INEQUALITY, GROWTH, AND OVERTAKING

Yoshiaki Sugimoto[*]

February 28, 2003

This research develops a theory about the role of inequality in the overtaking of growth performance across countries. The theory captures two opposing effects of inequality on factor accumulation and suggests that the qualitative change in their combined effect is a prime cause of overtaking. Due to the initial dominance of the positive effect of inequality, a less egalitarian economy undergoes a higher growth path in the short run, followed by a lower growth path in the long run. It is also shown that divergence or convergence may arise instead of overtaking, depending on the initial levels of development and inequality.

Keywords: Wealth Distribution, Overtaking, Divergence, Convergence

JEL Classification: J31, O11, O15, O40

---

[*] Department of Economics, Brown University, Providence, RI 02912, USA. Email: Yoshiaki_Sugimoto @brown.edu. I am grateful to Oded Galor, whose insightful comments and suggestions improved the paper considerably. Furthermore, I wish to thank Gareth Lee, Fumiko Tamura, Kenichi Ueda and participants at Brown University Macro Lunch for their helpful comments.

INTRODUCTION

Over the past two centuries, the evolution of the world income distribution has been characterized by shifts in the ranking of countries, as well as the great divergence and a recent convergence among industrialized economies. As documented by Maddison (2001), the Netherlands, whose per capita GDP had been the highest in Europe since 1600, was overtaken by the United Kingdom towards the end of the ninetieth century. The economic leadership was then replaced by the United states in the beginning of the twentieth century. Outside the Western world, Japan and newly industrializing countries (Hong Kong, Singapore, South Korea, and Taiwan) had overtaken Argentina and Chile during the second half of the twentieth century.[1]

Growth theorists have attempted to construct the theoretical foundations that account for these unpredictable phenomena, and three major approaches have been proposed so far. The first approach, taken by Brezis et al. (1993), suggests that overtaking reflects country-specific learning effects in the existing technology and the resulting comparative disadvantage of a leading economy in adopting a new technology. The second approach highlights the international flows of ideas. Goodfriend and McDermott (1998) argue that familiarity with a trading partner facilitates technological spillovers and enhances learning productivity, human capital accumulation, and growth. Thus, overtaking is caused by unilateral familiarization of a less developed country with the leading country.[2] The third approach focuses on the changing role of natural resources in the process of development. Galor et al. (2002) suggest that while land abundance is beneficial for the process of development in early stages, it hinders the execution of education reform that is a precondition for industrialization and sustained growth.

This research examines the role of income inequality in the overtaking of growth performance across countries. It employs a unified approach that captures two opposing effect of inequality on factor accumulation, and thereby examines the link between initial income distribution and the pattern of development. Inequality promotes physical capital accumulation by stimulating savings of the rich, whose marginal propensity to save is relatively high.[3] At the same time, in the presence of credit market imperfections, low-income households have limited access to loans to finance their education.[4]

---

[1] See Abramovitz (1986), Quah (1993, 1996), Jones (1997), and Pritchett (1997) for the disccusion on the long-run evolution of the world income distribution.

[2] See Fischer and Serra (1996) and Mountford (1998) for other studies that emphasize international links in overtaking. The former demonstrate that due to the local externality of human capital, a country with a highly equal income distribution may overtake leaders. The latter finds a source of overtaking in the dynamic structure of the standard Heckscher-Ohlin model.

[3] See Keynes (1936), Kaldor (1957), Stiglitz (1969), and Bourguignon (1981) for theretical considerations, and Mayer (1966), Carrol (2000), and Dynan et al. (2000) for empirical evidence. The last paper finds that higher lifetime income hoseholds save a larger fraction of their income for precauitonary saving and bequest motives.

[4] Flug et al. (1998) draw evidence from cross-country and panel regressions that credit market imperfections and unequal wealth distribution lead to lower average secondary enrollment. Perotti (1996) empirically supports the view that income equality encourages both male and female education attainments.

Since each individual's investment in human capital is subject to diminishing marginal returns, concentrating resources into a small part of the population retards the accumulation of aggregate human capital (Galor and Zeira, 1993).

The relationship between distribution of personal income and economic growth has been one of the most controversial topics in macroeconomics over the last decade. Despite considerable number of empirical investigations, little is known about the relationship between these two factors within a single country. Most studies in the 1990s support the view that inequality is a hindrance to growth, while some recent studies find that their relationship turns positive in the short run.[5] Although these puzzling results would reflect, to some extent, differences in estimation methods as well as data quality, it appears that this empirical ambiguity may reflect opposing forces that operate simultaneously.[6]

The proposed theory attributes the overtaking to a qualitative change in the combined effect of inequality on physical and human capital accumulation. The positive effect of inequality on physical capital accumulation is dominant in the early stages of development, where the return to skill is low relative to the return to physical capital and the saving-rate differential among the rich and poor is significant. Capital accumulation and the accompanying increase in wages, however, raises the return to skill and reduces the saving-rate differential, reversing the qualitative effect of initial inequality in later stages.[7]

The eventual dominance of the negative effect of inequality generates multiple steady-state equilibria. A country characterized by a substantially equal distribution will converge to the higher-level steady state, in terms of per capita output, where all individuals can equally acquire skills and accumulate wealth. On the other hand, a country characterized by a highly unequal distribution will converge to the lower-level steady state, where unskilled workers are unable to accumulate wealth and inequality exists persistently between skilled and unskilled workers. The initial distribution of wealth therefore plays a significant role in determining both long-run growth performance and individuals' welfare.[8] The lower-level steady state act as a development trap because, without a substantial increase in output or equality, countries in this state cannot permanently escape from stagnation. Despite the necessity of reducing inequality, those countries will encounter capitalists' strong resistance to redistribution. As will be

---

[5] See Barro (2000) as well as Benabou's (1996) careful overview of the empirical studies in the early 1990s. A recent empirical work by Forbes (2000, p. 885) concludes that "the relationship between inequality and growth is far from resolved."

[6] For instance, Banerjee and Duflo (2000a, 2000b) argue that the difference in previous estimates is partly due to the linearity of the models. Atkinson and Brandolini (1999) find inappropriate to simply use "high quality" observations in Deininger and Squire's (1996) data set on income inequality.

[7] Due to the complementarity between skill and physical capital in production, the return to skill rises as physical capital accumulates. See Goldin and Katz (1998) for empirical evidence. Perotti (1996) finds that income equality encourages investment in education more significantly in a group of high-income countries.

[8] See Galor and Zeira (1993) and Banerjee and Newman (1993).

shown later, capitalists can hold higher wealth in the lower-level steady state than in the others, and thus redistribution is undesirable from their short-term and long-term viewpoints.

Overtaking therefore results from the initial dominance of the positive effect of inequality and the multiplicity of steady-state equilibria. A less egalitarian economy would undergo a higher growth path in the short run, followed by a lower growth path in the long run. Once overtaken, the country can never catch up with the leader and the income gap remains open. Furthermore, divergence applies to countries starting out in the mature stages of development where the adverse effect of inequality on human capital outweighs the positive effect of inequality on aggregate saving.

Paradoxically, overtaking is likely to occur under the condition that the marginal productivity of physical capital is higher relative to human capital. In such a circumstance, a major fraction of national income is dominated by asset owners since production of final output relies largely on physical capital. This uneven system of factor payments slows down the increase in wages relative to output, intensifying the adverse effect of credit constraints on human capital accumulation. It should be noted that the share of the labor income is less important for egalitarian economies where many individuals obtain asset earnings as well as wages.

Although the analysis suggests that inequality in the early stages of development has a positive effect on the growth process, there are two prime forces that make overtaking less probable than divergence. The one is that, as noted above, the positive effect of inequality is the dominant factor only in underdeveloped stages, where the saving rate differential among individuals is large. The other is that globalization of international capital markets encourages the flow of capital across borders, which makes domestic saving less important for physical capital accumulation. The positive aspect of inequality would thus be more relevant to the process of development experienced by currently developed countries.

This general tendency of divergence is supported by some empirical evidence. Benabou (1996) examines the role of inequality in economic development of South Korea and the Philippines, which were similar with respect to all major macroeconomic variables, such as GDP per capita, population, urbanization, and secondary school enrollment in the early 1960s. As a key factor to interpret South Korea's superior growth performance over the next 25 years, Benabou points out the significant difference in their initial distributions of income and land ownership. Inequality was much lower in South Korea as a result of its successful land reform following World War II. In this respect, Birdsall et al. (1995) suggest that policies that reduced poverty and income inequality, such as improving the quality of basic education and augmenting labor demand, have stimulated East Asian economies' growth since the 1960s. Engerman and

Sokoloff (2002) and Galor et al. (2002) argue that the divergence in income levels between North and Latin America, observed in the second half of the twentieth century, may be attributed partly to their different distributions of land ownership.

Turning to income convergence, countries will converge to the same steady state provided that they possess similar economic structures and levels of initial inequality. In particular, a country with moderate inequality can evolve toward the higher-level steady state through the transition from selective to universal human capital accumulation. Even though inequality delays the spread of education in society, moderate inequality permits the economy to reach a stage of development where unskilled workers are wealthy enough to support their children.[9] The resulting universal investment in education fuels the process of development, and credit constraints become less binding among the poor over time with the improvement of inequality. The growth process of the economy with moderate inequality would be associated with the evolution of some currently developed countries, which might have experienced a rise and fall of inequality over the last century, as surveyed by Kuznets (1955).

The rest of the paper is organized as follows. Section 1 outlines the basic structure of the model, and short-run equilibrium is derived in Section 2. The analytical framework is based on Galor and Moav's (2001) unified growth model that features four fundamental elements: capital market imperfections, altruistic linkage, capital-skill complementarity, and the increasing marginal propensity to save as a result of consumers' optimal behaviors.

There are two clear aspects that distinguish the current paper from Galor and Moav's research. First, they do not address the issue of overtaking and divergence. Galor and Moav divide the process of industrialization into four stages, and examine the effect of inequality in one stage on subsequent growth within the same stage (i.e., short-run growth). This paper, by contrast, studies longer-term growth beyond the initial stage so as to observe the diverse patterns of development. Second, Galor and Moav's analysis executes the moderate redistribution of wealth so that the ex-ante state of the economy is maintained. Drastic wealth redistribution, however, brings the economy into a state of widespread education, and this situation would be associated with the experience of some East Asian countries that markedly improved inequality after World War II.[10] The analysis in Section 3 carries out both drastic and moderate redistributions of wealth.

In order to accomplish these two objectives, Section 3 first elucidates the global behavior of the dynamical system that governs the evolution of inequality and demonstrates the multiplicity of steady-state equilibria. Using these results, Section 4 analyzes the impact of wealth distribution on the behavior of output growth, by comparing the

---

[9] Alternatively, as argued by Galor and Moav (2000), capitalists would be willing to support the accumulation of human capital of workers in order to sustain their profit rates.

[10] Birdsall et al. (1995) cite the examples of land reforms in Korea and Taiwan, public housing in Hong Kong, and extensive investment in rural infrastructure in Malaysia and Tahailand.

growth paths of hypothetical economies that differ only in their initial wealth distributions. The last section summarizes the discussion and proposes future research. Proofs of technical results are placed in the Appendix.

1. THE MODEL

Consider a closed overlapping-generations economy with heterogeneous agents in which activities occur over infinite discrete time. People invest in assets and education in the presence of imperfect capital markets. Producers employ the resulting physical and human capital through perfectly competitive markets, and generate a single final good that can be consumed or relinquished to the next generation. Population and the level of technology are exogenously determined and constant over time.

1.1. *Producers*

The amount of aggregate output produced at time t, Y_t, is determined by the aggregate stock of physical capital, K_t, and that of human capital, H_t, which are available in the economy at time t. The production function takes the Cobb-Douglas form:

$$Y_t = F(K_t, H_t) \equiv f(k_t)H_t = Ak_t^{\alpha}H_t, \tag{1}$$

where $\alpha \in (0, 1)$, $k_t \equiv K_t/H_t$, and $A > 0$ stands for a level of technology. The market price of the final good is normalized to 1. Production operates by hiring workers and renting the services of physical capital from households through the competitive markets, without incurring any adjustment costs. In contrast to individuals' loans for education, firms can rent physical capital under no liquidity constraints on the grounds that physical capital is easily collateralized. Hence, they maximize their profits given the market wage per unit of human capital, w_t, and the rental price per unit of physical capital, r_t. Those firms are represented by a single producer who maximizes $f(k_t)H_t - w_tH_t - r_tK_t$ by choosing K_t and H_t. Then the first-order conditions are given by

$$\begin{aligned} r_t &= f'(k_t) = \alpha A k_t^{\alpha-1} \equiv r(k_t), \\ w_t &= f(k_t) - f'(k_t)k_t = (1-\alpha)Ak_t^{\alpha} \equiv w(k_t), \end{aligned} \tag{2}$$

which are the inverse demand functions for K_t and H_t in terms of the factor ratio k_t. The rate of return to human capital, w_t, increases with physical capital due to the complementarity between them. Physical capital depreciates at a constant rate of $\delta \in [0, 1]$ in every period.

1.2. *Households*

1.2.1. Environment. A new generation of individuals is born in every period, and individuals live for two periods. Thus there are two generations in society at each point in time. Each individual has a single parent when young, and has a single child when old. The number of members of a generation is normalized to one. Individuals may differ from each other in their initial endowments, yet they are homogeneous in all the other aspects. The preferences of an individual who is born in period t are defined over c_{t+1}, consumption of the individual and his/her child in period $t+1$, and b_{t+1}, the wealth transferred to the child in period $t+1$. They are represented by the utility function

$$u(c_{t+1}, b_{t+1}) = (1-\beta)\ln c_{t+1} + \beta \ln(\bar{\theta} + b_{t+1}), \tag{3}$$

where $\beta \in (0,1)$ and $\bar{\theta} > 0$. The underlying assumption of (3) is that intergenerational transfers are a luxury good and are motivated by parents' "joy of giving." The nonnegative amount of wealth is necessarily transferred to offspring since the utility is an increasing function of b_{t+1}.

In period 0, old people in society are divided into two groups, S (Small) and L (Large). Group i ($i = S, L$) involves a fraction λ^i of them, where $0 < \lambda^S = 1 - \lambda^L < 0.5$ and λ^i remains unchanged over time once determined in period 0. Members within a group are endowed with the same amount of wealth, I_0^i, whereas individuals between the two groups may differ in this regard. As follows from (3), this between-group inequality generates the difference in their intergenerational transfers. On the other hand, since members within one group are completely homogeneous, their descendents behave identically in all subsequent periods. Hence an individual who is descended from group i and is born in period t may be referred to as a member i of generation t.

In the first period, when young, individuals consume part of their parents' wealth to live and spend the entire time to acquire skills. The formation of human capital is augmented by investing physical resources in education, and the investment during the first period is the only way to enhance human capital. Hence, individuals allocate their inheritance between education (investment in human capital) and savings (investment in physical capital). The amount saved by a member i of generation t is thus

$$s_t^i = b_t^i - e_t^i,$$

where b_t^i is the transfer from his/her parent and e_t^i is the real expenditure on education.

In the second period, when old, the individual acquires the resulting human capital $h_{t+1}^i \equiv h(e_t^i)$, which is an increasing function with strict concavity and Inada conditions.[11] Individuals acquire only basic skills in the absence of the real expenditure on education, and the associated level of human capital is $h(0) = 1$. Wage income is obtained by supplying the acquired human capital inelastically in competitive labor markets. In addition, those who have savings rent out capital services to producers at the market price. Then the second-period wealth of a member i of generation t, I_{t+1}^i, is

$$\begin{aligned} I_{t+1}^i &= w_{t+1} h_{t+1}^i + s_t^i R_{t+1} \\ &= w_{t+1} h(e_t^i) + (b_t^i - e_t^i) R_{t+1}, \end{aligned} \tag{4}$$

---

[11] The assumption of Inada conditions, which deviates from the Galor-Moav (2001) model, is necessary to avoid a corner solution at which there is no incentive to invest in education. Alternative specifications complicate the exposition with no change in the basic results. See Footnote 12 as well.

where $R_{t+1} = 1 + r_{t+1} - \delta$. The budget constraint of the individual is given by

$$c_{t+1} + b_{t+1} \leq I^i_{t+1}. \tag{5}$$

1.2.2. Optimization. Each member of generation t maximizes his/her utility from (3) subject to (5). The optimal amount of transfer chosen by a member i of generation t is

$$b^i_{t+1} = \begin{cases} 0 & \text{if } I^i_{t+1} < \theta, \\ \beta(I^i_{t+1} - \theta) & \text{if } I^i_{t+1} \geq \theta, \end{cases} \tag{6}$$

where $\theta \equiv \bar{\theta}(1-\beta)/\beta > 0$ can be interpreted as the minimum need of current consumption. Since the bequest function and the associated saving function exhibit convexity, inequality in wealth augments national saving.

Observe that the indirect utility strictly monotonically increases with the second period's wealth, I^i_{t+1}. It then follows from (4) that utility is maximized by choosing the education costs that maximize I^i_{t+1}. Therefore, the optimal level of education in the face of no credit constraints, denoted as e_t, is

$$e_t = \arg\max_e [w_{t+1}h(e) - eR_{t+1}], \tag{7}$$

where individuals regard both wage and interest rate as given and predict these future variables accurately at time t. The first order condition is

$$w_{t+1}h'(e_t) = R_{t+1} \qquad \text{for } w_{t+1} > 0. \tag{8}$$

Also, $e_t = 0$ for $w_{t+1} = 0$ since zero wage yields no return on investment in education. Then noting $w_{t+1} = w(k_{t+1})$ and $R_{t+1} = 1 + r_{t+1} - \delta \equiv R(k_{t+1})$ from (2), there exists a continuous single-valued function

$$e_t = e(k_{t+1}), \tag{9}$$

such that $e(0) = 0$ and $e'(k_{t+1}) > 0$ for all $k_{t+1} > 0$.[12] The intuition of the positive reaction of education to the capital-labor ratio is straightforward: an increase in k_{t+1} enhances the return on human capital, w_{t+1}, while it reduces the return on savings, R_{t+1}. e_t is hence independent of b^i_t, meaning that it is most favored among all young members at time t. In other words, e_t is the education expenditure one is willing to pay if one can.

Next consider the optimal educational decision in light of credit constraints. It is assumed that imperfect capital markets completely limit access to credit, and young members are unable to make loans for education. Then a member i of generation t invests in education according to

$$e^i_t = \begin{cases} e_t & \text{if } b^i_t < e_t, \\ b^i_t & \text{if } b^i_t \geq e_t. \end{cases} \tag{10}$$

---

[12] In contrast to the present model, Galor and Moav (2001) propose Regime I, a state of the economy with no investment in education, which nullifies the adverse effect of inequality on production. Inequality therefore enhances the process of development in this regime. We will prove that this positive effect of inequality exists even in the absence of Regime I.

Contrary to the first case, e_t^i will be heterogeneous across the two groups, depending on their transfer amounts. The resultant savings are expressed as

$$s_t^i = b_t^i - e_t^i = \begin{cases} 0 & \text{if } b_t^i < e_t, \\ b_t^i - e_t & \text{if } b_t^i \geq e_t. \end{cases} \tag{11}$$

Substitution for b_t^i from the budget constraint into the above expression yields $s_t^i = I_t^i - c_t^i - e_t^i$, meaning that savings equal parent's wealth minus household's consumption and education costs. s_t^i can hence be viewed as the amount saved by a household in period t, rather than an individual of generation t. Taking (10) and (11) into account, the wealth in the second period (4) is modified to

$$I_{t+1}^i = \begin{cases} w_{t+1} h(b_t^i) & \text{if } b_t^i < e_t, \\ w_{t+1} h(e_t) + (b_t^i - e_t) R_{t+1} & \text{if } b_t^i \geq e_t. \end{cases} \tag{12}$$

This shows that members who inherit a larger amount of transfer will earn higher income. Hence, in light of (6), $b_t^i \geq b_t^j$ for all t if $b_0^i \geq b_0^j$, where $i, j = S, L$ and $i \neq j$.[13] In words, the initial ranking of wealth among people is never reversed in the future.

1.3. *The Formation of Aggregate Capital*

In a closed economy, the net amount of total investment equals total income minus total expenditures. Thus,

$$K_{t+1} - K_t = w_t H_t + (r_t - \delta) K_t - \lambda^S c_t^S - \lambda^L c_t^L - \lambda^S e_t^S - \lambda^L e_t^L. \tag{13}$$

Since people receive no income when young, the total wealth of the economy in period t is owned by old individuals in period t, meaning

$$w_t H_t + R_t K_t = \lambda^S I_t^S + \lambda^L I_t^L. \tag{14}$$

Substituting (14) into (13) and using $s_t^i = I_t^i - c_t^i - e_t^i$ with (11), aggregate physical capital is

$$K_{t+1} = \lambda^S s_t^S + \lambda^L s_t^L = \begin{cases} 0 & \text{for } b_t^j \leq b_t^i \leq e_t, \\ \lambda^i (b_t^i - e_t) & \text{for } b_t^j < e_t < b_t^i, \\ B_t - e_t & \text{for } e_t \leq b_t^j \leq b_t^i, \end{cases} \tag{15}$$

where $B_t \equiv \lambda^S b_t^S + \lambda^L b_t^L = \lambda^j b_t^j + \lambda^i b_t^i$. B_t can be best thought of as total resources the economy possesses in period t since they are allocated between investment in physical and human capital. Proceeding, it follows from (10) that the total stock of human capital is

$$H_{t+1} = \lambda^S h(e_t^S) + \lambda^L h(e_t^L) = \begin{cases} \lambda^j h(b_t^j) + \lambda^i h(b_t^i) & \text{for } b_t^j \leq b_t^i \leq e_t, \\ \lambda^j h(b_t^j) + \lambda^i h(e_t) & \text{for } b_t^j < e_t < b_t^i, \\ h(e_t) & \text{for } e_t \leq b_t^j \leq b_t^i. \end{cases} \tag{16}$$

---

[13] These rules apply throughout the present paper.

Then by definition, the future ratio of the two capital stocks is

$$k_{t+1} = \begin{cases} 0 & \text{for } b_t^j \leq b_t^i \leq e_t, \\ \dfrac{\lambda^i(b_t^i - e_t)}{(1-\lambda^i)h(b_t^j) + \lambda^i h(e_t)} & \text{for } b_t^j < e_t < b_t^i, \\ \dfrac{B_t - e_t}{h(e_t)} & \text{for } e_t \leq b_t^j \leq b_t^i. \end{cases} \tag{17}$$

2. SHORT-RUN EQUILIBRIUM

Since individuals at time t are capable of predicting k_{t+1} accurately, their expectation of k_{t+1} in (17) must coincides with the actual level of k_{t+1} in (9). In other words, k_{t+1} in equilibrium satisfies both (9) and (17) for a given pair (b_t^S, b_t^L).

2.1. *Existence and Uniqueness*

The equilibrium condition precludes the situation $\max(b_t^S, b_t^L) < e_t$, which applies to the first case of (17). Namely, members of the richer group are never credit-constrained in equilibrium. This is because, if $\max(b_t^S, b_t^L) < e_t$, there would be no savings and $k_{t+1} = e(k_{t+1}) = 0$, a contradiction to $(b_t^S, b_t^L) \in \mathbb{R}_+^2$. Similarly $\max(b_t^S, b_t^L) = e_t$ holds if and only if $b_t^S = b_t^L = 0$.

To examine the second and third cases of (17), one can unify them into a single form

$$\mathrm{k}(e_t; b_t^e, b_t^c, \lambda) \equiv \frac{\lambda(b_t^e - e_t)}{(1-\lambda)h(b_t^c) + \lambda h(e_t)}, \tag{18}$$

whereby $\mathrm{k}(e_t; b_t^i, b_t^j, \lambda^i)$ and $\mathrm{k}(e_t; B_t, 0, 1)$ respectively correspond to the second and third cases. λ is viewed as the number of members who are not credit-constrained. b_t^e is the average level of their transfers at time t. b_t^c is the average level of transfers at time t owned by members who are credit-constrained.

$\mathrm{k}(e_t; b_t^e, b_t^c, \lambda)$ is a continuous function on $\mathbb{R}_+^3 \times (0,1]$ characterized by (i) $\mathrm{k}(e_t; b^e, b^c, 0) = 0 \; \forall (e_t, b_t^e, b_t^c) \in \mathbb{R}_+^3$, (ii) $\partial \mathrm{k}(\cdot)/\partial e_t < 0 \; \forall (e_t, b_t^e, b_t^c, \lambda) \in \mathbb{R}_{++} \times \mathbb{R}_+^2 \times (0,1]$, (iii) $\mathrm{k}(0; b_t^e, b_t^c, \lambda) \geq 0$ and $\mathrm{k}(b_t^e; b_t^e, b_t^c, \lambda) = 0 \; \forall (b_t^e, b_t^c, \lambda) \in \mathbb{R}_+^2 \times (0,1]$.

The properties of $e(k_{t+1})$ and $\mathrm{k}(e_t; b_t^e, b_t^c, \lambda)$ assure the existence of a unique k_{t+1} that solves $e_t = e(k_{t+1})$ and $k_{t+1} = \mathrm{k}(e_t; b_t^e, b_t^c, \lambda)$ simultaneously for a given triplet $(b_t^e, b_t^c, \lambda) \in \mathbb{R}_+^2 \times (0,1]$. Hence there exists a continuous single-valued function

$$k_{t+1} = k(b_t^e, b_t^c, \lambda)$$

on $\mathbb{R}_+^2 \times (0,1]$, such that (i) $k(b_t^e, b_t^c, 0) = 0 \; \forall (b_t^e, b_t^c) \in \mathbb{R}_+^2$, (ii) $k(0, b_t^c, \lambda) = 0$ and $\lim_{b_t^e \to \infty} k(\cdot) = \infty \; \forall (b_t^c, \lambda) \in \mathbb{R}_+ \times (0,1]$, and (iii) $\partial k(\cdot)/\partial b_t^e > 0$, $\partial k(\cdot)/\partial b_t^c < 0$ and $\partial k(\cdot)/\partial \lambda > 0 \; \forall (b_t^e, b_t^c, \lambda) \in \mathbb{R}_{++}^2 \times (0,1]$.

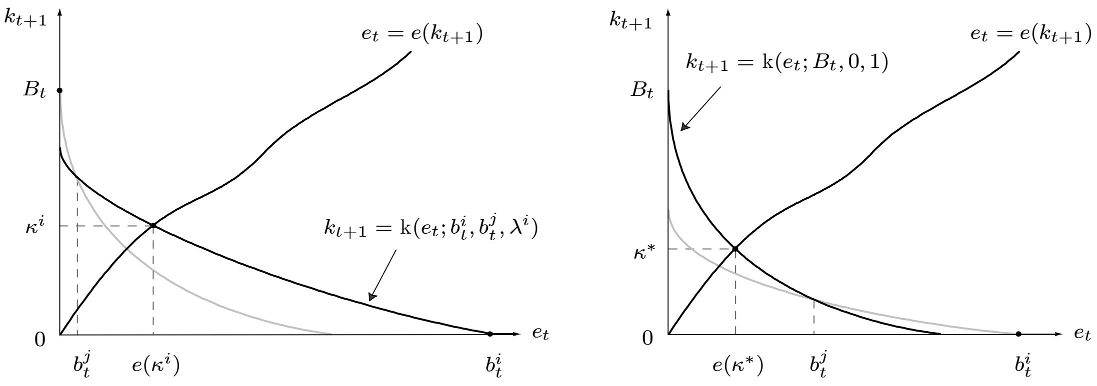

FIGURE 1. Inequality, Factor Intensity, and Education

Notes. Each of $k_{t+1} = \mathrm{k}(e_t; b_t^i, b_t^j, \lambda^i)$ and $k_{t+1} = \mathrm{k}(e_t; B_t, 0, 1)$ uniquely intersects with $e_t = e(k_{t+1})$. The equilibrium capital intensity, k_t, and desirable educational level, e_t, occur at one of the two intersections. A large gap between b_t^i and b_t^j characterize the equilibrium by $k_{t+1} = \kappa^i(b_t^S, b_t^L)$ and $b_t^j < e(\kappa^i(b_t^S, b_t^L)) < b_t^i$ (left panel), whereas a small gap results in $k_{t+1} = \kappa^*(b_t^S, b_t^L)$ and $e(\kappa^*(b_t^S, b_t^L)) < b_t^j < b_t^i$ (right panel).

The equilibrium level of the factor intensity, which solves (9) and (17), is now expressed as

$$\begin{aligned} k_{t+1} &= \begin{cases} k(b_t^i, b_t^j, \lambda^i) \equiv \kappa^i(b_t^S, b_t^L) & \text{for } (b_t^S, b_t^L) \in E^i \\ k(B_t, 0, 1) \equiv \kappa^*(b_t^S, b_t^L) & \text{for } (b_t^S, b_t^L) \in E^*. \end{cases} \\ &\equiv \kappa(b_t^S, b_t^L), \end{aligned} \tag{19}$$

where

$$\begin{aligned} E^i &\equiv \{(b_t^S, b_t^L) \in \mathbb{R}_+^2 \mid b_t^j < e(\kappa^i(b_t^S, b_t^L)) < b_t^i\}, \\ E^* &\equiv \{(b_t^S, b_t^L) \in \mathbb{R}_+^2 \mid e(\kappa^*(b_t^S, b_t^L)) \leq b_t^j \leq b_t^i\}. \end{aligned}$$

Note that superscript i $(= S, L)$ of $\kappa^i(b_t^S, b_t^L)$ indicates the richer group, and that whether λ equals λ^S, λ^L, or 1 is determined endogenously by a pair $(b_t^S, b_t^L) \in \mathbb{R}_+^2$.

Figure 1 geometrically represents the equilibrium levels of the factor intensity and education as an intersection of either $k_{t+1} = \mathrm{k}(e_t; b_t^i, b_t^j, \lambda^i)$ and $e_t = e(k_{t+1})$, or $k_{t+1} = \mathrm{k}(e_t; B_t, 0, 1)$ and $e_t = e(k_{t+1})$. The choice of the intersection is determined by the ratio of transfers between the two groups, b_t^S/b_t^L, as well as their total amount, B_t. The case in which $k_{t+1} = \kappa^i(b_t^S, b_t^L)$ and the poorer members are credit-constrained occurs if the gap between b_t^S and b_t^L are sufficiently large, otherwise the other case in which $k_{t+1} = \kappa^*(b_t^S, b_t^L)$ and nobody is constrained would be take place. The schedule $k_{t+1} = \mathrm{k}(e_t; B_t, 0, 1)$ is independent of the ratio b_t^S/b_t^L, and the same amount of B_t are used in both panels so as to capture the effect of inequality in transfers. Also, as shown in the diagram, $\mathrm{k}(e_t; b_t^i, b_t^j, \lambda^i)$ and $\mathrm{k}(e_t; B_t, 0, 1)$ uniquely intersects each other. Hence,

$$\kappa^i(b_t^S, b_t^L) = \kappa^*(b_t^S, b_t^L) \quad \text{if and only if} \quad b_t^j = e(\kappa(b_t^S, b_t^L)). \tag{20}$$

2.2. *The Credit Constraint Frontier*

In order to understand E^S, E^L, and E^* geometrically, we will introduce the credit constraint frontier, CC^i, on which members of group j spend the entire amount of transfers on the desirable level of education, leaving no savings. That is,

$$CC^i \equiv \{(b_t^S, b_t^L) \in \mathbb{R}_+^2 \mid b_t^i = e_t\},$$

where $e_t = e(\kappa(b_t^S, b_t^L))$.

Lemma 1. *There exists a single-valued function* $b_t^j = \varphi^j(b_t^i) \geq 0$ *on* $\mathbb{R}_+$ *such that* $(b_t^S, \varphi^L(b_t^S)) \in CC^S$, $(\varphi^S(b_t^L), b_t^L) \in CC^L$. *Furthermore, for* $(b_t^S, b_t^L) \in \mathbb{R}_+^2$,

i.

$$b_t^i - e(\kappa(b_t^S, b_t^L)) \begin{cases} < 0 & \text{if} \quad b_t^j > \varphi^j(b_t^i), \\ = 0 & \text{if} \quad b_t^j = \varphi^j(b_t^i), \\ > 0 & \text{if} \quad b_t^j < \varphi^j(b_t^i); \end{cases}$$

ii. $\varphi^j(0) = 0$ *and*

$$\frac{d\varphi^j(b_t^i)}{db_t^i} = 1 + \frac{h(e_t)}{\lambda^j e'(k_{t+1})} + \frac{k_{t+1}h'(e_t)}{\lambda^j} > 1.$$

Proof. Noting (19) and (20), let

$$G^i(b_t^i, b_t^j) \equiv b_t^i - e(k(b_t^j, b_t^i, \lambda^j)) = b_t^i - e(k(B_t, 0, 1)),$$

where $G^i(b_t^i, 0) \geq 0$, $\partial G^i(\cdot)/\partial b_t^j < 0$ and $\lim_{b_t^j \to \infty} G^i(\cdot) = -\infty \ \forall b_t^i \geq 0$. These properties assure the existence of a single-valued function $b_t^j = \varphi^j(b_t^i)$ on $\mathbb{R}_+$ such that $G^i(b_t^i, \varphi^j(b_t^i)) = 0$. (i) If $b_t^j > \varphi^j(b_t^i)$, then $b_t^i = e(k(\varphi^j(b_t^i), b_t^i, \lambda^j)) < e(k(b_t^j, b_t^i, \lambda^j)) = e_t$.[14] Likewise, if $b_t^j < \varphi^j(b_t^i)$, then $b_t^i = e(k(\lambda^i b_t^i + \lambda^j \varphi^j(b_t^i), 0, 1)) > e(k(B_t, 0, 1)) = e_t$. (ii) $\varphi^j(0) = 0$ because $G^i(0, 0) = 0$. $d\varphi^j(b_t^i)/db_t^i$ is obtained from a simple calculation. ∎

In light of the first result of the lemma, one can find

$$(b_t^S, b_t^L) \in \begin{cases} E^S & \text{if } b_t^S > \varphi^S(b_t^L), \\ E^L & \text{if } b_t^L > \varphi^L(b_t^S), \\ E^* & \text{otherwise.} \end{cases}$$

In words, whether a young individual is credit-constrained or not depends on the relative amount of b_t^S and b_t^L. If one group inherits a greater amount of wealth than the other, the desirable level of education is likely to cost more than the latter can afford to pay.

The results of Lemma 1 are summarized graphically in Figure 2. E^L is located above the CC^S line, E^S is below the CC^L line and E^* corresponds to the region between E^S and E^L. Moreover, $CC^S \cup CC^L = \{0\}$ and the infeasibility of the third case in (17) gives $E^S \cup E^L \cup E^* = \mathbb{R}_+^2$. Noting that the credit-constraint frontiers are the boundaries of E^* and the other sets, (20) assures the continuity of $\kappa(b_t^S, b_t^L)$ on $\mathbb{R}_+^2$. The following

---

[14] If $b_t^j > \varphi^j(b_t^i)$, then $b_t^i = e(k(\lambda^i b_t^i + \lambda^j \varphi^j(b_t^i), 0, 1)) < e(k(B_t, 0, 1))$. However, $\kappa(b_t^S, b_t^L) \neq k(B_t, 0, 1)$ if $b_t^i < e(k(B_t, 0, 1))$.

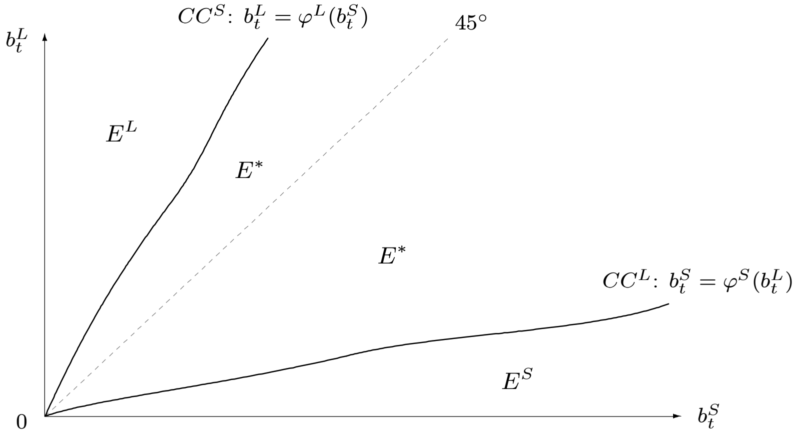

FIGURE 2. The Credit Constraint Frontiers

Notes. The frontiers are the boundaries that divide Regimes 2 and 3.

discussion sometimes uses the inverse function of φ^i, denoted as μ^j. In the diagram, for instance, the CC^L line is expressed by both $b_t^S = \varphi^S(b_t^L)$ and $b_t^L = \mu^L(b_t^S)$.

The discussion so far makes it clear that the economy falls under the category of one of the following regimes.[15]

Regime 1: Selective human capital accumulation
Regime 2: Universal human capital accumulation under constraints
Regime 3: Universal human capital accumulation under no constraints

Figure 2 associates a pair of transfers $(b_t^S, b_t^L) \in \mathbb{R}_+^2$ with the three regimes. For analytical convenience, the origin $(0,0)$ is defined as their intersection. By noting that the isoquant curve of B_t has a negative constant slope $-\lambda^S/\lambda^L$, one can find that the economy's regime is determined by distribution of B_t as well as the level of B_t. The economy with a pair of transfers $(b_t^S, b_t^L) \in \mathbb{R}_+^2 \backslash \mathbb{R}_{++}^2$ is in Regime 1 in period t, a state where young members of one group have no resources for investment. Such a pair can be regarded as a distribution extremely biased toward the other group. If $(b_t^S, b_t^L) \in \mathbb{R}_{++}^2 \backslash E^*$, then the economy belongs to Regime 2, where all young members receive transfers although those in the poorer group face credit constraints. Finally, the remaining region, E^*, corresponds to Regime 3, where B_t is most equally distributed among the three regimes, and all individuals can afford to attain the desirable level of education.

---

[15] Regimes 1-3 in the present paper are the counterparts of Stages I-III (of Regime II) in Galor and Moav (2001). As revealed below, the economy does not necessarily go through each stage in the process of development. For this reason, Stages I-III are renamed as Regimes 1-3. The counterpart of their Regime I does not exist in our model as discussed in Footnote 12.

2.3. *Inequality and Productivity*

It follows from (1), (15) and (16) that aggregate output, or equivalently output per worker, is

$$Y_{t+1} = Y(B_t, b_t^c, \lambda) = \begin{cases} Y(B_t, b_t^j, \lambda^i) & \text{for } (b_t^S, b_t^L) \in E^i \\ Y(B_t, 0, 1) & \text{for } (b_t^S, b_t^L) \in E^*, \end{cases} \tag{21}$$

where

$$Y(B_t, b_t^c, \lambda) \equiv A[B_t - (1-\lambda)b_t^c - \lambda e_t]^{\alpha}[(1-\lambda)h(b_t^c) + \lambda h(e_t)]^{1-\alpha}. \tag{22}$$

As follows from (2) and (7), e_t is the maximizer of the right hand side of (22), and thus the envelope theorem verifies the expression of $Y(B_t, b_t^c, \lambda)$. Observe $Y_B(B_t, b_t^c, \lambda) > 0$ and $Y_{BB}(B_t, b_t^c, \lambda) < 0$, $\forall (B_t, b_t^c, \lambda) \in \mathbb{R}_{++} \times \mathbb{R}_{+} \times (0, 1]$, and $Y(0, b_t^c, \lambda) = 0$ and $\lim_{B_t \to \infty} Y_B(B_t, b_t^c, \lambda) = 0$, $\forall (b_t^c, \lambda) \in \mathbb{R}_{+} \times (0, 1]$.

The optimality of e_t and the definition of E^i imply

$$\begin{aligned} &\partial Y(B_t, b_t^j, \lambda^i)/\partial b_t^c > 0 \qquad \text{for } (b_t^S, b_t^L) \in E^i, \\ &\partial Y(B_t, 0, 1)/\partial b_t^c = 0. \end{aligned} \tag{23}$$

Hence the isoquant of Y_{t+1} exhibits convexity in the (b_t^S, b_t^L) space; it is strictly convex to the origin in the regions E^S and E^L, and has a constant slope $-\lambda^S/\lambda^L$ in region E^*. Noting that the isoquant of B_t has a slope of $-\lambda^S/\lambda^L$, a great change in b_t^S/b_t^L not only worsens between-group inequality but also affects the level of output adversely.

Equally important,

$$\partial Y(B_t, b_t^c, \lambda)/\partial \lambda > 0 \qquad \text{for } (b_t^S, b_t^L) \in E^S \cup E^L, \tag{24}$$

meaning that the adverse effect of between-group inequality in transfers on output is more significant when the number of richer members is smaller.

These results reflect the fact that inequality lowers *intertemporal productivity* as a result of the constrained investment in human capital. Since personal human capital increases disproportionately with the cost of investment, concentrating resources towards a small portion of the population reduces the aggregate level of human capital. Note that redistribution of transfers in period t has no positive impact on physical capital and output in period $t+1$. By contrast, redistribution of income in period t generates a trade-off in period $t+1$ between the efficiency in production and the abundance of resources.

3. THE DYNAMICAL SYSTEM

This section explores the global behavior of the dynamical system that governs the evolution of wealth inequality in the process of development. Plugging (2) and (9) into

(12) and then using (19), we obtain

$$\begin{aligned} I^i_{t+1} &= I(b^i_t, k_{t+1}) \\ &\equiv \begin{cases} w(\kappa^i)h(e(\kappa^i)) + (b^i_t - e(\kappa^i))R(\kappa^i) & \text{for } (b^S_t, b^L_t) \in E^i \\ w(\kappa^*)h(e(\kappa^*)) + (b^i_t - e(\kappa^*))R(\kappa^*) & \text{for } (b^S_t, b^L_t) \in E^* \\ w(\kappa^j)h(b^i_t) & \text{for } (b^S_t, b^L_t) \in E^j. \end{cases} \\ &\equiv I^i(b^S_t, b^L_t), \end{aligned} \tag{25}$$

where κ^i, κ^* and κ^j, are the abbreviations for $\kappa^i(b^S_t, b^L_t)$, $\kappa^*(b^S_t, b^L_t)$ and $\kappa^j(b^S_t, b^L_t)$. Hence, member i's future income is affected by the current wealth level of the other group through changes in the capital ratio. Now define a set Z^i such that

$$Z^i \equiv \{(b^S_t, b^L_t) \in \mathbb{R}^2_+ \mid I^i_{t+1} < \theta\}. \tag{26}$$

It then follows from (6) that the evolution of intergenerational transfers within a group is

$$\begin{aligned} b^i_{t+1} &= \phi(b^i_t, k_{t+1}) \\ &\equiv \begin{cases} \beta[w(\kappa^i)h(e(\kappa^i)) + (b^i_t - e(\kappa^i))R(\kappa^i) - \theta] & \text{for } (b^S_t, b^L_t) \in E^i \backslash Z^i \\ \beta[w(\kappa^*)h(e(\kappa^*)) + (b^i_t - e(\kappa^*))R(\kappa^*) - \theta] & \text{for } (b^S_t, b^L_t) \in E^* \backslash Z^i \\ \beta[w(\kappa^j)h(b^i_t) - \theta] & \text{for } (b^S_t, b^L_t) \in E^j \backslash Z^i \\ 0 & \text{for } (b^S_t, b^L_t) \in Z^i. \end{cases} \\ &\equiv \psi^i(b^S_t, b^L_t). \end{aligned} \tag{27}$$

A trajectory $\{b^S_t, b^L_t\}^\infty_{t=0}$ is fully determined by the dynamical system that consists of the initial condition (b^S_0, b^L_0), the autonomous equations

$$\begin{aligned} b^S_{t+1} &= \psi^S(b^S_t, b^L_t); \\ b^L_{t+1} &= \psi^L(b^S_t, b^L_t), \end{aligned}$$

and the state space $\mathbb{R}^2_+$. Adult individuals at time 0 grant b^S_0 and b^L_0 to their children according to $b^i_0 = \max[\beta(I^i_0 - \theta), 0]$ from (6). The initial income I^i_0 consists of the earnings from the ownership of physical and human capital, and their levels and distributions are exogenously given.

In order to simplify the following analysis of the dynamical system, we assume complete capital depreciation, $\delta = 1$, so that $R_t = r_t$.[16] The stock variable I^i_t is then reduced to a flow variable, the income obtained in period t.

3.1. *The BB Locus*

In order to fully understand the dynamics of transfers, it is necessary to characterize the BB^S and BB^L loci. BB^i is defined as all pairs of (b^S_t, b^L_t) for which $b^i_{t+1} = b^i_t$. That is,

$$BB^i \equiv \{(b^S_t, b^L_t) \in \mathbb{R}^2_+ \mid b^i_{t+1} = b^i_t\}. \tag{28}$$

---

[16] Assuming $\delta \in [0, 1)$ would not alter any qualitative properties of the dynamcal system.

As is established in (27), the domain of the function $\psi^i(b_t^S, b_t^L)$ is divided into E^*, E^i, E^j and Z^i, depending on the relative amount of b_t^S to b_t^L, as well as their total amount. The following discussion characterizes BB^i and the dynamics of transfers in each of these four sets.

It would be plausible to assume that a nontrivial, locally stable, steady-state equilibrium exists when all members are free from credit constraints. This situation occurs if the technology is advanced enough to satisfy

$$A > A(\alpha, \beta, \theta), \tag{29}$$

where $A(\alpha, \beta, \theta)$ is a continuous single-valued function.[17]

3.1.1. $(b_t^S, b_t^L) \in E^i \backslash Z^i$ — Distribution Biased to the Group Itself. First, consider the dynamics of b_t^i when resource distribution is biased toward group i. In light of (27), define

$$\phi^R(b_t^i, \kappa^i) \equiv \beta[w(\kappa^i)h(e(\kappa^i)) + (b_t^i - e(\kappa^i))R(\kappa^i) - \theta] \equiv \psi^{Ri}(b_t^S, b_t^L), \tag{30}$$

where $\kappa^i = \kappa^i(b_t^S, b_t^L) = k(b_t^i, b_t^j, \lambda^i)$ and $(b_t^S, b_t^L) \in \mathbb{R}_+^2$. Then $b_{t+1}^i = \psi^{Ri}(b_t^S, b_t^L)$ for $(b_t^S, b_t^L) \in E^i \backslash Z^i$. One can find $\phi_\kappa^R(b_t^i, \kappa^i) < 0$ by noting that e_t is the maximizer of I_{t+1}^i, as shown in (7), and that $b_t^i - e_t = K_{t+1}/\lambda^i$ from (15). Thus, an increase in b_t^j raises richer members' transfer b_{t+1}^i by reducing the capital ratio.

Lemma 2. *Let* $b(\lambda) \equiv \{b_t^i \in \mathbb{R}_+ \,|\, \phi^R(b_t^i, k(b_t^i, 0, \lambda)) = b_t^i\}$, $\underline{b} \equiv \min b(1)$ *and* $\bar{b} \equiv \max b(1)$. *Under* (29),

i. $\{0\} \notin b(\lambda)$ *for all* $\lambda \in (0, 1]$;
ii. $b(1) = \{\underline{b}, \bar{b}\}$.

Proof. (i) The result follows from $\phi^R(0, k(0, 0, \lambda)) = -\beta\theta < 0$. (ii) If $\lambda = 1$, only group i exists and thus $I_{t+1}^i = Y_{t+1}$. Then (22) yields $\phi^R(b_t^i, k(b_t^i, 0, 1)) = \beta[Y(b_t^i, 0, 1) - \theta]$, and the properties of $Y(b_t^i, 0, 1)$ shown earlier establish the result. ∎

As will become evident, the following lemma shows that the smaller the number of the rich, the more wealth their offspring inherit in a steady state.

Lemma 3. *Under* (29), $\underline{b}(\lambda) \equiv \max\{b(\lambda) \cap [0, \bar{b})\}$ *and* $\bar{b}(\lambda) \equiv \min\{b(\lambda) \cap (\underline{b}, \infty)\}$ *are continuous single-valued functions on* $(0, 1]$ *such that* $\underline{b}(1) = \underline{b}$, $\bar{b}(1) = \bar{b}$, $\underline{b}'(\lambda) > 0$ *and* $\bar{b}'(\lambda) < 0$ $\forall \lambda \in (0, 1]$.

Proof. Using $\phi_\kappa^R(b_t^i, \kappa^i) < 0$, we get $\partial\psi^{Ri}(b_t^S, b_t^L)/\partial b_t^i < \phi_b^R(b_t^i, \kappa^i) = \beta R(\kappa^i)$, where $\lim_{b_t^i \to \infty} R(\kappa^i) = 0$. Since $k_\lambda(b_t^i, 0, \lambda) > 0$ and $\psi^{Ri}(0, 0) < 0$ as shown earlier, the results follow except for the continuity. For the continuity of $\bar{b}(\lambda)$, it is enough to prove

---

[17] $A(\alpha, \beta, \theta)$ is a level of A that guarantees a unique, nontrivial, steady-state equilibrium of the autonomous system $b_{t+1} = \beta[Y(b_t, 0, 1) - \theta]$.

that $\beta R(k(\bar{b}(\lambda), 0, \lambda)) \leq 1$, $\forall \lambda \in (0, 1]$. Since $\bar{b}$ is a steady-state value higher than $\underline{b}$, $\beta R(k(\bar{b}(1), 0, 1)) < 1$. Then the result follows from

$$dk(\bar{b}(\lambda), 0, \lambda))/d\lambda = \frac{\bar{b}'(\lambda)[\beta R(k(\bar{b}(\lambda), 0, \lambda)) - 1]}{-\phi^R_\kappa(b^i_t, \kappa^i)}, \tag{31}$$

$\forall \lambda \in (0, 1]$. The continuity of $\underline{b}(\lambda)$ is similarly proven. ∎

Let $\underline{b}^i \equiv \underline{b}(\lambda^i)$, $\bar{b}^i \equiv \bar{b}(\lambda^i)$, $b^i \equiv b(\lambda^i)$, $b^i_{\min} \equiv \min b^i$ and $b^i_{\max} \equiv \max b^i$. Then Lemma 3 asserts

$$0 < b^i_{\min} \leq \underline{b}^i < \underline{b} < \bar{b} < \bar{b}^i \leq b^i_{\max}.$$

Also, $\bar{b}^L \leq \bar{b}^S$ follows from $\lambda^S \leq \lambda^L$.

COROLLARY 1. *(a) If $h'''(e) \geq 0$ $\forall e > 0$, then $e''(k_{t+1}) < 0$ $\forall k_{t+1} > 0$; (b) If $e''(k_{t+1}) \leq 0$ $\forall k_{t+1} > 0$, then $b(\lambda) = \{\underline{b}(\lambda), \bar{b}(\lambda)\}$.*

Proof. (*a*) Follows from $e'(k_{t+1}) = -h'(e(k_{t+1}))/[k_{t+1}h''(e(k_{t+1}))]$. (*b*) Follows from the fact that $\psi^{Ri}(b^S_t, b^L_t)$ is strictly concave in b^i_t under the condition. ∎

It would be reasonable to suppose that the desirable level of education exhibits a concave reaction to the factor ratio, and this condition greatly simplifies the exposition of the following lemma. Recall that μ^j stands for the inverse function of φ^i.

LEMMA 4. *If (29) is satisfied and $e''(k) \leq 0$ for all $k > 0$, there exists a continuous single-valued function $b^j_t = \xi^j(b^i_t) \geq 0$ on $\mathbb{R}_{++} \backslash (\underline{b}^i, \bar{b}^i)$ such that*

i. *$\xi^j(b^i_t) = \mu^j(b^i_t)$ if and only if $b^i_t = b^i_a$ or b^i_b, where $b^i_a \in (0, \underline{b})$, $b^i_b \in (\bar{b}, \infty)$,*

$$\begin{aligned} d\xi^j(b^i_a)/db^i_t &< d\mu^j(b^i_a)/db^i_t, \\ d\xi^j(b^i_b)/db^i_t &> d\mu^j(b^i_b)/db^i_t; \end{aligned}$$

ii.

$$\psi^{Ri}(b^S_t, b^L_t) - b^i_t \begin{cases} < 0 & \text{if} \quad b^j_t < \xi^j(b^i_t) \quad \text{or} \quad b^i_t = 0 \\ = 0 & \text{if} \quad b^j_t = \xi^j(b^i_t) \\ > 0 & \text{otherwise.} \end{cases} \tag{32}$$

Proof. See the Appendix. ∎

Figure 3 depicts the properties of the function $\xi^L(b^S_t)$. Note that $(b^S_t, \xi^L(b^S_t)) \notin E^S \cap Z^S$ because $\psi^{RS}(b^S_t, \xi^L(b^S_t)) = b^S_t \geq 0$. It thus follows from (32) that any pair $(b^S_t, \xi^L(b^S_t)) \in E^S$ is an element of the BB^S locus, as indicated by the solid black lines. The slope of $\xi^L(b^S_t)$ is zero on the b^S_t axis. $b^S_t = b^S_a$ and b^S_b at point a and b respectively.

Without the concavity of $e(k_{t+1})$, it might be the case that $\bar{b}^i \neq b^i_{\min}$ or $\bar{b}^i \neq b^i_{\max}$ and that some discrete sets, on which $\xi^j(b^i_t)$ does not exist, emerge on the interval $[\bar{b}^i, b^i_{\max}]$ or $[\bar{b}^i, b^i_{\min}]$. Nevertheless, the basic results of the present paper can be maintained, as will be shown later.

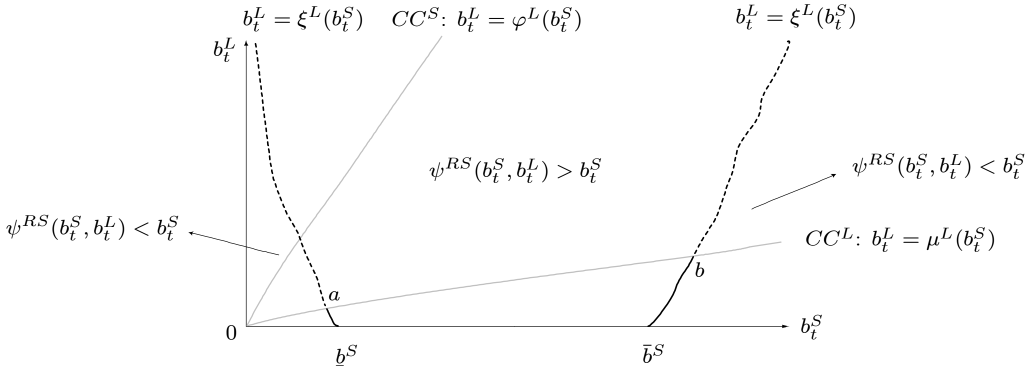

FIGURE 3. Discrete Portions of the BB^S Locus (a)

Notes. The solid black lines featured are part of the BB^S locus. The lines $\xi^L(b_t^S)$ divide the space into three regions: $\psi^{RS}(b_t^S, b_t^L) > b_t^S$ in one of them and $\psi^{RS}(b_t^S, b_t^L) < b_t^S$ in the others.

3.1.2. $(b_t^S, b_t^L) \in E^* \backslash Z^i$ — Egalitarian Distribution. We next describe the BB^i locus that belongs to $E^* \backslash Z^i$, where resources are distributed in a relatively equal fashion. In light of (27), we define

$$\phi^*(b_t^i, \kappa^*) \equiv \beta[w(\kappa^*)h(e(\kappa^*)) + (b_t^i - e(\kappa^*))R(\kappa^*) - \theta] \equiv \psi^{*i}(b_t^S, b_t^L), \tag{33}$$

where $\kappa^* = \kappa^*(b_t^S, b_t^L)$. Then $b_{t+1}^i = \psi^{*i}(b_t^S, b_t^L)$ for $(b_t^S, b_t^L) \in E^* \backslash Z^i$.

LEMMA 5. $\psi^{*i}(b_t^i, b_t^i) = \beta[Y(b_t^i, 0, 1) - \theta] = \phi^R(b_t^i, k(b_t^i, 0, 1))$.

Proof. If $b_t^S = b_t^L = b_t^i$, $I_{t+t}^S = I_{t+1}^L = Y_{t+1} = Y(b_t^i, 0, 1)$ from (21) and (39), noting $(b_t^i, b_t^i) \in E^*$. Thus (6) yields the first equality. The second equality is shown in the proof of Lemma 2. ∎

This lemma is quite meaningful. The evolution of transfers starting with $b_0^S = b_0^L$ is now expressed by the system $b_{t+1}^i = \phi^R(b_t^i, k(b_t^i, 0, 1))$. Thus, the steady-state values of this system, $(\underline{b}, \underline{b})$ and $(\bar{b}, \bar{b})$, are in fact the steady-state values for Regime 3.

LEMMA 6. *Under condition* (29)*, there exists a continuous function* $b_t^j = \zeta^j(b_t^i) \gtreqless 0$ *on* $\mathbb{R}_+ \backslash (\underline{b}, \bar{b})$ *such that*

i. *If* $b_t^i \neq 0, \underline{b}$ *or* $\bar{b}$*, then* $\zeta^j(b_t^i)$ *is double-valued and*

$$\min \zeta^j(b_t^i) < b_t^i < \max \zeta^j(b_t^i),$$
$$d\zeta^j(b_t^i)/db_t^i < \infty.$$

ii. *If* $b_t^i = 0, \underline{b}$ *or* $\bar{b}$*,* $\zeta^j(b_t^i)$ *is single-valued;* $\zeta^j(0) > 0$ *and* $\bar{b} = \zeta^j(\bar{b}) > \underline{b} = \zeta^j(\underline{b}) > 0$*. The pairs* $(\underline{b}, \zeta^j(\underline{b}))$ *and* $(\bar{b}, \zeta^j(\bar{b}))$ *are the bifurcation points such that*

$$\begin{aligned} \lim_{b_t^i \to \underline{b} - 0} d\zeta^j(b_t^i)/db_t^i &= \infty, \\ \lim_{b_t^i \to \bar{b} + 0} d\zeta^j(b_t^i)/db_t^i &= \infty. \end{aligned}$$

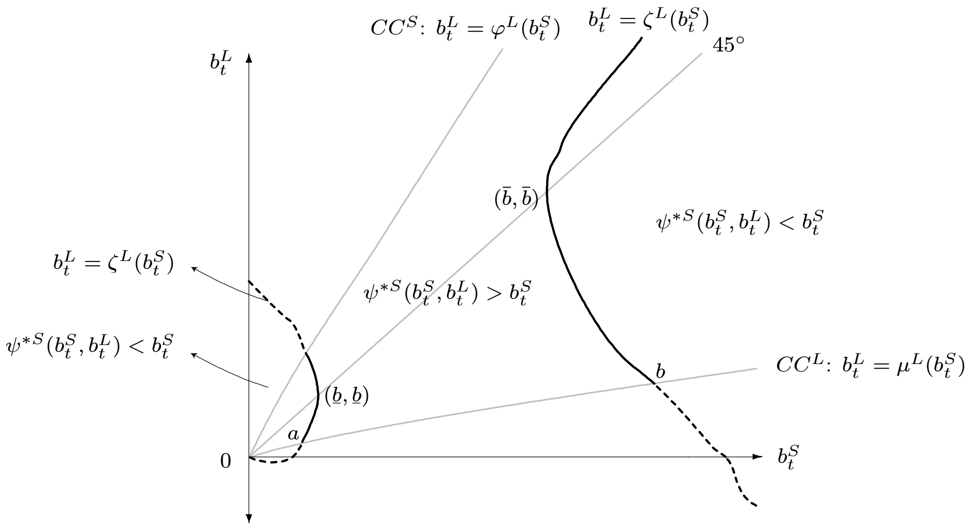

FIGURE 4. Discrete Portions of the BB^S Locus (b)

Notes. The solid black lines featured are part of the BB^S locus. The lines $\zeta^L(b_t^S)$ divide the space into three regions: $\psi^{*S}(b_t^S, b_t^L) > b_t^S$ in one of them and $\psi^{*S}(b_t^S, b_t^L) < b_t^S$ in the others. $\xi^L(b_t^S)$ and $\zeta^L(b_t^S)$ coincide at points a and b on CC^S.

iii. $\zeta^j(b_t^i) = \mu^j(b_t^i)$ *if and only if* $b_t^i = b_a^i \in (0, \underline{b})$ *or* $b_b^i \in (\bar{b}, \infty)$, *where*

$$\begin{aligned} d\zeta^j(b_a^i)/db_t^i &> d\mu^j(b_a^i)/db_t^i, \\ d\zeta^j(b_b^i)/db_t^i &< d\mu^j(b_b^i)/db_t^i. \end{aligned}$$

iv.

$$\psi^{*i}(b_t^S, b_t^L) - b_t^i \begin{cases} < 0 & \textit{if} \quad \min \zeta^j(b_t^i) < b_t^j < \max \zeta^j(b_t^i), \\ = 0 & \textit{if} \quad b_t^j = \zeta^j(b_t^i), \\ > 0 & \textit{otherwise}. \end{cases} \tag{34}$$

Proof. See the Appendix. ∎

The second property of the lemma implies that there exists a value of b_t^i in the interval $(0, \underline{b})$ such that $\varphi^j(b_t^i) = \zeta^j(b_t^i)$. Figure 4 depicts the properties of the function $\zeta^L(b_t^S)$. $(\underline{b}, \underline{b})$ and $(\bar{b}, \bar{b})$ are the bifurcation points, and the slope of $\zeta(b_t^S)$ is vertical at these points. Note that $(b_t^S, \zeta^L(b_t^S)) \notin E^* \cap Z^S$ because $\psi^{*S}(b_t^S, \zeta^L(b_t^S)) = b_t^S \geq 0$. Therefore, any pair $(b_t^S, \zeta^L(b_t^S)) \in E^*$ is an element of the BB^S locus, as illustrated by the solid black lines. $b_t^S = b_a^S$ and b_b^S at points a and b respectively. Recall (20) to get $\kappa^L(b_t^S, b_t^L) = \kappa^*(b_t^S, b_t^L)$ and $\psi^{RS}(b_t^S, b_t^L) = \psi^{*S}(b_t^S, b_t^L)$ on CC^L. Thus, if $\zeta(b_t^S)$ were to be depicted in the figure, $\xi^L(b_t^S)$ and $\zeta^L(b_t^S)$ would coincide at points a and b.

3.1.3. $(b_t^S, b_t^L) \in E^j \backslash Z^i$ — Distribution Biased to the Other Group. Thirdly, consider the dynamics of b_t^i when wealth distribution is biased toward group j. In light of (27), let

$$\phi^P(b_t^i, \kappa^j(b_t^S, b_t^L)) \equiv \beta[w(\kappa^j(b_t^S, b_t^L))h(b_t^i) - \theta] \equiv \psi^{Pi}(b_t^S, b_t^L), \tag{35}$$

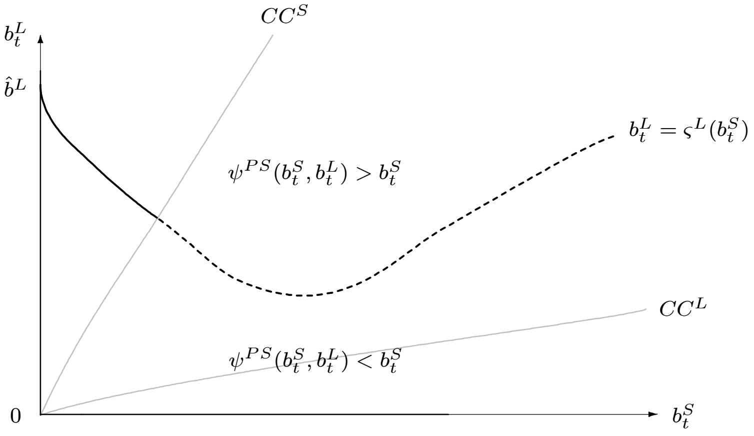

FIGURE 5. A Discrete Portion of the BB^S Locus (c)

Notes. The solid black line featured is part of the BB^S locus. The line $\varsigma^L(b_t^S)$ divides the space into two regions, where $\psi^{PS}(b_t^S, b_t^L) > b_t^S$ and $\psi^{PS}(b_t^S, b_t^L) < b_t^S$ respectively. $\varsigma^L(b_t^S)$ and $\zeta^L(b_t^S)$ intersect on CC^S.

where $(b_t^S, b_t^L) \in \mathbb{R}_+^2$. Then $b_{t+1}^i = \psi^{Pi}(b_t^S, b_t^L)$ for $(b_t^S, b_t^L) \in E^j \backslash Z^i$.

LEMMA 7. *Under* (29)*, there exists a continuous single-valued function* $b_t^j = \varsigma^j(b_t^i) > 0$ *on* $\mathbb{R}_+$ *such that*

i.

$$\lim_{b_t^i \to 0} d\varsigma^j(b_t^i)/db_t^i = -\infty.$$

ii. *There exists a value of* b_t^i *such that* $\varsigma^j(b_t^i) = \varphi^j(b_t^i)$.

iii.

$$\psi^{Pi}(b_t^S, b_t^L) - b_t^i \begin{cases} < 0 & \textit{if} \quad b_t^j < \varsigma^j(b_t^i), \\ = 0 & \textit{if} \quad b_t^j = \varsigma^j(b_t^i), \\ > 0 & \textit{if} \quad b_t^j > \varsigma^j(b_t^i). \end{cases} \tag{36}$$

Proof. ii. Lemma 6 implies that $b_t^i = \psi^{*i}(b_t^S, b_t^L)$ if $b_t^j = \zeta^j(b_t^i)$, and that there exists a value of b_t^i such that $\zeta^j(b_t^i) = \varphi^j(b_t^i)$. Also notice that, from (20), $\psi^{*i}(b_t^S, b_t^L) = \psi^{Pi}(b_t^S, b_t^L)$ if $b_t^j = \varphi^j(b_t^i)$. Hence $\psi^{Pi}(b_t^S, b_t^L) = b_t^i$ if $b_t^j = \varphi^j(b_t^i) = \zeta^j(b_t^i)$. This establishes the result. See the Appendix for the other parts. ▮

Let $\hat{b}^j \equiv \varsigma^j(0)$; that is, $w(k(\hat{b}^j, 0, \lambda^j)) = \theta$. $\hat{b}^j$ is a critical level for universal human capital accumulation. If members of group j, the rich, bequeath wealth more than $\hat{b}^j$, the resultant high wage level induces members of group i, the unskilled, to transfer wealth to their offspring. Figure 5 depicts the properties of the function $\varsigma^L(b_t^S)$ and indicates $\hat{b}^L$. Note that $(b_t^S, \varsigma^L(b_t^S)) \notin E^L \cap Z^S$ because $\psi^{PS}(b_t^S, \varsigma^L(b_t^S)) = b_t^S \geq 0$. Therefore, any pair $(b_t^S, \varsigma^L(b_t^S)) \in E^L$ is an element of the BB^S locus, as illustrated by the solid black line. The slope of $\varsigma^L(b_t^S)$ is vertical at point $(0, \hat{b}^L)$. Recall (20) to get

$\kappa^L(b_t^S, b_t^L) = \kappa^*(B_t)$ and $\psi^{PS}(b_t^S, b_t^L) = \psi^{*S}(b_t^S, b_t^L)$ on CC^S. Thus, if $\zeta(b_t^S)$ were to be represented in the diagram, $\varsigma^L(b_t^S)$ and $\zeta^L(b_t^S)$ would coincide on CC^S.

Lemma 8. *There exists a continuous single-valued function* $\hat{b}(\lambda) > 0$ *on* $(0, 1]$ *such that* $w(k(\hat{b}(\lambda), 0, \lambda)) = \theta$ *and* $\hat{b}'(\lambda) < 0 \ \forall \lambda \in (0, 1]$.

Proof. Noting (2), we define

$$Q(b, \lambda) \equiv A[k(b, 0, \lambda)]^\alpha - \theta/(1 - \alpha).$$

The properties of $k(\cdot)$ assert that $Q_b(b, \lambda) > 0 \ \forall (b, \lambda) \in \mathbb{R}_{++} \times (0, 1]$, and that $Q(0, \lambda) < 0$ and $\lim_{b \to \infty} Q(b, \lambda) = \infty \ \forall \lambda \in (0, 1]$. Then there exists a single-valued function $\hat{b}(\lambda) > 0$ such that $Q(\hat{b}(\lambda), \lambda) = 0$. $\hat{b}'(\lambda) < 0$ follows from $Q_\lambda(b, \lambda) > 0 \ \forall (b, \lambda) \in \mathbb{R}_{++} \times (0, 1]$. ∎

As follows from the above lemma, we get $\hat{b}^j = \hat{b}(\lambda^j)$ and $\hat{b}^L \le \hat{b}^S$.

The property $\hat{b}'(\lambda) < 0$ can be thought of in the following manner. An increase in the richer group's transfer raises the capital-skill ratio and thus the wage level, w_{t+1}. Hence a smaller number of richer members in society requires each of them to save a greater amount of transfers, b_t^j, in order to obtain $w_{t+1} = \theta$.

One can easily demonstrate that A and α respectively have negative and positive correlations with $\hat{b}(\lambda)$. Due to the complementarity between technology and skill, a more productive technology increases the wage for a given k_{t+1} and thus reduces $\hat{b}(\lambda)$ for all λ. An increase in α, the capital share, implies a greater dependance of production on physical capital, and human capital becomes less important in production. The wage level then declines for a given k_{t+1}, and accordingly a higher level of b_t is required to satisfy $w_{t+1} = \theta$. As α approaches 1, w_{t+1} goes to zero and therefore $\hat{b}(\lambda)$ must go to infinity for all λ.

3.1.4. $(b_t^S, b_t^L) \in Z^i$ — Absolute Poverty. We are now in a position to describe the BB^i locus for the fourth (and final) case, $(b_t^S, b_t^L) \in Z^i$. This is the situation in which members i of generation t are not wealthy enough to leave any amount of transfers to their descendants. Obviously,

$$b_{t+1}^i - b_t^i \begin{cases} < 0 & \text{if } (b_t^S, b_t^L) \in Z^i \text{ and } b_t^i > 0, \\ = 0 & \text{if } (b_t^S, b_t^L) \in Z^i \text{ and } b_t^i = 0. \end{cases} \tag{37}$$

3.2. *Conditional Dynamics*

Figure 6 constructs the entire BB^S locus by combining Figures 4-5. Once the set Z^i is characterized, the direction of the motion of b_t^i can be described by using (34), (32), (36) and (37). The following technical results are useful to identify Z^i on the (b_t^S, b_t^L) space.

Lemma 9. $(b_t^S, b_t^L) \in Z^i \cap BB^i$ *if and only if* $b_t^i = 0$ *and* $b_t^j \in [0, \hat{b}^j)$.

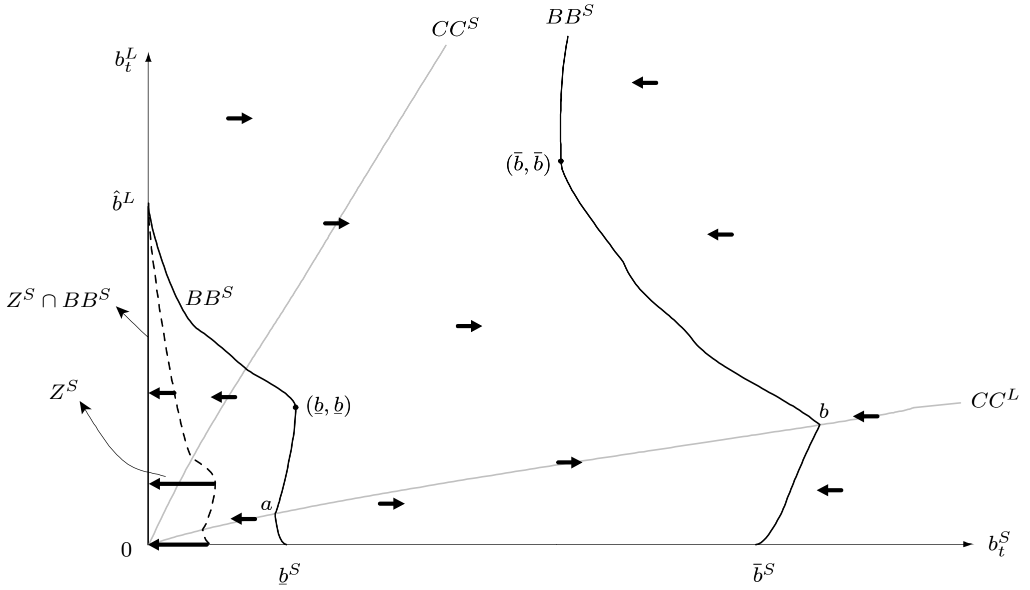

FIGURE 6. The Dynamics of b_t^S

Notes. The figure is depicted by integrating Figures 4-5. The black thick lines illustrate the entire BB^S locus, and $I_{t+1}^S = \theta$ on the dotted line. $(b_t^S, b_t^L) \in Z^S \cap BB^S$ if $b_t^S = 0$ and $b_t^L \in [0, \hat{b}^L)$. $b_t^S = b_a^S, b_b^S$ at points a and b respectively.

Proof. Sufficiency: Suppose $(b_t^S, b_t^L) \in Z^i \cap BB^i$. By definition, $I_{t+1}^i < \theta$ and $b_{t+1}^i = b_t^i$. Since (6) yields $b_{t+1}^i = 0$, we find $b_t^i = 0$ and $I_{t+1}^i = w(k(b_t^j, 0, \lambda^j)) < \theta$. Then $\partial k(\cdot)/\partial b_t^j > 0$ shows $b_t^j < \hat{b}^j$.

Necessity: Suppose $b_t^i = 0$ and $b_t^j \in [0, \hat{b}^j)$. Then $I_{t+1}^i = w(k(b_t^j, 0, \lambda^j)) < \theta$, meaning $b_{t+1}^i = 0$. Hence $(b_t^S, b_t^L) \in Z^i \cap BB^i$. ∎

LEMMA 10. *$\partial I^i(b_t^S, b_t^L)/\partial b_t^i > 0$ for all $b_t^i > 0$.*

Proof. See the Appendix. ∎

The two lemmas reveal that, in Figure 6, the dotted line on which $I_{t+1}^S = \theta$ exists between the b_t^L axis and the BB^S locus passing through $(\underline{b}, \underline{b})$, and Z^S appears on the left-hand side of the dotted line. Accordingly, one can indicate the directions of b_t^S by the short arrows and can use the long arrows to show that b_{t+1}^S jumps to zero. Further characterizations of Z^i provide no additional information on the qualitative nature of the dynamical system and one can neglect them without loss of generality.

3.3. *Global Dynamics*

The previous results enable us to analyze the global dynamic behavior of transfers. By utilizing the phase diagrams, we demonstrate that the initial distribution of transfers between the two groups determines both short-run and long-run economic performances. To conduct the analysis, it would be reasonable to assume

$$\lim_{\lambda \to 0} k(\bar{b}(\lambda), 0, \lambda) \leq \hat{k} < k(\bar{b}, 0, 1), \tag{38}$$

where $\hat{k} \equiv w^{-1}(\theta) = [\theta/(1-\alpha)A]^{1/\alpha}$, $\bar{b} = \bar{b}(1)$, and $k(\bar{b}(\lambda), 0, \lambda)$ is strictly increasing in λ as follows from (31). The assumption is feasible because $\bar{b}(\lambda)$ and thus $k(\bar{b}(\lambda), 0, \lambda)$ increase with A for all $\lambda \in (0, 1]$. The first inequality in (38) states that the technology is not productive enough to nullify the role of distribution; if the inequality is violated, parental transfers eventually occur in all households regardless of their initial asset holdings. The second inequality states that unskilled workers who own no assets would transfer wealth to their offspring if they were in the higher-level steady state in Regime 3.

It will become apparent that the magnitude of the gap between $\bar{b}^i \equiv \bar{b}(\lambda^i)$ and $\hat{b}^i \equiv \hat{b}(\lambda^i)$ changes the nature of the dynamical system.

Lemma 11. *Under* (38), *there exists a unique* $\lambda^\circ \in (0, 1)$ *such that*

$$\hat{b}(\lambda)/\bar{b}(\lambda) \begin{cases} > 1 & \textit{for } \lambda \in (0, \lambda^\circ) \\ = 1 & \textit{for } \lambda = \lambda^\circ \\ < 1 & \textit{for } \lambda \in (\lambda^\circ, 1]. \end{cases}$$

Proof. By noting $\hat{k} = k(\hat{b}(\lambda), 0, \lambda)$, (38) assures the existence of a λ such that $\hat{b}(\lambda) = \bar{b}(\lambda)$. The other results follow from $\hat{b}'(\lambda) < \bar{b}'(\lambda) < 0$ if $\hat{b}(\lambda) = \bar{b}(\lambda)$, which is shown by simple calculations. ∎

Remark 1. *By noting* $\lambda^S = 1 - \lambda^L < 0.5$, *the corollary rules out the simultaneous occurrence of* $\hat{b}^S/\bar{b}^S < 1$ *and* $\hat{b}^L/\bar{b}^L \geq 1$.

Among several possible cases, first consider $\hat{b}^S \geq \bar{b}^S$ and $\hat{b}^L < \bar{b}^L$ as a benchmark. Figure 7 illustrates the global behavior of the dynamical system. The diagram is the result of the combination of Figure 6, which features group S, and the hypothetical figure which features group L. Such a combination is formed by means of three simplifications. First, the diagram omits the range of absolute poverty, $Z^S \cup Z^L$, on the grounds that the omission does not affect the direction of motion of transfers, as shown in Figure 6. Second, the BB locus is drawn to be gradual, and this way of drawing rules out some steady states that otherwise would exist. Third, we assume the strict concavity of $e(\cdot)$ so that some potential steady states are neglected, as discussed previously. It will be clear that the qualitative nature of the system is robust to alternative assumptions.

Steady-state equilibria occur at the intersections of BB^S and BB^L. Among them, $(0, 0)$, $(\bar{b}^S, 0)$, and $(\bar{b}, \bar{b})$ are locally stable, whereas all the others are unstable. The system therefore exhibits the multiplicity of nontrivial, locally stable, steady-state equilibria, implying that the distribution of transfers determines the steady-state point to which the economy converges. Observe that steady-state equilibrium does not occur

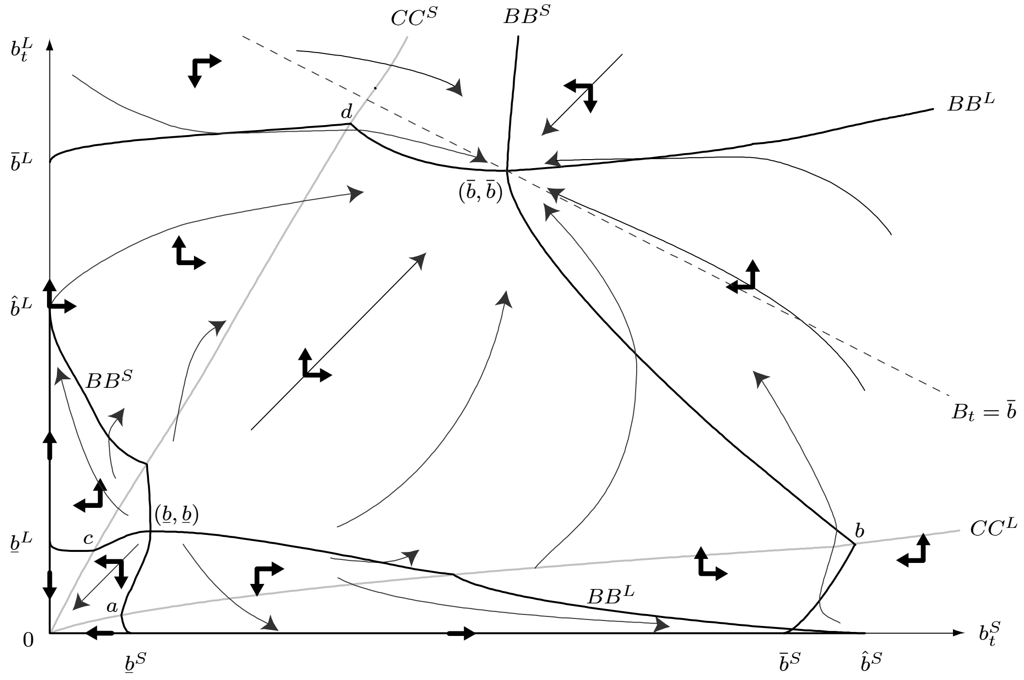

Figure 7. The Joint Evolution of Transfers ($\hat{b}^S \geq \bar{b}^S$ and $\hat{b}^L < \bar{b}^L$)

Notes. The figure describes the evolution of the transfers for the two groups whose sizes differ substantially. $b_t^S = b_a^S, b_b^S$ at the points a and b, and $b_t^L = b_a^L, b_b^L$ at c and d respectively. The system exhibits multiplicity of locally stable steady-state equilibria.

at $(0, \bar{b}^L)$. Also, note that the transitional behavior of b_t^i is not necessarily monotonic, and that the economy enter may enter Regime 2 from Regime 3.

It is worth while mentioning that the diagram illustrates the growth path proposed by Galor and Moav (2001). According to their scenario, the economy starts out with the pair of $b_0^L > \underline{b}^L$ and $b_0^S = 0$. Hence the initial state is Regime 1 where b_t^L increases over time and members in group S remain uneducated. When b_t^L exceeds $\hat{b}^L$, the economy enters Regime 2 and the level of b_t^S begins to ascend. At this stage, members in group S receive education in the presence of liquidity constraints. The economy finally reaches Regime 3 where the constraints are no longer binding, converging towards $(\bar{b}, \bar{b})$.

As established in Lemma 3, the economy is characterized by $\underline{b}^S < \underline{b}^L < \underline{b}$ and $\bar{b} < \bar{b}^L < \bar{b}^S$, which holds independently of the ratio $\hat{b}^i/\bar{b}^i$. In light of Figure 7, the property $\bar{b} < \bar{b}^L < \bar{b}^S$ implies that unequal distribution is beneficial for members of group S in the long run, as well as in the short run. However, it is proven below that unequal distribution is undesirable for the long-run performance of economic growth.

Proposition 1. *Under* (29), $\lambda^i \bar{b}^i < \bar{b}$ *and* $Y(\lambda^i \bar{b}^i, 0, \lambda^i) < Y(\bar{b}, 0, 1)$. *That is to say, the egalitarian steady-state pair,* $(\bar{b}, \bar{b})$, *generates higher levels of aggregate transfers and output than the non-egalitarian steady-state pairs,* $(\bar{b}^S, 0)$ *and* $(0, \bar{b}^L)$.

Proof. Consider Figure 7, where $\zeta^j(\bar{b}) = \bar{b}$ and $\zeta^j(b_t^i)$ is vertically sloped at $b_t^i = \bar{b}$, as proved by Lemma 6. Moreover, the proof of Lemma 2 implies $\beta R(k(\bar{b}, 0, 1)) < 1$, and the use of (59) and (63) in the Appendix establishes

$$d\zeta^j(b_t^i)/db_t^i < -\lambda^i/\lambda^j \qquad \text{if } \zeta^j(b_t^i) < b_t^i \text{ and } \beta R(k(\lambda^i b_t^i + \lambda^j \zeta^j(b_t^i), 0, 1)) < 1.$$

Combining these results reveals that in the (b_t^i, b_t^j) space, $(\bar{b}, \bar{b})$ is the unique intersection of $\zeta^j(b_t^i)$ and the isoquant for $B_t = \bar{b}$, whose slope is $-\lambda^i/\lambda^j$. Hence $\lambda^i b_b^i + \lambda^j \zeta^j(b_b^i) < \bar{b}$. Since Lemma 4 shows that all elements of b^i must be between b_a^i and b_b^i, we now find $\lambda^i \bar{b}^i < \lambda^i b_b^i < \bar{b}$. The second result thus follows from (23). ∎

The proposition requires condition (29) merely for the existence of the steady-state equilibria. To grasp the implication of the proposition, suppose that society is endowed with a sufficient amount of wealth $B_0 \in (\underline{b}, \lambda^S \bar{b}^S]$. It then follows from the proposition that the economy ends up stagnating in Regime 1 if all resources are owned by members of group S. Transfers in this case converge to $(\bar{b}^S, 0)$ and members of group L remain unskilled in all subsequent periods. Inequality hence hampers the onset of universal human capital accumulation and the process of development. On the other hand, under more equal distribution such as $b_0^S = 0$ and $b_0^L > 0$, the wage income eventually exceeds the critical level θ and the economy evolves toward the egalitarian steady state with higher aggregate transfers and output. Note that the economy does not even go through Regime 1 under even more equal distribution, $(b_0^S, b_0^L) \gg (\underline{b}, \underline{b})$.

The effect of distribution changes qualitatively in the case of scarce resources, $B_0 \in [\lambda^S \underline{b}^S, \underline{b}]$. Now one can see that egalitarian distributions may result in the recession toward no resources, $(0, 0)$, as all people would spend a large fraction of their wealth on consumption and education. While positive amounts of output are sustained in unequal society, high inequality tends to lead the economy to either $(\bar{b}^S, 0)$ or $(0, \bar{b}^L)$, rather than the egalitarian steady-state equilibrium, $(\bar{b}, \bar{b})$.

To summarize, high inequality lowers the growth performance of wealthy countries in the long run, whereas it would prevent less developed countries from degenerating. This finding is in correlation with recent studies that employ the capital market imperfection approach in the literature on inequality and growth.

Now turn to the second possible case, $\hat{b}^S \geq \bar{b}^S$ and $\hat{b}^L \geq \bar{b}^L$. Unlike $(0, \bar{b}^L)$ in Figure 7, a locally stable steady-state equilibrium occurs at $(0, \bar{b}^L)$ in this case. Although the long-run performance of the economy is more sensitive to the initial distribution than in the benchmark case, one can find the qualitatively same effect of distribution, which depends on the scarcity of initial resources.

For the last case, $\hat{b}^S < \bar{b}^S$ and $\hat{b}^L < \bar{b}^L$, neither $(\bar{b}^S, 0)$ nor $(0, \bar{b}^L)$ is a steady-state equilibrium. Hence, the role of distribution becomes less important; regardless of the initial distribution, any economy endowed with $B_0 > \underline{b}$ converges to the egalitarian steady-state equilibrium, $(\bar{b}, \bar{b})$.

3.4. *Redistribution and Welfare*

Consider now redistribution of wealth, B_0, by controlling λ^S as well as b_0^S. A few remarks deserve to be made at this point. First, by definition, λ^L and b_0^L are uniquely determined for a given set of λ^S, b_0^S and B_0. Second, the degree of inequality is changed by the relative size of the two groups as well as the distribution of wealth across the two groups. With a small value of λ^S, one can execute more drastic redistribution of wealth; distribution biased toward group S allots the small fraction of population the large portion of national wealth, generating high inequality. Third, and finally, changing the group size affects the ratio $\hat{b}^S/\bar{b}^S$ and $\hat{b}^L/\bar{b}^L$ and the nature of the dynamical system.

If a sufficiently small level of λ^S is chosen, Lemma 11 suggests that the system is characterized by the benchmark case presented previously, and the remaining cases do not take place. It is shown that the initial distribution of wealth determines the economy's long-run performance when the amount of wealth is neither extremely abundant nor scarce: $B_0 \in (\lambda^S \underline{b}^S, \lambda^S \bar{b}^S]$. Recall that the egalitarian steady-state level $(\bar{b}, \bar{b})$ is independent of the choice of λ^S.

On the other hand, if moderate redistribution is executed by choosing a sufficiently large level of λ^S, λ^S/λ^L approaches one, and the heterogeneity across the two groups generates the dynamical system characterized by either of the last two cases. However, the last case, $\hat{b}^S < \bar{b}^S$ and $\hat{b}^L < \bar{b}^L$, may not occur under (38). The feasibility of the last case is assured under a stronger condition on the structural parameters, such as a higher level of A.

The discussion in Subsections 3.3-3.4 can be summarized as follows.

THEOREM 1. *Under* (38), *highly unequal distribution of initial wealth is desirable from the viewpoint of the richer members and their offspring, but not from the long-term viewpoint of society.*

4. OUTPUT GROWTH

The preceding analysis has fully revealed the process of economic development in terms of intergenerational transfers. By considering the underlying evolution of output, one can examine the impact of initial wealth distribution on the behavior of output growth.

This section uses subscripts i and j, where $i, j = S, L$ and $i \neq j$, to indicate the richer and poorer groups, respectively. It is assumed that income distribution at time 0 is highly unequal in the absence of redistribution. More accurately, $h_0^i = h(e(k_0))$ and adult individuals of group j initially possess neither advanced skills nor physical capital.

4.1. *Aggregate Transfers*

At the outset, consider the economy is in Regimes 2 or 3 in period t, $(b_t^S, b_t^L) \in \mathbb{R}_{++}^2$. As follows from (1), (2) and (14),

$$Y_t = w_t H_t + R_t K_t = \lambda^S I_t^S + \lambda^L I_t^L. \tag{39}$$

Since either $(b_t^S, b_t^L) \gg 0$ or $(b_t^S, b_t^L) = (0, 0)$ is the case in both Regimes, (6) and (39) prove that aggregate transfers in Regimes 2 and 3 are

$$B_t = \max[\beta(Y_t - \theta), 0]. \tag{40}$$

Next suppose that the economy is in Regime 1 in period $t-1$. Then skilled workers with human capital $h(e(k_t))$ are a fraction λ^i of adult individuals in period t, and unskilled workers are $1-\lambda^i$ of adult individuals in period t. Noting $h_t^j = h(0) = 1$, the output in period t is

$$Y_t = Ak_t^{\alpha}[\lambda^i h(e(k_t)) + 1 - \lambda^i].$$

Given the properties of $e(k_t)$ and $h(k_t)$, there exists a unique k_t for a given pair (Y_t, λ^i). It follows that

$$k_t = \varkappa(Y_t, \lambda^i),$$

where $\varkappa(Y_t, \lambda)$ is a single-valued function on $\mathbb{R}_+ \times (0, 1]$ such that $\varkappa(0, \lambda) = 0$, $\lim_{Y_t \to \infty} \varkappa(Y_t, \lambda) = \infty \ \forall \lambda > 0$, and $\varkappa_Y(Y_t, \lambda) > 0$ and $\varkappa_\lambda(Y_t, \lambda) < 0 \ \forall (Y_t, \lambda) \in \mathbb{R}_{++} \times (0, 1]$.[18] Then the wage rate can be written as $w_t = w(\varkappa(Y_t, \lambda)) \equiv \omega(Y_t, \lambda)$, which preserves the above properties of $\varkappa(Y_t, \lambda)$.[19] Unskilled workers (the poorer members) receive $I_t^j = \omega(Y_t, \lambda^i)$, and as follows from (39), skilled workers receive

$$I_t^i = [Y_t - \lambda^j \omega(Y_t, \lambda^i)]/\lambda^i = \omega(Y_t, \lambda^i) h(e(\varkappa(Y_t, \lambda^i))) + \alpha Y_t/\lambda^i, \tag{41}$$

where αY_t is the return on savings, $R_t K_t$, noting (1) and (2). Hence, all workers' incomes strictly monotonically increase with output, and will be zero if no output is produced.

Corollary 2. *There exist single-valued functions* $\breve{Y}(\lambda)$ *and* $\hat{Y}(\lambda)$ *on* $\lambda \in (0, 1]$ *such that*

$$[\breve{Y}(\lambda) - (1 - \lambda)\omega(\breve{Y}(\lambda), \lambda)]/\lambda = \omega(\hat{Y}(\lambda), \lambda) = \theta. \tag{42}$$

Furthermore, noting (41)*,*

i. $\breve{Y}'(\lambda) > 0$ *and* $\hat{Y}'(\lambda) > 0$ *for all* $\lambda \in (0, 1]$;
ii. $\lim_{\lambda \to 0} \breve{Y} = 0$ *and* $\breve{Y}(1) = \theta$;
iii. $\breve{Y}(\lambda) < \hat{Y}(\lambda)$ *for all* $\lambda \in (0, 1]$;
iv. $\hat{Y}^i \equiv \hat{Y}(\lambda^i) = Y(\lambda^i \hat{b}^i, 0, \lambda^i)$.

$\breve{Y}(\lambda)$ and $\hat{Y}(\lambda)$ are respectively the minimum levels of output that induce skilled workers who own assets and unskilled workers to transfer wealth to their offspring at time t.

The last property of Corollary 2 suggests that $b_{t-1}^j = 0$ and $Y_t \leq \hat{Y}^i$ occur if and only if $b_t^j = b_{t-1}^i = 0$. Under these conditions, unskilled workers transfer no wealth to their offspring in period t and the economy is in Regime 1 in both periods t and $t-1$.

---

[18] For reasons that will be clear shortly, $\varkappa(Y_t, \lambda)$ is defined on $\mathbb{R}_+ \times (0, 1]$ rather than $\mathbb{R}_+ \times (0, 1)$.

[19] By construction, $\omega(Y_t, 1)$ does not describe the wage rate in Regime 2.

Substituting for I_t^i from (41) into (6) for Regime 1 and recalling (40) for Regimes 2 and 3, we obtain the expression for aggregate transfers in each of the three regimes:

$$B_t = \begin{cases} B(Y_t, \lambda^i) & \text{if } b_t^j = b_{t-1}^j = 0, \\ B(Y_t, 1) & \text{if } (b_t^S, b_t^L) \in \mathbb{R}^2_{++}, \end{cases} \tag{43}$$

where

$$B(Y_t, \lambda) \equiv \begin{cases} 0 & \text{if } Y_t \in [0, \breve{Y}(\lambda)] \\ \beta[Y_t - (1-\lambda)\omega(Y_t, \lambda) - \lambda\theta] & \text{if } Y_t \geq \breve{Y}(\lambda). \end{cases} \tag{44}$$

Note that $\breve{Y}(1) = \theta$ as shown in Corollary 2 and $B_Y(Y_t, \lambda) > 0$ as follows from (41). Furthermore, if $e''(k) \leq 0$ for all $k > 0$,

$$\begin{aligned} &B_\lambda(Y_t, \lambda) \begin{cases} < 0 & \text{for } Y_t \in (\breve{Y}(\lambda), \hat{Y}(\lambda)) \\ = 0 & \text{for } Y_t = \hat{Y}(\lambda) \\ > 0 & \text{for } Y_t > \hat{Y}(\lambda); \end{cases} \\ &B_{\lambda Y}(Y_t, \lambda) > 0, \end{aligned}$$

for all $(Y_t, \lambda) \in \mathbb{R}_{++} \times (0, 1]$. The first property suggests that the inequality generated by a small fraction of the rich stimulates resource accumulation in the early stages of development, and the qualitative effect is reversed at $Y = \hat{Y}$, the critical level dividing Regimes 1 and 2.[20] The second property implies that the positive effect of inequality on aggregate transfers is more conductive at lower levels of development.

4.2. *The Evolution of Output within Regimes*

We may now proceed to the description of the evolution of output in each regime. Substitution for B_t from (43) into (21) establishes

$$\begin{aligned} Y_{t+1} &= Y(B_t, b_t^c, \lambda) \\ &= \begin{cases} Y(B(Y_t, \lambda^i), 0, \lambda^i) \equiv \Phi^{1i}(Y_t) & \text{if } b_t^j = b_{t-1}^j = 0 \\ Y(B(Y_t, 1), b_t^j, \lambda^i) \equiv \Phi^{2i}(Y_t, b_t^j) & \text{if } (b_t^S, b_t^L) \in E^i \cap \mathbb{R}^2_{++} \\ Y(B(Y_t, 1), 0, 1) \equiv \Phi^3(Y_t) & \text{if } (b_t^S, b_t^L) \in E^*, \end{cases} \end{aligned} \tag{45}$$

where

$$Y(B_t, b_t^c, \lambda) \equiv A[B_t - (1-\lambda)b_t^c - \lambda e_t]^\alpha [(1-\lambda)h(b_t^c) + \lambda h(e_t)]^{1-\alpha}.$$

$\Phi^{1i}(Y_t)$ hence determines the output at time $t+1$ on the condition that the economy remain in Regime 1 for the previous two periods.

(45) shows that one can identify the economy's current regime not from the current level of output but from the levels of transfers. In order to visualize the entire evolution of output beyond one regime, we next examine the relationship between $\Phi^{1i}(Y_t)$, $\Phi^{2i}(Y_t, b_t^j)$, and $\Phi^3(Y_t)$.

---

[20] Without the concavity of $e(k)$, $B_\lambda(Y_t, \lambda) = 0$ may not occur at a unique level of Y_t, and $B_{\lambda Y}(Y, \lambda) > 0$ may not hold on the entire domain.

4.2.1. Regime 1 and Regime 3. First consider the evolution of output in Regimes 1 and 3. It is convenient to define

$$\Phi(Y_t,\lambda) \equiv A[B(Y_t,\lambda)-\lambda e_t]^{\alpha}[1-\lambda+\lambda h(e_t)]^{1-\alpha}$$

on $\mathbb{R}_+ \times (0,1]$, whereby $\Phi(Y_t,\lambda^i)=\Phi^{1i}(Y_t)$ and $\Phi(Y_t,1)=\Phi^3(Y_t)$. The function exhibits properties such that

$$\begin{aligned}\Phi(Y_t,\lambda)&\begin{cases}=0 & \text{for } Y_t \le \breve{Y}(\lambda)\\ >0 & \text{for } Y_t > \breve{Y}(\lambda);\end{cases}\\ \Phi_Y(Y_t,\lambda)&=R(k_{t+1})B_Y(Y_t,\lambda)>0,\end{aligned} \tag{46}$$

for all $(Y_t,\lambda)\in(\breve{Y}(\lambda),\infty)\times(0,1]$, and the second derivative is negative for $\lambda=1$.[21] As will become apparent, $\Phi_\lambda(Y_t,\lambda)$, the effect of equality, is negative at lower levels of Y_t and the sign is reversed at higher levels of Y_t. The use of (18) yields the capital-skill ratio

$$k_{t+1}=k(B_t/\lambda,0,\lambda)=k(B(Y_t,\lambda)/\lambda,0,\lambda), \tag{47}$$

implying k_{t+1} is strictly monotonically increasing with Y_t. Moreover, for all $\lambda\in(0,1]$,

$$\begin{aligned}\lim_{Y_t\to\breve{Y}(\lambda)}\Phi_Y(Y_t,\lambda)&=\infty,\\ \lim_{Y_t\to\infty}\Phi_Y(Y_t,\lambda)&=0.\end{aligned}$$

The last property stems from the fact that individuals' investment in human capital is subject to diminishing marginal returns.

Lemma 12. *Under* (15), $\forall\lambda\in(0,1]$,

i. $Y=\Phi(Y,\lambda)>0$ *if and only if* $Y=Y(\lambda b(\lambda),0,\lambda)$, *where* $b(\lambda)\in b(\lambda)$.
ii. $\Phi_Y(\underline{Y}(\lambda),\lambda)>1$, $\Phi_Y(\bar{Y}(\lambda),\lambda)<1$ *and* $Y_t<\Phi(Y_t,\lambda)$ $\forall Y_t\in(\underline{Y}(\lambda),\bar{Y}(\lambda))$, *where* $\underline{Y}(\lambda)\equiv Y(\lambda\underline{b}(\lambda),0,\lambda)$ *and* $\bar{Y}(\lambda)\equiv Y(\lambda\bar{b}(\lambda),0,\lambda)$.

Proof. i. A value of Y that satisfies $Y=\Phi(Y,\lambda)$ can be thought of as a steady-state level of output conditional on $b_t^c=0$ $\forall t$. Since $b(\lambda)$ is the steady-state level of the transfer conditional on $b_t^c=0$ $\forall t$, Y must coincide with $Y(\lambda b(\lambda),0,\lambda)$.

ii. As follows from Lemmas 2, 3, and 10, the steady state of the system $b_{t+1}^i=\phi^R(b_t^i,k(b_t^i,0,\lambda))$ is unstable at $\underline{b}(\lambda)$ and stable at $\bar{b}(\lambda)$. Hence part i establishes the result. ∎

The first result is not to say that $Y(\lambda b(\lambda),0,\lambda)$ is necessarily a steady-state level of output. The second result asserts that there is no steady-state level of output on the interval $(\underline{Y}^i,\bar{Y}^i)$. To understand the first point, note that

$$\hat{b}^i \lesseqgtr \bar{b}^i \qquad \text{if and only if} \qquad \hat{Y}^i \lesseqgtr \bar{Y}^i, \tag{48}$$

where $\hat{b}^i\equiv\hat{b}(\lambda^i)$, $\bar{b}^i\equiv\bar{b}(\lambda^i)$, $\hat{Y}^i\equiv\hat{Y}(\lambda^i)$ and $\bar{Y}^i\equiv\bar{Y}(\lambda^i)$. Therefore, $\bar{Y}^i$ is not a steady-state level of output when $\hat{b}^i<\bar{b}^i$, as shown by $\hat{b}^L<\bar{b}^L$ in Figure 7. Conversely,

---

[21] If $e''(k)\le 0$ $\forall k>0$, $\Phi_{YY}(Y_t,\lambda)<0$ and $\Phi_{Y\lambda}(Y_t,\lambda)>0$ $\forall(Y_t,\lambda)\in(\theta,\infty)\times(0,1]$.

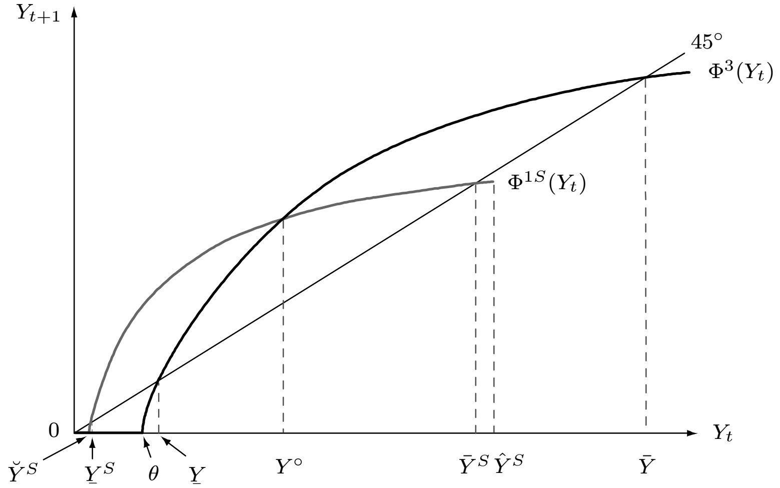

FIGURE 8. The Evolution of Output in Regimes 1 and 3 ($\bar{Y}^S \leq \hat{Y}^S$)

Notes. There exists a locally stable steady-state equilibrium in each of Regimes 1 and 3. In the early stages of development, the economy starting in Regime 1 produces higher output than the economy starting in Regime 3. However, the former economy is unable to take off from Regime 1 to Regime 2, converging to the lower-level steady state where $Y_t = \bar{Y}^S$.

if $\hat{b}^i \geq \bar{b}^i$, which corresponds to $\hat{b}^S \geq \bar{b}^S$ in Figure 7, $\bar{Y}^i$ is a nontrivial, locally stable, steady-state equilibrium that occurs in Regime 1.

In contrast to $\bar{b}^i$, both $\bar{b} \equiv \bar{b}(1)$ and $\underline{b} \equiv \underline{b}(1)$ are steady-state levels of transfers in Regime 3 regardless of the ratio $\hat{b}^i/\bar{b}^i$, and $\underline{Y} \equiv Y(\underline{b}, 0, 1)$ and $\bar{Y} \equiv Y(\bar{b}, 0, 1)$ are the corresponding steady-state levels of output. It should be noted that, by definition, $\underline{Y}$ and $\bar{Y}$ are independent of the fraction of richer members, λ^i; i.e., $\partial \underline{Y}/\partial \lambda^i = \partial \bar{Y}/\partial \lambda^i = 0$.

Figure 8 illustrates the evolution of output in Regime 1 and Regime 3 for the small fraction of richer members, $\hat{Y}^S \geq \bar{Y}^S$. As follows from Lemma 11, for a set of parameters that satisfy condition (38), $\hat{Y}^S \geq \bar{Y}^S$ whenever $\hat{Y}^L \geq \bar{Y}^L$, yet the reverse is not true. For this reason, we treat $\hat{Y}^S \geq \bar{Y}^S$ as a general case.

In the diagram, $\underline{Y}^S$ stands for the steady-state level of output for $\underline{b}^S$, $\underline{Y}^S \equiv Y(\lambda^S \underline{b}^S, 0, \lambda^S)$, and it is assumed that $b(\lambda^S) = \{\underline{b}^S, \bar{b}^S\}$ for the simplicity of the exposition. One can see the function $\Phi^{1S}(Y_t)$ crossing the 45° line at $\underline{Y}^S$ and $\bar{Y}^S$, and the function $\Phi^3(Y_t)$ crossing the 45° line at $\underline{Y}$ and $\bar{Y}$. As shown in Lemma 3, $\underline{b}^S < \underline{b}$ and hence $\underline{Y}^S < \underline{Y}$. In addition, $\bar{Y}^S < \bar{Y}$ follows from Proposition 1.

Figure 9 illustrates the evolution of output in Regime 1 and Regime 3 for the large fraction of the rich, $\hat{Y}^L < \bar{Y}^L$. Unlike $\bar{Y}^S$ in Figure 8, $\bar{Y}^L$ is not a steady-state level of output and the economy starting in Regime 1 automatically enters Regime 2.

4.2.2. Regime 2 and Regime 3. We are now in a position to examine the evolution of output in Regimes 2 and 3. One can find that $\Phi^{2i}(Y_t, b_t^j)$ and $\Phi^3(Y_t)$ exhibits the qualitatively same property with respect to Y_t since $B_t = B(Y_t, 1)$ in both Regimes,

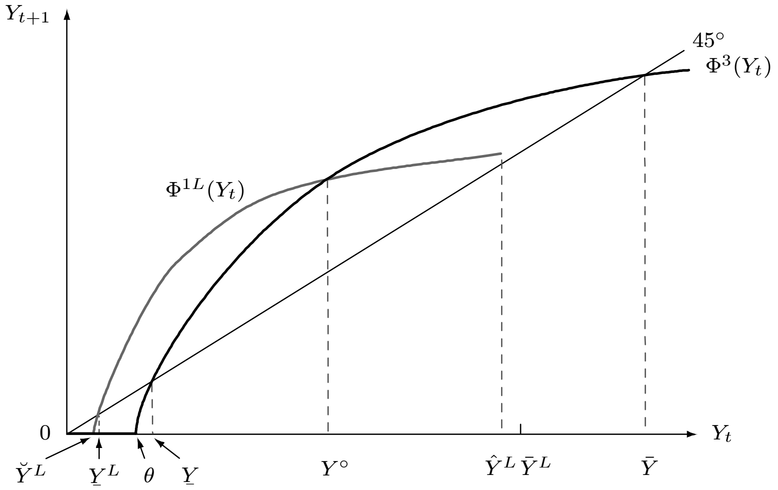

FIGURE 9. The Evolution of Output in Regimes 1 and 3 ($\hat{Y}^L < \bar{Y}^L$)

Notes. Unlike $\bar{Y}^S$ in Figure 8, $\bar{Y}^L$ is not a steady-state level of output. The economy starting in Regime 1 can take off from Regime 1 to Regime 2, converging to the steady state where $Y_t = \bar{Y}$.

as presented in (45). $\Phi^{2i}(Y_t, b_t^j)$ is hence strictly monotonically increasing and strictly concave in Y_t.

Recall that credit constraints are binding in the poorer group in Regime 2, whereas all individuals are free from the constraints and spend e_t in Regime 3. Since e_t is the optimal level of education to maximize the output at time $t+1$, $\Phi^{2i}(Y_t, b_t^j)$ is smaller than $\Phi^3(Y_t)$ when b_t^j, which is b_t^c, does not coincide with e_t. That is,

$$\begin{aligned} &\Phi^{2i}(Y_t, b_t^j) = \Phi^3(Y_t) \quad \text{if } Y_t \leq \theta \text{ or } Y_t = Y^{cc}, \\ &\Phi^{2i}(Y_t, b_t^j) < \Phi^3(Y_t) \quad \text{othewise,} \end{aligned} \tag{49}$$

where Y^{cc} is a value such that

$$b_t^j = e_t = e(k(B_t, 0, 1)) = e(k(B(Y^{cc}, 1), 0, 1))$$

by using (20). Hence Y^{cc} is a continuous function $Y^{cc}(b_t^j)$ on $\mathbb{R}_+$ such that $Y^{cc}(0) = [0, \theta]$, $Y^{cc}(b_t^j) > \theta$ and $Y^{cc\prime}(b_t^j) > 0$ for $b_t^j > 0$. Since $\partial\Phi^{2i}(Y_t, b_t^j)/\partial\lambda^i > 0$, $\Phi^{2S}(Y_t, b_t^j) < \Phi^{2L}(Y_t, b_t^j)$ and an increase in λ^i, the fraction of the richer members, narrows the gap between $\Phi^{2i}(Y_t, b)$ and $\Phi^3(Y_t)$. The result (49) implies that $\Phi^3(Y_t)$ is the envelope curve of $\Phi^{2i}(Y_t, b)$, where b is constant and greater than zero. As illustrated in Figure 10, the fact that $e(k_{t+1})$ increases with Y_t while b is constant creates a discrepancy between $e(k_{t+1})$ and b. Given that $Y^{cc\prime}(b) > 0$, an increase in b shifts $\Phi^{2i}(Y_t, b)$ gradually to the left along the envelope curve $\Phi^3(Y_t)$.

4.3. *The Short-Run Effect of Inequality*

We now examine how a change in λ, the fraction of people who are not credit-constrained, affects the dynamical system $Y_{t+1} = \Phi(Y_t, \lambda)$.

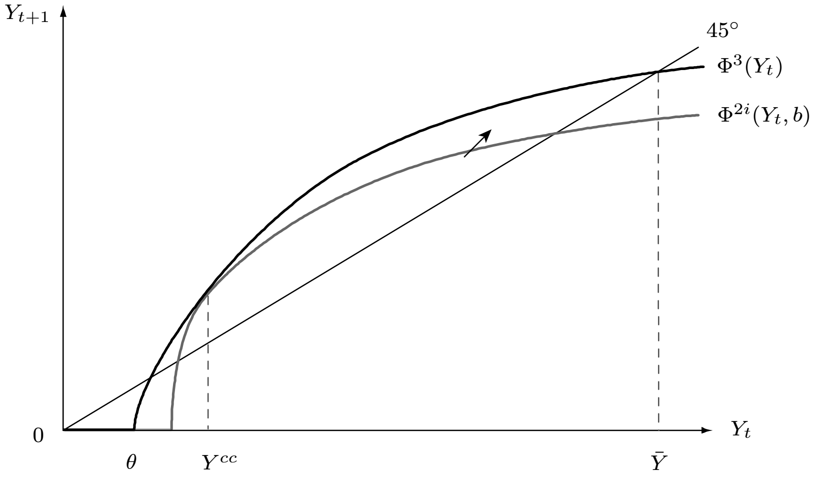

FIGURE 10. The Evolution of Output in Regimes 2 and 3

Notes. The figure depicts the conditional dynamics for Regime 2, $Y_{t+1} = \Phi^{2i}(Y_t, b)$, and the unconditional dynamics for Regime 3, $Y_{t+1} = \Phi^3(Y_t)$. $b < e_t$ if and only if $Y_t > Y^{cc}$.

LEMMA 13. $\forall(\lambda, \lambda') \in (0,1) \times (\lambda, 1]$,

$$\Phi(Y_t, \lambda) - \Phi(Y_t, \lambda') \begin{cases} > 0 & \textit{for } Y_t \in (\check{Y}(\lambda), \check{Y}(\lambda')] \\ < 0 & \textit{for } Y_t \geq \hat{Y}(\lambda). \end{cases}$$

Proof. The first result follows from (46) and $\check{Y}(\lambda) < \check{Y}(\lambda')$. For the second result, observe

$$B(Y_t, \lambda') - B(Y_t, \lambda) = \beta[(1-\lambda)\omega(Y_t, \lambda) - (1-\lambda')\omega(Y_t, \lambda') - (\lambda' - \lambda)\theta],$$

where, noting $\omega(Y_t, \lambda') < \omega(Y_t, \lambda)$,

$$(1-\lambda)\omega(Y_t, \lambda) - (1-\lambda')\omega(Y_t, \lambda') > (\lambda' - \lambda)\omega(Y_t, \lambda).$$

Since $\omega(Y_t, \lambda) \geq \theta \ \forall Y_t \geq \hat{Y}(\lambda)$, we find that $B(Y_t, \lambda') - B(Y_t, \lambda) \geq 0 \ \forall Y_t \geq \hat{Y}(\lambda)$. Then the result follows from $Y_\lambda(B_t, 0, \lambda) > 0 \ \forall (B_t, \lambda) \in \mathbb{R}_{++} \times (0, 1]$. ∎

The above results are illustrated in Figures 8 and 9, where they intersect each other at Y^e. Since $\Phi_Y(Y_t, \lambda) < \Phi_Y(Y_t, 1)$ for $(Y_t, \lambda) \in (\theta, \infty) \times (0, 1]$, Lemma 13 implies the uniqueness of Y°.[22]

COROLLARY 3. *If* $\hat{Y}^i \leq \bar{Y}^i$, *then* $\underline{Y} < Y^\circ < \bar{Y}^i$.

Proof. Follows from Lemma 13 and that $\underline{Y} < Y^\circ$ must occur because of $\underline{Y}^i < \underline{Y}$. ∎

We have therefore found that $\underline{Y}^i < \underline{Y} < Y^\circ < \bar{Y}^i < \bar{Y}$ if $\hat{Y}^i \leq \bar{Y}^i$.

One can view the distance $\Phi^{1i}(Y_t) - \Phi^3(Y_t)$ as the impact of inequality on Y_{t+1} at a given level of Y_t or, to put it another way, at a given stage of development. Furthermore, as long as the economy on each path remains in its initial regime, the discrepancy in the subsequent growth paths of output tells us the impact of the redistribution on the

---

[22] $\Phi_Y(Y_t, \lambda) < \Phi_Y(Y_t, 1)$ follows from (46), $k(B_t, 0, 1) < k(B_t/\lambda, 0, \lambda)$, and $B_Y(Y_t, \lambda) < B_Y(Y_t, 1)$.

growth process. Redistribution can be executed in a more drastic fashion by choosing a smaller number of the rich, λ^i.

Since intergenerational transfers are a luxury good, poorer households leave no wealth to their offspring at underdeveloped stages, which are associated with low wages. As noted earlier, as long as Y_t is below $\hat{Y}^i$, there is a discrepancy in personal saving rates between richer and poorer people, and inequality enhances aggregate levels of transfers, savings, and physical capital by suppressing aggregate consumption. Equally important, this positive effect is generally more significant at lower output levels. Although, less equal distribution causes educational inequality and low productivity of output as discussed in Section 2.3, the positive effect of inequality is proven to be dominant for $Y_t \in (\theta, Y^\circ]$.

At low levels of development, scarce amounts of physical capital together with the capital-skill complementarity give rise to lower returns on education relative to savings, which in turn discourages the private investment in education. Therefore, the scarcity of physical capital, rather than credit constraints, is the prime reason for low stocks of average human capital. Accordingly the positive impact of inequality outweighs the negative one for $Y_t \leq Y^\circ$.

Yet as illustrated in Figures 8 and 9, the difference $\Phi^{1i}(Y_t) - \Phi^3(Y_t)$ narrows down as output increases towards Y°. The accompanying accumulation of physical capital raises the return on skills and thereby intensifies the negative effect of inequality on human capital accumulation, as long as liquidity constraints are binding. The gap between $\Phi^{1i}(Y_t)$ and $\Phi^3(Y_t)$ therefore shrinks as output increases towards Y°.[23] At the same time, inequality becomes less conductive for aggregate saving at higher levels of Y_t. Since this positive effect disappears at $\hat{Y}^i$, where $w_t = \theta$, the two opposing effects of inequality offset each other at $Y^\circ (< \hat{Y}^i)$ to the point of negating their values, and the negative effect becomes dominant for $Y_t > Y^\circ$.

4.4. *Inequality and the Pattern of Development*

By integrating all the results demonstrated so far, one can analyze the impact of initial income distribution on the behavior of output growth. For this purpose, we employ the approach of comparing hypothetical economies that are identical in all respects except for their initial distributions of wealth.

Recall that $h_0^i = h(e(k_0))$ and adult individuals of group j initially possess neither advanced skills nor physical capital. In the absence of redistribution,

$$I_0^j = \omega(Y_0, \lambda^i), \tag{50}$$

and $I_0^i = (Y_0 - \lambda^j I_0^j)/\lambda^i$ by noting (39).

We limit the analysis to the case in which regardless of the initial regime, the economy ends up with a positive amount of resources, andunskilled workers initially leave no

---

[23] Recall that the function $\Phi^{1i}(Y_t)$ is based on the condition that liquidity constraints are binding among poorer individuals, whereas $\Phi^3(Y_t)$ is on that liquidity constraints are not binding.

bequests. This is the case if

$$\underline{Y} < Y_0 < \hat{Y}^i. \tag{51}$$

Note that $\hat{Y}^i < \hat{Y} < \bar{Y}$ from condition (38) and Corollary 2.

Under (51) and (50), the initial state of the economy is Regime 1. As shown in (48), the magnitude of the ratio $\hat{Y}^i/\bar{Y}^i$ determines the existence of the nontrivial, locally stable, steady-state equilibrium in Regime 1. Since a sufficiently small (large) λ^i can make $\hat{Y}^i$ larger (smaller) than $\bar{Y}^i$ under assumption (38), the level of λ^i substantially changes the growth path of the economy starting out in Regime 1.

PROPOSITION 2. *Consider the economy characterized by* (38)*,* (50)*, and* (51)*.*

i. *If λ^i is small enough to generate $\hat{Y}^i \geq \bar{Y}^i$, the economy remains in Regime 1 for all periods and Y_t monotonically converges to the steady-state level $\bar{Y}^i$.*

ii. *If λ^i is large enough to generate $\hat{Y}^i < \bar{Y}^i$, Y_t increases strictly monotonically over Regimes 1, 2 and 3 sequentially and converges to the steady-state level $\bar{Y}$.*

Proof. i. $\underline{b}^i < b_0^i$ follows from the fact that $\underline{b}^i < \underline{b}$ shown in Lemma 3 and that (50) generates a higher level of b_0^i than (52). Since b_t^i in the present case corresponds to b_t^S in Figure 7, one can find that $b_t^L = 0$ in all periods and b_t^S converges to $\bar{b}^S$, noting $b_t^S < \hat{b}^S$ $\forall t \geq 0$. Then the property $B_Y(Y_t, \lambda^i) > 0$ yields the behavior of Y_t that corresponds to $B_t = \lambda^S b_t^S$.

ii. The result is proven by combining the above proof with the one in Proposition 3 and noting (b_t^S, b_t^L) increases in Regime 2. ∎

REMARK 2. *In the first case of Proposition 2, Y_t may decrease over time.*

Figure 8 depicts the case $\hat{Y}^S \geq \bar{Y}^S$. The economy that starts off in Regime 1 fails to take off to Regime 2. Since individuals' investment in human capital is subject to diminishing marginal returns, less equal opportunities for education among individuals retards the accumulation of aggregate human capital, and this adverse effect increases as the cost of being unskilled rises with output. Furthermore, as argued previously, the increase in the wage income over time diminishes the positive effect of inequality on aggregate saving. Their net effect on factor accumulation eventually turns negative, and this qualitative change leads the economy starting out in Regime 1 to the lower-level steady state.

Figure 9 depicts the case $\hat{Y}^L < \bar{Y}^L$, which is brought about by a sufficiently large fraction of the richer members in society. Now that moderate inequality in education opportunity mitigates the adverse effect of selective human capital accumulation on output growth, the economy with less equal distribution does not end up with Regime 1. It endogenously enters Regime 2 when the wage rate reaches the level that induces unskilled workers to transfer wealth to their offspring. Although b_t^L rises concurrently with e_t over Regime 2, Figure 7 ensures that b_t^L eventually catches up with e_t, and once that occurs, the economy enters Regime 3. Consequently Y_t evolves toward $\bar{Y}$, and

(b^S_t, b^L_t) approaches the egalitarian steady-state point $(\bar{b}, \bar{b})$. Wealth inequality thus improves in the long run, albeit not necessarily monotonically, and the evolution of wealth inequality exhibits an inverted U-curve.

Now consider a redistribution of wealth (I^i_0, I^j_0) such that

$$b^j_0 = \beta(I^j_0 - \theta) > \underline{b}, \tag{52}$$

and $I^i_0 = (Y_0 - \lambda^j I^j_0)/\lambda^i$ as follows from (39). Due to assumption (51), one can achieve (52) by choosing a sufficiently small gap between I^S_0 and I^L_0. It follows from (52) and $I^i_0 > I^j_0$ that $(b^S_0, b^L_0) \gg (\underline{b}, \underline{b})$.

Proposition 3. *Under* (38), (51), *and* (52), Y_t *increases strictly monotonically in either Regime 2 or Regime 3 for all* $t \geq 0$ *and converges to the steady-state level* $\bar{Y}$ *in Regime 3.*

Proof. Consider Figure 7. If $(b^S_0, b^L_0) \gg (\underline{b}, \underline{b})$, the economy is in either Regime 2 or Regime 3 in all periods, and eventually enters Regime 3 to approach $(\bar{b}, \bar{b})$. Hence Y_t correspondingly converges to $\bar{Y}$.

Suppose that the economy enters Regime 2 in period $t > 0$. The diagram shows that $(b^S_t, b^L_t) \ll (b^S_{t+1}, b^L_{t+1})$, implying $B_t < B_{t+1}$. Thus $Y_t < Y_{t+1}$ follows from (40). Since Lemma 12 implies $Y_t < \Phi^3(Y_t)$ if and only if $Y_t \in (\underline{Y}, \bar{Y})$, (49) shows that $\underline{Y} < Y_t < Y_{t+1} < \bar{Y}$ if the economy is in Regime 2 at time $t > 0$.

The above result, together with $Y_0 \in (\underline{Y}, \bar{Y})$, proves $Y_t < Y_{t+1}$ if the economy is in Regime 3 at time $t \geq 0$. ∎

As follows from Propositions 2-3, the economy that has a high equality, (52), or a moderate inequality, (50) and $\hat{Y}^i < \bar{Y}^i$, evolves over time toward the highest steady-state level of output, $\bar{Y}$. On the other hand, the economy with a substantially unequal distribution, (50) and $\hat{Y}^i \geq \bar{Y}^i$, is unable to reach that level and undergoes a lower growth path in the long run.

Theorem 2. *Consider a group of countries that differ only in their initial wealth distributions. Under* (38) *and* (51),

i. *If a country has a highly unequal wealth distribution in an underdeveloped stage, it will attain higher levels of output in the short run yet converge to a lower-level steady-state equilibrium.*
ii. *If a country has a highly unequal distribution of wealth in a well-developed stage, it will attain lower levels of output in all subsequent periods and converge to a lower-level steady-state equilibrium.*
iii. *If countries have similar levels of initial inequality, they will converge to the same steady-state equilibrium, regardless of their initial regimes.*

Proof. The theorem follows from Corollary 11, Propositions 2-3 and Figures 8-10. ∎

Overtaking in growth performance results from the initial dominance of the positive effect of inequality and the multiplicity of steady-state equilibria. Furthermore, divergence applies to countries starting out in the mature stages of development where the adverse effect of inequality on human capital outweighs the positive effect of inequality on aggregate saving.

A few remarks deserve special emphasis at this point. First, overtaking takes place after the output of a lagging country reaches Y°. Second, the model abstracts from technological change, and the introduction of technological progress will permit steady-state growth with a positive rate. Regardless of this formulation, inequality will constrain output growth in the long run, as asserted by the first two results of the proposition. Yet the last result will not hold because the convergence in growth rates does not imply the convergence in income levels. Third and lastly, overtaking is likely to occur under the condition that the marginal productivity of physical capital is sufficiently high. This is because $\hat{b}(\lambda)$ goes to infinity as α approaches 1, as shown earlier, whereas this is not the case for $\bar{b}(\lambda)$.[24] Provided that α is sufficiently small, a major fraction of national income is dominated by asset owners since production of final output relies primarily on physical capital. This uneven system of factor payments slows down the increase in the wage income relative to output, intensifying the adverse effect of credit constraints on human capital accumulation.

4.5. *Factor Accumulation*

Countries starting out in different regimes experience different patterns of factor accumulation. Firstly, as follows from (15), the aggregate physical capital in Regimes 1 and 3 is

$$K(B_t, \lambda) \equiv B_t - \lambda e(k(B_t/\lambda, 0, \lambda)),$$

whereby $K_{t+1} = K(B_t, \lambda^i)$ for Regime 1 and $K_{t+1} = K(B_t, 1)$ for Regime 3. As proven in the Appendix,

$$K(B_t, \lambda^i) > K(B_t, 1) \qquad \text{for } B_t > 0. \tag{53}$$

Recalling $B(Y_t, \lambda^i) \geq B(Y_t, 1)$ for $Y_t \leq \hat{Y}^i$ and Theorem 2, we find that in the early stages of development, inequality promotes the accumulation of physical capital.

Secondly, the aggregate human capital in Regimes 1 and 3 is expressed as

$$H(B_t, \lambda) \equiv (1 - \lambda) + \lambda h(e(k(B_t/\lambda, 0, \lambda))),$$

whereby $H_{t+1} = H(B_t, \lambda^i)$ for Regime 1 and $H_{t+1} = H(B_t, 1)$ for Regime 3. As follows from (18),

$$\lim_{\lambda \to 0} H(B_t, \lambda) = 1.$$

---

[24] Since total output is finite for any $\alpha \in [0, 1]$, all individuals obtain finite incomes and therefore $\bar{b}(\lambda)$ remains finite for any $\alpha \in [0, 1]$.

This means that if the size of the rich is sufficiently small, their high educational attainments would not be reflected in aggregate human capital. Using this result, one can then find a small value λ^i such that

$$H(B'_t, \lambda^i) < H(B_t, 1) \qquad \text{for } 0 < B_t \leq B'_t.$$

Thus, despite the relative abundance of B_t, less egalitarian economies encounter a delay in human capital accumulation in the early stages of development. As shown earlier, this adverse effect of inequality becomes dominant in later stages. This is confirmed by

$$H(B_t, \lambda^i) < H(B_t, 1) \qquad \text{for } B_t > 0 \text{ and } 0 < \lambda^i < 1. \tag{54}$$

and by Proposition 1, which asserts that in the long run, economies in Regime 3 produce greater amounts of aggregate transfers than those in Regime 1.

This tendency of inequality generating overinvestment in physical capital and underinvestment in human capital implies an unbalanced ratio of physical to human capital (18) yields

$$k(B_t/\lambda^i, 0, \lambda^i) > k(B_t, 0, 1) \qquad \text{for } B_t > 0, \tag{55}$$

where, as shown in (19), $k_{t+1} = k(B_t/\lambda^i, 0, \lambda^i)$ in Regime 1 and $k_{t+1} = k(B_t, 0, 1)$ in Regime 3.

CONCLUDING REMARKS

This research develops a theory about the role of inequality in the overtaking of growth performance across countries. The proposed theory highlights two opposing effects of income inequality on factor accumulation and argue that the qualitative change in their combined effect is a prime cause of overtaking. Inequality enhances physical capital accumulation by concentrating wealth among individuals whose marginal propensity to save is relatively high. Conversely, in the presence of borrowing constraints, inequality acts as a barrier to the universal investment in human capital that is a prerequisite for sustained growth. Due to the eventual dominance of this negative effect of inequality, countries characterized by highly unequal wealth distributions will fail to take off and be overtaken in the long run.

It is also shown that divergence or convergence may arise instead of overtaking, depending on the initial levels of development and inequality. Sufficiently developed countries with different levels of inequality will diverge in income levels. On the other hand, countries that have similar levels of inequality will converge to the same steady state, regardless of their initial regimes.

Although the analysis finds that the effect of inequality in the early stages of development is positive, there are two major forces that make overtaking less probable than divergence. The one is that the positive effect of inequality is the dominant factor only in underdeveloped stages, where the saving rate differential among individuals is significant. The other is that globalization of international capital markets encourages the flow of capital across borders, and domestic saving becomes less important for physical capital accumulation. The positive aspect of inequality would thus be more relevant to the process of development experienced by currently developed countries.

APPENDIX

LEMMA 14. *Under* (29), $b_t^i = \psi^{*i}(b_t^S, b_t^L) = \psi^{Ri}(b_t^S, b_t^L)$ *if and only if* $(b_t^i, b_t^j) = (b_a^i, \eta^j(b_a^i))$, $(b_b^i, \eta^j(b_b^i))$, *where* $b_a^i \in (0, \underline{b})$, $b_b^i \in (\bar{b}, \infty)$.

Proof. Sufficiency: If $b_t^i = \psi^{*i}(b_t^S, b_t^L) = \psi^{Ri}(b_t^S, b_t^L)$, then $h(e_t) = H_{t+1}$ and $b_t^i - e_t = K_{t+1}/\lambda^i$. It thus follows from (1), (2), (21) and (33),

$$\psi^{*i}(b_t^S, b_t^L) = \beta\{[(1-\alpha) + \alpha/\lambda^i]Y_{t+1} - \theta\} = \beta[\Gamma Y(b_t^i, 0, 1) - \theta],$$

where $\Gamma \equiv [(1-\alpha) + \alpha/\lambda^i](\lambda^i)^\alpha > 1$. Thus, under (29), there exist exactly two steady-state values for the system $b_{t+1}^i = \psi^{*i}(b_t^S, b_t^L)$. By referring them as b_a^i and b_b^i ($b_a^i < b_b^i$) and noting $\Gamma > 1$, we obtain $b_a^i \in (0, \underline{b})$ and $b_b^i \in (\bar{b}, \infty)$. Also, we get $b_t^j = \eta^j(b_t^i)$ because $\kappa^*(b_t^S, b_t^L) = \kappa^i(b_t^S, b_t^L)$ follows from $\psi^{*i}(b_t^S, b_t^L) = \psi^{Ri}(b_t^S, b_t^L)$. The sufficiency hence holds.

Necessity: If $(b_t^i, b_t^j) = (b_a^i, \eta^j(b_a^i))$ or $(b_b^i, \eta^j(b_b^i))$, then $b_t^j = e_t$, $h(e_t) = H_{t+1}$, and $b_t^i - e_t = K_{t+1}/\lambda^i$. Hence the necessity holds similarly. ▮

As follows from $R_{t+1} = r_{t+1}$ and (2), the first order condition (8) is written as

$$k_{t+1}h'(e_t) = \frac{\alpha}{1-\alpha} \qquad \text{for } k_{t+1} > 0. \tag{56}$$

Using this result with (17) and (7) yields

$$\begin{aligned} k_e(b_t^e, b_t^c, \lambda) &\equiv \frac{\partial k(b_t^e, b_t^c, \lambda)}{\partial b_t^e} = \frac{\lambda(1-\alpha)}{(1-\alpha)H_{t+1} + \lambda e'(k_{t+1})}, \\ k_c(b_t^e, b_t^c, \lambda) &\equiv \frac{\partial k(b_t^e, b_t^c, \lambda)}{\partial b_t^c} = -\frac{(1-\lambda)(1-\alpha)h'(b_t^c)k_{t+1}}{(1-\alpha)H_{t+1} + \lambda e'(k_{t+1})}, \end{aligned}$$

where $0 < k_e(b_t^e, b_t^c, \lambda) < 1$ and $k_c(b_t^e, b_t^c, \lambda) < 0$ for $(b_t^e, b_t^c, \lambda) \in \mathbb{R}_{++}^2 \times (0, 1]$.

To simplify the exposition of the following proofs, we define $\kappa_i^* \equiv \partial\kappa^*/\partial b_t^i = \lambda^i k_e(B_t, 0, 1)$, $\kappa_j^* \equiv \partial\kappa^*/\partial b_t^j = \lambda^j k_e(B_t, 0, 1)$, $\kappa_i^i \equiv \partial\kappa^i/\partial b_t^i = k_e(b_t^i, b_t^j, \lambda^i)$, and $\kappa_j^i \equiv \partial\kappa^i/\partial b_t^j = k_c(b_t^i, b_t^j, \lambda^i)$.

Proof of Lemma 4. Consider $\Psi^{Ri}(b_t^i, b_t^j) \equiv \phi^R(b_t^i, \kappa^i(b_t^S, b_t^L)) - b_t^i$ as a function of b_t^j with a parameter $b_t^i \geq 0$. Recalling $\phi_\kappa^R(b_t^i, \kappa^i) < 0$, the properties of $\kappa^i(b_t^S, b_t^L)$ yield

$$\partial\Psi^{Ri}(b_t^i, b_t^j)/\partial b_t^j > 0 \qquad \forall (b_t^i, b_t^j) \in \mathbb{R}_{++}^2. \tag{57}$$

By the definitions of $b_{\min}^i$ and $b_{\max}^i$, $\Psi^{Ri}(b_{\min}^i, 0) = \Psi^{Ri}(b_{\max}^i, 0) = 0$ and

$$\begin{aligned} \partial\Psi^{Ri}(b_{\min}^i, 0)/\partial b_t^i &> 0; \\ \partial\Psi^{Ri}(b_{\max}^i, 0)/\partial b_t^i &< 0. \end{aligned}$$

Thus, there exists a single-valued function $b_t^j = \xi^j(b_t^i) \geq 0$ such that $\Psi^{Ri}(b_t^i, \xi^j(b_t^i)) = 0$, on $(0, b_{\min}^i]$ and $[b_{\max}^i, \infty)$. $\xi^j(b_t^i)$ does not exist if $\Psi^{Ri}(b_t^i, 0) > 0$, which occurs at least for $b_t^i \in (\underline{b}^i, \bar{b}^i)$. Also, $\xi^j(0)$ does not exist since $\Psi^{Ri}(0, b_t^j) = -\beta\theta$ for all $b_t^j \geq 0$. Part ii then follows, noting (57).

If $e''(k_{t+1}) \leq 0$, $\forall k_{t+1} > 0$, Corollary 1 asserts $\underline{b}^i = b^i_{\min}$ and $\bar{b}^i = b^i_{\max}$, meaning $\xi^j(b^i_t)$ exists on $\mathbb{R}_{++}\backslash(\underline{b}^i, \bar{b}^i)$. Without the concavity of $e(k_{t+1})$, $\xi^j(b^i_t)$ may not exist on the neighbor of $b^i \in b^i$.

i. Note that from Lemma 14,

$$\begin{aligned} &\Psi^{Ri}(b^i_t, \mu^j(b^i_t)) > 0 \quad \text{if } b^i_t \in (b^i_a, b^i_b), \\ &\Psi^{Ri}(b^i_t, \mu^j(b^i_t)) = 0 \quad \text{if } b^i_t = b^i_a, b^i_b, \\ &\Psi^{Ri}(b^i_t, \mu^j(b^i_t)) < 0 \quad \text{otherwise.} \end{aligned} \tag{58}$$

This, together with part ii, establishes the result. ∎

Proof of Lemma 6. Let

$$\begin{aligned} \Psi^{*i}(b^i_t, b^j_t) &\equiv \psi^{*i}(b^S_t, b^L_t) - b^i_t \\ &= \beta[w(k_{t+1})h(e_t) + (b^i_t - e_t)R(k_{t+1}) - \theta] - b^i_t, \end{aligned}$$

where $k_{t+1} = \kappa^*(b^S_t, b^L_t) = k(B_t, 0, 1)$. Consider $\Psi^{*i}(b^i_t, b^j_t)$ as a function of b^j_t with a parameter $b^i_t \geq 0$. Then $b^j_t \geq -\lambda^i b^i_t/\lambda^j \equiv b^*(b^i_t)$ because $k(B_t, 0, 1)$ is well-defined for $B_t \geq 0$. Recalling (7) and (17),

$$\begin{aligned} \partial\Psi^{*i}(b^i_t, b^j_t)/\partial b^j_t &= \beta[w'(k_{t+1})h(e_t) + (b^i_t - e_t)R'(k_{t+1})]\kappa^*_j \\ &= \beta\alpha(1-\alpha)Ak^{\alpha-2}_{t+1}\lambda^j(b^j_t - b^i_t)\kappa^*_j. \end{aligned} \tag{59}$$

Thus, noting $\kappa^*_j > 0$, $\Psi^{*i}(b^i_t, b^j_t)$ is minimized with respect to b^j_t at b^i_t. Moreover, since $\lim_{b^j_t \to \infty} w_{t+1} = \infty$ and $w_{t+1}h(e_t) - e_t R_{t+1} \geq w_{t+1}$,

$$\lim_{b^j_t \to \infty} \Psi^{*i}(b^i_t, b^j_t) = \infty \qquad \text{for } b^i_t \geq 0. \tag{60}$$

Therefore, there exists a function $b^j_t = \zeta^j(b^i_t)$ such that $\Psi^{*i}(b^i_t, \zeta^j(b^i_t)) = 0$ on the domain where b^i_t satisfies $\Psi^{*i}(b^i_t, b^i_t) < 0$ and $b^i_t \geq 0$. Lemmas 2 and 5 suggest that

$$\Psi^{*i}(b^i_t, b^i_t) < 0 \quad \text{if and only if} \quad b^i_t \in \mathbb{R}_+\backslash(\underline{b}, \bar{b}). \tag{61}$$

i. In addition to (60),

$$\lim_{b^j_t \to b^*+0} \Psi^{*i}(b^i_t, b^j_t) = \lim_{b^j_t \to b^*+0} b^i_t R(k(B_t, 0, 1)) = \infty \quad \text{for } b^i_t > 0. \tag{62}$$

$\zeta^j(b^i_t)$ is hence double-valued on $\mathbb{R}_{++}\backslash(\underline{b}, \bar{b})$ and $\min \zeta^j(b^i_t) < b^i_t < \max \zeta^j(b^i_t)$.

ii. First, $\zeta^j(0)$ is a positive single value because of $\Psi^{*i}(0, 0) - \beta\theta < 0$ and the restriction $b^j_t \geq b^*(0) = 0$. Second, since Lemmas 2 and 5 show $\Psi^{*i}(b^i_t, b^i_t) = 0$ for $b^i_t = \underline{b}$ and $\bar{b}$, where $\underline{b} < \bar{b}$, we get $\zeta^j(\underline{b}) = \underline{b}$ and $\zeta^j(\bar{b}) = \bar{b}$.

The results regarding $d\zeta^j(b^i_t)/db^i_t$ in parts i-ii are obtained from (59) and

$$\partial\Psi^{*i}(b^i_t, b^j_t)/\partial b^i_t = \beta\alpha(1-\alpha)Ak^{\alpha-2}_{t+1}\lambda^j(b^j_t - b^i_t)\kappa^*_i + \beta R_{t+1} - 1. \tag{63}$$

Part iii follows from (58), $\Psi^{Ri}(b^i_t, \mu^j(b^i_t)) = \Psi^{*i}(b^i_t, \mu^j(b^i_t))$, and part iv. The above discussion establishes part iv. ∎

Proof of Lemma 7. Let $\Psi^{Pi}(b_t^i, b_t^j) \equiv \phi^P(b_t^i, \kappa^j(b_t^S, b_t^L)) - b_t^i$. Given the properties of $\kappa^j(b_t^S, b_t^S)$, $\partial\Psi^{Pi}(b_t^i, b_t^j)/\partial b_t^j > 0 \ \forall (b_t^i, b_t^j) \in \mathbb{R}_+ \times \mathbb{R}_{++}$, $\lim_{b_t^j \to \infty} \Psi^{Pi}(b_t^i, b_t^j) = \infty$ and $\Psi^{Pi}(b_t^i, 0) = -\beta\theta \ \forall b_t^i \geq 0$. Thus there exists a single-valued function $\varsigma^j(b_t^i) > 0$ on $\mathbb{R}_+$ such that $\Psi^{Pi}(b_t^i, \varsigma^j(b_t^i)) = 0$. Then part iii follows, noting $\partial\Psi^{Pi}(b_t^i, b_t^j)/\partial b_t^j > 0$.

i. The result follows from the implicit function theorem, noting (64) and

$$\lim_{b_t^i \to 0} \partial\Psi^{Pi}(b_t^i, \varsigma^j(b_t^i))/\partial b_t^j < \infty.$$

∎

Proof of Lemma 10. Recall (25). $(b_t^S, b_t^L) \in E^i$: Using (7), $0 < \kappa_i^i < \lambda^i/H_{t+1}$ and $K_{t+1} = \lambda^i(b_t^i - e_t) > 0$ from (15),

$$\begin{aligned}
\partial I^i(b_t^S, b_t^L)/\partial b_t^i &= [w'(k_{t+1})h(e_t) + R'(k_{t+1})(b_t^i - e_t)]\kappa_i^i + R_{t+1} \\
&> R'(k_{t+1})(b_t^i - e_t)\lambda^i/H_{t+1} + R_{t+1} \\
&= (\alpha - 1)R_{t+1} + R_{t+1} > 0.
\end{aligned}$$

$(b_t^S, b_t^L) \in E^*$: Similarly, noting (7), $0 < \kappa_i^* < \lambda^i/H_{t+1}$ and $K_{t+1} > \lambda^i(b_t^i - e_t) > 0$,

$$\partial I^i(b_t^S, b_t^L)/\partial b_t^i > (\alpha - 1)R_{t+1}\lambda^i(b_t^i - e_t)/K_{t+1} + R_{t+1} > 0.$$

$(b_t^S, b_t^L) \in E^j$: Noting $w'(k_{t+1}) = \alpha w_{t+1}/k_{t+1}$ and $h(b_t^i)\kappa_i^j > -h'(b_t^i)k_{t+1}$,

$$\begin{aligned}
\partial I^i(b_t^S, b_t^L)/\partial b_t^i &= w'(k_{t+1})h(b_t^i)\kappa_i^j + w_{t+1}h'(b_t^i) \\
&> -\alpha w_{t+1}h'(b_t^i) + w_{t+1}h'(b_t^i) > 0.
\end{aligned} \tag{64}$$

These three complete the proof. ∎

Proofs of (53) *and* (54). Define

$$\mathrm{K}(B_t, b_t^j) \equiv B_t - \lambda^j b_t^j - \lambda^i e(k_{t+1}),$$

where $k_{t+1} \equiv \kappa^i(b_t^S, b_t^L)$. Noting $b_t^i = (B_t - \lambda^j b_t^j)/\lambda^i$,

$$\partial\mathrm{K}(B_t, b_t^j)/\partial b_t^j = \lambda^j[e'(k_{t+1})(\kappa_i^i - \kappa_j^i\lambda^i/\lambda^j) - 1].$$

If $b_t^j \leq e_t$, then $h'(b_t^j)k_{t+1} \leq \alpha/(1-\alpha)$ and

$$e'(k_{t+1})(\kappa_i^i - \kappa_j^i\lambda^i/\lambda^j) = \frac{\lambda^i e'(k_{t+1})(1-\alpha)[1 + h'(b_t^j)k_{t+1}]}{(1-\alpha)H_{t+1} + \lambda^i e'(k_{t+1})} < 1.$$

We thus find $\partial\mathrm{K}(B_t, b_t^j)/\partial b_t^j < 0$ for $(B_t, b_t^j) \in \mathbb{R}_+ \times (0, e_t]$. Then noting $e_t > 0$ for $B_t > 0$,

$$\mathrm{K}(B_t, e_t) < \mathrm{K}(B_t, 0) \qquad \text{for } B_t > 0.$$

The result then follows from $\mathrm{K}(B_t, e_t) = K(B_t, 1)$ and $\mathrm{K}(B_t, 0) = K(B_t, \lambda^i)$. (54) can be proved in the similar way. ∎

REFERENCES

Atkinson, A., and A. Brandolini (2001): “Promise and Pitfalls in the Use of "Secondary” Data-Sets: Income Inequality in OECD Countires as a Case Study,” *Journal of Economic Literature*, 41, 771–799.

Banerjee, A. V., and E. Duflo (2000a): “Inequality and Growth: What Can the Data Say?,” Working Paper 7793, National Bureau of Economic Research, Cambridge, MA.

——— (2000b): “A Reassessment of the Relationship Between Inequality and Growth: Comment,” Discussion paper, MIT.

Banerjee, A. V., and A. Newman (1993): “Occupational Choice and the Process of Development,” *Journal of Political Economy*, 101(2), 274–298.

Barro, R. J. (2000): “Inequality and Growth in a Panel of Countries,” *Journal of Economic Growth*, 5, 5–32.

Benabou, R. (1996): “Inequality and Growth,” in *NBER Macroeconomics Annual 1996*, ed. by B. S. Bernanke, and J. J. Rotemberg, pp. 11–74. MIT Press, Cambridge, MA.

Birdsall, N., D. Ross, and R. Sabot (1995): “Inequality and Growth Reconsidered: Lessons from East Asia,” *World Bank Economic Review*, 9(3), 477–508.

Bourguignon, F. (1981): “Pareto Superiority of Unegalitarian Equilibria in Stiglitz’ Model of Wealth Distribution with Convex Saving Function,” *Econometrica*, 49(6), 1469–1475.

Brezis, E. S., P. R. Krugman, and D. Tsiddon (1993): “Leapfrogging in International Competition: A Theory of Cycles in National Technological Leadership,” *American Economic Review*, 83(5), 1211–1219.

Carroll, C. D. (2000): “Why Do the Rich Save So Much?,” in *Does Atlas Shrug?: The Economic Consequences of Taxing the Rich*, ed. by J. B. Slemrod, pp. 465–484. Harvard University Press, Cambridge, Massachusetts.

Deininger, K., and L. Squire (1996): “A New Data Set Measuring Income Inequality,” *World Bank Economic Review*, 10(3), 565–91.

Dynan, K. E., J. Skinner, and S. P. Zeldes (2000): “Do the Rich Save More?,” Working Paper 7906, National Bureau of Economic Research, Cambridge, MA.

Engerman, S. L., and K. L. Sokoloff (2002): “Factor Endowments, Inequality, and Paths of Development Among New World Economies,” Working Paper 9259, National Bureau of Economic Research, Cambridge, MA.

Fischer, R. D., and P. J. Serra (1996): “Income Convergence Within and Between Countries,” *Internatinal Economic Review*, 37(3), 531–551.

Flug, K., A. Spilimbergo, and E. Wachtenheim (1998): “Investment in Education: Do Economic Volatility and Credit Constraints Matter?,” *Journal of Development Economics*, 55, 465–481.

Forbes, K. J. (2000): “A Reassessment of the Relationship Between Inequality and Growth,” *American Economic Review*, 90(4), 869–887.

Galor, O., and O. Moav (2001): “From Physical to Human Capital Accumulation: Inequality and the Process of Development,” mimeo.

——— (2002): “Das Human Kapital,” mimeo.

Galor, O., O. Moav, and D. Vollrath (2002): “Divergence and Overtaking: Land Abundance as a Hurdle for Education Reforms,” mimeo.

Galor, O., and J. Zeira (1993): “Income Distribution and Macroeconomics,” *Review of Economic Studies*, 60(1), 35–52.

Goldin, C., and L. F. Katz (1998): “The Origins of Technology-Skill Complementarity,” *Quarterly Journal of Economics*, 113(3), 693–732.

Goodfriend, M., and J. McDermott (1998): “Industrial Development and the Convergence Question,” *American Economic Review*, 88(5), 1277–1289.

Jones, C. I. (1997): “On the Evolution of the World Income Distribution,” *Journal of Economic Perspective*, 11(3), 19–36.

Kalor, N. (1957): “A Model of Economic Growth,” *Economic Journal*, 67(268), 591–624.

KEYNES, J. M. (1936): *The General Theory of Employment, Interest, and Money.* Harcourt, Brace, New York.

KUZNETS, S. (1955): "Economic Growth and Income Inequality," *American Economic Review*, 45(1), 1–28.

MADDISON, A. (2001): *The World Economy: A Millennial Perspective.* the OECD Development Centre.

MAYER, T. (1966): "The Propensity to Consumer Permanent Income," *American Economic Review*, 56(5), 1158–1177.

MOUNTFORD, A. (1998): "Trade, Convergence and Overtaking," *Journal of International Economics*, 46, 167–182.

PRITCHETT, L. (1997): "Divergence: Big Time," *Journal of Economic Perspective*, 11(3), 3–17.

QUAH, D. (1993): "Galton's Fallacy and Tests of the Convergence Hypothesis," *Scandinavian Journal of Economics*, 95, 427–43.

——— (1996): "Twin Peaks: Growth and Convergence in Models of Distribution Dynamics," *Economic Journal*, 106, 1045–55.

STIGLITZ, J. E. (1969): "Distribution of Income and Wealth Among Individuals," *Econometrica*, 37(3), 382–397.