Knowledge Spillover and Growth Promoting Redistribution

DEBASIS BANDYOPADHYAY* University of Auckland

XUELI TANG† Deakin University

Abstract

Redistribution of income turns out to be sometimes essential for fueling long-run growth in a non-political model of endogenous growth with knowledge spillover but without a credit market. As the per capita income grows, income inequality increases but a higher inequality also slows down the income growth rate. If the income inequality exceeds a threshold then to sustain positive economic growth income redistribution would be essential. A redistributive policy that facilitates accumulation and spilllover of knowledge boosts labour productivity and that, in turn, could increase the net profit of the taxpayers, by raising the exchange value of their resources. We characterize the properties of the welfare-maximizing progressivity as a function of the degree of knowledge spillover and provide a range of numerical estimates. Examining their non-linear relationship, we discover that in order to maximize the long-run growth rate, the average marginal income tax rate may be as large as 20%. Also, facilitating the medium of knowledge spillover such as the Internet and income redistribution appear to be complementary up to a point beyond which to promote economic growth, spillover reduces the potential macroeconomic benefits from income redistribution to make it harmful for growth. Our discovery calls for future research on estimating the extent of knowledge spillover for determining the growth-maximising progressivity of income distribution.

Keywords: Knowledge Externality, Endogenous Growth, Average Marginal Income Tax Rate, Intergenerational mobility, Investment Reallocation Effect.

JEL: E25, E62, O11, O15, O23, O33, O41

*Contact address: Dept. of Economics, OGGB (Level 6), 12 Grafton Road, University of Auckland, PB 92019, Auckland, New Zealand; e-mail: debasis@auckland.ac.nz.

†Contact address: School of Accounting, Economics and Finance, 221 Burwood Highway, Burwood East, Deakin University, Victoria 3125, Australia; email: xtang@deakin.edu.au.

1 Introduction

This paper contributes to a newly emerging vision of sustaining perpetual economic growth on the foundation of a strong middle class. We argue that a growing economy typically widens income inequality which, in turn, lowers the growth rate of income per capita. If income inequality exceeds a threshold then some income redistribution would be necessary for sustaining long-run economic growth. We discover the above fact by focussing on a well-known, but relatively less explored, feature of a growing economy. In particular, we highlight that only through accumulation and spillover of non-rival knowledge, we can overcome the curse of scarce resources, accompanied by the law of diminishing returns to rival inputs that chokes off economic growth. However, the cultivation of knowledge in human minds within and across generations requires markets with human collateral that cannot exist in a decent society. The lack of opportunity to invest in the fertile minds of the impoverished sector of the economy, leaves the super rich with idle resources due to their low rate of returrn at the margin. A prudent government can redistribute such idle resource to prevent the growth engine from slowing down, by carefully reallocating them to reduce illiteracy, high-school dropouts and to provide basic health to those who may be lagging far behind. Such policy may, ironically, help also the wealthy entrepreneurs who would receive an upskilled and more productive labour force which could augment their profit by an amount more than the tax that they pay. Thus we argue the choice between growth and equity may sometimes be a false choice, because inequality retards growth and in order to prevent inequality from becoming too high to choke off economic growth altogether, some redistribution of income would be necessary. Moreover, we identify the characteristics of the optimal redistributive policy package that maximizes economic welfare to ensure that redistribution benefits even the net tax payers in the long-run.

Prior to 1990s we believed in a growth-equity trade-off with the presumption that a greater income inequality is sometimes a necessary price to pay for raising output. A large body of literature in the 1990s changed that wisdom altogether. The 1992 George Seltzer lecture in Minnesota by Robert Solow outlined a new hypothesis that more "equity" could actually promote more growth. Subsequently, wide varieties of dynamic general equilibrium models (see, for example, Galor and Zeira (1993)) have emerged to make growth and inequality interdependent and then to establish alternative versions of 'growth with equity' hypothesis (see for example, Fernandez and Rogerson (1998), Cooper (1998), Benabou (2000, 2002) and Seshadri and Yuki (2004)). In most of those models incomplete markets restrict parental investments in children of heterogeneous ability to a fraction of their disposable income and that gives rise to inefficiently rigid interpersonal differences in productivity. In presence of a convex technology such disparity provides scopes for a redistributive policy to generate output gains. We argue that if income inequality and the associated disparity in productivity are sufficiently large then the output gains from redistribution would be more than the output loss implied by redistribution induced economic distortions. Naturally, the question arises: how large the income inequality has to be before it causes economic waste or, how much inequality is too much for a production efficient economy? We argue that it all depends on a few specific characteristics of an economy. In other words, a large degree of inequality by itself does not call for redistribution while under certain circumstances even a small degree of inequality does so.[1] We design a new technique to compute two indices from the income distribution data: the degree of inequality that would prevail with no redistribution, the laissez-faire income inequality, and the 'threshold inequality' for ensuring an efficient allocation of resources in the credit-constrained economy. We prove that a strategy of no redistribution creates economic waste if and only if the laissez-faire income inequality exceeds the threshold income inequality for the economy. We report a surprising discovery regarding the characteristic of that threshold that the presence of knowledge spillover as the engine of growth is a critical factor in determining if the above threshold would be finite; or, equivalently, if economic efficiency would require income inequality to have an upper limit.

This notion of such threshold for income inequality was implicit in Becker and Tomes (1979), Glomm and Ravikumar (1992) and Benabou (2000).[2] Benabou (2002) alludes to its existence more directly while exploring the answer to a closely related question. In his pioneering work, Benabou (2002) highlights positive roles of various redistributive policies in promoting economic growth and reports that a significantly large and positive degree of redistribution is necessary to maximize long run output per capita in the US economy.

Some discussion on the intuition behind the notion of threshold inequality is in order. In our model a greater degree income inequality corresponds to a wider interpersonal difference

[1] We leave aside welfare and equity considerations, which typically strengthen the case for redistribution.

[2] Implicit in Becker and Tomes (1979) is a notion of threshold inequality such that if income inequality exceeds that threshold then redistribution would raise average income in the long run, provided the fraction of 'family income' that parents spent on their children exceeds a critical value. Glomm and Ravikumar (1992) and Benabou (2000) make relatively explicit statement regarding when inequality would be large enough such that a redistributive policy would promote growth. However, the focus of their papers differ significantly from ours and they leave out consideration of complementarity between physical and human capital and the role of a market for capital and labor, which constitute two important ingredients of our unifying framework for the existing literature.

in productivity that negatively correlates with a country's income per capita. Intergenerational persistence of these systematic differences provide potentials for output gains through redistribution that reallocates funds for investment to reduce variance of productivity in the economy. The potential gains from redistribution can be mapped to a metric for the overall macroeconomic distribution of income which monotonically increases with the parameter representing the variance of idiosyncratic shock. The loss of output, on the other hand, primarily arises from the typical microeconomic adverse effect of redistribution on the supplies of inputs and hence can be identified by a single statistic independent of the distribution of income. It turns out that we can define this statistic purely by the microeconomic fundamentals such as the preference, technology, policy and institution. Moreover, this separation of macro and micro factors corresponding to gains and losses from redistribution aids us to develop an explicitly transparent algorithm to determine a unique threshold for income inequality such that output loss from redistribution would break even with the output gains if the income inequality prior to any redistribution equals this threshold.

We prove that a comparison of the threshold for efficient income inequality and the prevailing laissez-faire income inequality would be necessary and sufficient to determine if a redistributive policy has the potential to promote economic growth or not. In particular, we report that if the laissez-faire inequality exceeds the threshold for efficient income inequality then income redistribution could promote growth; otherwise not. The presence of knowledge spillover as an engine of economic growth lower that threshold and its absence raises the threshold to infinity and, thereby, providing a reason for any income inequality to be efficient without any scope for knowledge spillover.

From numerical simulations we find that the threshold for efficient inequality would more likely to be higher than the laissez-faire inequality in a country with a smaller degree of capital-skill complementarity, with a poorer quality of educational system, and with less segregated communities such that family connections play a smaller role in determining a child's human capital. At the same time, between two countries with similar economic fundamentals, the one with relatively diverse population-mix would be more likely to benefit from redistribution. A typically developed and industrialized country attracts a wider variety of people from around the world than the less developed countries. Also, the rich industrialized countries typically rely on skill intensive technology and high-tech capital goods that requires complementary skill. Consequently, we argue that a less developed country that primarily relies on unskilled labor and is populated by relatively homogeneous group of people would be likely to incur a net loss of output from redistribution while a developed country that heavily relies on modern machines and complementary skilled labor and, is populated by a relatively diverse group of people, would be likely to make a net gain of output from some redistribution.

Section 2 develops the model with no credit market, knowledge spillover and progressive redistribution with income taxes. Section 3 characterizes the model's steady state and the condition for endogenous growth. Section 4 examines how knowledge spillover affects inequality and economic growth while Section 5 characterizes the welfare maximizing progressivity and the critical minimum threshold for income inequality beyond which income redistribution would promote growth. Section 6 includes a few concluding remarks followed by the Appendix and the list of references.

2 Model and Equilibrium

The model considers a continuum of infinitely lived dynasties $i \in [0, 1]$. Following Loury (1981), each dynasty is made of a sequence of families consisting of individuals who live for two periods or two generations, first as a child and then as an adult. In each period t, the dynasty is represented by a family of an adult and a child. The adult, in period t represents the dynasty from that period onward and makes all decisions for that period subject to the constraint that she cannot pass on her debt to her child. We call this adult of the dynasty i in period t the dynastical agent i or simply agent i. The preference of the dynastical agent i at period t is given by:

$$\ln U_t^i = E_t \left[\sum_{n=0}^{\infty} \rho^n \left(\ln c_{t+n}^i - \left(l_{t+n}^i \right)^{\eta} \right) \right], \eta > 1, \tag{1}$$

where $c_t^i \geq 0$ and $l_t^i \in [0, 1]$ denote, respectively, consumption and labor supply by the adult of the dynasty i in period t; $\rho \in (0, 1)$ is the discount factor.[3]

Following Galor and Zeira (1993) and Benabou (2002), we choose the relevant environment for studying issues of income distribution and economic growth to be the one without credit market. Also, to avoid extraneous issues, we assume that everyone operates with a backyard technology but allow both physical and human capital to affect output as complementary inputs in the same way as Barro, Mankiw and Sala-i-Martin (1995) such that the

[3]Note that the intertemporal elasticity of substitution of labor, $\varrho = \frac{1}{\eta - 1}$. We assume $\varrho > 0$ or, equivalently, $\eta > 1$.

output of the self-employed agent i as a function of her physical and human capital k_t^i, h_t^i and labor l_t^i satisfies

$$y_t^i = \left(k_t^i\right)^{\lambda} \left(h_t^i\right)^{\mu} \left(l_t^i\right)^{\epsilon}, \text{ where, } \epsilon = 1 - \lambda - \mu. \tag{2}$$

The disposable income $\hat{y}_t^i$ of the agent i must equal the total expenditure on consumption c_t^i, private education expenditure e_t^i and bequest b_t^i. In other words,

$$\hat{y}_t^i = c_t^i + e_t^i + b_t^i. \tag{3}$$

2.1 Knowledge Spillover

In the following period her grown up child's human capital h_{t+1}^i as a function of her innate ability ξ_{t+1}^i, external effects arising from neighborhood or family as proxied by parental human capital h_t^i, and the private investment on her education e_t^i, is given by,

$$h_{t+1}^i = \kappa_t \xi_{t+1}^i \left(h_t^i\right)^{\alpha} \left(e_t^i\right)^{\beta}, \alpha + \beta < 1. \tag{4}$$

where

$$\kappa_t \equiv \kappa \left(\int_0^1 \left(h_t^i\right)^{\omega} di \right)^{\delta/\omega}, \kappa > 0 \text{ and } \omega > 0. \tag{5}$$

The idiosyncratic shocks ξ_t^i that arise from discrepancies in innate ability or in efficiency of human capital usage are $i.i.d.$ with $\ln \xi_t^i \sim N\left(-\sigma^2/2, \sigma^2\right)$, where σ^2 is constant. The parameter $\alpha \in (0,1)$ measures the child's human capital elasticity of "neighborhood externality," a phrase explored originally in Benabou (1996) in the context of human capital inequality and the parameter $\beta \in (0,1)$ measures the same elasticity of the education expenditure, which is primarily determined by the quality of the education system. The the externality is following Lucas (1988)'s basic idea of externality of human capital and considering the aggregate effect of efficiency unit of human capital, $\left(h_t^i\right)^{\omega}$, as in Benabou (2002). The intuition of this formulation is that individuals learn from the knowledge of others. The parameter δ measures the degree of externality and κ denotes the unit of coversion from educational expenditure to human capital.

Note, the only source of uncertainty in our model comes from the shocks to efficiency of human capital usage such as the inborn talent shock, as described above, in the production of human capital, and not from an income shock. To promote this assumption, following Becker and Tomes (1979), we argue that “market luck”, which performs like a shock to income, would not be as important as "endowment luck", corresponding to the random assignment of natural ability in determining agents’ income. This conclusion follows from the fact that there are competitive markets readily available to insure the income shock, while the difficulty of verifying talent or the inborn ability of a child precipitates moral hazard and adverse selection problems, ruling out a similar insurance market for the talent shocks.[4] A policy of income redistribution does provide a cushion against the implied disutility for the risk-averse agents. Consequently, our modelling of the uninsurable risk that naturally arises in the human capital production process, provides a welfare improving role for the government’s income redistribution policy.

Capital goods are complementary to human capital and become obsolete at the end of each generation. A tool loses value when its user dies. Parents buy new tools for their children and leave them as bequest. To capture this feature we assume that they depreciate completely in the production process. Consequently, in the generation $t+1$, the agent i’s physical capital k_{t+1}^{i} consists only of her parent’s bequest b_{t}^{i} such that

$$k_{t+1}^{i} = b_{t}^{i}. \tag{6}$$

Initial endowments of physical and human capital k_{0}^{i} and h_{0}^{i} are jointly, lognormally distributed and the adult receives one unit of labor endowment in each period.

2.2 Progressivity

By assumption, the government cannot detect individual innate ability ξ_{t}^{i} and neighborhood or family effects h_{t}^{i}, but does observe individual incomes y_{t}^{i} and their expenditure on education e_{t}^{i}. Following Benabou (2002), the disposable income of a typical agent at a date t satisfies

$$\hat{y}_{t}^{i} \equiv \left(y_{t}^{i}\right)^{1-\tau} \left(\tilde{y}_{t}\right)^{\tau}, \tag{7}$$

such that those with income higher than $\tilde{y}_{t}$ pay net tax while those with income below $\tilde{y}_{t}$ receive net transfers and the balanced-budget constraint is

$$\int_{0}^{1} \left(y_{t}^{i}\right)^{1-\tau} \left(\tilde{y}_{t}\right)^{\tau} di = y_{t}, \tag{8}$$

[4]Note that unemployment insurance and income protection insurance are readily available to provide protection against income shock, whereas no such insurance exists to protect someone against ability shock.

where $y_t \equiv \int_0^1 y_t^i di$ denotes per-capita income, $\tilde{y}_t$ represents the break-even level of income and $0 < \tau < 1$ measures the average marginal tax rate and is identified as the degree of redistribution or *progressivity* in fiscal policy.[5] The parameter τ captures the notion of progressivity that we consider in this paper. Earlier work of Benabou (2002), Jakobsson (1976) and Kakwani (1977) posits the appropriateness of such parameterization. Note the parameter τ is equal to the income weighted average marginal tax (and transfer) rate: $\tau = \int_0^1 T'\left(y_t^i\right) \cdot \left(y_t^i / y_t\right) di$, where $T\left(y_t^i\right) = y_t^i - \hat{y}_t^i$ is the net tax paid at income level y_t^i.

At each date t, let m_{ht}, m_{kt} denote the means and $\Delta_{ht}^2, \Delta_{kt}^2$ denote the variances of $\ln h_t^i$ and $\ln k_t^i$, respectively, and let cov_t denote the covariance between $\ln h_t^i$ and $\ln k_t^i$. Suppose $M_t \equiv (m_{ht}, m_{kt}, \Delta_{ht}^2, \Delta_{kt}^2, cov_t)$. Then for the agent's dynamic optimization problem, the state variables are $(h_t^i, k_t^i, M_t; \tau)$, the control variables are $(c_t^i,\ l_t^i,\ e_t^i,\ b_t^i)$ and the Bellman equation is as follows

$$\ln U\left(h_t^i, k_t^i, M_t; T\right) = \max_{c_t^i, l_t^i, e_t^i, b_t^i} \left\{ \begin{array}{l} (1-\rho)\left[\ln c_t^i - (l_t^i)^\eta\right] \\ +\rho E_t\left[\ln U(h_{t+1}^i, k_{t+1}^i, M_{t+1}; \tau)\right] \end{array} \right\}, \tag{9}$$

subject to (2), (3), (4), (6) and (7).

The first order conditions associated with the Bellman equation described by (9) yield complete solutions to the agent's problem which can be found in the Appendix.

Together with the government's budget constraint (8), a unique sequence of aggregate state variables $\{M_t\}$ coincides with what agent i takes as given in solving (9) in the equilibrium.

In other words, at each date t the economy's output, measured by the sum of the output produced by the self-employed dynastical agents, equals the aggregate demand for consumption and investment.

Together with the government's budget constraints (8), the above decision rules imply a unique sequence of aggregate state variables $\{M_t\}$ that coincides with what the agent i takes as given in (9) such that at each date $t = 0, 1, 2, ..,$ the following aggregate consistency condition holds:

$$\int_0^1 y_t^i di = \int_0^1 c_t^i di + \int_0^1 e_t^i di + \int_0^1 b_t^i di. \tag{10}$$

In other words, at each date t the economy's output, measured by the sum of the output

[5] Note that on a logarithmic scale τ denotes the proportional tax rate on the log of personal income and we only focus on redistributive policies that transfer resources only from high income to low income people.

produced by the self-employed dynastical agents, equals the aggregate demand for consumption and investment in education and bequest.

In line with Benabou (2002) we define for each date t an index of inequality Λ_t as the logarithm of the ratio of mean to median income, which equals the variance of logarithmic earnings of agents.

LEMMA 2: *At each date t, inequality index Λ_t equals the variance of logarithmic earnings of agents such that $\Lambda_t = \left(\lambda^2 \Delta_{kt}^2 + \mu^2 \Delta_{ht}^2 + 2\lambda\mu cov_t\right)/2$ and the evolution of earnings of adults is governed by a lognormal distribution such that*
$\ln y_t^i \sim N\left(\lambda m_{kt} + \mu m_{ht} + \epsilon \ln l, 2\Lambda_t\right)$. *The break-even level of income $\tilde{y}_t$ at which an agent's net tax obligation is zero satisfies:*

$$\begin{aligned} \ln \tilde{y}_t &= \ln y_t + (1-\tau)\,\Lambda_t, \\ &= \lambda m_{kt} + \mu m_{ht} + \epsilon \ln l\,(\tau) + (2-\tau)\,\Lambda_t. \end{aligned} \tag{11}$$

Proof: see appendix.

3 Steady State

3.1 Inequality

In the long run, by (60), (62) and (63), the following lemma claims that in the presence of knowledge spillover (i.e., $\delta > 0$) the sequence of income inequality converges to a stationary state.

LEMMA 3: *Irrespective of the initial conditions, income inequality converges monotonically to its unique ergodic limit Λ, where,*

$$\Lambda = \frac{\mu^2\left(1+\lambda\alpha\,(1-\tau)\right)}{\left(1-\lambda\alpha\,(1-\tau)\right)\left(\left(1+\lambda\alpha\,(1-\tau)\right)^2 - \left(\alpha + (\lambda+\beta\mu)\,(1-\tau)\right)^2\right)}\frac{\sigma^2}{2}. \tag{12}$$

And differentiating (12) w.r.t τ gives

$$\frac{\partial \Lambda}{\partial \tau} = -z\Lambda, \tag{13}$$

where, $z > 0$ and its expression is given by (A.27) in Appendix.

Proof: see appendix.

Importance of the knowledge spillover externality The absence of a credit market creates rigid and inefficient interpersonal differences in marginal productivity of human and physical capital. In the presence of diminishing returns technology for accumulating human capital and production of output as the variance of productivity increases the average output decreases. Consequently, in such an environment any economic factor that increases heterogeneity among the production units would lower per capita output and hence its growth rate in the present context. Our assumption of a positive knowledge spillover parameterized by $\delta > 0$ turns out to be critical for generating a negative growth-inequality relationship. If $\delta = 0$ then our endogenous growth condition, as stated in Lemma 3, would require that either the production technology or the human capital accumulation technology must exhibit increasing returns; but that would offset the possibility of a negative growth-inequality relationship, which typically requires a unique combination of assumptions of a missing credit market and technologies with diminishing returns.

3.2 Endogenous Growth

If the knowledge spillover remains bounded such that $\delta < 1 - \alpha$, then from any arbitrary initial condition, the economy reaches a balanced growth path as described in the following Proposition:

PROPOSITION 1: *The model economy exhibits endogenous growth if, and only if, the following condition holds:*

$$\frac{\beta}{1-\alpha-\delta} = \frac{1-\lambda}{\mu}. \tag{14}$$

Then the economy follows a balanced growth path such that the growth rate of per capita income is constant, $\gamma_t \equiv \ln y_{t+1} - \ln y_t = \gamma$, where,

$$\gamma = \frac{1}{1-\lambda\left(\alpha+\delta\right)}\left(\Phi - \Psi\Lambda\right), \tag{15}$$

where $\Phi \equiv \mu \ln \kappa + \mu\beta \ln s_1(\tau) + \lambda(1-\alpha-\delta)\ln s_2(\tau) - \left(\mu - (\lambda+\mu)^2\right)\sigma^2/2 + (1-\alpha-\delta)\,\epsilon \ln l(\tau)$, *depends on the education investment rate* $s_1(\tau)$*, bequest rate* $s_2(\tau)$ *and labor supply* $l(\tau)$*, and the harmful effect of inequality on growth, denoted by* Ψ*, is:*

$$\Psi = \Psi_1 + \Psi_2 + \Psi_3 + \Psi_4, \tag{16}$$

where

$$\Psi_1 \equiv \left(1 - \lambda^2\alpha^2(1-\tau)^2\right)\left(1 + \frac{\lambda}{\mu}\right)^2 > 0,$$

$$\Psi_2 \equiv -(1-\alpha-\delta)(2-\tau)\tau < 0,$$

$$\Psi_3 \equiv -\frac{(\lambda+\mu)^2(1-\lambda\alpha(1-\tau))(\alpha + (\lambda + (1-\alpha-\delta)(1-\lambda))(1-\tau))^2}{(1+\lambda\alpha(1-\tau))\mu^2} < 0,$$

$$\Psi_4 \equiv -\delta\omega \frac{\begin{pmatrix} (1-\lambda(1-\tau)(\alpha + 2(1-\alpha-\delta)(1-\lambda)(1-\tau)))\left(1-\lambda^2(1-\tau)^2\right) \\ -2(1-\alpha-\delta)(1-\lambda)\lambda^3(1-\tau)^4 \end{pmatrix}}{\mu(1+\lambda\alpha(1-\tau))} > 0.$$

Proof: see appendix.

We note from above that the terms Ψ_i, $i = 1,..,4$, captures the partial effect of a reduction in income inequality. It turns out that two of those effects, Ψ_1 and Ψ_4, are positive while the other two Ψ_2 and Ψ_3 are negative. Also, the positive parial effect through the channel of Ψ_1 works irrespective of the presence of knowledge spillover parameter δ while the negative effect through works though the channel of accumulation of human capital irrespective of the presence of physical capital, decreases with knowledge spillover and can be eliminitated by setting the progressivity parameter $\tau = 0$, as discussed in Bandyopadhyay and Tang (2011). The negative partial effect of a lower income inequality on growth that works though the channel captured by Ψ_3 diminishes with a greater spillover of knowledge but depends non-trivially on the process of accumulation of physical and human capital, Interestingly, only in the absence of knowledge spillover such that $\delta = 0$, this negative effect of inequality reduction on growth can be eliminitated by avoiding income redistribution, that is, by setting $\tau = 0$. The channel Ψ_4 captures the positive partial effect on growth only if $\delta > 0$. Thus, the presence of knowledge spillover in the economy brings additional beneficial effect on economic growth from a reduction on income inequality in a way which has remained a relatively less explored area in the literature.

4 Knowledge Spillover and Growth

On the balanced growth path, by (14) and Lemma 1, we can rewrite the function of growth rate of per capita income, optimal labour supply and saving rates as follows,

$$\gamma = \frac{1}{1-\lambda(\alpha+\delta)}(\Phi - \Psi\Lambda), \tag{17}$$

where $\Phi = \mu \ln \kappa + (1-\alpha-\delta)(1-\lambda)\ln s_1(\tau) + \lambda(1-\alpha-\delta)\ln s_2(\tau) - \left(\mu - (\lambda+\mu)^2\right)\sigma^2/2 + (1-\alpha-\delta)\epsilon \ln l(\tau)$,

$$l(\tau) = \left(\frac{(\epsilon/\eta)(1-\rho\alpha)(1-\tau)}{(1-\rho\alpha)(1-\rho\lambda(1-\tau)) - \rho(1-\alpha-\delta)(1-\lambda)(1-\tau)}\right)^{1/\eta}, \tag{18}$$

$$s_1(\tau) = \frac{\rho(1-\alpha-\delta)(1-\lambda)(1-\tau)}{1-\rho\alpha}, \tag{19}$$

$$s_2(\tau) = \rho\lambda(1-\tau). \tag{20}$$

The term Φ captures the positive contribution of private investment in physical and human capital as well as work effort on economic growth that we can trace at the individual level without considering any distributional issues that arise from interactiona among a heterogeneous population in the economy. In presence of income inequality the country forgoes $\Psi\Lambda$ units of growth where $\Psi > 0$ measures partial negative effect of inequality on the rate of growth at the margin. Clearly, if the total adverse effect of inequality $\Psi\Lambda$ exactly offsets the total contribution Φ on economic growth from individual investments in human and physical capital and work effort then growth disappears. In presence of externality and physical capital as a complementary input to human capital, any growth spur from individual investment creats an additional mutiplier effect. In particular, the contribution of additional investment to output gets multiplied as the spillover of knowledge within and across dynasty coupled with the mutual augmenting of marginal productivity between physical and human capital help to keep that growth spur going for many successive generations to result in the above mutiplier effect.

Equation (18) and (19) show that optimal labour supply and saving rate for education expenditure go down as δ goes up.

By (14), we can rearrange the income inequality function (12) and get

$$\Lambda = \frac{\mu^2 (1+\lambda\alpha(1-\tau))}{(1-\lambda\alpha(1-\tau))\left((1+\lambda\alpha(1-\tau))^2 - (\alpha+(\lambda+(1-\alpha-\delta)(1-\lambda))(1-\tau))^2\right)} \frac{\sigma^2}{2}. \tag{21}$$

The equation (21) shows that income inequality goes down as δ increases.

In order to discuss the effect of δ on γ, we differentiate γ w.r.t δ. For simplicity, we set $\tau = 0$. Then, by (16), (18), (19), and (21), differentiating (17) w.r.t δ yields

$$\begin{aligned} \frac{\partial\gamma}{\partial\delta} \quad | \quad {}_{\tau=0} = \frac{1}{1-\lambda(\alpha+\delta)} \left(\frac{\partial\Phi}{\partial\delta} - \left(\frac{\partial\Psi}{\partial\delta}\Lambda + \Psi\frac{\partial\Lambda}{\partial\delta} \right) + \lambda\gamma \right) \qquad (22) \\ = \quad \frac{1}{1-\lambda(\alpha+\delta)} \left(\frac{\partial\Phi}{\partial\delta} + \Gamma\Lambda + \lambda\gamma \right), \text{ denoting } \Gamma \equiv a\Psi - \frac{\partial\Psi}{\partial\delta} \end{aligned}$$

where $\dfrac{\partial\Lambda}{\partial\delta} = -a\Lambda < 0$, since,

$$a \equiv \frac{2((1-\delta)+\lambda(\alpha+\delta))}{\delta((1+\lambda)(1+\alpha)+(1-\lambda)(1-\alpha-\delta))} > 0; \tag{23}$$

and

$$\begin{aligned} \frac{\partial\Phi}{\partial\delta} \quad = \quad & -(1-\lambda)(1+\ln s_1) - \lambda\ln s_2 \qquad (24) \\ & -\epsilon\ln l - \frac{1}{\eta}\frac{\rho(1-\alpha-\delta)\epsilon(1-\lambda)}{(1-\rho\alpha)(1-\rho\lambda)-\rho(1-\alpha-\delta)(1-\lambda)}, \end{aligned}$$

$$\begin{aligned} \frac{\partial\Psi}{\partial\delta} \quad = \quad & \frac{2}{\alpha\lambda+1}\left(\frac{\lambda}{\mu}+1\right)^2 (1-\lambda)(1-\alpha\lambda)(1+\alpha\lambda-\delta(1-\lambda)) \qquad (25) \\ & -\frac{\omega}{\mu}\frac{1-\lambda}{\alpha\lambda+1}((\alpha\lambda+1)(1-\lambda)+4\delta\lambda). \end{aligned}$$

The term $\Gamma\Lambda$ represents the potential gain from spillover. It becomes larger with higher inequality. The term $\lambda\gamma$ shows that the economy would benefit more from externality if it has inital higher growth rate.

Lemma 4a: There exists $\rho^* \in (0,1)$ and $\lambda^* \in (0,1)$ such that if $\rho < \rho^*$ and $\lambda \leq \lambda^*$ then $\frac{\partial\Phi}{\partial\delta} > 0$. Or, there exists $\rho^* \in (0,1)$, $\alpha^* \in (0,1)$ and $\lambda^* \in (0,1)$ such that if $1 > \rho > \rho^*$, $\alpha > \alpha^*$ and $\lambda \leq \lambda^*$, then $\frac{\partial\Phi}{\partial\delta} > 0$.

Proof: see appendix.

Lemma 4b: If $\alpha\mu < 0.5$, then there exists $\lambda^* \in (0,1)$ and ω_1^* such that if $\lambda < \lambda^*$ then $\omega_1^* \in (0,1)$. Then if $\omega > \omega_1^*$, there exists a $\delta_1^* \in (0, 1-\alpha)$ such that $\frac{\partial\Psi}{\partial\delta} \lesseqqgtr 0$ if $\delta \gtreqless \delta_1^*$, where

$$\delta_1^* \equiv (1+\alpha\lambda)\frac{2(1-\alpha\lambda)\left(1+\frac{\lambda}{\mu}\right)^2 - \dfrac{\omega}{\mu}(1-\lambda)}{2\left(2\lambda\frac{\omega}{\mu} + (1-\lambda)(1-\alpha\lambda)\left(1+\frac{\lambda}{\mu}\right)^2\right)}. \tag{26}$$

Proof: see appendix.

By (16), evaluating Ψ at $\tau = 0$, we can get

$$\Psi = \frac{\delta(1-\lambda)}{(1+\lambda\alpha)}\left(\begin{array}{c} (2-(1-\lambda)\delta + 2\alpha\lambda)(1-\lambda\alpha)\left(1+\dfrac{\lambda}{\mu}\right)^2 \\ \\ -\dfrac{\omega}{\mu}((1+\alpha\lambda)(1-\lambda)+2\lambda\delta) \end{array}\right). \tag{27}$$

Lemma 4c: If $\alpha\mu < 1-\mu$, then there exists $\lambda^{**} \in (0,1)$ and ω_2^*. If $\lambda < \lambda^{**}$ then $\omega_2^* \in (0,1)$. Then, if $\omega_2^* < \omega \Leftrightarrow \delta_2^* < 1-\alpha$, then we can say $\Psi \gtreqless 0$ is equivalent to

$$\delta \lesseqqgtr \delta_2^* = 2\delta_1^*, \text{ where } 0 < \delta_2^* < 1-\alpha. \tag{28}$$

Proof: see appendix.

By Lemma 4a-4c, to have $\frac{\partial\gamma}{\partial\delta}\mid_{\tau=0} > 0$, we need the following Lemma.

LEMMA 4: If $\alpha\mu < 0.5$, $\alpha > \alpha^* \in (0,1)$ and $\lambda \leq \lambda^* \in (0,1)$, $\delta_1^* < \delta < \delta_2^*$ then $\frac{\partial\gamma}{\partial\delta}\mid_{\tau=0} > 0$.

Proof: By Lemma 4a, we know that if $\alpha > \alpha^*$ and $\lambda \leq \lambda^*$, then $\frac{\partial \Phi}{\partial \delta} > 0$. We know that $\Lambda > 0$ and $\gamma > 0$. Thus, if $\Psi > 0$ and $\frac{\partial \Psi}{\partial \delta} < 0$ then $\Gamma > 0$. By Lemma 4b and 4c, we know that if $\delta_1^* < \delta < \delta_2^*$, $\Gamma > 0$. $\square$

If $\lambda = 0$, we can have a similar Lemma as below.

LEMMA 5: If $\sigma^2 < \sigma^*$, there exists a $\delta_3^* \in (0, 1-\alpha)$ such that $\frac{\partial \gamma}{\partial \delta} = 0$ if $\delta = \delta_3^*$. For any $\delta > \delta_3^*$, $\frac{\partial \gamma}{\partial \delta} \mid_{\tau=0} > 0$.

Proof: see appendix.

The above result can also be found implicitly in Zhang (2005) that for a range of δ, any increase of δ can increase the balanced growth rate.

In summing up this section we note that a greater degree of knowledge spillover does not automatically help economy to grow faster because it generates non-trivial disincentives at the micro level which offsets the positive macroeconomic effect on economic growth due to reductions in income inequality. Consequently, we would not necessarily expect an inequality reducing income redistribution to promote growth even though knowledge spillovers fuel economic growth.

5 Progressivity and Growth

After getting the function of endogenous growth rate, in this section, we introduce the definition of threshold inequality and prove that if the actual inequality is larger than the threshold inequality then the economy would waste resources without redistribution.

Differentiating (15) w.r.t τ yields

$$\frac{\partial \gamma}{\partial \tau} = \frac{1}{1 - \lambda(\alpha + \delta)} \left(\frac{\partial \Phi}{\partial \tau} - \left(\frac{\partial \Psi}{\partial \tau} \Lambda + \Psi \frac{\partial \Lambda}{\partial \tau} \right) \right). \tag{29}$$

The term $\frac{\partial \Phi}{\partial \tau} < 0$ represents the negative microeconomic effect of redistribution, the term $\frac{\partial \Psi}{\partial \tau}$ represents the ambiguous macroeconomic effect of redistribution and the term $\frac{\partial \Lambda}{\partial \tau}$ measures the negative effect of redistribution on income inequality. Overall, the equation (29) captures three distinct chanells through which redistribution of income affects economic growth.

5.1 Threshold Inequality

By Lemma 3, $\frac{\partial \Lambda}{\partial \tau} = -z\Lambda$, where $z > 0$. It follows, therefore, from (29) that

$$\begin{aligned} \frac{\partial \gamma}{\partial \tau} &= \frac{1}{1-\lambda(\alpha+\delta)} \left(\frac{\partial \Phi}{\partial \tau} - \left(\frac{\partial \Psi}{\partial \tau} - \Psi z \right) \Lambda \right), \\ &= \frac{1}{1-\lambda(\alpha+\delta)} \left(\frac{\partial \Phi}{\partial \tau} + \Omega\Lambda \right), \text{ denoting } \Omega \equiv \Psi z - \frac{\partial \Psi}{\partial \tau}. \end{aligned} \tag{30}$$

LEMMA 6: *If* $\Omega \leq 0$ *then* $\frac{\partial \gamma}{\partial \tau} < 0$ *since* $\frac{\partial \Phi}{\partial \tau} < 0$ *and* $\Lambda > 0$.

Clearly, if $\Omega \leq 0$, then the optimal redistributive tax rate $\tau^* = 0$.

PROPOSITION 2: *If* $\Omega > 0$, *there exists a real* $\Lambda^* > 0$ *such that* $\frac{\partial \gamma}{\partial \tau} |_{\tau=0} > 0$ *if and only if* $\Lambda > \Lambda^*$, *where*

$$\Lambda^* = -\frac{\partial \Phi}{\partial \tau} / \Omega. \tag{31}$$

We define Λ^* to be a critical minimal **threshold income inequality**. Following Lemma 6 and Proposition 2, it can be seen that the condition for $\tau^* > 0$ crucially depends on the sign of Ω. Therefore, in the following Section, we seek sufficient condition to determine the sign of Ω.

Recall that $\Omega \equiv \Psi z - \frac{\partial \Psi}{\partial \tau}$ where $z > 0$. It means that the sign of Ω depends on the sign of Ψ and $\frac{\partial \Psi}{\partial \tau}$. By Lemma 6 and Proposition 2, we can get the following Lemmas:

L1). *If* $\Psi > 0$ *and* $\frac{\partial \Psi}{\partial \tau} > 0$ *then there exists a* $z^* \equiv \left(\frac{\partial \Psi}{\partial \tau} / \Psi \right) |_{\tau=0} > 0$ *such that if* $z < z^*$ *then* $\Omega < 0$.

L2). *If* $\Psi < 0$ *and* $\frac{\partial \Psi}{\partial \tau} < 0$ *then there exists a* $z^* > 0$ *such that if* $z > z^*$ *then* $\Omega < 0$.

L3). *If* $\Psi < 0$ *and* $\frac{\partial \Psi}{\partial \tau} > 0$ *then* $\Omega < 0$.

L4). *If* $\Psi > 0$ *and* $\frac{\partial \Psi}{\partial \tau} < 0$ *then* $\Omega > 0$.

L5). *If* $\Psi = 0$ *and* $\frac{\partial \Psi}{\partial \tau} > 0$ *then* $\Omega < 0$.

L6). *If* $\Psi = 0$ *and* $\frac{\partial\Psi}{\partial\tau} = 0$ *then* $\Omega = 0$.

PROPOSITION 3: *In our model, the presence of knowledge spillover as an engine of growth is a necessary condition for a finite threshold income inequality and growth promoting progressive redistribution.*

To prove the Proposition 3, we first check the sign of Ψ and $\frac{\partial\Psi}{\partial\tau}$ and then Ω under two sub-cases 1) $\delta = 0$ with $\lambda = 0$ or $\lambda > 0$, 2) $\delta > 0$ with $\lambda = 0$ or $\lambda > 0$.

Sub-Case 1: $\delta = 0$ **with** $\lambda = 0$ **or** $\lambda > 0$

(i) If $\lambda = 0$

By (16), we get

$$\begin{aligned}\Psi_1 &= 1, \\ \Psi_2 &= -(1-\alpha)(2-\tau)\tau, \\ \Psi_3 &= -(\alpha + (1-\alpha)(1-\tau))^2,\end{aligned}$$

and then,

$$\Psi = \alpha(1-\alpha)\tau^2 > 0. \tag{32}$$

Then differentiating (32) w.r.t τ gives

$$\frac{\partial\Psi}{\partial\tau} = 2\alpha(1-\alpha)\tau > 0. \tag{33}$$

When $\tau = 0$, both $\Psi = 0$ and $\frac{\partial\Psi}{\partial\tau} = 0$. Thus, by Lemma L6, we can conclude that the optimal redistributive tax rate $\tau^* = 0$. This is consistent with the conclusion in Bandyopadhyay and Tang (2011).

(ii) If $\lambda > 0$

By (16), we have

$$\begin{aligned}\Psi_1 &= \left(1 - \lambda^2\alpha^2(1-\tau)^2\right)\left(1 + \frac{\lambda}{\mu}\right)^2 > 0, \qquad (34) \\ \Psi_2 &= -(1-\alpha)(2-\tau)\tau < 0, \\ \Psi_3 &= -\frac{(1-\lambda\alpha(1-\tau))(\alpha + (1-\alpha(1-\lambda))(1-\tau))^2}{1+\lambda\alpha(1-\tau)}\left(1 + \frac{\lambda}{\mu}\right)^2 < 0,\end{aligned}$$

and it follows,

$$
\begin{aligned}
\Psi &= \left(1-\lambda^{2}\alpha^{2}\left(1-\tau\right)^{2}\right)\left(1+\frac{\lambda}{\mu}\right)^{2}-\left(1-\alpha\right)\left(2-\tau\right)\tau \\
&\quad -\frac{\left(\lambda+\mu\right)^{2}\left(1-\lambda\alpha\left(1-\tau\right)\right)\left(\alpha+\left(1-\alpha\left(1-\lambda\right)\right)\left(1-\tau\right)\right)^{2}}{\left(1+\lambda\alpha\left(1-\tau\right)\right)\mu^{2}}, \\
&= 0, \text{ when } \tau=0.
\end{aligned} \tag{35}
$$

Differentiating (34) w.r.t τ gives

$$
\begin{aligned}
\frac{\partial\Psi_1}{\partial\tau} &= 2\alpha^{2}\lambda^{2}\left(1-\tau\right)\left(1+\frac{\lambda}{\mu}\right)^{2}>0, \\
\frac{\partial\Psi_2}{\partial\tau} &= -2\left(1-\alpha\right)\left(1-\tau\right)<0, \\
\frac{\partial\Psi_3}{\partial\tau} &= -\left(1+\frac{\lambda}{\mu}\right)^{2}A\frac{\partial\ln A}{\partial\tau},
\end{aligned} \tag{36}
$$

where $A\equiv\dfrac{\left(1-\lambda\alpha\left(1-\tau\right)\right)\left(\alpha+\left(1-\alpha\left(1-\lambda\right)\right)\left(1-\tau\right)\right)^{2}}{1+\lambda\alpha\left(1-\tau\right)}>0$, and

$$
\frac{\partial\ln A}{\partial\tau}=2\left(\frac{\lambda\alpha}{1-\lambda^{2}\alpha^{2}\left(1-\tau\right)^{2}}-\frac{1-\alpha\left(1-\lambda\right)}{\alpha+\left(1-\alpha\left(1-\lambda\right)\right)\left(1-\tau\right)}\right).
$$

The sufficient condition to have $\frac{\partial\ln A}{\partial\tau}\mid_{\tau=0}<0$ is

$$
\frac{\alpha^{2}\lambda^{2}}{1-\alpha\lambda}<1-\alpha.
$$

The sufficient condition to have $\frac{\partial\Psi}{\partial\tau}>0$ is

$$
\frac{\partial\Psi_1}{\partial\tau}+\frac{\partial\Psi_2}{\partial\tau}+\frac{\partial\Psi_3}{\partial\tau}>0.
$$

By (36), when $\tau=0$, we have

$$
\frac{\partial\Psi}{\partial\tau}=2\frac{\lambda}{\mu^{2}}\left(1-\alpha\right)\left(\left(\lambda+\mu\right)\left(1-\alpha\left(\lambda+\mu\right)\right)+\mu\right)>0. \tag{37}
$$

Then, by (35) and (37), we have $\Psi=0$ and $\frac{\partial\Psi}{\partial\tau}>0$. Thus, by Lemma L5, we can say that the optimal redistributive tax rate $\tau^{*}=0$.

Sub-Case 2: $\delta > 0$ **with** $\lambda = 0$ **or** $\lambda > 0$**.**

(i) If $\lambda = 0$

By (16), we get

$$\begin{aligned}
\Psi_1 &\equiv 1 > 0, \\
\Psi_2 &\equiv -(1-\alpha-\delta)(2-\tau)\tau < 0, \\
\Psi_3 &\equiv -(\alpha+(1-\alpha-\delta)(1-\tau))^2 < 0, \\
\Psi_4 &\equiv -\delta\frac{\omega}{\mu} < 0,
\end{aligned}$$

and then,

$$\begin{aligned}
\Psi &= 1-(1-\alpha-\delta)(2-\tau)\tau-(\alpha+(1-\alpha-\delta)(1-\tau))^2-\delta\frac{\omega}{\mu}, \qquad (38)\\
&= \delta\left(2-\delta-\frac{\omega}{\mu}\right) > 0, \text{ when } \tau = 0 \text{ and } \delta < 2-\frac{\omega}{\mu}.
\end{aligned}$$

Differentiating (38) w.r.t τ gives

$$\begin{aligned}
\frac{\partial\Psi}{\partial\tau} &= 2(1-\alpha-\delta)(\alpha-(\alpha+\delta)(1-\tau)), \qquad (39)\\
&\gtreqqless 0, \text{ if } \delta \lesseqqgtr \frac{\alpha\tau}{1-\tau}, \\
&= -2\delta(1-\alpha-\delta) < 0, \text{ when } \tau = 0.
\end{aligned}$$

By (38) and (39), we can get

$$\begin{aligned}
\Omega = {} & \left(1-(1-\alpha-\delta)(2-\tau)\tau-(\alpha+(1-\alpha-\delta)(1-\tau))^2-\delta\frac{\omega}{\mu}\right)z \qquad (40)\\
& -2((\alpha+\delta)\tau-\delta)(1-\alpha-\delta).
\end{aligned}$$

To have $\Psi > 0$, by (38), the condition, $\delta < 2-\frac{\omega}{\mu}$, needs to be satisfied. In (5), we assume that $\delta < 1-\alpha$. Therefore, we need to discuss the cases when i) $2-\frac{\omega}{\mu} > 1-\alpha$; ii) $2-\frac{\omega}{\mu} < 1-\alpha$. It is summarized in the following Lemma.

First we define $\varepsilon_\tau^\Lambda \equiv \frac{\partial\Lambda}{\Lambda}/\frac{\partial\tau}{\tau}$ and $\varepsilon_\tau^\Psi \equiv \frac{\partial\Psi}{\Psi}/\frac{\partial\tau}{\tau}$ as the elasticity of inequality and Ψ w.r.t τ.

LEMMA 7: *If* $0 < \omega < \mu(1+\alpha)$, *there exists a* $\delta^* \equiv 1-\alpha$ *and a real* Λ^* *such that* $\tau^* > 0$ *if* $\delta < \delta^*$ *and* $\Lambda > \Lambda^*$*; else if* $\omega > \mu(1+\alpha)$, *then there exists a* $\delta^{**} \equiv 2 - \frac{\omega}{\mu}$ *and a real* Λ^* *such that* $\tau^* > 0$ *if* $\delta < \delta^{**}$ *and* $\Lambda > \Lambda^*$, *or if* $\delta^{**} < \delta < 1-\alpha$, $-\varepsilon_\tau^\Lambda < \varepsilon_\tau^\Psi$ *and* $\Lambda > \Lambda^*$*; else if* $\delta = 0$ *or* $\delta^{**} < \delta < 1-\alpha$ *and* $-\varepsilon_\tau^\Lambda > \varepsilon_\tau^\Psi$, *then* $\tau^* = 0$.

Proof: First, we discuss case i), $2 - \frac{\omega}{\mu} > 1-\alpha \Rightarrow 0 < \omega < \mu(1+\alpha)$. By assumption, we know that $\delta < 1-\alpha$. Since $2 - \frac{\omega}{\mu} > 1-\alpha$, by (38), we can get $\Psi > 0$ if $\delta < \delta^* \equiv 1-\alpha$. Thus, at $\tau = 0$, by (38) and (39), we can get $z^* < 0$ if $\delta < \delta^*$. Since $z > 0$, by Lemma L4, we can say that $\Omega > 0$ when $\delta < \delta^*$. If $\omega = \mu$ as in Benabou (2002), then, Ω is always positive. It implies that there exists a real Λ^* such that if $\Lambda > \Lambda^*$ then $\tau^* > 0$.

Next, we discuss case ii), $2 - \frac{\omega}{\mu} < 1-\alpha$, i.e, $\omega > \mu(1+\alpha)$. Then, by (38), we can get $\Psi > 0$ if $\delta < \delta^{**} \equiv 2 - \frac{\omega}{\mu}$. Thus, at $\tau = 0$, by (38) and (39), we can get $z^* < 0$ if $\delta < \delta^{**}$. Since $z > 0$, by Lemma L4, we can say that $\Omega > 0$ when $\delta < \delta^{**}$. It implies that there exists a real Λ^* such that if $\Lambda > \Lambda^*$ then $\tau^* > 0$.

If $\delta^{**} < \delta < 1-\alpha$, by (38), we can get $\Psi < 0$. Then, to have $\Omega > 0$, we need $\Psi z > \frac{\partial\Psi}{\partial\tau}$. Since $\Psi < 0$ and $z = -\frac{\partial\Lambda}{\partial\tau}/\Lambda$, $\Psi z > \frac{\partial\Psi}{\partial\tau} \Rightarrow -\varepsilon_\tau^\Lambda < \varepsilon_\tau^\Psi$ where $\varepsilon_\tau^\Lambda < 0$. Then, we can say that if $\delta^{**} < \delta < 1-\alpha$ and $-\varepsilon_\tau^\Lambda < \varepsilon_\tau^\Psi$, Ω is positive. It implies that there exists a real Λ^* such that if $\Lambda > \Lambda^*$ then $\tau^* > 0$. Otherwise, if $\delta^{**} < \delta < 1-\alpha$ and $-\varepsilon_\tau^\Lambda > \varepsilon_\tau^\Psi$, $\tau^* = 0$.

As shown in sub-case 1 that if $\delta = 0$ then $\tau^* = 0$. Thus, the proof is completed. $\square$

(ii) If $\lambda > 0$

Setting $\tau = 0$, by (16), we can get,

$$\Psi_1 \equiv \left(1 - \lambda^2\alpha^2\right)\left(1 + \frac{\lambda}{\mu}\right)^2 > 0,$$

$$\Psi_2 \equiv 0,$$

$$\Psi_3 \equiv -\frac{(\lambda+\mu)^2 (1-\lambda\alpha)(1-\delta+(\alpha+\delta)\lambda)^2}{(1+\lambda\alpha)\mu^2} < 0,$$

$$\Psi_4 \equiv -\delta\omega\frac{\left((1-\lambda(\alpha+2(1-\alpha-\delta)(1-\lambda)))\left(1-\lambda^2\right) - 2(1-\alpha-\delta)(1-\lambda)\lambda^3\right)}{\mu(1+\lambda\alpha)}.$$

Differentiating (16) w.r.t τ and evaluating it at $\tau = 0$ give

$$\frac{\partial \Psi}{\partial \tau} \mid_{\tau=0} = \frac{\partial \Psi_1}{\partial \tau} \mid_{\tau=0} + \frac{\partial \Psi_2}{\partial \tau} \mid_{\tau=0} + \frac{\partial \Psi_3}{\partial \tau} \mid_{\tau=0} + \frac{\partial \Psi_4}{\partial \tau} \mid_{\tau=0}, \tag{41}$$

where

$$\frac{\partial \Psi_1}{\partial \tau} \mid_{\tau=0} = 2\alpha^2\lambda^2 \left(1 + \frac{\lambda}{\mu}\right)^2 > 0, \tag{42}$$

$$\frac{\partial \Psi_2}{\partial \tau} \mid_{\tau=0} = -2\left(1 - \alpha - \delta\right) < 0, \tag{43}$$

$$\begin{aligned} \frac{\partial \Psi_3}{\partial \tau} \mid_{\tau=0} &= -\frac{\begin{pmatrix} 2\left(\lambda + \mu\right)^2 \left(1 - \delta + \left(\alpha + \delta\right)\lambda\right) \\ * \left(\lambda\alpha^2 + \left(\alpha\lambda\left(1 + \alpha\lambda\right) - 1\right)\left(1 - \delta - \alpha + \left(\alpha + \delta\right)\lambda\right)\right) \end{pmatrix}}{\mu^2\left(1 + \alpha\lambda\right)^2}, \\ &< 0, \text{ if } \alpha\lambda\left(1 + \alpha\lambda\right) > 1, \\ &\geq 0, \text{ if } \alpha\lambda\left(1 + \alpha\lambda\right) \leq 1 - \frac{\lambda\alpha^2}{1 - \delta - \alpha + \left(\alpha + \delta\right)\lambda}, \end{aligned} \tag{44}$$

$$\frac{\partial \Psi_4}{\partial \tau} \mid_{\tau=0} = -\frac{2\omega\lambda\delta}{\mu\left(1 + \alpha\lambda\right)^2} \begin{pmatrix} \lambda\left(1 - \alpha^2\lambda^2\right) + \alpha\left(1 - \lambda^2\right) \\ + \left(1 - \alpha - \delta\right)\left(1 - \lambda\right)\left(2 + \alpha\lambda\right) \end{pmatrix} < 0. \tag{45}$$

Thus, by the definition of Ω and (27), (41) and z from (A.27), and we can get

$$\Omega = \Omega_1 \left(\mu\Omega_2 - \omega\Omega_3\right), \tag{46}$$

where

$$\Omega_1 \equiv \left(\frac{2}{\mu\left(1 - \alpha^2\lambda^2\right)}\right)\left(\frac{1 - \alpha - \delta}{2\left(1 + \alpha\lambda\right) - \left(1 - \lambda\right)\delta}\right) > 0,$$

$$\Omega_2 \equiv \left(1 - \alpha^2\lambda^2\right)\left(2\left(1 + \alpha\lambda\right) - \left(1 - \lambda\right)\delta\right) > 0,$$

$$\Omega_3 \equiv \left(1 - \lambda\right)\left(1 + \alpha\lambda - \alpha^2\lambda^2 - \alpha^3\lambda^3 - \left(1 - \alpha^2\lambda^2 + \left(1 + \alpha\lambda\right)2\alpha\lambda^2 - \left(1 - \lambda\right)\alpha\lambda^2\delta\right)\delta\right).$$

Then, a sufficient condition to have $\Omega > 0$ is $\Omega_3 < 0$ as shown in the following Lemma.

The following Proposition gives the necessary and sufficient condition for $\Omega \lesseqqgtr 0$, i.e., $\mu\Omega_2 \lesseqqgtr \omega\Omega_3$.

Note a lower value of $\omega > 0$ increases the likelihood of Ω to be positive and hence the liklihood of the optimal progressivity to be positive. In particular, we have the following lemma regarding the existence of a strictly positive threshold inequality:

Lemma: If $\omega < \mu(1+\alpha)$; or, if $\omega > \mu(1+\alpha)$ and $\delta < \delta^* = 2 - \omega/\mu$ then there exists $\Lambda^* > 0$.

PROPOSITION 4: *There exists* $\delta_4^* < \delta_5^*$, *if and only if* $\delta \leq \delta_4^*$ or $\delta \geq \delta_5^*$ *then* $\tau^* = 0$, *else* $\tau^* > 0$, *where*

$$\begin{aligned} \delta_4^* &\equiv \frac{1+\alpha\lambda}{2\alpha\lambda^2\omega(1-\lambda)} \left(\begin{array}{c} (1-\alpha\lambda)(\omega-\mu) + 2\alpha\lambda^2\omega \\ -\sqrt{(1-\alpha\lambda)^2(\mu-\omega)^2 + 4\alpha\omega\lambda^2(\lambda(\omega-\alpha\mu)+\mu)} \end{array} \right), \quad (47) \\ &> 0, \text{ if } \omega(1-\lambda) > 2\mu. \end{aligned}$$

$$\delta_5^* \equiv \frac{1+\alpha\lambda}{2\alpha\lambda^2\omega(1-\lambda)} \left(\begin{array}{c} (1-\alpha\lambda)(\omega-\mu) + 2\alpha\lambda^2\omega \\ +\sqrt{(1-\alpha\lambda)^2(\mu-\omega)^2 + 4\alpha\omega\lambda^2(\lambda(\omega-\alpha\mu)+\mu)} \end{array} \right) > 0. \quad (48)$$

Proposition 4 implies that to have $\tau^* > 0$, we need $\delta > \delta_4^*$.

LEMMA 9: Given any δ, there exists a λ^* such that for any $\lambda > \lambda^*$ then $\tau^* = 0$.

Proof: i). By solving (47)$= 0$, we can get λ^*. Moreover, $\delta_4^* < 0$ if $\lambda > \lambda^*$. Any positive δ can be greater than it. It means that $\tau^* > 0$ if $\lambda > \lambda^*$.

ii). Equation (48) shows that as λ increases, δ_5^* can go up even to infinity. That is that there exists a $\lambda^* \to 1$ such that for any $\lambda > \lambda^*$ then $\delta_5^* \to \infty$. Any reasonable δ can not be greater than it. It means that $\tau^* = 0$ if $\lambda > \lambda^*$. $\square$

To eliminate the microeconomic distortion on the saving rate for education expenditure, $s_1(\tau)$, caused by any income based redistribution, following Benabou (2002), we introduce education subsidy in the following sub-section.

6 Redistribution with Education Subsidy

As discussed in Bandyopadhyay and Tang (2011) and sub-case 1 that if there is no externality, the optimal income tax rate is zero. If externality is included, then, the degree of optimal progressivity can be bigger if we consider education subsidy following Benabou (2002).

One may also reasonably wonder if the microeconomic disincentive of redistribution on the investment of human capital drives the above result. In fact, by Lemma 1, we do allow such negative effect of redistribution. However, contrary to what we find in Maoz and Moav (1999) and others, the negative effect of redistribution on the parental investment in the child's human capital does not drive the result described in Lemma 7. To make this specific point clear as well as to avoid the inessential details discussed above, we now expand the policy package for redistribution to include an education subsidy which offsets completely the above negative effect of redistribution described in Lemma 1 and thereby switch off its negative effect on growth via the microeconomic channel of individual optimization.

Suppose to offset the negative effect of redistribution on the optimal rate of parental investment in education, the government distributes d units of subsidy to each parent i per unit of the date t parental investment e_t^i in education and finances all of it with a non-distortionary tax, at a rate $\theta \in [0,1]$, on consumption c_t^i such that the post-subsidy expenditure on education $\hat{e}_t^i = (1+d)\, e_t^i$ and the government's choice of tax and subsidy satisfies

$$\theta \int_0^1 c_t^i di = d \int_0^1 e_t^i di. \tag{49}$$

The new date t budget constraint for a parent i becomes,

$$\hat{y}_t^i = (1+\theta)\, c_t^i + e_t^i + b_t^i. \tag{50}$$

The modified human capital production function becomes,

$$h_{t+1}^i = \kappa_t \xi_{t+1}^i \left(h_t^i\right)^{\alpha} \left(\hat{e}_t^i\right)^{\beta}. \tag{51}$$

Next, for any given progressivity rate τ, to switch off completely the negative impact of redistribution on the rate $s_1(\tau)$ of parental investment in education, we require that the subsidy rate d must satisfy

$$(1+d)\, s_1(\tau) = s_1(0), \tag{52}$$

where $s_1(\tau)$ continues to be given by (54) and $s_1(0)$ denotes its value at the laissez-faire state corresponding to zero progressivity.

Thus the government ensures that with appropriate education subsidy, a redistributive policy does not distort the parental decision rule for the child's education.

7 Knowledge Spillover and Optimal Progressivity

In this section, we numerically show some optimal tax rate when education subsidy is and is not included. The benchmark parameters are from Barro, Mankiw and Sala-i-Martin (1995) and Benabou (2002), For example, the share of physical and human capital are Barro, Mankiw and Sala-i-Martin (1995), $\lambda = 0.3$, $\mu = 0.5$. The other parameters are from Benabou (2002), such as $\alpha = 0.35$, $\rho = 0.4$, $\eta = 6$, $\sigma^2 = 1$.

Specifically, in this section, we show how the optimal redistributive income tax rate changes with δ and ω. By estimating the optimal tax rate, we use (17), (18), (19), (55) and (21).

The following Figure shows the optimal redistributive income tax rate when education subsidy is included.

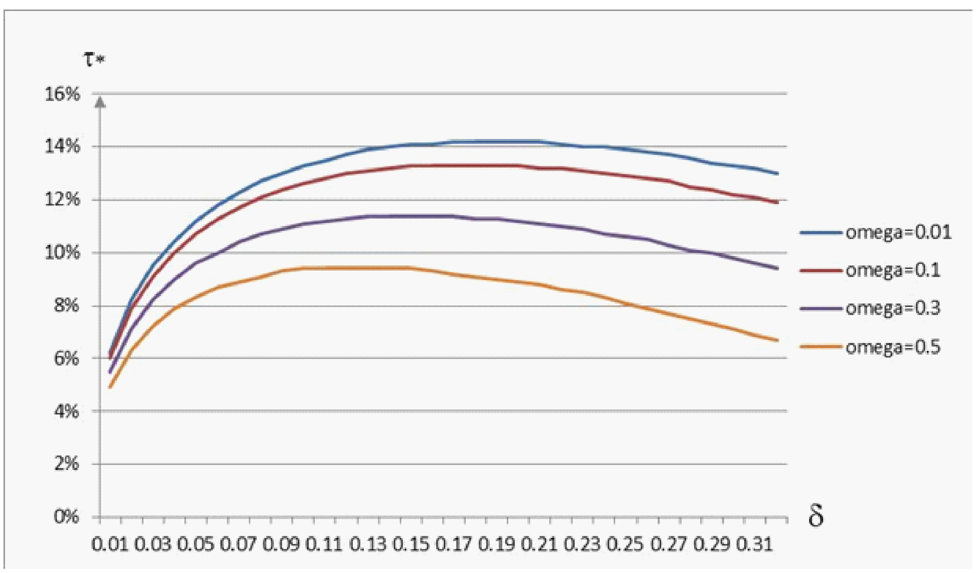

Figure 1: optimal redistributive income tax rate, δ and ω when education subsidy is included.

The following Figure shows the optimal redistributive income tax rate when education subsidy is not included.

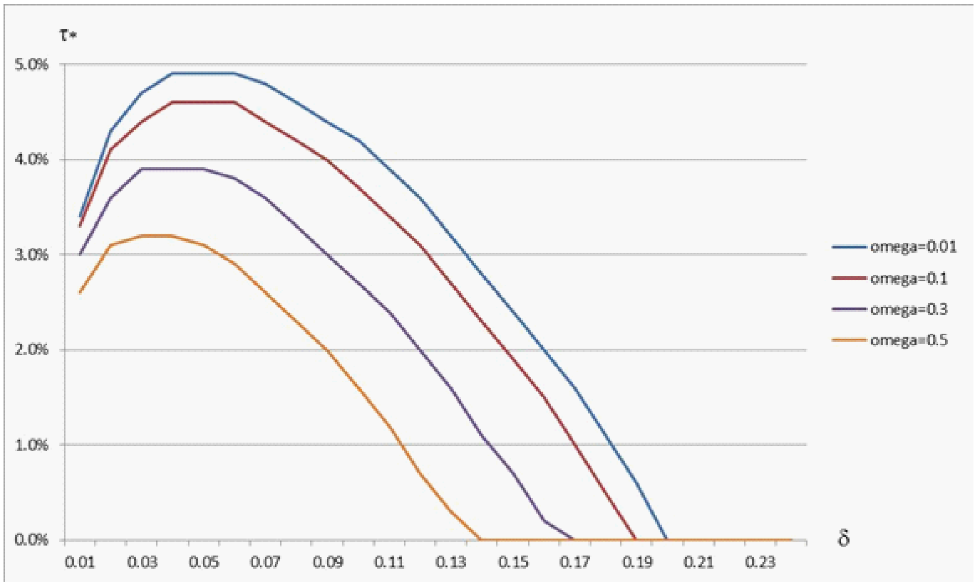

Figure 2: optimal redistributive income tax rate, δ and ω when education subsidy is not included.

8 Concluding Remarks

In this paper, we first present a theoretical model and show that there exists a unique upper bound or a threshold for efficient income inequality such that if the laissez-faire income-inequality exceeds this threshold then positive income redistribution would be consistent with a welfare-maximizing policy, irrespective of any equity considerations. We prove analytically that the degree of external spillover of knowledge plays a critical role in determining the effectiveness of a policy of progressive income redistribution in uplifting the growth rate of per capita income. We characterize the properties of the optimal progressivity as a function of knowledge spillover and other parameters of the model. Examining their non-linear relationship that emerges in our model, we discover that while the average marginal income tax rate may be as large as 20%, too little or too much spillover of knowledge would make any income redistribution bad for economic growth. Our discovery calls for future research on estimating the extent of knowledge spillover for determining the growth-maximising progressivity of income distribution. Nevertheless, our findings confirm that too much inequality can kill economic growth. Therefore, to sustain perpetual economic growth, the government must maintain a strong middle class with suitable redistribution of income. We showed with an analytical proof that the choice between growth and equity may sometimes be a false choice. Because inequality retards growth and we need some redistribution of income to prevent inequality from becoming too high to sustain economic growth.

Appendix

LEMMA 1: *The optimal labor supply which remains invariant to time and personal characteristics and decreases with the average marginal income tax rate τ such that:*

$$l_t^i = \left(\frac{(\epsilon/\eta)(1-\rho\alpha)(1-\tau)}{(1-\rho\alpha)(1-\rho\lambda(1-\tau)) - \rho\beta\mu(1-\tau)} \right)^{1/\eta} \equiv l(\tau), \, l'(\tau) < 0, \tag{53}$$

and the optimal saving rate $s_{1t}^i \equiv e_t^i / \hat{y}_t^i$ for investment in education and optimal saving rate for bequest, $s_{2t}^i \equiv b_t^i / \hat{y}_t^i$ are time invariant constant and decreases with the average marginal income tax rate τ:

$$s_{1t}^i = \frac{\rho\beta\mu(1-\tau)}{1-\rho\alpha} \equiv s_1(\tau), \, s_1'(\tau) < 0, \tag{54}$$

$$s_{2t}^i = \rho\lambda(1-\tau) \equiv s_2(\tau), \, s_2'(\tau) < 0. \tag{55}$$

Lemma 1 explicitly spells out the typical disincentives or negative effects, measured by $s'(\tau)$, of changing the rate τ of progressivity of redistribution on the optimal saving rate $s(\tau)$ for the child's education. Also, as expected, the rate of parental saving for the child's education increases with the thriftiness parameter ρ and with the parameters α and β, which indicates, respectively, the quality of parental nurturing and educational institutions.

PROOF OF LEMMA 1: By (3), (4) and (6), we rewrite (9) as follows:

$$\ln U\left(h_t^i, k_t^i, M_t; T\right) = \max_{s_{1t}^i, s_{2t}^i, l_t^i} \left\{ \begin{array}{c} (1-\rho)\left[\begin{array}{c} \ln\left(1 - s_{1t}^i - s_{2t}^i\right) \\ + \ln \hat{y}_t^i - \left(l_t^i\right)^\eta \end{array} \right] \\ + \rho E_t \left[\ln U\left(h_{t+1}^i, k_{t+1}^i, M_{t+1}; T\right) \right] \end{array} \right\}. \tag{A.1}$$

Agent solves (A.1) subject to (2) and (7) and

$$h_{t+1}^{i}=\kappa_t\left(s_{1t}^{i}\right)^{\beta}\xi_{t+1}^{i}\left(k_t^{i}\right)^{\beta\lambda(1-\tau)}\left(h_t^{i}\right)^{\alpha+\beta\mu(1-\tau)}\left(l_t^{i}\right)^{\beta\epsilon(1-\tau)}\left(\tilde{y}_t\right)^{\beta\tau}\text{, and} \tag{A.2}$$

$$k_{t+1}^{i}=s_{2t}^{i}\left(k_t^{i}\right)^{\lambda(1-\tau)}\left(h_t^{i}\right)^{\mu(1-\tau)}\left(l_t^{i}\right)^{\epsilon(1-\tau)}\left(\tilde{y}_t\right)^{\tau}. \tag{A.3}$$

We guess the value function as: $\ln U\left(h_t^i,k_t^i,M_t;T\right)=Z_1\ln h_t^i+Z_2\ln k_t^i+B_t$. Then by substituting this value function into (A.1), we get

$$\begin{aligned}
Z_1\ln h_t^i+Z_2\ln k_t^i+B_t= \; & (1-\rho)\begin{pmatrix}\ln\left(1-s_{1t}^{i}-s_{2t}^{i}\right)\\ +\epsilon\left(1-\tau\right)\ln l_t^i+\tau\ln\tilde{y}_t-\left(l_t^i\right)^{\eta}\end{pmatrix} \\
& +\left(1-\rho+\rho\beta Z_1+\rho Z_2\right)\lambda\left(1-\tau\right)\ln k_t^i \\
& +\left(\left(1-\rho+\rho\beta Z_1+\rho Z_2\right)\mu\left(1-\tau\right)+\rho\alpha Z_1\right)\ln h_t^i \\
& +\rho\begin{pmatrix} Z_1\begin{pmatrix}\ln\kappa_t+\beta\ln s_{1t}^i-\sigma^2/2\\ \\ +\beta\epsilon\left(1-\tau\right)\ln l_t^i+\beta\tau\ln\tilde{y}_t\end{pmatrix}\\ \\ +Z_2\begin{pmatrix}\ln s_{2t}^i+\epsilon\left(1-\tau\right)\ln l_t^i\\ \\ +\tau\ln\tilde{y}_t\end{pmatrix}\\ +B_{t+1}\end{pmatrix}.
\end{aligned} \tag{A.4}$$

Taking partial differentials with respect to $\ln k_t^i$ and $\ln h_t^i$ yields

$$Z_1=\left(1-\rho+\rho\beta Z_1+\rho Z_2\right)\mu\left(1-\tau\right)+\rho\alpha Z_1, \tag{A.5}$$

$$Z_2=\left(1-\rho+\rho\beta Z_1+\rho Z_2\right)\lambda\left(1-\tau\right). \tag{A.6}$$

Rearranging equations (A.5) and (A.6) yields

$$Z_1=\frac{(1-\rho)\mu\left(1-\tau\right)}{\left(1-\rho\alpha\right)\left(1-\rho\lambda\left(1-\tau\right)\right)-\rho\beta\mu\left(1-\tau\right)}, \tag{A.7}$$

$$Z_2 = \frac{(1-\rho\alpha)(1-\rho)\lambda(1-\tau)}{(1-\rho\alpha)(1-\rho\lambda(1-\tau)) - \rho\beta\mu(1-\tau)}. \tag{A.8}$$

The values of human and physical capital, as expressed by their utility elasticities, are given by Z_1 and Z_2, respectively. Note that the tax rate τ can alter these values individually, but does not alter the relative value of human to physical capital, $\frac{\mu}{\lambda(1-\rho\alpha)}$, which increases with the output elasticity of human capital μ, the neighborhood effect α and patience ρ, but remains unaffected by the quality of education β. We can then verify the guess and confirm the existence of (A.4).

The first-order conditions of (A.1) with respect to the saving rates and labor supply are

$$\frac{1-\rho}{1-s_{1t}^i - s_{2t}^i} = \rho\left(\frac{\partial \ln U_{t+1}^i}{\partial \ln h_{t+1}^i}\frac{\partial \ln h_{t+1}^i}{\partial s_{1t}^i} + \frac{\partial \ln U_{t+1}^i}{\partial \ln k_{t+1}^i}\frac{\partial \ln k_{t+1}^i}{\partial s_{1t}^i}\right), \tag{A.9}$$

$$\frac{1-\rho}{1-s_{1t}^i - s_{2t}^i} = \rho\left(\frac{\partial \ln U_{t+1}^i}{\partial \ln h_{t+1}^i}\frac{\partial \ln h_{t+1}^i}{\partial s_{2t}^i} + \frac{\partial \ln U_{t+1}^i}{\partial \ln k_{t+1}^i}\frac{\partial \ln k_{t+1}^i}{\partial s_{2t}^i}\right), \tag{A.10}$$

$$\begin{aligned}(1-\rho)\,\eta\left(l_t^i\right)^{\eta-1} = {} & (1-\rho)\,\epsilon(1-\tau)/l_t^i \\ & + \rho\left(\frac{\partial \ln U_{t+1}^i}{\partial \ln h_{t+1}^i}\frac{\partial \ln h_{t+1}^i}{\partial l_t^i} + \frac{\partial \ln U_{t+1}^i}{\partial \ln k_{t+1}^i}\frac{\partial \ln k_{t+1}^i}{\partial l_t^i}\right),\end{aligned} \tag{A.11}$$

where $\partial \ln k_{t+1}^i/\partial s_{1t}^i = 0$, $\partial \ln k_{t+1}^i/\partial s_{2t}^i = 1/s_{2t}^i$, $\partial \ln h_{t+1}^i/\partial s_{1t}^i = \beta/s_{1t}^i$, $\partial \ln h_{t+1}^i/\partial s_{2t}^i = 0$, $\partial \ln k_{t+1}^i/\partial l_t^i = \epsilon(1-\tau)/l_t^i$ and $\partial \ln h_{t+1}^i/\partial l_t^i = \beta\epsilon(1-\tau)/l_t^i$.

The above optimization problem (A.4) is strictly concave. Consequently, (A.9)-(A.11) are sufficient for the optimization exercise. After substituting (A.7) and (A.8) into (A.9)-(A.11), we get (53)-(55).

PROOF OF LEMMA 2: The logarithm of (6), combining with (53) and (55) yields the dynamics of physical capital for the dynasty i,

$$\begin{aligned}\ln k_{t+1}^i = {} & \ln s_2(\tau) + \epsilon(1-\tau)\ln l(\tau) + \lambda(1-\tau)\ln k_t^i \\ & + \mu(1-\tau)\ln h_t^i + \tau \ln \tilde{y}_t.\end{aligned} \tag{56}$$

The logarithm of (4), combining with (53) and (54) yields the dynamics of human capital for the dynasty i,

$$\begin{aligned}\ln h_{t+1}^{i} &= \ln \kappa_t + \beta \ln s_1(\tau) + \beta\epsilon(1-\tau)\ln l(\tau) + \ln \xi_{t+1}^{i} \\ &\quad + \beta\lambda(1-\tau)\ln k_t^i + (\alpha + \beta\mu(1-\tau))\ln h_t^i + \beta\tau \ln \tilde{y}_t. \end{aligned} \tag{57}$$

Substituting (56) and (57) into the logarithm of (2) yields the equilibrium path of income for agent i such that

$$\begin{aligned}\ln y_{t+1}^{i} &= \psi(\tau) + \mu \ln \kappa_t + \mu \ln \xi_{t+1}^{i} \\ &\quad + (\lambda + \beta\mu)\tau \ln \tilde{y}_t - \alpha\lambda\tau \ln \tilde{y}_{t-1} \\ &\quad + (\alpha + (\lambda+\beta\mu)(1-\tau))\ln y_t^i - \alpha\lambda(1-\tau)\ln y_{t-1}^i, \end{aligned} \tag{58}$$

where $\psi(\tau) \equiv \mu\beta \ln s_1(\tau) + \lambda(1-\alpha)\ln s_2(\tau) + (1-\alpha)\epsilon \ln l(\tau)$ is a time-invariant constant, and $\tilde{y}_t$ is given by (11). Denote the intergenerational persistence of income, $p(\tau) \equiv \alpha + (\lambda+\beta\mu)(1-\tau)$, which decreases with the rate τ of progressivity, and hence a policy of redistribution enhances intergenerational social mobility.

Given the joint initial distribution of human and physical capital, by (56) and (57), physical and human capital remain jointly lognormally distributed over time, such that at each date t, M_t satisfies

$$\begin{aligned} m_{kt+1} &= \ln s_2(\tau) + \epsilon \ln l(\tau) + \lambda m_{kt} + \mu m_{ht} \\ &\quad + \tau(2-\tau)\left(\lambda^2\Delta_{kt}^2 + \mu^2\Delta_{ht}^2 + 2\lambda\mu cov_t\right)/2, \end{aligned} \tag{59}$$

$$\Delta_{kt+1}^2 = (1-\tau)^2\left(\lambda^2\Delta_{kt}^2 + \mu^2\Delta_{ht}^2 + 2\lambda\mu cov_t\right), \tag{60}$$

$$\begin{aligned} m_{ht+1} &= \ln\kappa - \sigma^2/2 + \beta \ln s_1(\tau) + \beta\epsilon \ln l(\tau) + \beta\lambda m_{kt} \\ &\quad + (\alpha + \beta\mu + \delta)m_{ht} + \delta\omega\Delta_{ht}^2/2 \\ &\quad + \beta\tau(2-\tau)\left(\lambda^2\Delta_{kt}^2 + \mu^2\Delta_{ht}^2 + 2\lambda\mu cov_t\right)/2, \end{aligned} \tag{61}$$

$$
\begin{aligned}
\Delta_{ht+1}^{2} = & \, \sigma^{2}+\beta^{2}\lambda^{2}\left(1-\tau\right)^{2}\Delta_{kt}^{2}+\left(\alpha+\beta\mu\left(1-\tau\right)\right)^{2}\Delta_{ht}^{2} \\
& +2\beta\lambda\left(1-\tau\right)\left(\alpha+\beta\mu\left(1-\tau\right)\right)cov_{t},
\end{aligned}
\tag{62}
$$

$$
\begin{aligned}
cov_{t+1} = & \, \beta\lambda^{2}\left(1-\tau\right)^{2}\Delta_{kt}^{2}+\mu\left(1-\tau\right)\left(\alpha+\beta\mu\left(1-\tau\right)\right)\Delta_{ht}^{2} \\
& +\lambda\left(1-\tau\right)\left(\alpha+2\beta\mu\left(1-\tau\right)\right)cov_{t}.
\end{aligned}
\tag{63}
$$

Given M_0, (59)-(63) yield a unique sequence $\{M_t\}_{t=1,2,..\infty}$ that characterizes the key equilibrium dynamics of our model.

Aggregating the equilibrium path of all agents' income, given by (58), and combining with (11), we get the dynamic equation of per capita income as below,

$$
\begin{aligned}
\ln y_{t+1}-\ln y_{t} = & \, \psi\left(\tau\right)+\mu\ln\kappa-\mu\sigma^{2}/2-\delta\epsilon\ln l\left(\tau\right)-\delta\lambda\ln s_{2}\left(\tau\right) \\
& -\left(\left(1-\lambda\right)\left(1-\alpha-\delta\right)-\beta\mu\right)\ln y_{t}+\left(\alpha+\delta\right)\lambda\left(\ln y_{t}-\ln y_{t-1}\right) \\
& +\Lambda_{t+1}+\left(\left(\lambda+\beta\mu\right)\tau\left(1-\tau\right)-\delta-p\left(\tau\right)\right)\Lambda_{t}+\alpha\lambda\left(1-\tau\right)^{2}\Lambda_{t-1} \\
& +\delta\left(\lambda\Delta_{kt}^{2}+\mu\omega\Delta_{ht}^{2}\right)/2.
\end{aligned}
\tag{64}
$$

The above growth-inequality relationship implies that an economy forgoes its per capita income growth in proportion to the size of its existing income inequality, partially reflecting the forgone TFP due to interpersonal differences in the marginal product of human capital implied by the existing inequality. At the same time, future inequality impacts positively on the per capita income growth reflecting its positive influence in the growth rate of average human capital due to the presence of increasing returns in the education technology.

By assumption, at the initial date $t=0$, the physical and human capitals are lognormally distributed. By (56) and (57), it follows that k_t^i and h_t^i remain lognormally distributed over time, and hence, by (2), y_t^i is lognormal and is given by,

$$
\ln y_{t}^{i}=\lambda\ln k_{t}^{i}+\mu\ln h_{t}^{i}+\epsilon\ln l. \tag{A.12}
$$

By (53), it follows that the mean of the lognormal distribution of y_t^i is given by,

$$\int_0^1 \ln y_t^i di = \lambda m_{kt} + \mu m_{ht} + \epsilon \ln l. \tag{A.13}$$

The variance of $\ln y_t^i$ is the sum of the variances of $\ln k_t^i$ and $\ln h_t^i$ plus the covariance of these two variables,

$$\text{var}\left[\ln y_t^i\right] = \lambda^2 \Delta_{kt}^2 + \mu^2 \Delta_{ht}^2 + 2\lambda\mu cov_t. \tag{A.14}$$

The income per capita y_t is

$$y_t = \int_0^1 y_t^i di = \exp\left(\int_0^1 \ln y_t^i di + \frac{1}{2}\text{var}\left[\ln y_t^i\right]\right). \tag{A.15}$$

The median income is

$$y_{t,median} = \exp\left(\int_0^1 \ln y_t^i di\right). \tag{A.16}$$

In line with Benabou (2002), therefore, the inequality index is

$$\Lambda_t \equiv \log\left(\frac{y_t}{y_{t,median}}\right) = \frac{1}{2}\text{var}\left[\ln y_t^i\right] = \left(\lambda^2 \Delta_{kt}^2 + \mu^2 \Delta_{ht}^2 + 2\lambda\mu cov_t\right)/2. \tag{A.17}$$

The break-even income is defined in (7). We note that the mean of y_t^i in logarithm, according to equation (A.15), satisfies

$$\ln y_t = \lambda m_{kt} + \mu m_{ht} + \epsilon \ln l + \Lambda_t, \tag{A.18}$$

and the mean of $\left(y_t^i\right)^{1-\tau}$ in logarithm satisfies

$$\ln \int_0^1 \left(y_t^i\right)^{1-\tau} di = (1-\tau)\left(\lambda m_{kt} + \mu m_{ht} + \epsilon \ln l\right) + (1-\tau)^2 \Lambda_t. \tag{A.19}$$

Taking the difference between the income before and after tax yields

$$\ln \tilde{y}_t = \lambda m_{kt} + \mu m_{ht} + \epsilon \ln l + (2-\tau)\Lambda_t. \tag{A.20}$$

We know that the physical and human capital are distributed lognormally. Then, from the property of the moment generating function for lognormal distribution, we get

$$\ln k = m_k + \Delta_k^2/2 \text{ and } \ln h = m_h + \Delta_h^2/2. \tag{A.21}$$

Then substituting (A.21) into (A.20) yields (11).

By (5) and (A.21), we can get

$$\begin{aligned} \ln \kappa_t &= \ln \kappa + \delta \left(\begin{array}{l} \frac{1}{\mu}\left(\ln y_t - \epsilon \ln l - \left(\Lambda_t - \left(\lambda \Delta_{kt}^2 + \mu \Delta_{ht}^2\right)/2\right)\right) \\ + \left(\omega - 1\right) \Delta_{ht}^2/2 \end{array} \right) \\ & -\frac{\delta \lambda}{\mu}\left(\ln s_2 + \ln y_{t-1}\right). \end{aligned} \tag{A.22}$$

Thus, the proof of Lemma 2 is completed. □

PROOF OF LEMMA 3: Writing the system of linear equations (60), (62) and (63) in a matrix form, we get

$$\Delta_{t+1} = A_0 + A_1 * \Delta_t, \tag{A.23}$$

where

$$\Delta_{t+1} \equiv \begin{bmatrix} \Delta_{kt+1}^2 \\ \Delta_{ht+1}^2 \\ \mathrm{cov}_{t+1} \end{bmatrix}, A_0 \equiv \begin{bmatrix} 0 \\ \sigma^2 \\ 0 \end{bmatrix},$$

$$A_1 \equiv \begin{bmatrix} \lambda^2 \left(1-\tau\right)^2 & \mu^2 \left(1-\tau\right)^2 & 2\lambda\mu \left(1-\tau\right)^2 \\ \beta^2 \lambda^2 \left(1-\tau\right)^2 & \left(\alpha + \beta\mu \left(1-\tau\right)\right)^2 & 2\beta\lambda \left(1-\tau\right) \left(\alpha + \beta\mu \left(1-\tau\right)\right) \\ \beta\lambda^2 \left(1-\tau\right)^2 & \mu \left(1-\tau\right)\left(\alpha + \beta\mu \left(1-\tau\right)\right) & \lambda \left(1-\tau\right)\left(\alpha + 2\beta\mu \left(1-\tau\right)\right) \end{bmatrix},$$

or,

$$A_1 \equiv \begin{bmatrix} x^2 & y^2 & 2xy \\ \beta^2 x^2 & z^2 & 2\beta xz \\ \beta x^2 & yz & x\left(z + \beta y\right) \end{bmatrix},$$

where $x \equiv \lambda\left(1-\tau\right)$, $y \equiv \mu\left(1-\tau\right)$, $z \equiv \alpha + \beta\mu\left(1-\tau\right)$.

Then, rearranging (A.23) yields

$$\left(I - A_1 L\right) \Delta_{t+1} = A_0, \tag{A.24}$$

where *I* is an identity matrix and *L* is a lag operator.

All eigenvalues of A_1 in (A.23) are positive real and less than unity. Consequently, $\{\Delta_t\}$ monotonically converges to Δ,[6] where

$$\Delta = [I - A_1]^{-1} A_0. \tag{A.25}$$

By Lemma 2, it follows, therefore, that income inequality $\Lambda_t = w\Delta_t$, where, $w \equiv 0.5\left(\lambda^2,\ \mu^2,\ 2\lambda\mu\right)$, monotonically converges to $\Lambda = w\Delta$. Thus by setting $\Delta_{t+1} = \Delta_t = \Delta$, we can get (12). Taking logarithm on (12) and differentiating it yield:

$$\frac{\partial\Lambda}{\partial\tau} = -z\Lambda, \tag{A.26}$$

where

$$z \equiv \left(\begin{array}{c} \dfrac{2\alpha\lambda}{\left(1+\alpha\lambda\left(1-\tau\right)\right)\left(1-\alpha\lambda\left(1-\tau\right)\right)} \\ \\ +\dfrac{\left(-2\alpha\lambda\left(1+\alpha\lambda\left(1-\tau\right)\right)+2\left(\alpha+\left(1-\tau\right)\left(\lambda+\beta\mu\right)\right)\left(\lambda+\beta\mu\right)\right)}{\left(1+\alpha\lambda\left(1-\tau\right)\right)^2-\left(\alpha+\left(1-\tau\right)\left(\lambda+\beta\mu\right)\right)^2} \end{array} \right). \tag{A.27}$$

Then, the proof of Lemma 3 is completed. □

PROOF OF LEMMA 4a: By (24), we define

$$LHS \equiv -\left(1-\lambda\right)\ln s_1 - \lambda \ln s_2 - \epsilon \ln l, \tag{65}$$

and

$$RHS \equiv 1-\lambda+\frac{1}{\eta}\frac{\rho\left(1-\alpha-\delta\right)\epsilon\left(1-\lambda\right)}{\left(1-\rho\alpha\right)\left(1-\rho\lambda\right)-\rho\left(1-\alpha-\delta\right)\left(1-\lambda\right)}. \tag{66}$$

Then, from equation (24), we can see that to have $\frac{\partial\Phi}{\partial\delta} > 0$ we need to have $LHS\left(\delta=0\right) > RHS\left(\delta=0\right)$ since $\partial RHS\left(\delta\right)/\partial\delta < 0$. Then, evaluating (65) and (66) at $\delta = 0$, we get

$$\begin{aligned} LHS\left(\delta=0\right) &= -\left(1-\lambda\right)\ln\frac{\rho\left(1-\alpha\right)\left(1-\lambda\right)}{1-\rho\alpha} - \lambda\ln\rho\lambda \\ &\quad -\frac{\epsilon}{\eta}\ln\left(\frac{\left(\epsilon/\eta\right)\left(1-\rho\alpha\right)}{\left(1-\rho\alpha\right)\left(1-\rho\lambda\right)-\rho\left(1-\alpha\right)\left(1-\lambda\right)}\right), \end{aligned} \tag{67}$$

[6] See Galor (2007) for details about discrete dynamic convergence.

$$RHS\left(\delta=0\right)\equiv 1-\lambda+\frac{\epsilon}{\eta}\frac{\rho\left(1-\alpha\right)\left(1-\lambda\right)}{\left(1-\rho\alpha\right)\left(1-\rho\lambda\right)-\rho\left(1-\alpha\right)\left(1-\lambda\right)}. \tag{68}$$

It is not clear to tell whether $LHS\left(\delta=0\right)$ is greater or smaller than $RHS\left(\delta=0\right)$. Since both of them is a function of λ, we can check how they change w.r.t λ and then discuss which one is bigger starting from $\lambda=0$. Differentiating (67) and (68) w.r.t λ gives

$$\begin{aligned}\frac{\partial}{\partial\lambda}LHS\left(\delta=0\right) &= \ln\frac{\left(1-\alpha\right)\left(1-\lambda\right)}{1-\rho\alpha}-\ln\lambda \\ &\quad -\frac{\epsilon}{\eta}\frac{\alpha\rho\left(1-\rho\right)}{\left(1-\rho\alpha\right)\left(1-\rho\lambda\right)-\rho\left(1-\alpha\right)\left(1-\lambda\right)},\end{aligned} \tag{69}$$

$$\frac{\partial}{\partial\lambda}RHS\left(\delta=0\right)=-1-\frac{\epsilon}{\eta}\frac{\rho}{\left(1-\alpha\lambda\rho\right)^{2}}\frac{1-\alpha}{1-\rho}\left(1-\alpha\rho\right)<0. \tag{70}$$

From (69), it can be seen that as $\lambda\rightarrow 0$, $-\ln\lambda\rightarrow+\infty$ while as $\lambda\rightarrow 1$, $\ln\frac{(1-\alpha)(1-\lambda)}{1-\rho\alpha}\rightarrow-\infty$. Then, we can say that $\frac{\partial}{\partial\lambda}LHS\left(\delta=0\right)$ is positive for small λ but negative for big λ such that there exists a $0<\lambda^{*}<1$ such that if $\lambda\leq\lambda^{*}$, $\frac{\partial}{\partial\lambda}LHS\left(\delta=0\right)\geq 0$. Equation (70) shows that $RHS\left(\delta=0\right)$ always goes down with λ.

Now, we evaluate $LHS\left(\delta=0\right)$ and $RHS\left(\delta=0\right)$ at $\lambda=0$ and get

$$LHS\left(\delta=0,\lambda=0\right)=-\ln\frac{\rho\left(1-\alpha\right)}{1-\rho\alpha}-\frac{\epsilon}{\eta}\ln\left(\frac{\left(\epsilon/\eta\right)\left(1-\rho\alpha\right)}{1-\rho}\right), \tag{71}$$

$$RHS\left(\delta=0,\lambda=0\right)\equiv 1+\frac{\epsilon}{\eta}\frac{\rho\left(1-\alpha\right)}{1-\rho}. \tag{72}$$

By (71), we can see that $\frac{\partial}{\partial\alpha}LHS\left(\delta=0,\lambda=0\right)>0$ and $LHS\left(\delta=0,\lambda=0\right)\rightarrow\infty$ as $\alpha\rightarrow 1$. In contrast, $RHS\left(\delta=0,\lambda=0\right)$ goes down with α. If $LHS\left(\delta=0,\lambda=0,\alpha=0\right)$ is less than $RHS\left(\delta=0,\lambda=0,\alpha=0\right)$, then we can find a $0<\alpha^{*}<1$ such that if $\alpha>\alpha^{*}$, $LHS\left(\delta=0,\lambda=0\right)>RHS\left(\delta=0,\lambda=0\right)$. Otherwise, if $LHS\left(\delta=0,\lambda=0,\alpha=0\right)$ is greater than $RHS\left(\delta=0,\lambda=0,\alpha=0\right)$, then $\alpha^{*}<0$ for $LHS\left(\delta=0,\lambda=0\right)=RHS\left(\delta=0,\lambda=0\right)$. Then any positive α is greater than it.

By $RHS\left(\delta=0,\lambda=0,\alpha=0\right)<RHS\left(\delta=0,\lambda=0,\alpha=0\right)$, we can find that there exists a $0<\rho^{*}<1$ such that if $\rho<\rho^{*}$, then $\alpha^{*}<0$, otherwise $0<\alpha^{*}<1$.

Thus, we can conclude that if $\rho<\rho^{*}$ and $\lambda\leq\lambda^{*}$ or $1>\rho>\rho^{*}$, $\alpha>\alpha^{*}$ and $\lambda\leq\lambda^{*}$, then $LHS\left(\delta=0\right)>RHS\left(\delta=0\right)$.

Next, we discuss how LHS and RHS change with δ. As $\delta\rightarrow 1-\alpha$, (65) and (66) show that

$$
\begin{aligned}
LHS\left(\delta=1-\alpha\right) &= -\left(1-\lambda\right)\ln 0-\lambda\ln\rho\lambda-\frac{\epsilon}{\eta}\ln\frac{\epsilon/\eta}{\left(1-\rho\lambda\right)} \\
&\rightarrow +\infty,
\end{aligned}
\tag{73}
$$

$$
RHS\left(\delta=1-\alpha\right)\equiv 1-\lambda. \tag{74}
$$

Since $\partial RHS\left(\delta\right)/\partial\delta<0$, we can say that if $\rho<\rho^{*}\in\left(0,1\right)$ and $\lambda\leq\lambda^{*}\in\left(0,1\right)$ or $1>\rho>\rho^{*}$, $\alpha>\alpha^{*}\in\left(0,1\right)$ and $\lambda\leq\lambda^{*}$, then, $LHS\left(\delta\right)>RHS\left(\delta\right)$ for any δ. $\square$

PROOF OF LEMMA 4b: By setting (25)$=0$, we can get δ_1^{*}. $\delta_1^{*}>1-\alpha$ is equivalent to

$$
\omega<\frac{2}{\mu}\frac{\left(1-\alpha\lambda\right)\left(\alpha+\lambda\right)\left(\lambda+\mu\right)^{2}}{3\lambda\left(1-\alpha\right)+1-\alpha\lambda^{2}}\equiv\omega_1^{*}. \tag{75}
$$

When $\lambda=0$, we can get

$$
\omega_1^{*}\left(\lambda=0\right)=2\alpha\mu>0. \tag{76}
$$

Thus, if $\alpha\mu<0.5$, then $\omega_1^{*}\left(\lambda=0\right)<1$.

When $\lambda\rightarrow1$, we can get

$$
\omega_1^{*}\left(\lambda\rightarrow1\right)=\frac{\left(\alpha+1\right)\left(1+\mu\right)^{2}}{2\mu}>1 \tag{77}
$$

By (76) and (77), we can say that if $\alpha\mu<0.5$, then there exists a λ^{*} such that if $\lambda<\lambda^{*}$ then $\omega_1^{*}<1$. Hence, by (75), we can get that if $\omega>\omega_1^{*}$, $\delta_1^{*}<1-\alpha$. Thus, by (25), the Lemma is proved. $\square$

PROOF OF LEMMA 4c: By setting (27)=0, we can get δ_2^{*}. And $\delta_2^{*}<1-\alpha$ is equivalent to

$$
\left(1+\alpha\lambda\right)\frac{2\left(1-\alpha\lambda\right)\left(1+\frac{\lambda}{\mu}\right)^{2}-\dfrac{\omega}{\mu}\left(1-\lambda\right)}{2\lambda\frac{\omega}{\mu}+\left(1-\lambda\right)\left(1-\alpha\lambda\right)\left(1+\frac{\lambda}{\mu}\right)^{2}}<1-\alpha
$$

$\Leftrightarrow$

$$
\omega>\left(2\left(1+\alpha\lambda\right)-\left(1-\alpha\right)\left(1-\lambda\right)\right)\frac{\mu}{1+\lambda}\left(1+\frac{\lambda}{\mu}\right)^{2}\equiv\omega_2^{*}
$$

When $\lambda=0$, we have $\omega_2^{*}\left(\lambda=0\right)=\left(1+\alpha\right)\mu$. It means that if $\alpha\mu<1-\mu$ then $\omega_2^{*}\left(\lambda=0\right)<1$. When $\lambda\rightarrow1$, we have $\omega_2^{*}\left(\lambda\rightarrow1\right)=\left(1+\alpha\right)\left(\mu+\frac{1}{\mu}+2\right)>1$. Thus, we can say that if $\alpha\mu < 1 - \mu$ then there exists a $\lambda^{**} \in (0, 1)$ such that if $\lambda < \lambda^{**}$ then $\omega_2^* < 1$. That is that if $\alpha\mu < 1 - \mu$ and $\lambda < \lambda^{**}$ then we can have $\omega_2^* < \omega < 1$ such that $\delta_2^* < 1 - \alpha$. $\square$

PROOF OF LEMMA 5: If $\lambda = 0$, $\tau = 0$, by (21) and (16), differentiating (17) w.r.t δ yields

$$\begin{aligned} \frac{\partial \gamma}{\partial \delta} \mid_{\tau=0} &= \frac{\partial \Phi}{\partial \delta} - \left(\frac{\partial \Psi}{\partial \delta} \Lambda + \Psi \frac{\partial \Lambda}{\partial \delta} \right), \\ &= \frac{\partial \Phi}{\partial \delta} + \Upsilon \Lambda, \text{ denoting } \Upsilon \equiv \Psi q - \frac{\partial \Psi}{\partial \delta}, \end{aligned} \tag{78}$$

where

$$\Psi = \delta \left(2 - \delta - \frac{\omega}{\mu} \right), \tag{79}$$

$$\frac{\partial \Psi}{\partial \delta} = 2 - 2\delta - \frac{\omega}{\mu}, \tag{80}$$

$$\Lambda = \frac{\mu^2}{\delta (2 - \delta)} \frac{\sigma^2}{2}, \tag{81}$$

$$\frac{\partial \Lambda}{\partial \delta} = -q\Lambda, \text{ where } q \equiv \frac{2 (1 - \delta)}{\delta (2 - \delta)}, \tag{82}$$

$$\Upsilon \equiv \frac{\omega}{\mu} \frac{\delta}{2 - \delta} > 0, \tag{83}$$

$$\frac{\partial \Phi}{\partial \delta} = -\ln s_1 - 1 - \epsilon \ln l - \frac{1}{\eta} \frac{(1 - \alpha - \delta) \epsilon \rho}{1 - \rho\alpha - \rho (1 - \alpha - \delta)}. \tag{84}$$

By (19) we can show that as $\delta \to 1 - \alpha$, $\ln s_1 \to -\infty$. By (18), (19) and (84), we can show that $-\frac{\partial \Phi}{\partial \delta}$ goes down from a positive value to $-\infty$ as δ increases from 0 to approaching $1 - \alpha$.

And by (81) and (83), we can show

$$\frac{\partial}{\partial \delta} (\Upsilon \Lambda) = \frac{2}{(2 - \delta)^3} \frac{\omega}{\mu} \frac{\mu^2 \sigma^2}{2} > 0. \tag{85}$$

It means that $\Upsilon\Lambda$ goes up as δ increases.

When $\delta = 0$, $\Upsilon\Lambda (\delta = 0) = \frac{\omega\mu\sigma^2}{8}$. When $\delta \to 1 - \alpha$, $\Upsilon\Lambda (\delta \to 1 - \alpha) = \frac{\omega\mu\sigma^2}{2}$. Thus, by evaluating $\Upsilon\Lambda$ and $\frac{\partial\Phi}{\partial\delta}$ at $\delta = 0$, we can find a σ^* such that if $\sigma^2 < \sigma^*$, there exists a $\delta_3^* \in (0, 1-\alpha)$ such that $\frac{\partial\gamma}{\partial\delta} = 0$ if $\delta = \delta_3^*$. For any $\delta > \delta_3^*$, $\frac{\partial\gamma}{\partial\delta} |_{\tau=0} > 0$. □

PROOF OF PROPOSITION 1: We write the system of linear equations (59) and (61) in a matrix form, and get

$$m_{t+1} = Bm_t + D\Lambda_t + d\Delta_{ht}^2/2 + E, \tag{86}$$

where

$$m_{t+1} \equiv \begin{bmatrix} m_{kt+1} \\ m_{ht+1} \end{bmatrix},$$

$$B \equiv \begin{bmatrix} \lambda & \mu \\ \beta\lambda & \alpha + \beta\mu + \delta \end{bmatrix}, D \equiv \begin{bmatrix} \tau(2-\tau) \\ \beta\tau(2-\tau) \end{bmatrix}, d \equiv \begin{bmatrix} 0 \\ \delta\omega \end{bmatrix},$$

$$E \equiv \begin{bmatrix} \ln s_2 + \epsilon \ln l \\ \ln \kappa - \sigma^2/2 + \beta \ln s_1 + \beta\epsilon \ln l \end{bmatrix}.$$

Rearranging (86), we get,

$$(I - BL)\, m_{t+1} = E + D\Lambda_t + d\Delta_{ht}^2/2, \tag{87}$$

$\Leftrightarrow$

$$m_{t+1} = [I - B]^{-1} E + [I - BL]^{-1} D\Lambda_t + [I - BL]^{-1} d\Delta_{ht}^2/2. \tag{88}$$

It follows from the above equation that, if $|I - B| \neq 0$, then m_t converges to a stationary state. Thus, the endogenous growth of m_t requires $|I - B| = 0$, i.e.,

$$\left| \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & \mu \\ \beta\lambda & \alpha + \beta\mu + \delta \end{bmatrix} \right| = 0.$$

$\Leftrightarrow$

$$\left| \begin{bmatrix} 1-\lambda & -\mu \\ -\beta\lambda & 1 - (\alpha + \beta\mu) \end{bmatrix} \right| = 0.$$

$\Leftrightarrow (1 - \alpha - \delta)(1 - \lambda) - \beta\mu = 0$; or, equivalently, $\dfrac{\mu}{1-\lambda} = \dfrac{1-\alpha-\delta}{\beta}$.

By (60), (62) and (63), at steady state, we can get

$$\Delta_k^2 = \frac{\mu^2 (1-\tau)^2 (1+\lambda\alpha(1-\tau))}{(1-\lambda\alpha(1-\tau))\left((1+\lambda\alpha(1-\tau))^2 - (\alpha+(\lambda+\beta\mu)(1-\tau))^2\right)}\sigma^2, \tag{A.28}$$

$$\Delta_h^2 = \frac{(1-\lambda(1-\tau)(\alpha+2\beta\mu(1-\tau)))\left(1-\lambda^2(1-\tau)^2\right) - 2\mu\beta\lambda^3(1-\tau)^4}{(1-\lambda\alpha(1-\tau))\left((1+\lambda\alpha(1-\tau))^2 - (\alpha+(\lambda+\beta\mu)(1-\tau))^2\right)}\sigma^2, \tag{A.29}$$

$$cov = \frac{\beta\lambda^2\mu^2(1-\tau)^4 + \mu(1-\tau)(\alpha+\beta\mu(1-\tau))\left(1-\lambda^2(1-\tau)^2\right)}{(1-\lambda\alpha(1-\tau))\left((1+\lambda\alpha(1-\tau))^2 - (\alpha+(\lambda+\beta\mu)(1-\tau))^2\right)}\sigma^2. \tag{A.30}$$

Then, by (64), Lemma 3, (A.28), and (A.29), the proof of Proposition 1 is completed. □

References

Acemoglu, D., P. Aghion and F. Zilibotti (2006). "Distance to Frontier, Selection, and Economic Growth," *Journal of the European Economic Association*, 4, 37–74.

Acemoglu, D. and J. A. Robinson. (2006). "Economic Backwardness in Political Perspective," *American Political Science Review*, 100, 115–131.

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

Altonji, J. G., and U. Doraszelski. (2005). "The Role of Permanent Income and Demographics in Black/White Differences in Wealth," *Journal of Human Resources*, 40, 1–30.

Barro, R., G. Mankiw, and X. Sala-i-Martin. (1995). "Capital Mobility in Neoclassical Models of Growth," *American Economic Review*, 85, 103–115.

Becker, G. S., and N. Tomes. (1979). " An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy* 87, 1153–1189.

Benabou, R. (1996). "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review* 86, 584–609.

—— (2002). "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica*, 70, 481–517.

Cremer, H., and P. Pestieau. (2001). "Non-linear Taxation of Bequests, Equal Sharing Rules and the Tradeoff Between Intra- and Inter-family Inequalities," *Journal of Public Economics*, 79, 35–53.

Crow, E. L., and K. Shimizu. (1988). "Lognormal Distributions: Theory and Applications," Marcel Dekker Inc., New York.

Easterly, W., and S. Rebelo. (1993). "Marginal Income Tax Rates and Economic Growth in Developing Countries," *European Economic Review*, 37, 409–417.

Foster, A. D. and M. R. Rosenzweig. (1995). "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 103, 1176–1209.

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

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

Garcia-Penalosa, C. and J. Wen (2008). "Redistribution and Entrepreneurship with Schumpeterian Growth," Journal of Economic Growth, 13, 57–80.

Glewwe, P., H. G. Jacoby and E. M. King (2001): "Early Childhood Nutrition and Academic Achievement: A Longitudinal Analysis," *Journal of Public Economics*, 81, 345–368.

Glomm, G., and B. Ravikumar. (1992). "Public versus Private Investment in Human Capital: Endogenous Growth and Income Inequality," *Journal of Political Economy*, 100, 818–834.

Gordon, R., and W. Li. (2009). "Tax structures in Developing Countries: Many Puzzles and a Possible Explanation," *Journal of Public Economics*, 93, 855–866.

Jha, S. K. (1999). "Fiscal Policy, Income Distribution, and Growth," *EDRC Report Series*, 67, 1–26.

Kakwani, N. (1977). "Applications of Lorenz Curves in Economic Analysis," *Econometrica*, 45, 719–727.

Kotlikoff, L. J., and L. H. Summers. (1981). "The Role of Intergenerational Transfers in Aggregate Capital Accumulation," *Journal of Political Economy*, 89, 706–732.

Laitner, J. (1979a). "Bequests, Golden-age Capital Accumulation and Government Debt," *Economica*, 46, 403–414.

——(1979b). "Household Bequest Behaviour and the National Distribution of Wealth," *Review of Economic Studies*, 46, 467–483.

——(1979c). "Household Bequests, Perfect Expectations, and the National Distribution of Wealth," *Econometrica*, 47, 1175–1193.

Lucas, R. E. (1988). "On the Mechanics of Economic Development," *Journal of Monetary Economics,* 22, 3–42.

Maoz, Y. D. and O. Moav (1999). “Intergenerational Mobility and the Process of Development,” *Economic Journal*, 109, 677–697.

Momota, A. (2009). "A Population-macroeconomic Growth Model for Currently Developing Countries," *Journal of Economic Dynamics and Control*, 33, 431–453.

Mookherjee, D. and D. Ray. (2003). "Persistent Inequality," *Review of Economic Studies*, 70, 369-393.

Park, K. H. (1998). "Distribution and Growth: Cross-Country Evidence," *Applied Economics,* 30, 943–949.

Perotti, R. (1993). "Political Equilibrium, Income Distribution, and Growth," *Review of Economic Studies*, 60, 755–776.

——(1996). "Growth, Income Distribution, and Democracy: What the Data Say," *Journal of Economic Growth*, 1, 149–187.

Persson, T., and G. E. Tabellini. (1994). "Is Inequality Harmful for Growth? Theory and Evidence," *American Economic Review,* 84, 600–621.

Rebelo, S. (1991). "Long-Run Policy Analysis and Long-Run Growth," *Journal of Political Economy*, 99, 500–521.

Romer, P. M. (1986). "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, 94, 1002–1037.

Rosenzweig, M. R., and K. I. Wolpin. (1994). "Are There Increasing Returns to the Intergenerational Production of Human Capital? Maternal Schooling and Child Intellectual Achievement," *Journal of Human Resources,* 29, No. 2, Special Issue: Women’s Work, Wages, and Well-Being, 670–693.

Saint-Paul, G. and T. Verdier. (1993): "Education, Democracy and Growth," *Journal of Development Economics*, 42, 399–407.

Sicat, G. P., and A. Virmani. (1988). "Personal Income Taxes in Developing Countries," *The World Bank Economic Review,* 2, 123–138.

Solow R. M. (1957). “Technical Change and the Aggregate Production Function,” *Review of Economics and Statistics,* 39, 312–320.

Stewart, F. and E. Ghani. (1991). "How Significant Are Externalities for Development?" *World Development*, 19, 569–594.

Tamura, R. (1991). "Income Convergence in an Endogenous Growth Model," *Journal of Political Economy,* 99, 522–540.

Trostel, P. A. (2004). "Returns to Scale in Producing Human Capital From Schooling," *Oxford Economic Papers,* 56, 461–484.

Zhang, J. (2005). "Income Ranking and Convergence with Physical and Human Capital and Income Inequality," *Journal of Economic Dynamics & Control*, 29, 547–566.