}\frac{Cov(X_0,R)^2}{\widetilde{\Delta}^2\sigma_x^2}\cdot\delta\left(\alpha_R\right)\right], \tag{99}$$

$$\delta\left(\alpha_R\right) = \kappa(\alpha_R)\left[\kappa(\alpha_R)-\alpha_R\right] \tag{100}$$

Rearranging terms, we obtain the following:

$$\mu_{X_0/R>0} = E\left[X_0/R>0\right] = \mu_x - \widetilde{\Delta}\kappa(\alpha_R), \tag{101}$$

$$\kappa(\alpha_R) = \frac{\phi\left(-\alpha_R\right)}{1-\Phi\left(-\alpha_R\right)}, \alpha_R = \frac{\beta+\psi-\vartheta}{\beta+\psi}\frac{\mu_\infty-\widetilde{\mu}}{\widetilde{\Delta}} \tag{102}$$

$$\sigma^2_{X_0/R>0} = Var\left[X_0/R>0\right] = \sigma_x^2\left[1-\frac{\widetilde{\Delta}^2}{\sigma_x^2}\cdot\delta\left(\alpha_R\right)\right], \tag{103}$$

$$\delta\left(\alpha_R\right) = \kappa(\alpha_R)\left[\kappa(\alpha_R)-\alpha_R\right]. \tag{104}$$

A9. Detailed expressions of $E\left(\ln h_\infty/B^u\right)$ and $E\left(\ln h_\infty/B^d\right)$

The terms $E\left(\ln h_\infty/B^u\right)$ and $E\left(\ln h_\infty/B^d\right)$ correspond to the first moment of a truncated trivariate normal distribution. From the literature we know that the trivariate normal density with average vector $\boldsymbol{\mu}\in\mathbb{R}^3$ and covariance matrix $\boldsymbol{\Sigma}$ is:

$$\phi_3^{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\mathbf{x}) = (2\pi)^{-3/2}\left|\boldsymbol{\Sigma}\right|^{-1/2}\exp\left\{-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}\right)'\boldsymbol{\Sigma}^{-1}\left(\mathbf{x}-\boldsymbol{\mu}\right)\right\},\ \mathbf{x}\in\mathbb{R}^3. \tag{105}$$

Accordingly, the trivariate truncated normal density, truncated at $\mathbf{a}\in\mathbb{R}^3$ is defined as

$$\phi_3^{\boldsymbol{\mu},\boldsymbol{\Sigma},\alpha}(\mathbf{x}) = \begin{Bmatrix} \frac{1}{\alpha}\phi_3^{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\mathbf{x}), & \text{for } \mathbf{x}\geq\mathbf{a} \\ 0, & \text{otherwise} \end{Bmatrix} \tag{106}$$

where $\alpha = \Pr\left[X_1>a_1, X_2>a_2, X_3>a_3\right]$. By applying the moment generating function to (106), Manjunath and Wilhelm (2009) obtained the first moment of the trivariate truncated normal distribution:[39]

$$
\begin{aligned}
E\left(X_i|X_1 > a_1 \cap X_2 > a_2 \cap X_3 > a_3\right) &= \mu_i \\
+\frac{1}{\alpha}\sum_{k=1}^{3}\sigma_{ik}\int_{\mathbf{a}_{(\neq k)}}^{\infty}\phi_3^{\mathbf{\Sigma}}(\mathbf{x}_{(\neq k)};x_k &= a_k)d\mathbf{x}_{(\neq k)}
\end{aligned} \tag{107}
$$

where $(\neq k)$ denotes that the k-th array component has been dropped from the vector and

$$
\phi_3^{\mathbf{\Sigma}}(\mathbf{x}) = (2\pi)^{-3/2}\left|\mathbf{\Sigma}\right|^{-1/2}\exp\left\{-\frac{1}{2}\mathbf{x}'\mathbf{\Sigma}^{-1}\mathbf{x}\right\}. \tag{108}
$$

In order to compute the integral of the last term in (107) and following Tallis (1961), we use without loss of generality a z-transformation for all variates $\mathbf{x}' = (x_1,x_2,x_3)$ as well as for all lower truncation points $\mathbf{a}' = (a_1, a_2, a_3)$, resulting in a $N(0, \mathbf{R})$ distribution with correlation matrix $\mathbf{R}$ for the standarized untruncated variates. Then, we have

$$
\int_{\frac{\mathbf{a}-\boldsymbol{\mu}}{\boldsymbol{\sigma}}_{(\neq k)}}^{\infty}\phi_3^{\mathbf{R}}(\mathbf{z}_{(\neq k)};z_k = \frac{a_k-\mu_k}{\sigma_k})d\mathbf{z}_{(\neq k)} = \frac{1}{\sigma_k}\phi(\frac{a_k-\mu_k}{\sigma_k})\Theta_2(\mathbf{A}_{kq};\mathbf{R}_k) \tag{109}
$$

where $q = 1, 2, 3$ with $q \neq k$, $\mathbf{R}_k$ is the matrix of first-order partial correlation coefficients and

$$
\Theta_2(\mathbf{A}_{kq};\mathbf{R}_k) = \int_{\mathbf{A}_{kq}}^{\infty}\phi_2^{\mathbf{R}_k}(\mathbf{z}_{(\neq k)})d\mathbf{z}_{(\neq k)}, \tag{110}
$$

with

$$
A_{kq} = \frac{(\frac{a_q-\mu_q}{\sigma_q}) - \rho_{qk}\cdot(\frac{a_k-\mu_k}{\sigma_k})}{\sqrt{1-\rho_{qk}^2}}. \tag{111}
$$

Then, noting that $\rho_{ik} = \frac{\sigma_{ik}}{\sigma_i\sigma_k}$, the first moment of the trivariate truncated normal distribution can be rewritten as follows:

$$
E\left(X_i|X_1 > a_1 \cap X_2 > a_2 \cap X_3 > a_3\right) = \mu_i + \frac{\sigma_i}{\alpha}\sum_{k=1}^{3}\rho_{ik}\phi(\frac{a_k-\mu_k}{\sigma_k})\Theta_2(\mathbf{A}_{kq};\mathbf{R}_k). \tag{112}
$$

Notice that all terms involving ϕ, where ϕ is a function of any $a_k = -\infty$, are zero.

Focusing now on the average human capital for those dynasties who experience upward mobility, $E\left(\ln h_\infty / B^u\right)$, assume that $(\ln h_\infty, \ln\overline{h} - \ln\widetilde{h}, \ln\widetilde{h} - \ln h_{-1})$ follows a trivariate normal distribution. Then, noting that α is actually p^u and the lower truncation array $\mathbf{a}$ is $(-\infty, 0, 0)$, we have

[39] Actually, these authors found the first and second moments of the rectangularly doubly truncated multivariate normal density.

$$E\left(\ln h_\infty /B^u\right) = \mu_\infty + \frac{\sigma_{12}}{p^u\sigma_2}\phi(-\frac{\mu_2}{\sigma_2})\int_{-\infty}^{+\infty}\int_{A_{23}}^{+\infty}\phi_2\left(z_1,z_3;\rho_{13.2}\right)dz_1dz_3 \quad (113)$$
$$+\frac{\sigma_{13}}{p^u\sigma_3}\phi(-\frac{\mu_3}{\sigma_3})\int_{-\infty}^{+\infty}\int_{A_{32}}^{+\infty}\phi_2\left(z_1,z_2;\rho_{12.3}\right)dz_1dz_2.$$

where $\rho_{rq.k}$ is the partial correlation coefficient between X_r and X_q for fixed X_k. Moreover, we find that $\sigma_{12} = \frac{\vartheta}{(\beta+\psi-\vartheta)}\widetilde{\Delta}^2$; $\sigma_2 = \widetilde{\Delta}$; $\mu_2 = \ln\overline{h} - \widetilde{\mu}$; $\sigma_{13} = -\sigma_{12}$; $\sigma_3 = \left(\widetilde{\Delta}^2 + \Delta_{-1}^2\right)^{\frac{1}{2}}$; $\mu_3 = \widetilde{\mu} - \left(\ln\widehat{h} - \frac{\Delta_{-1}^2}{2}\right)$; $A_{23} = \frac{\left(-\frac{\mu_3}{\sigma_3}+\frac{\mu_2}{\sigma_2}\cdot\rho_{32}\right)}{\sqrt{1-\rho_{32}^2}}$; $A_{32} = \frac{\left(-\frac{\mu_2}{\sigma_2}+\frac{\mu_3}{\sigma_3}\cdot\rho_{23}\right)}{\sqrt{1-\rho_{23}^2}}$; $\rho_{32} = \rho_{23} = \frac{\sigma_{23}}{\sigma_2\sigma_3}$; and $\sigma_{23} = -\widetilde{\Delta}^2$. It can be seen that the result in (81) is obtained as a particular case of the last expression when $a_2 = -\infty$.

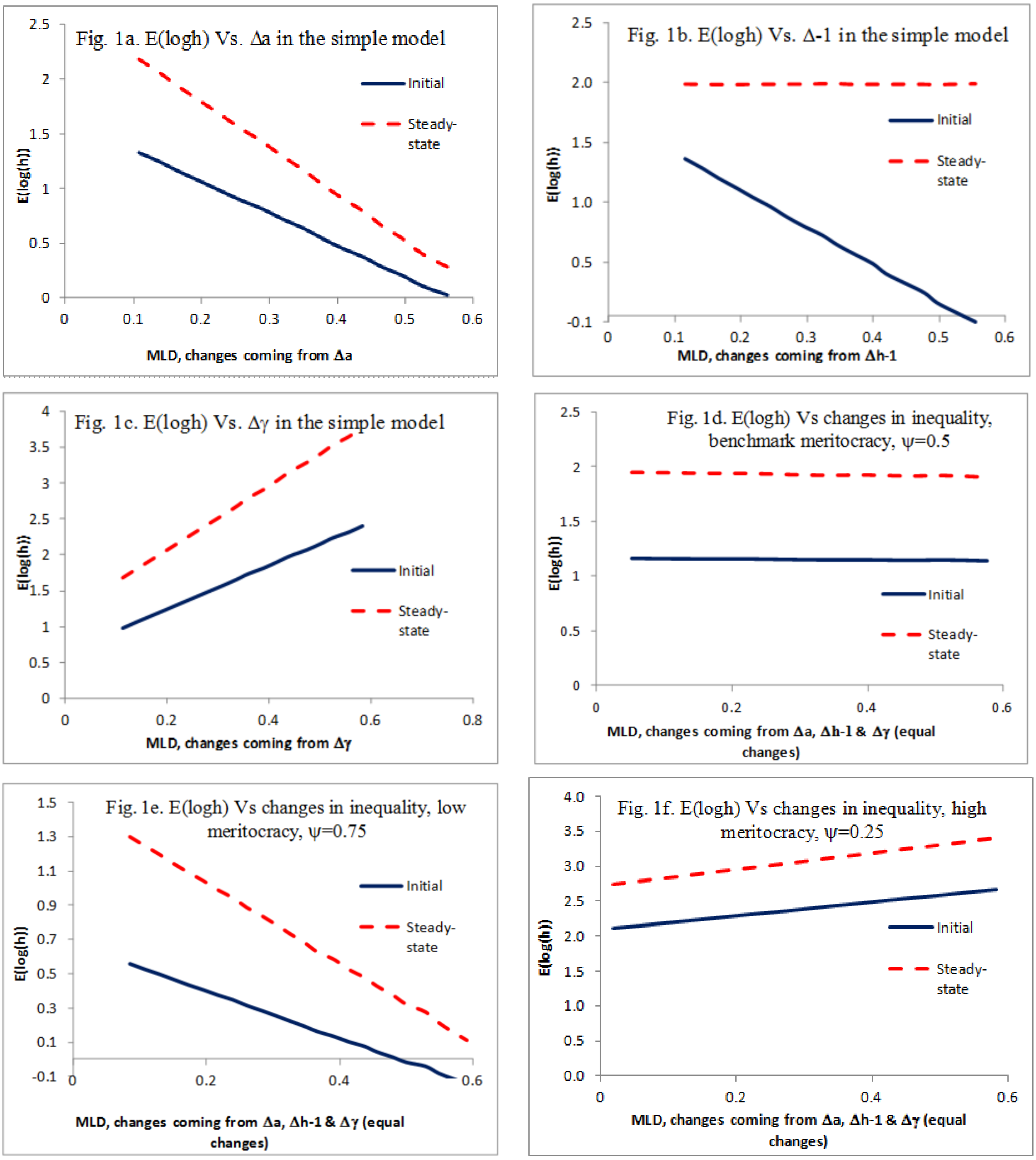
In a similar manner, we can obtain the average human capital for those dynasties who experience downward mobility, $E\left(\ln h_\infty /B^d\right)$. In particular, assuming that $(\ln h_\infty, \ln\widetilde{h} - \ln h_{-1}, \ln\widetilde{h} - \ln h_\infty)$ follows a trivariate normal distribution, we have

$$E\left(\ln h_\infty /B^d\right) = \mu_\infty - \frac{\sigma_{13}}{p^d\sigma_3}\phi(-\frac{\mu_3}{\sigma_3})\int_{-\infty}^{+\infty}\int_{-\infty}^{A_{34}}\phi_2\left(z_1,z_4;\rho_{14.3}\right)dz_1dz_4$$
$$+\frac{\sigma_{14}}{p^d\sigma_4}\phi(-\frac{\mu_4}{\sigma_4})\int_{-\infty}^{+\infty}\int_{A_{43}}^{+\infty}\phi_2\left(z_1,z_3;\rho_{13.4}\right)dz_1dz_3, \quad (114)$$

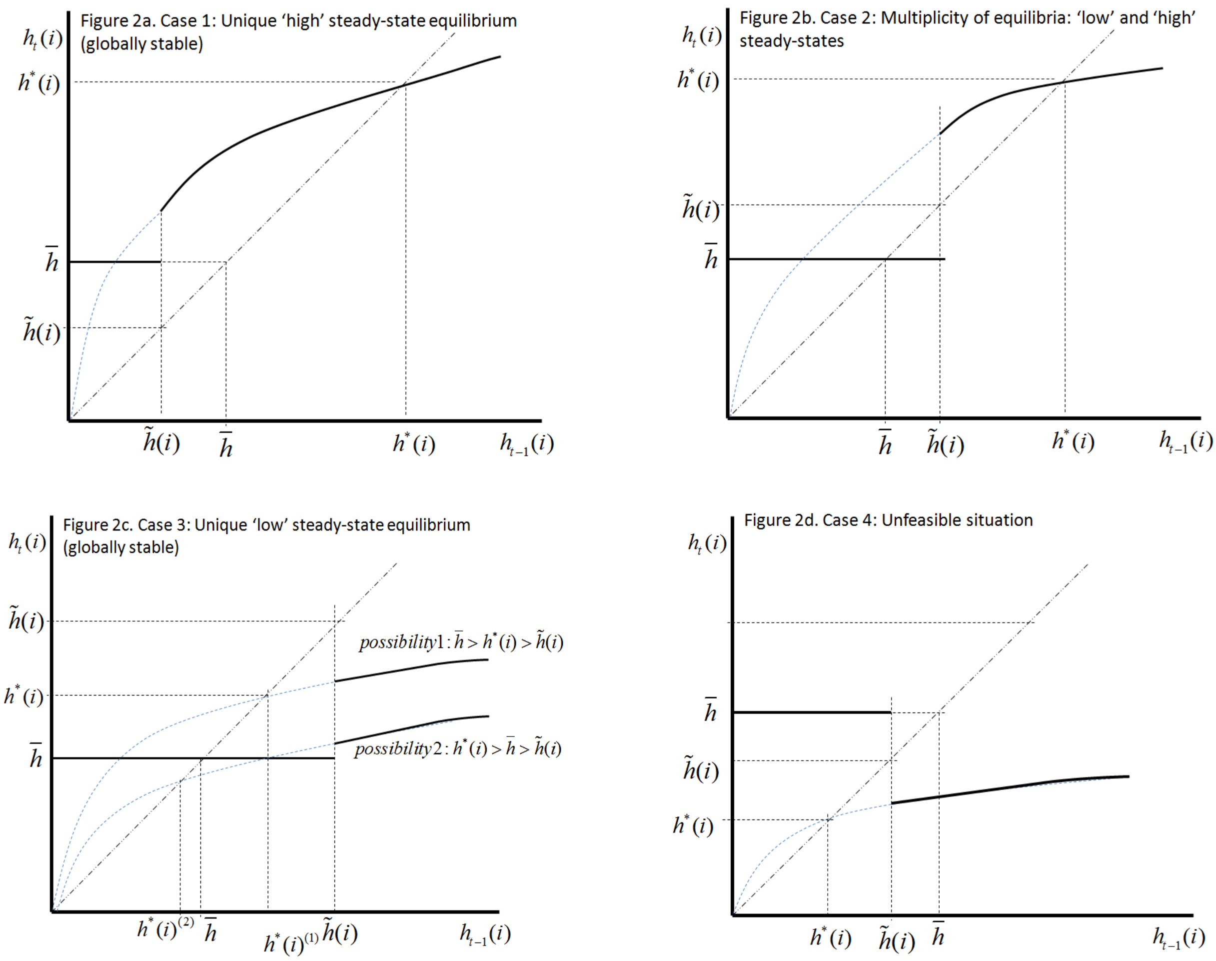
where now $\sigma_{14} = -\frac{(\beta+\psi)\cdot\vartheta}{(\beta+\psi-\vartheta)^2}\widetilde{\Delta}^2$; $\sigma_4 = \frac{\beta+\psi}{\beta+\psi-\vartheta}\widetilde{\Delta}$; $\mu_4 = \widetilde{\mu} - \mu_\infty$; $A_{34} = \frac{\left(-\frac{\mu_4}{\sigma_4}+\frac{\mu_3}{\sigma_3}\cdot\rho_{43}\right)}{\sqrt{1-\rho_{43}^2}}$; $A_{43} = \frac{\left(-\frac{\mu_3}{\sigma_3}+\frac{\mu_4}{\sigma_4}\cdot\rho_{34}\right)}{\sqrt{1-\rho_{34}^2}}$; $\rho_{43} = \rho_{34} = \frac{\sigma_{34}}{\sigma_3\sigma_4}$; and $\sigma_{34} =\beta + \psi_{\beta+\psi-\vartheta\widetilde{\Delta}^2.}$

FIGURES

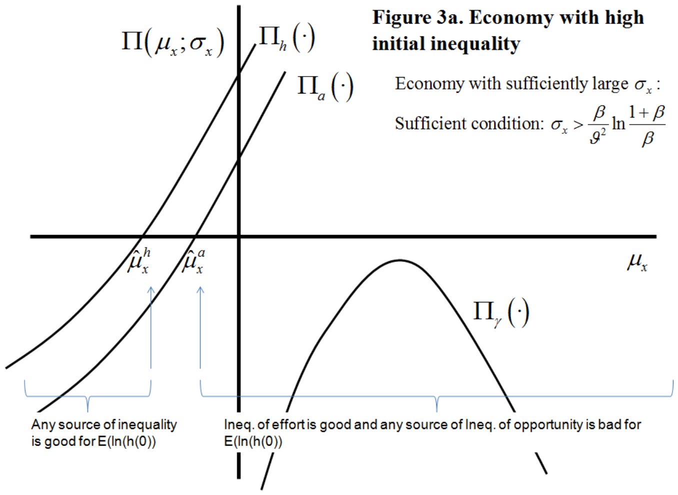
Figures 1a-1f. Average human capital, inequality and meritocracy

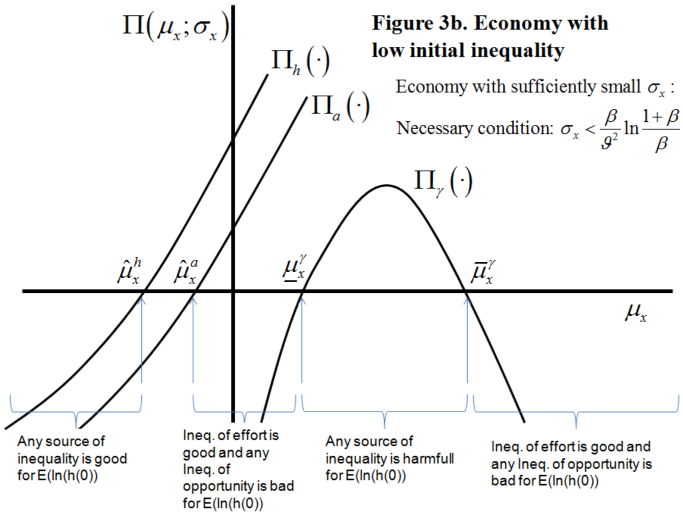


Figures 2a-2d. Human capital dynamics of alternative dynasties

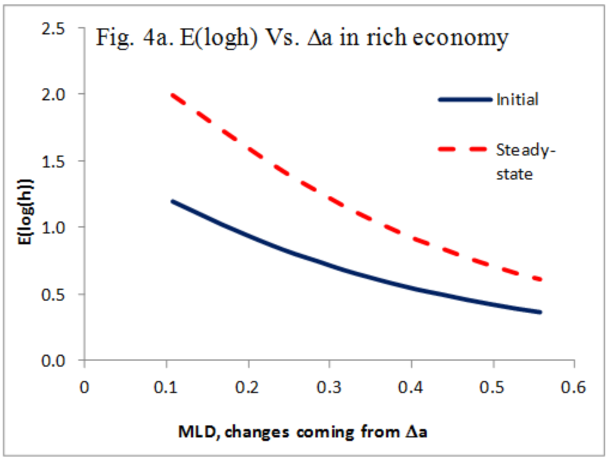


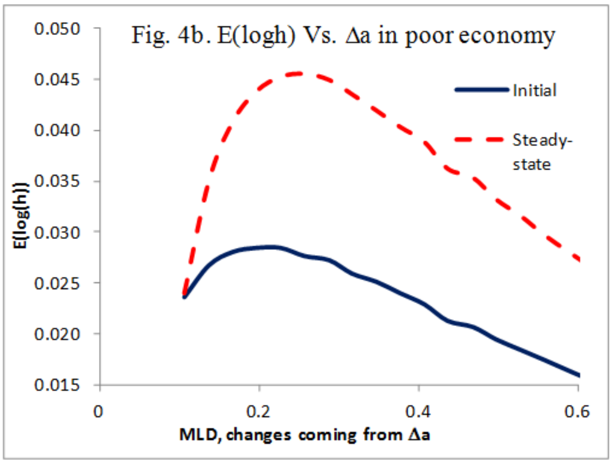
Figures 3a-3b. Inequality (opportunity and free-will) and initial human capital

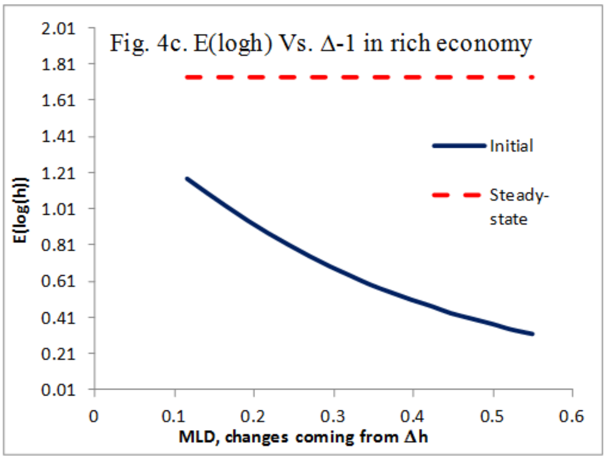


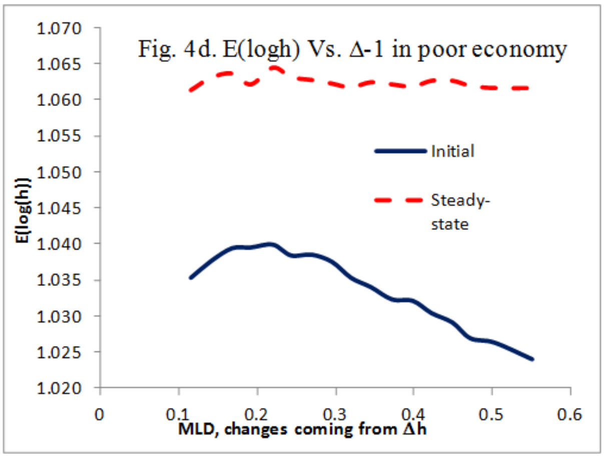


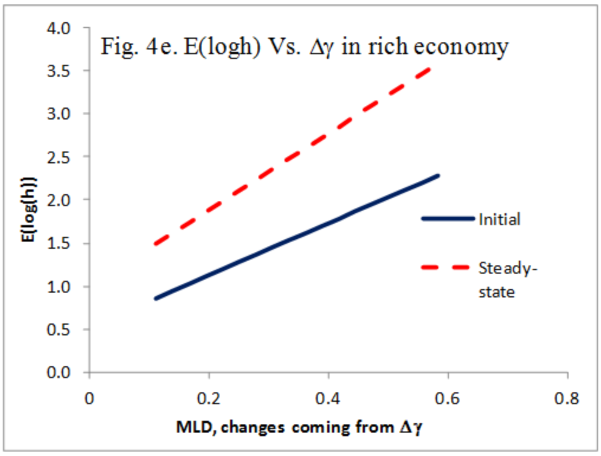
Figures 4a-4f.

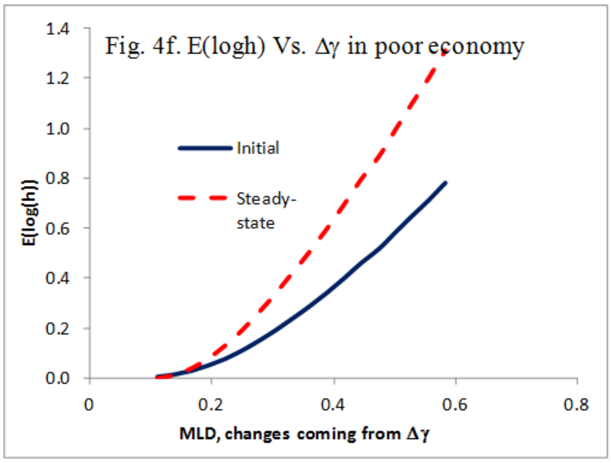












TABLES

Table 1. Average level of education and indices of inequality

Country	Year	Ypc	ln Ypc	h	ln h	I	IO
Austria	2005	37,048	10.5200	75.6	4.3255	0.1181	0.0060
Belgium	2005	36,234	10.4978	77.6	4.3516	0.1031	0.0123
Cyprus	2005	22,298	10.0123	67.2	4.2077	0.1250	0.0244
Denmark	2005	47,546	10.7695	62.2	4.1304	0.0689	0.0013
Finland	2005	37,331	10.5276	70.1	4.2499	0.1160	0.0038
France	2005	34,002	10.4342	82.0	4.4067	0.1096	0.0097
Germany	2005	33,514	10.4197	92.0	4.5218	0.1305	0.0027
Greece	2005	21,468	9.9743	68.0	4.2195	0.2127	0.0230
Iceland	2005	54,884	10.9130	62.0	4.1271	0.1090	0.0052
Ireland	2005	48,761	10.7947	81.8	4.4043	0.1611	0.0242
Italy	2005	30,446	10.3237	73.1	4.2918	0.1874	0.0220
Luxembourg	2005	82,370	11.3190	70.0	4.2485	0.1613	0.0379
Netherlands	2005	39,157	10.5753	87.6	4.4728	0.0884	0.0041
Norway	2005	65,767	11.0939	99.2	4.5971	0.1169	0.0048
Portugal	2005	18,196	9.8090	41.1	3.7160	0.2264	0.0503
Spain	2005	26,058	10.1681	73.2	4.2932	0.2144	0.0286
Sweden	2005	41,042	10.6224	87.5	4.4716	0.1095	0.0087
UK	2005	38,135	10.5489	65.4	4.1805	0.1952	0.0199
Czec R.	2005	12,726	9.4514	86.6	4.4613	0.1196	0.0070
Estonia	2005	10,330	9.2428	94.5	4.5486	0.1985	0.0218
Hungary	2005	10,937	9.2999	94.8	4.5518	0.1314	0.0152
Latvia	2005	6,913	8.8412	95.3	4.5570	0.2995	0.0213
Lithuania	2005	7,641	8.9413	93.6	4.5390	0.2482	0.0358
Poland	2005	7,963	8.9826	81.7	4.4031	0.2649	0.0272
Slovakia	2005	8,844	9.0875	83.7	4.4272	0.1251	0.0045
Slovenia	2005	17,840	9.7892	95.3	4.5570	0.0873	0.0084
Brazil	1996	5,109	8.5388	47.7	3.8649	0.6920	0.2230
Colombia	2003	2,268	7.7267	47.8	3.8670	0.5720	0.1330
Ecuador	2006	3,058	8.0255	45.0	3.8067	0.5800	0.1500
Guatemala	2000	1,530	7.3330	18.1	2.8959	0.5930	0.1990
Panamá	2003	4,138	8.3280	62.1	4.1287	0.6300	0.1900
Perú	2001	2,057	7.6290	64.8	4.1713	0.5570	0.1560
Chile	2009	10,179	9.2281	68.6	4.2283	0.4005	0.0542
Turkey	2003/04	5,833	8.6713	36.9	3.6082	0.3620	0.0948
Egypt	2006	1,426	7.2626	51.2	3.9357	0.3627	0.0477
India	2004/05	735	6.5999	38.7	3.6558	0.4088	0.0837
Ghana	1998	656	6.4862	53.8	3.9853	0.4000	0.0450
Ivory Coast	1985/88	893	6.7946	22.2	3.1001	0.3700	0.0500
South Africa	2008/10	5,689	8.6463	67.1	4.2062	0.6750	0.1690
Uganda	1992	179	5.1874	14.9	2.7014	0.4300	0.0400

Note: Data source for the per capita real income (in US dollars, 2005): the UN Statistics Division; Data source for the average human capital: Barro and Lee Database (v.1.3, 04/13). Data source for I and IO: Marrero and Rodríguez (2012a); Ferreira and Gignoux (2011); Piraino (2012); Ferreira et al. (2011); Baéz et al. (2011); Belhaj-Hassine (2012); Singh (2011); and Cogneau and Mesple-Somps (2008). See Section 4 for a description of the main characteristics of the data sources considered in this Table.

Table 2. The effect of different sources of inequality on the average human capital

Average Human Capital (μ_t)					
	(1)	(2)	(3)	(4)	(5)
I	0.1924	0.3404	1.8708**	--	2.0951*
	(0.7816)	(0.4378)	(0.8982)		(1.2263)
ln Ypc	--	0.3388***	0.3685***	--	0.3587**
		(0.0867)	(0.0916)		(0.1345)
IO	--	--	-4.4642*	--	--
			(2.6561)		
IO x **EU**	--	--	--	-8.7809**	-10.9958**
				(4.1825)	(4.2847)
IO x **East**	--	--	--	1.0946	-4.8942
				(2.1097)	(7.1985)
IO x **Lat**	--	--	--	-2.5490	-4.2105
				(2.6637)	(2.8644)
IO x **Afr**	--	--	--	6.5195**	-4.6354
				(2.8024)	(5.1694)
Regional Dummies	Yes	Yes	Yes	Yes	Yes
Intercept	Yes	Yes	Yes	Yes	Yes
Obs.	40	40	40	40	40
Adj_R^2	0.53	0.71	0.74	0.64	0.75

Note: Standard errors in parenthesis; * significant at 10%, ** significant at 5%, *** significant at 1%.

See note in Table 1.