Economic Theory (2005)
DOI 10.1007/s00199-005-0021-2

SPECIAL ISSUE ARTICLE

Debraj Ray

On the dynamics of inequality

Received: 18 October 2004 / Accepted: 5 August 2005

Abstract The dynamics of inequality are studied in a model of human capital accumulation with credit constraints. This model admits a multiplicity of steady state skill ratios that exhibit varying degrees of inequality across households. The main result studies equilibrium paths. It is shown that an equilibrium sequence of skill ratios must converge monotonically to the *smallest* steady state that exceeds the initial ratio for that sequence. Convergence is "gradual" in that the steady state is not achieved in finite time. On the other hand, if the initial skill ratio exceeds the largest steady state, convergence to a steady state is immediate.

Keywords Inequality · Income distribution · Intertemporal equilibrium

JEL Classification Numbers D31 · D90 · O15

This paper is based on unpublished notes from 1990; see `http://www.econ.nyu.edu/user/debraj/DevEcon/Notes/incdist.pdf`. Two considerations suggest that these results may be worth reporting in print. First, the existence of a sizeable recent literature indicates that these relatively early notes may have value outside a filing cabinet or a private webpage. Second, Mukul Majumdar's own research on economic growth with a nonconvex technology is an even earlier precursor to some of this literature, so the current outlet – a special issue in his honor – seems appropriate. Conversations with Glenn Loury simplified the proof of the main result. I thank Dilip Mookherjee for many useful discussions, and two anonymous referees for helpful comments on an earlier draft. Funding from the National Science Foundation under grant number 0241070 is acknowledged. This paper is dedicated with much affection and warm admiration to Mukul Majumdar – or to Mukulda, as I always think of him – on the occasion of his 60th birthday.

D. Ray
Department of Economics, New York University,
New York, NY 10003, USA
E-mail: debraj.ray@nyu.edu

1 Introduction

In this paper, I study the dynamics of inequality in a model of human capital accumulation with credit constraints. The particular framework I study is possibly the simplest representation of what might be called an *interactive model* of wealth distribution, in which the intertemporal decisions made by individual dynasties affect macroeconomic variables and so cannot be viewed as isolated dynamic processes. Below, I describe the specific setting that I study (Section 3.4 contains bibliographical references in a broader context).

There are an infinite number of generations, one at each date. In each period t a member of generation receives an income w, which depends on the proportions of the population in one of two different skill categories, as well as the skill category to which this member belongs. This income is partly consumed and partly used in educating the offspring of this member. Depending on the education level of the child, the child receives an income next period, and the entire process repeats itself.

Each generation is altruistic – nonpaternalistically so – towards its descendants, and maximizes the discounted sum of utilities of consumption starting from that generation onwards.

The following observations are well known. If both skill categories are necessary in production, wages *must* adjust so as to force separation in educational choices even if all individuals are ex-ante identical. That is, both skill categories must be inhabited in every generation. In turn, this necessitates the emergence of (utility) inequality within generations. There must be individuals in low-paying professions that involve low training costs, whose parents invested little. And there must be others in high-paying high-training-cost professions whose parents invested a lot. In particular, *every steady state of this model must exhibit persistent inequality*, not just in current wages, but also in lifetime utilities.

The endogenous emergence of inequality in these models stands in contrast to an earlier literature on income distribution in which the degree of inequality is a compromise between the tendency towards convergence – as in any convex model of intertemporal accumulation – and ongoing stochastic shocks. In the model studied here, inequality is a necessary outcome even without stochastic shocks. Notice, too, that a nonconvex model of accumulation, while more conducive to the study of inequality (initial conditions would determine the subsequent paths), does not *necessitate* its emergence. Both the interactive structure (wages determined endogenously) as well as the imperfect capital markets assumption are needed to drive this result.

The arguments above apply to steady states, of which there are many. Indeed, a continuum of steady states is possible, each fully described by the wage structure, or at a more primitive level, by the proportion of individuals in the high-skill category. (All the steady states display inequality.) Which of these are possible attractors for paths commencing from non-steady-state initial conditions? Do all such paths converge to some steady state? The purpose of this paper is to address these questions. In a more general setting than the one here (and I will presently discuss such settings), these questions must be classified as open. However, the particular focus of the model allows me to prove that

1. If the economy starts from a steady state skill distribution, it must remain there in every subsequent period.

2. If the initial skill proportion is smaller than *some* steady state skill proportion, the economy must climb monotonically to the *nearest* steady state proportion that exceeds the initial proportion.
3. If the initial skill proportion is larger than *every* steady state skill proportion, convergence to a steady state occurs in one period.

Section 2 presents the model, and states and proves the main results. Section 3 discusses generalizations, extensions and open questions. In particular, Section 3.4 provides brief bibliographical notes.

2 The dynamics of inequality

2.1 Preliminaries

Time is discrete, running $t = 0, 1, 2 \ldots$. A *dynasty* is represented by an infinite sequence of individuals, each individual living for a single period. There is a continuum of dynasties so that a unit mass of atomless individuals belongs to a *generation* at each date.

There are two skill categories, “high" and “low", which are combined via a production function f to produce a single final output, which we take to be the numeraire. An individual in the high-skill category (or a “high individual" for short) earns a wage $\bar{w}_t$ at date t. Likewise, a low individual earns $\underline{w}_t$. Whether or not an individual is high or low depends on the investment made by her parent. Being low at any date requires no investment by the parent; being high requires an exogenous investment of x.

Earned income is partly consumed and partly used in educating the individual's offspring. Depending on the education level of the child, the child receives an income next period, and the entire process repeats itself without end.

Now I turn to the determination of wages. Assume that the production function f for final output is smooth, CRS in its two inputs, strictly concave in each input and satisfies the Inada end-point conditions. Given a unit mass of individuals, if a fraction λ of them is high at some date, then the high wage is given by

$$\bar{w}(\lambda) \equiv f_1(\lambda, 1 - \lambda),$$

while the low wage is given by

$$\underline{w}(\lambda) \equiv f_2(\lambda, 1 - \lambda),$$

where these subscripts represent partial derivatives. We will call these wages the *wages associated with* λ.

It is easy to see that $\bar{w}(\lambda)$ is decreasing and continuous in λ, with $\bar{w}(\lambda) \to \infty$ as $\lambda \to 0$. Likewise, $\underline{w}(\lambda)$ is increasing and continuous in λ, with $\underline{w}(\lambda) \to \infty$ as $\lambda \to 1$. These observations imply, in particular, that there exists a threshold $\tilde{\lambda}$ such that $\bar{w}(\tilde{\lambda}) = \underline{w}(\tilde{\lambda})$.

To complete the description of the model, we presume that each generation t maximizes an additive function of the one-period utility u_t from its own consumption, and the lifetime utility (V_{t+1}) felt by generation $t + 1$, discounted by $\delta \in (0, 1)$. The utility function u will be assumed to be increasing, smooth and strictly concave in consumption, and defined at least on $[-x, \infty)$. This last requirement is innocuous but serves to simplify notation and exposition. Moreover, the idea that consumption can go negative captures the idea that the borrowing constraint is never absolute, but that the investment of x at lower wealth levels entails ever greater utility losses (by strict concavity of u).

2.2 Equilibrium

Suppose, now, that an infinite sequence of wages is given, one for each skill category. We may denote this by the path $\{\bar{w}_t, \underline{w}_t\}_{t=0}^{\infty}$. With such a sequence given, consider the maximization problem of generation t. Denote by $\bar{V}_t$ the lifetime utility for a high member of that generation, and by $\underline{V}_t$ the corresponding lifetime utility for a low member. Standard arguments tell us that the sequence $\{\bar{V}_t, \underline{V}_t\}_{t=0}^{\infty}$ is connected over time in the following way: for each date t,

$$\bar{V}_t = \max u(c_t) + \delta V_{t+1} \tag{1}$$

subject to the conditions that

$$c_t + x_t = \bar{w}_t, \tag{2}$$

and

$$\begin{aligned} V_{t+1} &= \bar{V}_{t+1} \text{ if } x_t \geq x \\ &= \underline{V}_{t+1} \text{ if } x_t < x. \end{aligned} \tag{3}$$

In exactly the same way, $\underline{V}_t = \max u(c_t) + \delta V_{t+1}$, subject to the analogous budget constraint $c_t + x_t = \underline{w}_t$ and equation (3).

These maximization problems describe how education levels change from generation to generation, given some sequence of wage rates. To complete the equilibrium setting, we remind ourselves that the wages are endogenous; in particular, they will depend on the proportion of high individuals at each date.

Formally, for given $\lambda_0 \in (0, 1)$, a *competitive equilibrium* is a sequence $\{\bar{w}_t, \underline{w}_t, \lambda_t\}_{t=0}^{\infty}$ such that

1. Given λ_0, the path $\{\lambda_t\}$ is generated by the maximization problems described above.
2. For each t, $\bar{w}_t$ and $\underline{w}_t$ are the wages associated with λ_t.

Standard fixed-point arguments suffice to show that a competitive equilibrium exists, but we will not pursue such matters here.

Notice that our definition of competitive equilibrium assigns wages on the presumption that skilled labor *must* carry out skilled tasks. Alternatively, we could restate the definition so that if the "natural wages" (as given by marginal product) for a skilled worker falls short of that of an unskilled worker, the former will move into the sector of the latter so that the two wages will be *ex post* equalized. None of this matters much anyway because of the following easy observation, which holds no matter what definition we use:

Observation 1 *Recalling that $\tilde{\lambda}$ solves $\bar{w}(\tilde{\lambda}) = \underline{w}(\tilde{\lambda})$, $0 < \lambda_t < \tilde{\lambda}$ for all $t \geq 1$ along any competitive equilibrium.*

The formalities of the (obvious) proof are omitted.[1] From now on I shall also presume that $\lambda_0 \in (0, \tilde{\lambda})$ as well. There is no great mystery in this: it saves the expositional trouble of having to qualify several arguments for the initial value of λ.

For later use, I also record a familiar single-crossing observation.

Observation 2 *Under a competitive equilibrium, there is no date at which a low person creates a high child while simultaneously, a high person creates a low child.*

Proof If a low person creates a high child, then

$$u(\underline{w}_t - x) + \delta \bar{V}_{t+1} \geq u(\underline{w}_t) + \delta \underline{V}_{t+1},$$

or

$$u(\underline{w}_t) - u(\underline{w}_t - x) \leq \delta[\bar{V}_{t+1} - \underline{V}_{t+1}].$$

By strict concavity and the fact that $\lambda_t < \tilde{\lambda}$ for all t, we may conclude that

$$u(\bar{w}_t) - u(\bar{w}_t - x) < \delta[\bar{V}_{t+1} - \underline{V}_{t+1}].$$

But this means that a high person has a *strict* incentive to create a high child, and we are done. □

2.3 Steady states

A fraction λ is called a *steady state* if there exists a competitive equilibrium $\{\bar{w}_t, \underline{w}_t, \lambda_t\}_{t=0}^{\infty}$ from λ with $(\bar{w}_t, \underline{w}_t, \lambda_t) = (\bar{w}, \underline{w}, \lambda)$ for all t, where $\bar{w}$ and $\underline{w}$ are the wages associated with λ.

The single-crossing property in the previous section yields a simple characterization of steady states. Let $\bar{w} \equiv \bar{w}(\lambda)$ and $\underline{w} \equiv \underline{w}(\lambda)$ be the associated wages, and let $\bar{V}$ and $\underline{V}$ be the lifetime utilities associated with being (initially) high and low, respectively. By Observation 2, the following two conditions are necessary and sufficient for λ to be a steady state:

$$\bar{V} = u(\bar{w} - x) + \delta \bar{V} \geq u(\bar{w}) + \delta \underline{V},$$

while

$$\underline{V} = u(\underline{w}) + \delta \underline{V} \geq u(\underline{w} - x) + \delta \bar{V}.$$

Combining these two expressions, we may conclude that

$$u(\bar{w}) - u(\bar{w} - x) \leq \delta(\bar{V} - \underline{V}) \leq u(\underline{w}) - u(\underline{w} - x)$$

is a necessary and sufficient condition for λ to be a steady state. Combining this expression for the values of $\bar{V}$ and $\underline{V}$, we have established

[1] That $\lambda_t > 0$ for all t follows from the fact that the difference between skilled and unskilled wages would be infinitely high otherwise, so that some educational investment would have taken place prior to that period. (Here we use the assumption that u is defined on $[-x, \infty)$.) On the other hand, λ_t cannot exceed $\tilde{\lambda}$ for any $t \geq 1$, for in that case high and low wages are equalized, and no one in the previous generation would then have invested in high skills.

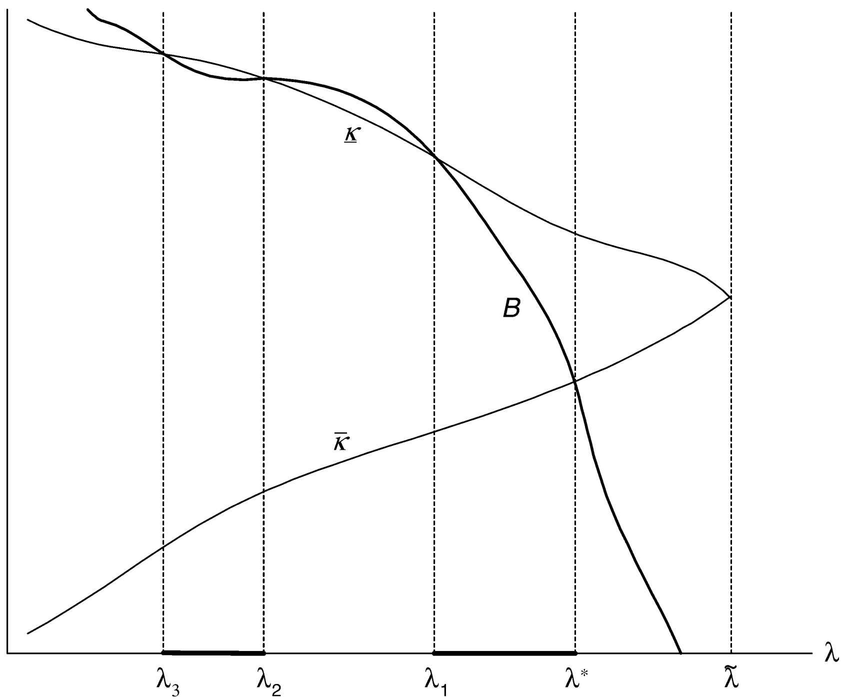

Fig. 1. The set of steady states

Theorem 1 *The fraction* λ *(with associated wages* $(\bar{w}, \underline{w})$*) is a steady state if and only if*

$$u(\bar{w}) - u(\bar{w} - x) \leq \frac{\delta}{1-\delta}[u(\bar{w} - x) - u(\underline{w})] \leq u(\underline{w}) - u(\underline{w} - x). \qquad (4)$$

Figure 1 plots the three terms in equation (4) as a function of λ. The left hand side, denoted by $\bar{\kappa}$, is just the utility cost to a high parent of acquiring skills for her child. This lies below the right hand side, denoted by $\underline{\kappa}$, which tracks the same utility cost to a low parent. Finally, the middle term, denoted by B, is the present value of benefits to being high rather than low. Of course, the $\bar{\kappa}$ and $\underline{\kappa}$ lines meet at $\tilde{\lambda}$, because wages are equalized there.

Note, moreover, that as $\lambda \to \tilde{\lambda}$, B turns negative while $\bar{\kappa}$ is positive. On the other hand, as $\lambda \to 0$, B grows unboundedly large while $\bar{\kappa}$ is bounded above. Because the changes are monotone, there is a unique $\lambda^* \in (0, \tilde{\lambda})$ such that the first inequality in equation (4) holds with equality. Observe, moreover, that at $\lambda = \lambda^*$, the *second* inequality in equation (4) must hold as well, because of the strict concavity of the utility function. Thus the set of steady states contains some interval to the left of λ^*, and must be a subset of $(0, \lambda^*]$.

Beyond this last observation, the set of steady states may be complicated. In particular, the set need not be connected. For instance, in Fig. 1, the set of steady states is the union of the two intervals $[\lambda_3, \lambda_2]$ and $[\lambda_1, \lambda^*]$.

2.4 Dynamics

Recall from the previous section that λ^* is a steady state, and it is the largest possible steady state. We now state and prove the following.

Theorem 2 *If $\lambda_0 > \lambda^*$, then there exists a unique competitive equilibrium from λ_0. It goes to a steady state in one period: $\lambda_0 > \lambda_1 = \lambda_t$ for all $t \geq 1$.*

If $\lambda_0 < \lambda^*$, then along a competitive equilibrium λ_t converges monotonically to the smallest steady state no less than λ. Convergence is never attained in finite time unless λ happens to be a steady state to start with, in which case $\lambda_t = \lambda_0$ for all subsequent t.

A discussion of the theorem is postponed to Section 3.1, but the reader uninterested in the proof (to which the remainder of this section is devoted) can go there right away.

Define $\bar{\kappa}(\lambda) \equiv u(\bar{w}(\lambda)) - u(\bar{w}(\lambda) - x)$; this is the utility cost of acquiring education when the parent is high. Define a similar utility cost for the low parent: $\underline{\kappa}(\lambda) \equiv u(\underline{w}(\lambda)) - u(\underline{w}(\lambda) - x)$. Let $b(\lambda) \equiv u(\bar{w} - x) - u(\underline{w})$ be the one-period gain to being high (assuming that the high parent also invests in her child and the low parent does not), and define $B(\lambda) \equiv (1 - \delta)^{-1} b(\lambda)$.

Finally, for any sequence $\{\lambda_s\}$ and for any date t, define

$$B_t \equiv \sum_{s=t}^{\infty} \delta^{s-t} b(\lambda_s).$$

This is the lifetime gain between a currently high and a currently low dynasty (starting from any date t), *assuming* that dynasties never switch their skill status.

The reason why B_t acquires salience is given by the following simple observation, which states that at every date, the *equilibrium* lifetime utility of the high (and low) must be equal to the utility they would have received *were their descendants never to switch status*. (To be sure, along the equilibrium path, switching of status will generally occur nevertheless.)

Lemma 1 *If $\{\bar{w}_t, \underline{w}_t, \lambda_t\}_{t=0}^{\infty}$ is a competitive equilibrium, then for each date t,*

$$\bar{V}_t = \sum_{s=t}^{\infty} \delta^{s-t} u(\bar{w}_s - x) \tag{5}$$

and

$$\underline{V}_t = \sum_{s=t}^{\infty} \delta^{s-t} u(\underline{w}_s), \tag{6}$$

so that in particular,

$$\bar{V}_t - \underline{V}_t = B_t \text{ for all } t. \tag{7}$$

Proof It suffices to show that for each $t \geq 0$,

$$\bar{V}_t = u(\bar{w}_t - x) + \delta \bar{V}_{t+1}$$

and

$$\underline{V}_t = u(\underline{w}_t) + \delta \underline{V}_{t+1}.$$

To prove this, apply Observations 1 and 2. By Observation 1 and our restriction on λ_0, $\lambda_t \in (0, \tilde{\lambda})$ for all $t \geq 0$. Now using Observation 2, we may conclude that at all dates, some of the high people stay high, while some of the low people stay low. This is enough to establish the result. □

Thus along any competitive equilibrium, no dynasty will strictly prefer to switch skills, though it may well be the case that it strictly prefers to stay where it is. Lemma 1 yields, in turn

Lemma 2 *If* $\{\bar{w}_t, \underline{w}_t, \lambda_t\}_{t=0}^{\infty}$ *is a competitive equilibrium, then for every* t, $(\bar{w}_t, \underline{w}_t)$ *are the wages associated with* λ_t, *and*

$$\bar{\kappa}(\lambda_t) \leq \delta B_{t+1} \leq \underline{\kappa}(\lambda_t), \tag{8}$$

with

$$\lambda_{t+1} > \lambda_t \text{ only if } \delta B_{t+1} = \underline{\kappa}(\lambda_t) \tag{9}$$

and

$$\lambda_{t+1} < \lambda_t \text{ only if } \delta B_{t+1} = \bar{\kappa}(\lambda_t). \tag{10}$$

Proof Let $\{\bar{w}_t, \underline{w}_t, \lambda_t\}$ be a competitive equilibrium. Then by definition, $(\bar{w}_t, \underline{w}_t)$ must be the wages associated with λ_t for every t. Using equation (5) and utility maximization, we see that

$$\bar{V}_t = \sum_{s=t}^{\infty} \delta^{s-t} u(\bar{w}_s - x) \geq u(\bar{w}_t) + \delta \underline{V}_{t+1},$$

so that (using equation (5) again)

$$u(\bar{w}_t) - u(\bar{w}_t - x) \leq \delta[\bar{V}_{t+1} - \underline{V}_{t+1}]$$

with equality holding whenever a switch from "high" to "low" does occur along the equilibrium path. Invoking equation (7) of Lemma 1, we get half of equation (8) as well as equation (10). The same argument applied to a currently low dynasty gets us the other half of equations 8 and 9. □

The next step is central:

Lemma 3 *If* $\{\bar{w}_t, \underline{w}_t, \lambda_t\}_{t=0}^{\infty}$ *is a competitive equilibrium, then for every* t,

$$\max\{B(\lambda_t), \frac{1}{\delta}\bar{\kappa}(\lambda_t)\} \geq B_{t+1} \geq \min\{B(\lambda_t), \frac{1}{\delta}\underline{\kappa}(\lambda_t)\}. \tag{11}$$

Proof It suffices to prove the result for $t = 0$. I first show that

$$\max\{B(\lambda_0), \frac{1}{\delta}\bar{\kappa}(\lambda_0)\} \geq B_1. \tag{12}$$

Suppose this assertion is false. Then I claim that there exists a *first* date $T \geq 0$ such that $\lambda_0 = \cdots = \lambda_T$ and

$$B_{T+1} \geq \frac{1}{\delta}\underline{\kappa}(\lambda_T). \tag{13}$$

Of course, if $B_1 \geq (1/\delta)\underline{\kappa}(\lambda_0)$, the claim is automatically true; otherwise $B_1 < (1/\delta)\underline{\kappa}(\lambda_0)$. We also have (by virtue of the presumption that equation (12) does not hold) that $B_1 > (1/\delta)\bar{\kappa}(\lambda_0)$, so that

$$\frac{1}{\delta}\bar{\kappa}(\lambda_0) < B_1 < \frac{1}{\delta}\underline{\kappa}(\lambda_0)$$

which – by Lemma 2 – implies that $\lambda_0 = \lambda_1$. Moreover, $B_1 = b(\lambda_1) + \delta B_2 = b(\lambda_0) + \delta B_2$, while $B(\lambda_0) = b(\lambda_0) + \delta B(\lambda_0)$, so that

$$B_2 - B(\lambda_0) = \frac{1}{\delta}[B_1 - B(\lambda_0)],$$

and simple manipulation of this equality shows that

$$B_2 = B_1 + \frac{1-\delta}{\delta}\epsilon, \tag{14}$$

where $\epsilon \equiv B_1 - B(\lambda_0) > 0$ (again, by the failure of equation (12)).

Following the same reasoning leading up to equation (14), as long as $B_{t+1} < (1/\delta)\underline{\kappa}(\lambda_t)$ – and as long as this is also true of all dates before t – we have $\lambda_0 = \cdots = \lambda_t = \lambda_{t+1}$, and

$$B_{t+1} = B_t + \frac{1-\delta}{\delta}\epsilon, \tag{15}$$

It follows that there must be a first date T such that $\lambda_0 = \cdots = \lambda_T$, and equation (13) holds, as claimed.

On the other hand, equation (13) cannot hold with strict inequality, as this would surely violate equation (8) of Lemma 2, so it must be that

$$B_{T+1} = \frac{1}{\delta}\underline{\kappa}(\lambda_T). \tag{16}$$

In turn, this means that $\lambda_{T+1} \geq \lambda_T$ (see equation (10)). Moreover, $\delta B_{T+2} = B_{T+1} - b(\lambda_{T+1})$, so that

$$\begin{aligned} \delta B_{T+2} - \underline{\kappa}(\lambda_{T+1}) &= B_{T+1} - b(\lambda_{T+1}) - \underline{\kappa}(\lambda_{T+1}) \\ &\geq B_{T+1} - b(\lambda_T) - \underline{\kappa}(\lambda_T) \\ &= B_{T+1} - b(\lambda_T) - \delta B_{T+1} \\ &= B_{T+1} - B_T > 0, \end{aligned} \tag{17}$$

which contradicts equation (8) at date $T + 1$. This establishes equation (12).

Now we prove that

$$B_1 \geq \min\left\{B(\lambda_0),\ \frac{1}{\delta}\underline{\kappa}(\lambda_0)\right\}. \tag{18}$$

The argument runs closely parallel to the previous one. Suppose equation (18) is false. Then I claim that there exists a *first* date $T \geq 0$ such that $\lambda_0 = \cdots = \lambda_T$ and

$$B_{T+1} \leq \frac{1}{\delta}\bar{\kappa}(\lambda_T). \tag{19}$$

The steps are very similar to those used to establish equation (13), and are omitted.

Continuing the parallel argument, equation (19) cannot hold with strict inequality, so we have

$$B_{T+1} = \frac{1}{\delta}\bar{\kappa}(\lambda_T). \tag{20}$$

In turn, this means that $\lambda_{T+1} \leq \lambda_T$. These two observations establish, however, that

$$\delta B_{T+2} - \bar{\kappa}(\lambda_{T+1}) < 0,$$

(following steps parallel to those establishing equation (17)), which contradicts equation (8) and completes the proof of the lemma. □

The following observation is a simple consequence of Lemma 3:

Lemma 4 *Along any equilibrium, if $B(\lambda_t) \leq (1/\delta)\bar{\kappa}(\lambda_t)$, then $B_{t+1} = (1/\delta)\bar{\kappa}(\lambda_t)$. Similarly, if $B(\lambda_t) \geq (1/\delta)\underline{\kappa}(\lambda_t)$, then $B_{t+1} = (1/\delta)\underline{\kappa}(\lambda_t)$.*

Proof We prove the first part; the second part uses a completely analogous argument. If $B(\lambda_t) \leq (1/\delta)\bar{\kappa}(\lambda_t)$, then $\max\{B(\lambda_t), (1/\delta)\bar{\kappa}(\lambda_t)\} = (1/\delta)\bar{\kappa}(\lambda_t)$, so that by equation (11), $(1/\delta)\bar{\kappa}(\lambda_t) \geq B_{t+1}$. On the other hand, Lemma 2 tells us that $(1/\delta)\bar{\kappa}(\lambda_t) \leq B_{t+1}$, and the proof is complete. □

With these steps in hand, we may complete the proof of the theorem. There are three possibilities to consider (each a restriction on the initial value λ_0):

I. $(1/\delta)\bar{\kappa}(\lambda_0) < B(\lambda_0) < (1/\delta)\underline{\kappa}(\lambda_0)$. Then by (11), $B(\lambda_0) = B_1$. In particular,

$$(1/\delta)\bar{\kappa}(\lambda_0) < B_1 < (1/\delta)\underline{\kappa}(\lambda_0)$$

so that by Lemma 2, $\lambda_0 = \lambda_1$. Continuing the argument recursively, we see that $\lambda_t = \lambda_0$ for all t.

II. $B(\lambda_0) \leq (1/\delta)\bar{\kappa}(\lambda_0)$. Then by Lemma 4, $B_1 = (1/\delta)\bar{\kappa}(\lambda_0)$ and so by Lemma 2, $\lambda_1 \leq \lambda_0$. Suppose, in fact, that strict inequality holds. Then $(1/\delta)\bar{\kappa}(\lambda_1) < (1/\delta)\bar{\kappa}(\lambda_0)$ and $(1/\delta)\underline{\kappa}(\lambda_1) > (1/\delta)\underline{\kappa}(\lambda_0)$, so that

$$(1/\delta)\underline{\kappa}(\lambda_1) > B_1 > (1/\delta)\bar{\kappa}(\lambda_1). \tag{21}$$

Now I claim that in fact,

$$(1/\delta)\underline{\kappa}(\lambda_1) > B(\lambda_1) > (1/\delta)\bar{\kappa}(\lambda_1). \tag{22}$$

Suppose not. First suppose that $B(\lambda_1) \le (1/\delta)\bar{\kappa}(\lambda_1)$. Then by Lemma 4, $B_2 \le (1/\delta)\bar{\kappa}(\lambda_1)$, so that $B_1 = (1-\delta)B(\lambda_1) + \delta B_2 \le (1/\delta)\bar{\kappa}(\lambda_1)$, which contradicts equation (21). In exactly the same way, one can rule out the possibility that $B(\lambda_1) \ge (1/\delta)\underline{\kappa}(\lambda_1)$, so equation (22) is established.

Now we are in Case I, and λ must remain constant thereafter. So in Case II, we move to a steady state *in at most one step*.

III. $B(\lambda_0) \ge (1/\delta)\underline{\kappa}(\lambda_0)$. Then by Lemma 4, $B_1 = (1/\delta)\underline{\kappa}(\lambda_0)$ and so by Lemma 2, $\lambda_1 \ge \lambda_0$.

To bring out the contrast with Case II, I claim that $B(\lambda_1) \ge (1/\delta)\underline{\kappa}(\lambda_1)$, with strict inequality if the corresponding inequality at date 0 also holds strictly.

Suppose on the contrary that $B(\lambda_1) < (1/\delta)\underline{\kappa}(\lambda_1)$. Then

$$B_1 = (1/\delta)\underline{\kappa}(\lambda_0) \ge (1/\delta)\underline{\kappa}(\lambda_1) > B(\lambda_1), \tag{23}$$

while it is also true that

$$B_1 = (1/\delta)\underline{\kappa}(\lambda_0) \ge (1/\delta)\underline{\kappa}(\lambda_1) \ge B_2. \tag{24}$$

But equations (23) and (24) together contradict the fact that $B_1 = (1-\delta)B(\lambda_1) + \delta B_2$.

If strict inequality holds at date 0 – $B(\lambda_0) > (1/\delta)\underline{\kappa}(\lambda_0)$ – then we can arrive at a contradiction simply under the weaker condition $B(\lambda_1) \le (1/\delta)\underline{\kappa}(\lambda_1)$. For then $\lambda_1 \ne \lambda_0$ and therefore (because $\lambda_1 \ge \lambda_0$) it must be that $\lambda_1 > \lambda_0$. Therefore the weak inequality in equation (23) holds strictly, and we obtain the same contradiction.

So the claim is established and we can apply Case III repeatedly to argue that $\lambda_{t+1} \ge \lambda_t$ for all t in this case.

Moreover, if $B(\lambda_0) > (1/\delta)\underline{\kappa}(\lambda_0)$, then $B(\lambda_t) > (1/\delta)\underline{\kappa}(\lambda_t)$ for all t subsequently, so convergence *cannot ever occur in finite time*, in contrast to the "one-step" property of Case II.

Finally, observe that in the case $B(\lambda_0) > (1/\delta)\underline{\kappa}(\lambda_0)$, in which $\lambda_{t+1} > \lambda_t$ for all t, there is no t and no $\lambda \in [\lambda_t, \lambda_{t+1}]$ which is a steady state. For suppose there were; then in particular, $\underline{\kappa}(\lambda) \ge B(\lambda)$, so that

$$\underline{\kappa}(\lambda_t) \ge \underline{\kappa}(\lambda) \ge B(\lambda) \ge B(\lambda_{t+1}) > B_{t+1}, \tag{25}$$

where the very last inequality follows from the fact that $\lambda_s < \lambda_{s+1}$ for all s. But equation (25) contradicts the equilibrium condition equation (8). This proves that in Case III, convergence occurs to the *smallest* steady state to the right of λ_0.

3 Discussion

3.1 Theorem 2

The theorem provides a full account of the behavior of skill proportions over time, starting from any initial condition. If that initial condition happens to be a steady state, the theorem rules out any equilibrium path other than the steady state path itself. More interesting is the asymmetry of equilibrium behavior under the two remaining kinds of initial conditions. When λ_0 is larger than the largest conceivable steady state, convergence to a steady state occurs in a single unit of time.

When λ_0 is such that there are steady states “above" it, convergence is gradual in that the process is never completed in finite time. This asymmetry may have interesting implications for unanticipated technical changes which, once realized, are expected to stay in place thereafter. Changes that call for a reduction in steady state skill proportions take place quickly and dramatically, whereas a climb to a higher steady state is more gradual and drawn out.

Moreover, in the case that convergence is “up" to a steady state, the theorem asserts that it will occur to the *nearest* steady state to the right of the initial conditions. Put another way, only the left-most steady state in each interval of steady states can be an attractor for initial conditions that are distinct from that steady state, and the basin of attraction is precisely the set of initial conditions that lie between it and the next lower interval of steady states (if any). Thus, despite the multiplicity of steady states, final outcomes can be tagged to initial conditions in a unique way, allowing us in principle to perform comparative dynamics. Tanaka (2003), which I discuss in more detail below, uses Theorem 2 in exactly this way.

To provide some intuition for these results, first study a non-steady-state value of λ that is smaller than λ^*. Why is this not a steady state? Surely, the “no-deviation" condition for the skilled (the first inequality in equation (4)) is satisfied; after all, it was satisfied at λ^*, and now at the smaller value of λ, *both* the wage differential is higher *and* the utility cost of education is lower for the skilled. So, the reason why λ fails to be a steady state is that the “no-deviation" condition for the *unskilled* – the second inequality in equation (4) – is violated; their utility cost of education is high, but not as high as the wage differential. To maintain equilibrium incentives, then, the economy-wide skill ratio must rise, compressing the sequence of wage differentials until the unskilled are exactly indifferent between acquiring and not acquiring skills. (We have already seen in Lemma 1 that this compression to indifference is a necessary feature of non-steady-state equilibrium, otherwise no one would stay unskilled.)

Now here is the main point: *the skill ratio achieved in the very next period cannot be a steady state.* For if it were, then the no-deviation condition of the unskilled must be satisfied here, so that the new implied wage differential generates no incentive for them to acquire skills. But the very same wage differential created indifference at date 0, when the utility cost of acquiring education was higher for the unskilled! This is a contradiction. By an obvious recursive argument, it follows that the upward movement in the skill ratio must be gradual and perennial.

Exactly the opposite is true for non-steady-state values of λ that exceed λ^*. For such values, the “no-deviation" condition for the skilled surely fails; the wage differential is too small relative to the utility cost of maintaining skills. So in equilibrium, λ falls. This fall along the equilibrium path raises wage differentials so that the skilled are now indifferent between maintaining and relinquishing skills.

We claim that *the new skill ratio one period later must be a steady state.* Suppose the claim is false. Then the new skill ratio is not a steady state, and this can happen for one of two reasons. First, the no-deviation condition for the unskilled fails – the wage differential at date 1 is too attractive. In that case, we already know that the remaining sequence of wage differentials (counting from date two on) must render the unskilled indifferent at date 1. But this means that the sequence of wage differentials counting from date *one* on is still attractive for the unskilled, and would have been *a fortiori* attractive relative to the skill-acquisition cost for the unskilled at the higher skill ratio prevailing at date 0. But this, in turn, exceeds the cost of skill maintenance for the skilled, which contradicts the fact that the equilibrium creates indifference for the skilled at date 0 (see last sentence of preceding paragraph).

The second reason why the new skill ratio may fail to be a steady state is that the no-deviation condition for the skilled fails again (just as it did at date 0). This means that the skill ratio must fall even further in succeeding periods to create indifference for the skilled at date 1. However, because the wage differential at date 1 is not attractive enough, this means that the entire sequence of wage differentials, *counting* the one at date 1, would not have been attractive enough for the very same skilled individuals, had they been located at date 0, but with the utility costs they possess at date 1. This would be *a fortiori* true if we were to replace the utility costs with the true utility costs of maintaining skills at date 0, which are higher. But now we have a contradiction again (see the last sentence two paragraphs above). This completes our intuitive description.

Notice that the observations above are in "sequence space"; i.e., they pertain to the infinite sequence $\{\lambda_t\}$ that makes up a competitive equilibrium. One might equivalently conduct the analysis in terms of the mapping that takes current values of λ to next period's equilibrium value of λ. More formally, define a correspondence G by

$$G(\lambda) \equiv \{\lambda' | \text{ There is an equilibrium with } \lambda_0 = \lambda \text{ and } \lambda_1 = \lambda'\}.$$

Our main theorem implies that G coincides with the 45° line whenever λ is restricted to be a steady state, lies strictly *above* the 45° line whenever λ is not a steady state but is nevertheless smaller than the largest steady state, and lies below the 45° line when λ exceeds the largest steady state. In this last case, there are additional restrictions: the image of G must be a subset of the set of steady states.

3.2 Generalizations

Our main result asserts convergence from every initial condition to some steady state. Moreover, the particular steady state to which convergence occurs can be identified, and this is of interest because there are multiple steady states.

How general are these results; in particular, do they extend to the case of several occupations? The simple answer is that I do not know. Certainly the methods I use rely heavily on the assumption that there are only two skill levels. One can perhaps go further with the dynamics if one is willing to assume generational utility functions that are somewhat more shortsighted. For instance, one might assume (though in my opinion this is an unsatisfactory approach) that utility is a "warm-glow" function of consumption and educational bequests. Less drastically, one might presume that utility is a function of current consumption and descendants' wealth.[2] But perhaps the most satisfactory formulation is the current one, based entirely on nonpaternalistic altruism. Apart from its conceptual attractiveness, such a formulation is also closely related to the literature on optimal growth

[2] The reason this is less drastic is that the relationship between bequest and descendant wealth is endogenous.

theory. Unfortunately, as far as the study of dynamics goes, the most satisfactory formulation also appears to be the hardest. It is true that versions of the turnpike theorem formulated for competitive economies are available (see Bewley 1982; Coles 1985; or Yano 1984) but these theorems do not apply here, as the arguments rely on the equivalence between competitive equilibria and full Pareto-optimality. (Such equivalence does not obtain in our setting because the credit market is missing.)

At the same time, as Mookherjee and Ray (2002, 2003) have shown, the basic setting of persistent inequality is very general. With endogenous price determination and imperfect credit markets, steady-state inequality among ex-ante identical agents is inevitable under broad conditions. It would, therefore, be of great interest to see if the dynamic counterparts of the results in this paper also carry over.

3.3 Extensions

3.3.1 Financial bequests

A major research question concerns an extension to the case in which financial bequests can supplement human capital investments. Formally, it is possible to view this as a special case of the model with multiple occupations and no financial bequests, because each different monetary value of the bequest can be treated as a separate occupation. However, as Mookherjee and Ray (2002) argue, the conditions required for the emergence of persistent inequality in steady state are now stronger. If those conditions are met, all steady states will display inequality as before, but if they are not, then perfect equality and inequality will co-exist.

Now the burden on the dynamics is a weightier one, for the dynamic model acts as a selection device for steady states. Which initial conditions are conducive to the (ultimate) emergence of perfect equality, and which are not? The need for a dynamic analysis is even stronger here. (Again, an analysis based on “simple" bequest motives is more likely to be tractable.)

3.3.2 Technical progess

Rigolini (2004) extends the general version of this model (as described in Mookherjee and Ray 2003) to analyze incentives to acquire skills under technological progress. In his setup, the costs of acquiring various skills also drift with technical progress, so that he is able to describe steady state balanced growth path. In his model, higher rates of technical progress *decrease* inequality in steady state; that is, in the “long run". yet we know that the “short-run" effects of technical progress can be quite the opposite; indeed, a substantial recent literature lays the blame for increasing wage differentials in OECD countries at the door of technical change. It is impossible to reconcile these two findings without a thorough understanding of the (non steady-state) dynamics of the model.

3.3.3 Trade

The model developed here can also bear on issues in trade and trade policy. Tanaka (2003) provides another extension of this framework in which the dynamics described here are put to an interesting application. Consider two copies of this model running side by side, with each skill producing a different commodity. View these as two countries under autarky. Then initial steady-state conditions determine comparative advantage in the two commodities (with the more unequal country having relative advantage in the “unskilled commodity"). A trade opening will then drive the “world economy" to a steady state, but this permits predictions to be made regarding inequality *across* countries. Tanaka uses the dynamic analysis presented here to show that the size of the long-run income gap between the two countries depends on the difference in domestic income inequality when they open up to trade. Based on these results, he also analyzes the effects of redistributive policy within a country, showing that redistribution in one country may increase the income of its trading partner if it is undertaken in steady state, while the opposite is true if the policy is undertaken during transition.

One may also use the model somewhat differently to understand the effects of economic integration on inequality across countries, by interpreting each family to represent a different country. For instance, one could view “occupational" setup costs as infrastructural investments made by the planners to facilitate a particular mix of economic activities in each country (e.g., a country may decide to subsidize agriculture, promote exports, or invest in high technology production capabilities). Then – in the absence of a perfect international capital market to finance these investments – global inequality must emerge, with historical events determining the subsequent fate of individual countries.

3.4 Bibliographical notes

Champernowne (1953), Becker and Tomes (1979), and Loury (1981) were among the first to study the evolution of inequality in a dynamic model. These models are noninteractive, in that it is sufficient to trace the stochastic behavior of a single dynasty without studying interactions across dynasties. They are also based on an underlying model of convergence (equivalently, the technology set in these models is convex). The theory of growth with nonconvex technologies, as described, for instance, in Clark (1971), Dechert and Nishimura (1983), and Majumdar and Mitra (1982, 1983) remains largely unexplored for the study of endogenous inequality, though see the first part of Galor and Zeira (1993). Aghion and Bolton (1997), Banerjee and Newman (1993), Freeman (1996), Galor and Zeira (1993), Ghatak and Jiang (2002), Matsuyama (2000), Ljungqvist (1993), Mani (2001), Mookherjee and Ray (2002, 2003), Piketty (1997), Ray (1990), and Ray and Streufert (1993) (among others) study the interactive model from different points of view and at various levels of generality. However, to my knowledge, there is no systematic study of the dynamics of this framework away from steady state when each generation has standard nonpaternalistic preferences. The present paper takes a step in that direction.

References

Aghion, P., Bolton, P.: A theory of trickle-down growth and development. Rev Econ Stud **64**, 151–172 (1997)

Bandyopadhyay, D.: Distribution of human capital, income inequality and rate of growth, PhD thesis, University of Minnesota (1993)

Banerjee, A., Newman, A.: Occupational choice and the process of development. J Political Econ **101**, 274–298 (1993)
Becker, G., Tomes, N.: An equilibrium theory of the distribution of income and intergenerational mobility. J Political Econ **87**, 1153-1189 (1979)
Bewley, T.: An integration of equilibrium theory and turnpike theory. J Math Econ **10**, 233–267 (1982)
Champernowne, D.: A model of income distribution. Econ J **63**, 318–351 (1953)
Clark, C.: Economically optimal policies for the utilization of biologically renewable resources. Math Biosci **12**, 245–260 (1971)
Coles, J.: Equilibrium turnpike theory with constant returns to scale and possibly heterogeneous discount factors. Int Econ Rev **26**, 671–679 (1985)
Dechert, W.D., Nishimura, K.: A complete characterization of optimal growth paths in an aggregated model with a non-concave production function. J Econ Theory **31**, 332–354 (1983)
Freeman, S.: Equilibrium income inequality among identical agents. J Political Econ **104**(5), 1047–1064 (1996)
Galor, O., Zeira, J.: Income distribution and macroeconomics. Rev Econ Stud **60**(1), 35–52 (1993)
Ghatak, M., Jiang, N.: A simple model of inequality, occupational choice and development. J Dev Econ **69**, 205–226 (2002)
Ljungqvist, L.: Economic underdevelopment: the case of missing market for human capital. J Dev Econ **40**, 219–239 (1993)
Loury, G.C.: Intergenerational transfers and the distribution of earnings. Econometrica **49**, 843–867 (1981)
Majumdar, M., Mitra, T.: Intertemporal allocation with a non-convex technology: the aggregative framework. J Econ Theory **27**, 101–136 (1982)
Majumdar, M., Mitra, T.: Dynamic optimization with a non-convex technology: the case of a linear objective function. Rev Econ Stud **50**, 143–151 (1983)
Mani, A.: Income distribution and the demand constraint. J Econ Growth **6**, 107–133 (2001)
Matsuyama, K.: Endogenous inequality. Review of Economic Studies **67**, 743–759; Constraint. J Econ Growth **6**, 107–133 (2000)
Mookherjee, D., Ray, D.: Is equality stable? Am Econ Rev **92**, 253–259 (2002)
Mookherjee, D., Ray, D.: Persistent inequality. Rev Econ Stud **70**, 369–393 (2003)
Piketty, T.: The dynamics of the wealth distribution and the interest rate with credit rationing. Rev Econ Stud **64**, 173–189 (1997)
Ray, D.: Income distribution and macroeconomic behavior. Mimeo, New York University, `http://www.econ.nyu.edu/user/debraj/DevEcon/Notes/incdist.pdf` (1990)
Ray, D., Streufert, P.: Dynamic equilibria with unemployment due to undernourishment. Econ Theory **3**, 61–85 (1993)
Rigolini, J.: Educational technologies, wages, and technological progress. J Dev Econ **75**, 55–77 (2004)
Tanaka, R.: Timing of trade liberalization. Mimeo, Department of Economics, New York University (2003)
Yano, M.: The turnpike of dynamic general equilibrium paths and its insensitivity to initial conditions. J Math Econ **13**, 235–254 (1984)