Persistent Inequality

Dilip Mookherjee
Boston University

Debraj Ray
New York University and Instituto de Análisis Económico (CSIC)

`This version: December 2000`

Abstract

This paper studies the persistence of inequality and lack of intergenerational mobility in a human capital accumulation model with perfect certainty, missing credit markets, endogenous training costs and dynastic utility maximization by households. Persistent inequality and immobility in steady states arise under general conditions, irrespective of the divisibility of human capital. The multiplicity of steady states, however, does require indivisibilities in human capital, and widens in the presence of alternative forms of nonhuman wealth. Convergence to steady states and the comparative dynamic effects of redistributive policies are studied in a two occupation context.

We thank Abhijit Banerjee, Michael Kremer, Glenn Loury and Andy Newman for helpful discussions, and seminar participants at Columbia University and the July 2000 NBER Income Distribution workshop for their comments. This research was supported by the National Science Foundation (NSF grant no. SBS-9709254), and Mookherjee was additionally funded by the MacArthur Foundation. Address correspondence to `dilipm@bu.edu` and `debraj.ray@nyu.edu`.

1 Introduction

The neoclassical growth model forms an important benchmark in most economists' thinking about income inequality and mobility. The central prediction of the model is that the market mechanism intrinsically promotes a tendency towards *convergence* of incomes of different agents, families or countries, a process by which historical inequality tends to vanish and cease to matter in the long run.[1]

Reformulations of this model in the context of intergenerational mobility (Becker and Tomes [1979], Loury [1981] and Mulligan [1997]) therefore rely on the presence of random factors ("luck") in explaining the persistence of inequality;[2] in the absence of such extrinsic randomness the models predict that all families will eventually converge to the same level of well-being. For this reason, the focus in much of this literature is on the value of "persistence coefficients" (such as the sensitivity of bequests to a unit increase in wealth) in determining the extent to which shocks might echo into the future before finally dying out. If the coefficient is less than unity (as it is in the bulk of these frameworks), inequality must be the result of constant stochastic perturbations. And even in the presence of these random factors, income distribution converges to a unique steady state distribution, independent of initial conditions — so historical endowements do not matter in the long run.

In contrast, a recent literature (see Banerjee and Newman [1993], Galor and Zeira [1993], Ljungqvist [1993], and Ray and Streufert [1993]) generates opposite predictions concerning the significance and persistence of inequality (in consumption and per period utilities, rather than earned incomes). These models, and others of a similar vein (see, e.g., Aghion and Bolton [1997], Bandopadhyay [1997], Freeman [1996], Mani [1997], Maoz and Moav [1999], Piketty [1997]), show how historical inequality may amplify and persist in the long run, despite the absence of any random shocks to technology or endowments.[3] While set in a wide variety of contexts, such as labor markets, occupational choice, human capital and financial markets, they all generate multiple steady states with varying inequality and per capita income. In particular, historical inequality can

---

[1] The original formulation of the model by Solow [1956] generates the convergence predictions in the presence of a number of additional assumptions, such as constant savings rates, and the absence of any uncertainty in technology or endowments. Loury's formulation of this model clarified that the longrun convergence predictions of the model survive in the presence of random shocks in technology and endowments, and with "strategic" savings behavior (i.e., dynastic utility maximization by households). In particular, the results extend to the context where credit markets are entirely absent.

[2] This same notion finds earlier grounding in the formulations of Champernowne [1953] and the subsequent literature on "stochastic" models of income distribution.

[3] It should be noted that there is an early literature on economic growth with nonconvex technologies that can be modified to yield similar observations regarding the persistence of inequality (see, e.g., Clark [1971], Skiba [1978], Majumdar and Mitra [1982, 1983], Dechert and Nishimura [1983], and Mitra and Ray [1984]). Papers such as Dasgupta and Ray [1986, 1987] argue that inequality has strong and persistent *functional* effects, though no dynamic model is studied there.

account for economic underdevelopment, and one shot redistributive policies can have permanent macroeconomic effects. With few exceptions, their underlying assumptions differ in a variety of respects from the neoclassical framework: there are indivisibilities in the investment technology (which is thereby nonconvex); the returns to investment may be subject to pecuniary externalities (thus not determined by technology alone), and households save a constant fraction of their income (owing to a 'warm glow' bequest motive).[4] Perhaps the central feature of all these models is that they posit imperfect credit markets.

The aim of this paper is to study the persistence and history-dependence of income inequality in a setting general enough to encompass several different modeling approaches. This exercise generates some new insights, and at the same time makes explicit the source of disparate conclusions arrived at by previous authors. We first describe the main features of our framework, and then summarize the main results.

First, our formulation of the production and investment technologies allows for indivisibilities (a hallmark of most nonergodic models) as well as an array of occupations or investment choices that are almost perfectly divisible (as in the neoclassical model). Our model is cast in the human capital setting, where different stocks of capital correspond either to differing amounts of skill within a given occupation, or to different occupations altogether. This allows us to explore the role of indivisibilities in the generation and persistence of inequality. While it is arguable that some indivisibilities are inherent in human capital, it is too extreme to postulate that there are only two or three possible occupations or levels of skill with differing training costs. For any two occupations with sharply varying levels of income, one can typically find a large number of intermediate occupations — e.g., between unskilled blue collar occupations such as janitors and highly skilled occupations such as managers, doctors and engineers, there are clerical services, plumbers, electricians, builders, hairdressers, nurses, bus drivers etc., that require varying levels of vocational training. And within any occupation there are various levels of training and experience: between a nurse and a brain surgeon there are paramedics, chiropractors, general practitioners and medical specialists. Accordingly, we wish to allow for a large number of possible levels of human capital, and explore the effects of decreasing the degree of indivisibility, with a continuum of possible levels of human capital forming a limit to the case of vanishing indivisibility.

Second, we allow for the returns to human capital as well as the costs of acquiring human capital to be subject to pecuniary externalities, i.e., be interdependent across different households, rather than determined by technology alone. In particular, skill or occupation premia will depend on the distribution of the existing population across different skills and occupations. Since the acquisition of skill frequently involves a process

---

[4] The papers by Freeman and Lundqvist referred to above, however, also consider dynastic utility maximization as the bequest motive.

of being trained by skilled personnel, the costs of acquiring human capital will also depend on existing skill premia. Combined with the possible divisibility of human capital, this implies that the generation of inequality in our model does not rely on the existence of a set of exogenous investment thresholds built into the formulation of the technology. If inequality emerges and persists owing to the existence of investment thresholds or nonconvex returns, these owe entirely to skill premia patterns that are endogenously generated. The endogeneity of the thresholds also implies that inequality can persist despite the presence of income growth (e.g., arising from technical progress), since this raises both benefits and costs of skill acquisition. In contrast, thresholds arising from technological indivisibilities can be eventually overcome by the entire population once income growth —slow though it may be — percolates throughout the economy.

Third, in common with much of the recent literature, we posit that opportunities to finance human capital investments by borrowing are limited. Indeed, for the most part, we follow Loury [1981] by assuming that credit markets are entirely absent. While it is possble to argue that the difficulty of using human capital (particularly of one's children) as collateral make this a natural and realistic assumption, our primary goal is to span the two classes of models described above and we therefore continue to make similar assumptions. Indeed, it will become evident that similar conclusions follow when credit markets are present but are imperfect, e.g., when borrowing rates of interest exceed the lending rate of interest.

Fourth, we assume that every household is motivated by dynastic utility maximization in consumption-investment decisions, and seeks to smooth consumption across successive generations (as in Loury [1981]). This is in contrast to formulations that postulate a constant savings rate, a constant bequest rate, or a utility function that depends on consumption and bequests (without taking stock of the *impact* of those bequests). The primary motivation for this is not realism, but rather to explore the role of alternative formulations of intergenerational altruism in inequality persistence or history-dependence. For instance, it might be argued that many existing formulations of persistent inequality (such as Banerjee and Newman [1993], Galor and Zeira [1993], or Piketty [1997]) leave little or no room for optimizing behavior in the presence of adversity.[5] If a family is locked into a poverty trap because they do not possess the wherewithal to pay for fixed costs that would lift them out of poverty, what is there to prevent them from saving for this eventuality, bit by bit, so that their descendants will at least be able to break free of such traps? This is not to argue that optimization will inevitably result in such liberation — indeed, we show otherwise in this paper — but to ensure that such considerations do not drive the results. We shall also argue that the principal qualitative results extend to the case of 'warm glow' bequest motives, so they do not really depend on the precise

---

[5] For examples of fully optimizing models (with endogenous prices), see, for instance, Bandyopadhyay [1997], Freeman [1996], Ljungqvist [1993] or Ray and Streufert [1993].

formulation of altruism.

Finally, we assume no uncertainty of any kind, either in technology or in preferences. Again this is not motivated by considerations of realism, for clearly randomness in ability and incomes is pervasive and an important source of inequality and mobility.[6] But the extent to which this may be so has been extensively studied in previous literature. Our purpose instead is to reconsider the question whether the market mechanism intrinsically tends to generate or dissolve inequality, a question most fruitfully addressed in a context *sans* uncertainty, as in the original formulation of the neoclassical model by Solow [1956]. Moreover, uncertainty may create the impression of mobility or ergodicity when — for all practical purposes — there is none.[7]

Our paper is organized as follows. Section 2 explores the generality of the result concerning persistence of inequality, after introducing the model, and explaining how different existing models of intergenerational mobility can be embedded as special cases. If there are no forms of nonhuman capital, we show that indivisibilities in human capital are entirely inessential to the result that all steady states must involve persistent utility differences across *ex ante* identical households, and lack of intergenerational mobility. All that is required is a minimal degree of specialization: at least two distinct occupations (or levels of skill) involving differential training cost will be selected by some households in the economy.

Section 3 subsequently explores sources of the multiplicity of steady states, which underlies possible history dependence. We start with a context where there are just two occupations (or levels of skill) — one involving some training and the other not requiring any training at all — and then allow for intermediate occupations (or skills). It turns out that the 'thicker' the space of possible levels of human capital, the more restricted is the extent of *multiplicity* of steady states. In the limit as there is a continuum of possible skill levels that are occupied, the steady state is unique. Hence the degree of divisibility of human capital levels affects the extent of possible history dependence of

---

[6] Indeed, in the absence of any uncertainty, all steady states in our model are characterized by zero occupational mobility. So we would interpret whatever mobility is observed in the world as arising from the effect of random shocks.

[7] Notice that this feature, in a sense, is almost opposite in spirit to the first — inequality *creating* — aspect of uncertainty. For instance, Galor and Zeira [1993] show that when there are significant investment thresholds, then final outcomes for an individual may depend on initial conditions. This sort of path dependence is easy enough to knock out — at least formally — by introducing some uncertainty, however small, as long as the support of the uncertainty is quite large (for instance, if there is always some probability that any given person will win the State Lottery). The peculiar laws of stochastic processes running in infinite time then dictate that there is full ergodicity of the Galor-Zeira process. But this conveys a misleading impression that the Galor-Zeira model invariably tends to promote long run equality. One sensible way of dealing with this problem is to use expected utility evaluations of all future streams, which would deal with low-probability events by giving them insignificant utility weight. This is exactly the approach we shall follow in evaluating welfares of households; consequently the results will not depend in any intrinsic manner on the existence of small amounts of randomness.

steady states, but not the fact that all steady states must involve persistent inequality. Accordingly, explanations for varying levels of long-run per capita income in terms of historical inequality, or arguments for permanent macroeconomic effects of one shot redistributive policies, must rely on some degree of indivisibility of investments.

Section 4 addresses the question of convergence to steady states, which is essential to any model that explains the role of historical inequality. We focus on a particular example which is a special case of the models developed earlier, involving two levels of skill and exogenous training costs. In this example we establish that the model has a unique competitive equilibrium which always converges to some steady state. It helps clarify how inequality emerges and persists, even if the economy starts (out of steady state) in perfect equality. Moreover, we are able to describe the mapping from initial inequality to long run inequality and per capita output, and derive comparative dynamic effects of one shot redistributive policies.

Finally, Section 5 enlarges the model to incorporate alternative forms of capital – physical or financial – which can be bequeathed by parents to children, besides human capital investments. In the context of the simple example with two skill levels, we show that the long run persistence of inequality is no longer inevitable. In particular, it is possible for ownership of human and nonhuman capital to be negatively correlated in steady states in such a way to equalize welfares across households. At the same time the unequal steady states arising in the world without nonhuman capital also exist, where the distributions of human and nonhuman capital are positively correlated. Hence the degree of multiplicity of steady states tends to be widen in the presence of alternative forms of capital.

2 The General Argument Underlying Persistent Inequality

The purpose of this section is to construct a framework wide enough to encompass most existing models of persistent inequality, in order to identify the elements common to them, and understand how far they extend. Our formulation in this section accordingly allows for multiple commodities and forms of human capital, and for a general specification of the technology of production and training. However, the model is special by omitting any form of nonhuman capital, an issue which will be addressed in Section 5.

2.1 Preferences and Technology

We consider a continuum of agents indexed by i on $[0, 1]$. In our interpretation, each agent will actually form a dynasty, with i at date t indexing a member of generation t belonging to dynasty i. Dynasties are linked by fully altruistic preferences as in Barro [1974], so we might equivalently think of i as an infinitely lived individual.

There are a finite set of goods. The consumption space of goods is just the corresponding nonegative Euclidean space, which we denote by $\mathcal{G}$. Each agent has identical (within-period) preferences on consumption bundles $g \in \mathcal{G}$, which we denote by the (strictly increasing and continuous) utility function v. Normalize payoffs such that the lower bound of v is 0.

If $\{g_s\}$ is an infinite sequence of consumptions, then generation t's payoff is given by the "tail sum"

$$\sum_{s=t}^{\infty} \delta^{s-t} v(g_s). \tag{1}$$

where $\delta \in (0, 1)$ is the discount factor, assumed common to all agents.

There is some set $\mathcal{H}$ of *professions* or possible levels of human capital, which individuals in each generation take up. We assume that the set $\mathcal{H}$ is a compact subset of a Euclidean space, and so do not impose any restriction on the divisibility of human capital.

Entry into a profession may or may not entail a setup cost, which is borne by the parents of the individuals concerned (more detail will be forthcoming below). Thus (the cost of aquiring) a profession is bequeathed by parents to children (as in Loury [1981]). In a later section we augment the model by allowing physical capital to be bequeathed as well.

A typical profession will be denoted by h. The notation $\lambda_t = \{\lambda_t(h)\}_{h \in \mathcal{H}}$ will denote a population distribution over professions at date t (specifically, a measure over the compact set $\mathcal{H}$).

Notice that two sorts of commodities are produced in this model: consumption goods and professions. The inputs are just people in various professions. While this formulation allows for all sorts of intermediate inputs that are used up in production, notice that it does not allow for durable physical capital, which we introduce in Section 5.

The technology is represented by means of a closed convex cone $\mathcal{T}$, which contains various combinations of the form:

$$(\mu, g, \sigma),$$

where the first entry μ represents the input vector (simply a population distribution over professions), the second entry describes the output of material goods, and the last entry σ denotes the supply of trained professionals (in each profession), given μ and g. To clarify, the technology combines a CRS production sector (which produces consumption goods from other goods and the services of professionals from different occupations), with an educational sector (which trains 'young' members of generation $t+1$ to acquire the skills of different occupations, using material goods and 'teachers' from different occupations from generation t), also characterized by constant returns to scale. Both production and educational sectors will be governed by profit maximization.

2.2 Prices

Three sets of prices are relevant at each date. First, there are prices for final goods, which we denote by the vector p. Next, there are the returns to professions, which we denote by the function $w(h)$, or more compactly by w. Finally, there are the training costs of *acquiring* professional skills for different occupations, denoted by the function $x(h)$ (more compactly by x). Denote by q the combined vector of prices (p, w, x).

Given prices q_t at any date, the economy generates (input) demands for professions (μ_t) , as well as supplies of final goods at date t (g_t) and new professional capacities (σ_{t+1}) for the *next* generation at period $t+1$. Profit maximization in the production and educational sector implies that $(\mu_t, g_t, \sigma_{t+1})$ must solve

$$\max p_t g + x_t.\sigma - w_t.\mu \tag{2}$$

subject to $(\mu, g, \sigma) \in \mathcal{T}$. Clearly in equilibrium, the prices of consumer goods will equal their respective unit costs of production, while tuition charges will equal unit training costs.

Now turn to household responses. Given some intertemporal sequence $\mathbf{q} \equiv \{q_s\}$, a generation t household i with "going" profession $h(i)$ will choose a sequence $\{h_s, g_s\}_{s \geq t}$ to solve

$$\max \sum_{s=t}^{\infty} \delta^s v(g_s) \tag{3}$$

subject to the constraints

$$h_t = h(i) \tag{4}$$

and

$$w_s(h_s) = p_s g_s + x_s(h_{s+1}) \tag{5}$$

for all $s \geq t$. Because preferences are perfectly altruistic, there is no conflict across generations, so we may as well restrict ourselves to the choices made by generation 0, with initial "endowment" of professions given by $\{h_0(i)\}_{i \in [0,1]}$, or equivalently, by the population distribution λ_0 on $\mathcal{H}$. Denote by $\{g_t(i), h_t(i)\}$ the consumption and professional choices made at every date, assuming these exist.[8]

We normalize so that goods prices lie on the unit simplex S at every date. Notice that for any such vector $p \in S$ (at some date) and any budget y allocated to overall consumption (at that date), a particular implication of the individual's overall optimization problem is that current consumption g will be chosen to

$$\max v(g) \tag{6}$$

---

[8] We omit the technical discussion of existence of these best choices (and of equilibrium) in this paper.

subject to

$$pg \leq y. \tag{7}$$

To be sure, because u is increasing, (7) will hold with equality for any price vector. Denote by $u(y,p)$ the maximum value thus obtained. Obviously, u is strictly increasing in y for given p. We assume further that

[U] For each $p \in S$, $u(y,p)$ is strictly concave in y.

This assumption captures the agent's desire to smooth "aggregate consumption" over time (at constant prices).

Observe that the optimization problem (3) formulated for an individual (or dynasty) incorporates the simplest description of a missing market for the accumulation of human capital. Generation $t+1$'s professional standing must be paid for by generation t; no loans are possible. If preferences are strictly convex, this means that self-finance has different implications for people depending on their current economic status. Specifically, the poor have a higher marginal cost of finance. This is the immediate and most direct implication of a missing capital market, and is captured in our formulation.[9]

2.3 Equilibrium

For given initial distribution λ, an *equilibrium* is a collection $\{q_t, \lambda_t, g_t\}$ (with $\lambda_0 = \lambda$) such that the following conditions are satisfied:

[1] At each date t, $(\lambda_t, g_t, \lambda_{t+1})$ is a solution (μ, g, σ) to the problem (2), under the price vector q_t.

[2] There exists $\{h_t(i), g_t(i)\}$ (for $i \in [0,1]$ and $t = 0, 1, 2, \ldots$) such that for all individuals i, $\{h_t(i), g_t(i)\}_{t=0}^{\infty}$ solves (3) starting from $h_0(i)$, and such that

$$g_t = \int_{[0,1]} g_t(i)di \tag{8}$$

and

$$\lambda_t(B) = \text{Measure}\{i : h_t(i) \in B\} \tag{9}$$

for every Borel subset of H.

A particular type of equilibrium is a *steady state*, one in which all prices and aggregate quantities remain the same over time. Formally, a triple (q, λ, g) is a steady state if there exists an equilibrium $\{q_t, \lambda_t, g_t\}$ with $(q_t, \lambda_t, g_t) = (q, \lambda, g)$ for all t.

---

[9] We eschew more sophisticated variants for simplicity, for they lead to the same basic form. For instance, one might proceed as in Galor and Zeira [1993] by assuming that there is a borrowing rate of interest which is higher than the lending rate. Or one might actually write down a simple model of credit rationing, as Banerjee and Newman [1993] do. Nothing of substance will be altered thereby.

2.4 Examples

The general description of the previous section houses several special cases of independent interest.

[1] Models of Skill Acquisition without Interaction. In the simplest models that display a link between inequality and other features of economic development (e.g., the introductory model in Galor and Zeira [1993]), some exogenous setup cost has to be paid to acquire a skilled profession. It is easy to mimic this setup by assuming a single final good, two professions: 1 (unskilled) and 2 (skilled), so that $\mathcal{H} = \{1, 2\}$, some constant cost x of acquiring the skill, and constant wage rates $w(1)$ and $w(2)$ to unskilled and skilled labor respectively. For a formal embedding that we do not repeat in all the examples, see this footnote.[10]

The formulation of Loury [1981] for the case of perfect certainty corresponds to the case where the set of professions $\mathcal{H}$ is an interval of the real line, the wage function $w(h)$ is an exogenously given concave increasing function, and required investment $x(h)$ by parents to ensure their children acquire skills of type h are also specified by an exogenously given convex increasing function. Indeed, this is equivalent to the one good Ramsey model of optimal growth and a convex technology.

[2] Models of Skill Acquisition with Interaction. More sophisticated models (such as the extended version in Galor and Zeira [1993] as well as Banerjee and Newman [1993], Ray and Streufert [1993], Bandyopadhyay [1997], Maoz and Moav [1999] and others) also display interaction across agents. One way of doing this is to suppose that the returns to skilled and unskilled labor in the previous example depend on the aggregate supplies of these two forms of labor.[11]

[3] Entrepreneurship. To be sure, we need not stick to the interpretation that professions represent different grades of *labor*. As in Banerjee and Newman [1993] or Freeman [1996], we might conceive of one profession as standing for "worker", the other for "entrepreneur". Postulate some fixed investment I that must be made to set up a business. [Notice that our technology set which describes the "supply" of professions has a novel

---

[10] Formally, derive the technology set from the following specifications. First, gross output $g = w(1)\mu(1) + w(2)\mu(2)$, where $\mu(h)$ is the input of labor of skill h. Next, a potential supply of skilled labor is created using g' units of the final good as input: $\sigma(2) = (1/x)g'$, while the potential supply of unskilled labor $\sigma(1)$ can be set to any nonnegative value (compactify this by setting some irrelevant upper bound $m > 1$). This creates the technology set $\mathcal{T} = \{(\mu, g, \sigma) \geq 0 | \; g = w(1)\mu(1) + w(2)\mu(2) - x\sigma(2)$ and $0 \leq \sigma(1) \leq m\}$.

[11] Formally, we can modify the final output production function from a linear specification to any constant returns concave specification: $g = f(\mu(1), \mu(2))$. Then the returns $w(h)$ are obtained as the value of the partial derivatives of this function.

interpretation in this case: I units of output must be sacrificed in order to "produce" a single unit of the "entrepreneur profession".]

As in example 2, this model creates a natural form of interaction. To see this, assume that (conditional on making the fixed investment I), an entrepreneur produces final output g through the use of a production function F that depends only on labor (L). Then each entrepreneur chooses L to

$$\max F(L) - w(1)L,$$

where $w(1)$ is the wage rate for labor (or the return to the first profession). In equilibrium, L is just the employment per entrepreneur, which is $\lambda(2)/\lambda(1)$. So $w(1)$ is given by

$$F'\left(\frac{\lambda(2)}{\lambda(1)}\right) = w(1),$$

while $w(2)$ — the return to entrepreneurship — is just profit:

$$w(2) = F\left(\frac{\lambda(2)}{\lambda(1)}\right) - F'\left(\frac{\lambda(2)}{\lambda(1)}\right)\frac{\lambda(2)}{\lambda(1)}.$$

[4] DEMAND. Now we consider a variant in which the multiplicity of goods matter, and, in particular, the *composition* of demand for those goods. To this end, suppose that each profession is a good-specific labor type. Then the composition of demand for goods translates directly into the composition of demand for professions. In general, inequality will affect the composition of demand (as in Baland and Ray [1991] and Matsuyama [1999]), and in particular, as Mani [1997] has observed, it may induce endogenous barriers to the acquisition of skills.

To embed the demand model, let $\mathcal{H}$ be identified with $\mathcal{G}$. In particular, assume that a person of profession h can produce one unit of the specialized commodity h. Notice that this implies $p_t(h) = w_t(h)$ in any equilibrium, and at any date. It is in this sense that the return to a profession depends intimately on the composition of demand, and therefore (if preferences are nonhomothetic) on inequality.

To complete the embedding, we need to specify how setup costs are determined. This will depend on the exact form of the technology. An important instance of this is presented in the next example.

[5] TEACHING. It is useful to recognize that, in general, setup costs should not be regarded as exogenous, even in a one-good model. For instance, if startup costs for a business are partly denominated in terms of human capital, these will move systematically with the skilled wage rate. Likewise, if skilled people are needed to create other skilled people, the cost of education will depend on the skilled wage rate (see, e.g., Ljungqvist

[1993]). As an example of the endogeneity of setup costs (which this general framework can easily capture), return to the skill acquisition model of Examples 1 and 2.

Again, $H = \{1, 2\}$, where 1 stands for unskilled worker and 2 stands for skilled worker. As in Examples 1 and 2, there is a production function F defined on unskilled and skilled labor. $(a(1), a(2))$. But $a(h)$ is now the number of *production workers* of type h (and is to be distinguished from $\lambda(h)$).

$x(h)$, the cost function, will now depend on the wage function: it is 0 if $h = 1$, and is $\alpha w(2)$ if $h = 2$. The idea is that to acquire a single unit of type 2 skills, you must be trained by α units of type 2 people (presumably $\alpha < 1$, otherwise this would not be sustainable).

Finally, we determine the wage function in equilibrium. To this end, let λ_t and λ_{t+1} be two "adjacent" population distributions over professions. Then notice that

$$a_t(1) = \lambda_t(1),$$

while

$$a_t(2) = \lambda_t(2) - \alpha\lambda_{t+1}(2),$$

so that the wage function at date t is given by

$$w_t(h) = F_h\left(\lambda_t(1), \lambda_t(2) - \alpha\lambda_{t+1}(2)\right)$$

for $h = 1, 2$.

2.5 Utility Inequality

The purpose of this section is to briefly show just how general are the ideas originally put forward by Ljungqvist [1993] and Freeman [1996] regarding the emergence and persistence of inequality (of utilities, not just incomes). In the discussion following the results below we relate our observations to the work of these two authors.

2.5.1 Inequality at Steady States

Let (q, λ, g) be a steady state. Say that two professions h and h' are *indistinguishable* (relative to this steady state) if $w(h) = w(h')$ and $x(h) = x(h')$. Otherwise, they are *distinct*.

Note a simple sufficient condition for two professions to be distinct in any equilibrium: if training someone for occupation h requires more of every material good and every kind of teacher than training someone for occupation h', then irrespective of the precise set of prices, occupation h will involve a higher training cost than h'. More generally, with distinct training technologies for two professions, they will turn out to be distinct *generically*, though we do not pursue the exact conditions required to make this claim precise.

Our first result states that even though a steady state is defined in terms of the stationarity of aggregates (such as the population distribution over professions, or the total production of final goods), it also involves stationarity at the individual level. Notice that this result does not automatically follow from the definition of a steady state. There is no reason why a steady state cannot involve a constant fraction of the population in each profession, while at the same time there are individuals constantly moving from one profession to another (think, for instance, of the ergodic distribution of a Markov chain).

PROPOSITION **1** *Let (q, λ, g) be a steady state. Then no positive measure of individuals will switch across distinct professions.*

Proof. Say that an occupation h is *dominated* if there is a distinct occupation p such that $x(p) \leq x(h)$ and $w(p) \geq w(h)$. It should be obvious that there is no set of dominated occupations which enjoys positive measure under λ.

Now suppose that the proposition is false. Then there is a set of individuals of positive measure such that for each individual in this set, a professional switch (to a distinct profession) takes place at some date. Then — because there are only a countable infinity of dates — there is some *common* date at which a professional switch takes place for a positive measure of individuals.

CLAIM. There exist undominated professions h, h', p and p' such that a person with occupation h moves to p, one with h' moves to p' and the following property is satisfied: $x(h) < x(h')$ and $x(p) > x(p')$.

To prove this claim, note that if a positive measure of people switch professions (say "up" from h to p or "down" from h' to p'), then to maintain the steady state distribution there must be flows in the opposite direction. Moreover, all these professions must be undominated, because no set of dominated professions exhibits postive measure under λ.

The Claim implies that there exist initial professions h and h' such that $w(h) < w(h')$, but with the property that the optimal choice of professions (p and p' respectively) satisfies $x(p) > x(p')$. Define $u^*(c) \equiv u(c, p)$, and let $V(h)$ denote the value function of starting at h under the going steady state. Then, because p' is feasible for h (after all, $x(p') < x(p)$),

$$u^*\left(w(h) - x(p)\right) + \delta V(p) \geq u^*\left(w(h) - x(p')\right) + \delta V(p'),$$

while because p is feasible under $w(h')$ (because p is feasible under $w(h)$ and $w(h) < w(h')$),

$$u^*\left(w(h') - x(p')\right) + \delta V(p') \geq u^*\left(w(h') - x(p)\right) + \delta V(p).$$

Combining these two inequalities and cancelling common terms, we see that

$$u^*\left(w(h') - x(p')\right) - u^*\left(w(h) - x(p')\right) \geq u^*\left(w(h') - x(p)\right) - u^*\left(w(h) - x(p)\right). \qquad (10)$$

However, given that $w(h) < w(h')$ and $x(p') < x(p)$, (10) contradicts the strict concavity of u^* (see [U]). ■

This "zero-mobility" result is based on a single-crossing property that stems from the convexity of preferences and the absence of credit markets (i.e., the fact that parents must pay for their children's education). In steady state, the present value utility of a generation currently occupying occupying occupation h and contemplating a permanent deviation to occupation p is given by

$$u^*\left(w(h) - x(p)\right) + \delta V(p)$$

and the strict concavity of u^* implies that richer families must endure a smaller utility sacrifice in educating their children, hence must be willing to invest more in education. Accordingly the children of families occupying the richest occupation (which must also entail the highest training costs) must be trained for the same occupation — otherwise this occupation would not be filled at subsequent dates, contradicting the steady state assumption. The same argument applies then to the next richest occupation, and so on to all occupations. The zero-mobility result is in deliberate contrast to theories that generate (or maintain) heterogeneity via the use of random shocks. In such models (e.g., Loury [1981]), the empirical distribution of economic characteristics at any date (in a steady state) will be experienced at different dates by a single dynasty.

This statement requires, of course, that in our model there are significant economic differences across individuals in a steady state, to begin with. That this is generally true is confirmed in

Proposition **2** [Inequality of Steady States.] *Let (q, λ, g) be a steady state. Consider any pair of distinct undominated professions h and h' with $x(h) > x(h')$. Then a dynasty with starting profession h is strictly better off at every date (in the sense of higher per period income, net of educational investment expenditure for children) than a dynasty with starting profession h'.*

Proof. Write $u^*(c) \equiv u(c, p)$. Let h and h' be distinct undominated professions with $x(h) > x(h')$. Then $w(h) > w(h')$. Now we know by Proposition 1 that for a person at h, choosing h represents the best continuation. It follows that

$$\begin{aligned} \frac{u^*\left(w(h) - x(h)\right)}{1-\delta} &\geq u^*\left(w(h) - x(h')\right) + \frac{\delta u^*\left(w(h') - x(h')\right)}{1-\delta} \\ &> u^*\left(w(h') - x(h')\right) + \frac{\delta u^*\left(w(h') - x(h')\right)}{1-\delta} \\ &= \frac{u^*\left(w(h') - x(h')\right)}{1-\delta}, \end{aligned}$$

which shows that a person at h has higher lifetime utility than a person at h'. Because no person switches professions at a steady state (Proposition 1), the person at h must have a higher utility at every date compared to the person at h'. ■

Say that a good is *necessary* if without it, individual utility must rest at its lower bound (zero, by our normalization). Now say that a profession is *necessary* if without it some necessary good cannot be produced.[12] Next, say that two professions are *ordered* if one requires greater input of every kind of labor and material goods in training. For instance, if two professions differ in the number of years of schooling required, then they are ordered: being trained for one profession requires a person to first acquire the training for the other profession, and some supplementary training thereafter. Clearly, these occupations will be distinct in every steady state. Finally, define a steady state to be nontrivial if there is a positive measure of individuals with strictly positive utility. Now Proposition 2 can be invoked to immediately yield the following

COROLLARY TO PROPOSITION 2. *Suppose that there are at least two ordered professions that are necessary. Then every nontrivial steady state must involve inequality, in the sense that there are at least two groups of individuals who enjoy different per-period utility.*

2.5.2 Discussion

Proposition 2 (and its corollary) states that inequality is an endemic feature of every steady state satisfying a minimar requirement of specialization in different professions. The latter condition requires two or more professions ordered in terms of their training requirements to be inhabited in any steady state. This is closely connected, in turn, to the endogeneity of market prices: if several professions are needed for economic activity, the behavior of prices must guarantee that each of those professions are actually chosen. While this, in itself, is not sufficient to create (utility) inequality, the subsequent bequests that arise from these choices will do so. The subsequent evolution of inequality is inevitable, even if there is perfect equality to start with.

The examples of Ljunqgvist [1993] and Freeman [1996] drive this point home. In both cases, there are two professions (skilled and unskilled labor in Ljungqvist, and managers and workers in Freeman). Consider the Ljungqvist scenario in which there are two skills, and both types of labor enter as inputs in a concave production function satisfying Inada conditions. Now suppose all individuals in a particular generation have equal wealth. Is it possible for all of them to make the same *choices*? The answer is no. If all of them choose to leave their descendants unskilled, then the return to skilled labor will become

---

[12] In the case where the set of professions forms a continuum, say that a profession h is necessary if for every $\epsilon > 0$, an input vector which is identically zero in the ϵ-ball around h must fail to produce *some* necessary good.

enormously high, encouraging some fraction of the population to educate their children. Similarly, it is not possible for all parents to educate their children, if unskilled labor is also necessary in production. Thus identical agents are forced to take nonidentical actions, precisely because of the interdependence of decisions made by different families.

To be sure, at this stage there are no implications for inequality. There is inequality of (earned) *incomes*, but no utility differences as far as the original generation is concerned. But utility differences do arise from the descendants onward. Suppose the economy converges to a steady state (which is verified in Section 4 below) in which both occupations are occupied. By Proposition 2, such a steady state must display (utility and consumption) inequality. This inequality is fundamental to the market mechanism and does not rely on stochastic shocks.

We conclude this section by sketching a model with 'warm glow' bequests which will exhibit similar results. In the standard formulation of this model (e.g., Galor and Zeira [1993] or Maoz and Moav [1999]), each individual lives for two periods denoted 1 (when young and acquiring skills) and 2 (when working), and has a lifetime utility function of the form $u_1(c_1) + u_2(c_2, b)$, where $c_i, i = 1, 2$ denotes consumption in period i, and b is the bequest left to his children. Given an inheritance of E, the individual selects an occupation h while young, and incurs training cost $x(h)$, so $c_1 = E - x(h)$. And $c_2 = w(h) - b$. Clearly the optimal bequest will be a function of income w: let this be denoted by the function $b(w)$, and assume that this is strictly increasing. Also let $v(w) = u_2(w - b(w), b(w))$ denote the date 2 indirect utility function. Then an individual with parent in occupation h will have a lifetime utility of $u_1(b(w(h)) - x(p)) + v(w(p))$ upon selecting occupation p. If u_1 is strictly concave, occupational preferences again satisfy a single crossing property: children of wealthier parents are more willing to invest in training. Hence in a steady state there can be no occupational mobility, parallel to Proposition 1. And Proposition 2 extends too, since lifetime utility must be strictly increasing in inheritance.

What is fundamental, then, is the single-crossing-property of occupational preferences, which ensure that family wealth, investment and utility are positively correlated. It is evident from this observation that the results will also extend to the case where credit markets are present but in an imperfect form, e.g., with borrowing rates of interest that exceed the lending rate of interest (as in Galor and Zeira [1993]) and possibly rise with loan size. The assumption that credit markets are entirely missing is therefore inessential, and justified as a convenient simplication.

3 Multiple Steady States and Individual Mobility

3.1 Multiplicity and Policy

Typically, there may be several steady-state configurations for the economy. This is apparent in several papers on inequality, such as Banerjee and Newman [1993], Lundqvist [1993], Galor and Zeira [1993], Ray and Streufert [1993], Mani [1997], Piketty [1997]. In this literature, the same economic parameters are generally consistent with numerous steady state outcomes, which display varying degrees of inequality, output, unemployment, productive efficiency, and so on. Historical inequality can cause convergence to steady states with lower per capita income, and hence can be viewed as a 'cause' of underdevelopment. Indeed, the multiplicity of long run outcomes may simply reflect the possible multiplicity of initial conditions; given initial conditions there may be no multiplicity of resulting long run outcomes.

Indeed, this forms the scope for redistributive shocks or policies to generate persistent effects. By changing initial conditions, the policy intervention may change the *particular* steady state that forms the attractor for the process and thereby generate permanent effects; there is no need to change the steady states themselves. This view is to be contrasted with a more classical notion in which convergence to some unique limit obtains. In that case, only policies that are permenently in place can have persistent effects: in the absence of the policy reversion to the unique limit distribution would occur. Thus an exploration of multiplicity is important, in the sense that it tells us what sort of shocks to policy interventions are likely to have lasting impact.

It will be useful to distinguish between two notions of multiplicity. *Individual or micro-* multiplicity refers to the case where initial endowments or perturbations *at the level of a household* significantly shapes the long run outcome of that household. Then there is *societal or macro-* multiplicity, in which initial conditions significantly affect the final destiny (in terms of one or more macro-indicators) of the economy as a whole. The references cited at the start of this section contain numerous instances of macro-multiplicity. These may or may not coexist with micro-multiplicity. For instance, the Galor and Zeira framework is an example of both. In contrast, in the Piketty model, there is macro-multiplicity, but given a particular societal steady state, individual behavior does *not* depend on initial conditions: there is no micro-multiplicity. Finally, theories such as those in Aghion and Bolton [1997] and Loury [1981] are examples of situations in which there is no multiplicity of either kind: there is a single ergodic distribution, *and* the members of each dynasty experience (over time) all the outcomes in the support of that distribution.

We show below that our model generates an entirely novel possibility, where there will be no macro-multiplicity at all, while there is micro-multiplicity! Moreover (as we have already explained) the unique societal steady state displays inequality. Perhaps this is the starkest display of inequality: at the individual level, economic destinies appear as whimsical outcomes which can be changed through a re-specification of initial conditions. Yet *long run macroeconomic outcomes cannot be simultaneously affected by any redistributive policy.* Societal equilibrium necessitates that there be occupants of various professional slots — even winners and losers — in certain proportions, and these proportions are invariant in the aggregate.

3.2 Exploring Multiplicity

3.2.1 Ingredients

Consider a situation in which there is a single final output, the price of which is normalized to unity. Suppose that training costs for different professions are exogenously given. Then professions can be indexed in order of their training costs. So let the universe of all *potential* professions be represented by an interval $[0,1]$, and that there is some exogenous increasing smooth function $x(h)$ describing the cost of acquiring each profession h (this is denominated in units of the final good). Suppose $x(0)=0$ and $x(1)$ is finite. We assume that the set of *available* professions — the set that we've so far labeled $\mathcal{H}$ — is given by some finite subset of $[0,1]$. For each n (which denotes the number of available professions), denote this set of professions by $\mathcal{H}(n)$.

The reason we distinguish between potential and available professions is that we are interested in parametrically "thickening" the space of professions by changing n. Each value n corresponds to a different specification of the model (the remaining potential profesions do not exist for that specification). It is useful (but not necessary) to nest the profession sets so that $\mathcal{H}(n) \subset \mathcal{H}(n+1)$ for all n.[13] In particular, the maximum distance between any adjacent pair of professions converges to zero as n tends to infinity, representing asymptotically vanishing indivisibility of human capital.

We now describe how the production function for final output changes from specification to specification. It turns out that a precise formula is not necessary.[14] However, as n expands we do not want the technology to change in some arbitrary way. To capture this, imagine the continuum limit with all inputs available and assume that there is some technology set $\mathcal{T}^*$ (a convex cone) defined for this limit to which the finite-dimensional technology sets "converge".[15]

---

[13] For instance, one could keep taking midpoints of intervals to create a nested grid: thus $\mathcal{H}(2)=\{0,1\}$, $\mathcal{H}(3)=\{0,1/2,1\}$, $\mathcal{H}(4)=\{0,1/4,1/2,3/4,1\}$, and so on. The exact description will turn out to be unimportant, assuming that u is smooth.

[14] For a concrete example the reader may want to keep in mind the CES form $F^n\left(\{\mu(h)\}_{h\in\mathcal{H}(n)}\right) = \left[\sum_{h\in\mathcal{H}(n)}\{e(h)\mu(h)\}^{\alpha}\right]^{\frac{1}{\alpha}}$, for each n and some $\alpha \in (0,1)$, where $e(h)$ represents the 'efficiency' units of a unit of labor of type h.

[15] Because all technology sets are cones, consider the restriction to the unit simplex of inputs, on which

3.2.2 Characterizing Steady States

Let us look at the collection of steady states for which *every* available profession is occupied. This is the relevant set to study, provided different professions supply distinct inputs, and the marginal product of any given input becomes very high at zero supply.[16] For then an unoccupied profession would have an infinitely high wage associated with it. If training into a profession requires teachers from the same profession, training costs could be prohibitely expensive, so it is logically possible for certain professions to be unoccupied. However, such steady states would rely on the assumption of totally missing capital markets and a closed economy. With a slight perturbation of these assumptions — allowing teachers to be imported and borrowing at a higher rate than the lending rate — such steady states would no longer survive. We therefore ignore such steady states. In what follows, we refer to the collection of steady states where every occupation is occupied, as *the* set of steady states (for each n).

Let w be the wage function across professions in some steady state. By Proposition 1, there is no individual mobility along any steady state. Therefore the only deviations that we need to consider — starting from any profession — is a single-generation move to another occupation which is occupied indefinitely thereafter. In other words, we are at a steady state if and only if for every pair of professions h and h' in $\mathcal{H}(n)$,

$$u\left(w(h)-x(h)\right) \geq (1-\delta)u\left(w(h)-x(h')\right)+\delta u\left(w(h')-x(h')\right), \tag{11}$$

where it is understood that the comparison need be made only over those h and h' for which $w(h)-x(h')$ is nonnegative.

Now recall that x is increasing over the set of professions, so that in any steady state, the same must be true of w. It follows from an easy single-crossing argument that all the inequalities in (11) we need to check are those for "neighboring" profession-pairs. To this end, let $h>0$ be some available profession from n onwards. Denote by h_n the adjacent available profession that immediately precedes it: call this the *lower n-adjacent profession.* Then we are at a steady state if and only if for each $h>0$ in $\mathcal{H}(n)$,

$$u\left(w(h_n)-x(h_n)\right)-u\left(w(h_n)-x(h)\right) \geq \frac{\delta\left[u\left(w(h)-x(h)\right)-u\left(w(h_n)-x(h_n)\right)\right]}{1-\delta}, \tag{12}$$

---

the technology sets are compact. Then, for concreteness, we require convergence in the Hausdorff metric on sequences of compact sets.

[16] It may be argued that this is an unnecessarily strong assumption: certain professions may be inessential, either because the input they supply can be supplied by some other profession, or if the input has a small enough marginal product at zero supply. For instance, suppose there are only two inputs, and a large number of possible professions, each of which supplies one of the two inputs. Then any profession which is not (training-)cost-effective in delivering its input relative to some other profession will be unoccupied. In this case there are effectively only two professions — those which deliver the respective inputs cost-effectively — that will be occupied. We are interested instead in the case where there may be a large and 'thick' set of occupied professions, as is commonly observed in reality. So for this purpose the assumption we make seems appropriate.

while

$$u\left(w(h)-x(h_n)\right)-u\left(w(h)-x(h)\right)\leq\frac{\delta\left[u\left(w(h)-x(h)\right)-u\left(w(h_n)-x(h_n)\right)\right]}{1-\delta},\tag{13}$$

where h_n is lower n-adjacent to h.

3.2.3 A Small Number of Professions

First study (12) and (13) for a small number of professions. In fact, assume that there are only two professions; call them "skilled" and "unskilled" (as in the work of Ljungqvist [1993] and others). Let λ denote the fraction of the population at any date that is skilled. If some well-behaved production function f determines the wage to skill categories, the skilled wage at that date will be given by

$$v(\lambda)\equiv f_1(\lambda,1-\lambda),$$

while the unskilled wage will be given by

$$w(\lambda)\equiv f_2(\lambda,1-\lambda).$$

where subscripts denote appropriate partial derivatives.[17] If f satisfies the usual curvature and Inada end-point conditions, we can conclude that $v(\lambda)$ is decreasing and continuous in λ, with $v(\lambda)\to\infty$ as $\lambda\to 0$. Likewise, $w(\lambda)$ is increasing and continuous in λ, with $w(\lambda)\to\infty$ as $\lambda\to 1$.

Assume that the exogenous setup cost is zero for the unskilled profession, and equal to some positive amount x for the skilled profession. Then (12) and (13) specializes to the following simple characterization: a fraction λ of skilled people is compatible with a steady state if and only if

$$u(v)-u(v-x)\leq\frac{\delta}{1-\delta}[u(v-x)-u(w)]\leq u(w)-u(w-x)\tag{14}$$

The left hand side of (14) represents the utility sacrifice (hereafter denoted $c_s(\lambda)$) of a skilled parent in educating its child, while the right hand side is the corresponding sacrifice (denoted $c_u(\lambda)$) for an unskilled parent. The term in the middle is the present

---

[17] This applies only in the unrealistic event that skilled workers cannot perform unskilled tasks. More generally if skilled workers can perform unskilled tasks, then the skilled wage cannot ever fall below the unskilled wage. So when the skill intensity λ is large enough that $f_1 < f_2$, the wages will not be given by f_1 and f_2, but will be equalized (as a result of skilled workers filling unskilled positions whenever the latter pay higher wages). We omit this complication here because a competitive equilibrium with a positive fraction of skilled workers will never give rise to skill intensities that are so large, as they would be incompatible with incentives for parents to educate their children.

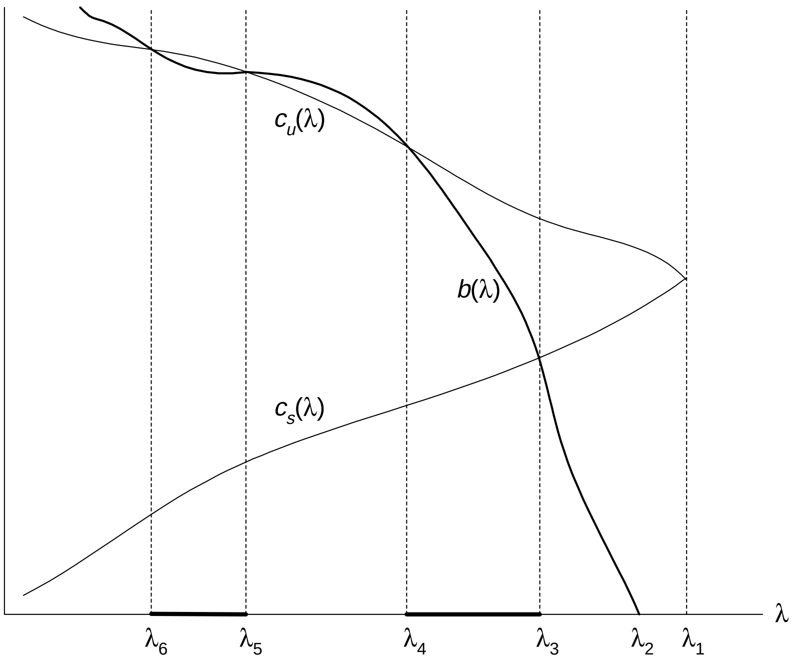

Figure 1: Education Costs and Benefits in Two-Profession Model

value benefit of all successive descendants being skilled rather than unskilled (which we shall denote by $b(\lambda)$). These benefit and sacrifice functions are illustrated in Figure 1.

It is easy to check that all three lines in that figure slope and intersect as depicted. For instance, $\lambda_1 \in (0, 1)$ denotes the skill intensity of the population at which the skill premium just disappears and the wages of the skilled and unskilled are equal. So c_s and c_u intersect there. Likewise, λ_2 is the point at which the wages of the skilled *net* of training equal those of the unskilled. So b drops to zero there. These observations can be used in conjunction with (14) to establish

Proposition **3** *There is a continuum of steady states in the two-profession model, and total output net of training costs unambiguously rises as the skill proportion in steady state rises.*

Proof. By the Inada conditions, there exists λ_3 such that $b(\lambda)$ and $c_s(\lambda)$ are equalized. Notice that λ_3 must be strictly less than λ_2, which in turn is less than λ_1. So, using the strict concavity of the utility function, it must be the case that $c_u(\lambda_3) > c_s(\lambda_3) = b(\lambda_3)$. Thus (14) is satisfied at λ_3 and we have a steady state.

Now use the slopes of these curves to argue that for all $\lambda < \lambda_3$ but sufficiently close to it,

$$c_u(\lambda) \geq b(\lambda_3) \geq c_s(\lambda_3),$$

which establishes that there must be a continuum of steady states.

To see that the steady states are ordered in terms of net output, consider the following maximization problem for net output:

$$\max_{\lambda \geq 0} f(\lambda, 1-\lambda) - x\lambda. \tag{15}$$

This is a strictly concave problem in λ and attains a unique maximum when $f_1 - f_2 = x$. Recalling that $v = f_1$ while $w = f_2$, we conclude that this is the point λ such that $v(\lambda) - x = w(\lambda)$, which is precisely λ_2 in Figure 1. Because every steady state lies to the left of λ_2 and the maximization problem (15) is strictly concave, the result follows. ■

Proposition 3 tells us that multiplicity is endemic for a small number of professions. While stated only for the two-profession case, it is easy enough to extend the argument for any finite number of distinct professions, though — as we shall soon see — the multiplicity may tend to vanish as the set of professions "thickens".

Notice that the *structure* of the set of steady states may be complicated. In particular, the set need not be connected. For instance, in Figure 1, the set of steady states is the union of the two intervals (λ_6, λ_5) and (λ_4, λ_3).

The proposition also states that steady states are ordered not only in terms of skill premium but also per capita income: a steady state with a higher λ and lower skill premium corresponds to higher per capita income net of training costs. At the same time, we must be careful not to confuse this finding with the Pareto-efficiency of a given steady state. It is true that there may be steady states "above" it that yield higher per-capita net output (and therefore higher per-capita utility) at every date. But this does not imply that the first steady state is Pareto-dominated *starting from the same initial conditions.* The question of Pareto-efficiency is subtle and we postpone it to a separate paper.

The societal multiplicity described in Proposition 3 is very much in line with the existing literature. We now turn to the question of how this multiplicity is modified as we enrich the space of professions.

3.2.4 Many Professions

Return to the more general formulation in (12) and (13). The argument can be illustrated most simply in the limiting case where there is a continuum of professions $h \in [0, 1]$, and where we make the provisional assumption that the wage and training functions are differentiable. Then a steady state with zero mobility requires a family in occupation h at any date to select the same occupation for its children, i.e., $p = h$ must maximize

$V(p|h) \equiv u(w(h) - x(p)) + \frac{\delta}{1-\delta}u(w(p) - x(p))$. The first-order condition for this requires for every h:

$$u'(w(h) - x(h))x'(h) = \frac{\delta}{1-\delta}u'(w(h) - x(h))\{w'(h) - x'(h)\} \tag{16}$$

generating the condition that

$$w'(h) = \frac{1}{\delta}x'(h) \tag{17}$$

so the wage function must take the form

$$w(h) = \frac{1}{\delta}x(h) + w(0) \tag{18}$$

i.e., is uniquely determined upto a constant intercept, the wage of the profession $h = 0$. This pins down the wage distribution (upto a constant of integration), and also (given the production function) the distribution of the population across different professions.

This uniqueness property can be interpreted as arising from the requirement that the holder of every profession now be indifferent between selecting the same profession for their children, relative to neighboring professions located arbitrarily close to it. In the language of incentive theory, the local incentive constraints must bind, whereas they need not bind with a discrete set of professions. For instance, in the case of two professions, there was room for local variations in the steady state distribution that continued to preserve the incentive constraints. Such room is no longer available with a large number of neighboring professions.

The preceding argument is hardly rigorous, because it assumed that the wage function was differentiable, rather than deriving that as a conclusion. It used only the first-order condition for optimal occupational choice, leaving open the question of global incentive compatibility. Moreover, the wage function was shown to be uniquely determined only up to a constant of integration. There is also the question whether equilibria in the continuum case are interpretable as the limit of equilibria in a sequence of progressively finer finite classification of professions.

So we turn to the discrete economy directly, and take a sequence of such economies that "converge" to some well-defined continuum limit. We construct a nested sequence of feasible professions$\mathcal{H}(n)$ as described above in Section 3.2.1 and solve for the set of steady states for each value of n.

Pick any profession h that lies in $\mathcal{H}(n)$ (from some n onwards); the inequality in (12) and the concavity of u tell us that

$$u'\left(w(h_n) - x(h)\right)[x(h) - x(h_n)] \geq \frac{\delta}{1-\delta}u'\left(w(h) - x(h)\right)[\{w(h) - w(h_n)\} - \{x(h) - x(h_n)\}$$

while the inequality in (13) and the concavity of u tell us that

$$u'\left(w(h)-x(h_n)\right)[x(h)-x(h_n)] \leq \frac{\delta}{1-\delta}u'\left(w(h_n)-x(h_n)\right)[\{w(h)-w(h_n)\}-\{x(h)-x(h_n)\}$$

Let us rearrange these inequalities. Define

$$a_n \equiv \frac{u'\left(w_n(h_n)-x(h)\right)}{u'\left(w_n(h)-x(h)\right)}, \tag{19}$$

and

$$b_n \equiv \frac{u'\left(w_n(h)-x(h_n)\right)}{u'\left(w_n(h_n)-x(h_n)\right)}, \tag{20}$$

where h_n is the lower n-adjacent profession to h. Applying these definitions to the inequalities above, we see that for each n and available profession h,

$$\frac{(1-\delta)b_n(h)+\delta}{\delta}[x(h)-x(h_n)] \leq w_n(h)-w_n(h_n) \leq \frac{(1-\delta)a_n(h)+\delta}{\delta}[x(h)-x(h_n)], \tag{21}$$

and summing these inequalities over all available professions h' in the grid (for n) such that $0 < h' \leq h$,

$$\sum_{h'>0;h'\leq h}\frac{(1-\delta)b_n(h')+\delta}{\delta}[x(h')-x(h'_n)] \leq w_n(h)-w_n(0) \leq \sum_{h'>0;h'\leq h}\frac{(1-\delta)a_n(h')+\delta}{\delta}[x(h')-x(h'_n)]. \tag{22}$$

Thus the first and third terms in (22) "sandwich" $w_n(h)-w_n(0)$. The details in Appendix 1 show that these terms *also* sandwich the expression $(1/\delta)x(h)$.[18] It isn't hard to see why. For large values of n, the terms $a_n(h)$ and $b_n(h)$ are very close to one (the precise statement of this is in Appendix 1). But then both the sums in (22) are arbitrarily close to $(1/\delta)x(h)$ (recalling that $x(0)=0$ by assumption).

We have shown thus far that for each h,

$$w_n(h)-w_n(0) \to (1/\delta)x(h) \tag{23}$$

as $n \to \infty$. Hence equilibrium wage functions must converge to a limit of the form described exactly by condition (18), which is uniquely defined once $w(0)$ is determined.

But there is an additional restriction, which pins down $w(0)$. The solution $w(h)$ must also satisfy profit maximization for the production sector of the economy. But observe that *there is at most one value of the free parameter* $w(0)$ *that can achieve this.* The

---

[18] The arguments in the appendix assume that the sequence $w_n(0)$ is bounded above and below by strictly positive numbers. It is possible to show that under the assumptions we place on the limit technology (see subsequent discussion in the main text), these restrictions on $w_n(0)$ can indeed be derived from profit-maximization. The details are omitted.

reason is that any two distinct solutions to (18) must be monotonically ordered: if w and $\hat{w}$ are two distinct solutions with $w(0) > \hat{w}(0)$, then it must be the case that $w(h) > \hat{w}(h)$ for all h. Two such functions cannot *simultaneously* support profit maximization in a single-output model, when the limit technology is a cone (i.e., displays constant returns to scale) and when the price of the final output is normalized to some given constant.[19]

There is, then, at most *one* wage function of the form (18) that *also* supports profit maximization. To make sure that there is *some* such function, we need an assumption on the limit technology set. We need the following end-point condition: let $\tilde{w}$ be the function in (18) when $w(0) = 0$. Then there is some input-output combination $\{\mu(h), y\} \in \mathcal{T}^*$ such that $y - \tilde{w}\mu > 0$. It can be shown that this guarantees existence (in addition to the uniqueness property); we omit the details. Under this condition, moreover, $w(0) > 0$.

These observations can also be used to pin down further the wage functions along the sequence of actual economies. With the profit-maximization requirement added to this sequence, it must be the case that $w_n(0)$ converges to the unique value of $w(0)$ described in the previous paragraph. We have therefore established that *every* equilibrium wage function for every n must converge to the unique wage function described by (18), together with (limit) profit maximization.

We may summarize this discussion in the form of a proposition.

PROPOSITION **4** *Suppose that professions are indexed on* $[0, 1]$ *in a single-output model, and that some smooth training cost function* x *is exogenously defined on* $[0, 1]$*. Consider a nested sequence of economies, each with a finite number of occupations that converge to the continuum limit as described in Section 3.2.1. Then there is a unique function* $w^*(h)$ *of the form*

$$w^*(h) = \frac{1}{\delta}x(h) + w \tag{24}$$

for some nonnegative constant w*, such that if* w_n *is any sequence of steady state wage functions, then*

$$w_n(h) \to w^*(h) \ asn \to \infty. \tag{25}$$

We conclude, then, that when the set of professions is "thick" (in the sense illustrated by this extended example), the multiplicity of steady state solutions asymptotically vanishes. Yet the steady state that does remain displays persistent inequality. It follows that initial conditions determine individual fates, yet macroeconomic outcomes are fully pinned down in steady state!

---

[19] Suppose that $w > \hat{w}$ in the sense described in the main text. Consider some input-output pair (μ, g) that maximizes profits under w; then, because $\mathcal{T}$ is a cone, $g - w\mu = 0$. It follows that $g - \hat{w}\mu > 0$. But then $\hat{w}$ cannot support profit-maximization at any point of the technology set.

3.2.5 Generalizations

Consider the important features that drive the results above.

First, we used the single-crossing property to argue that a local condition is sufficient for the characterization of steady states (recall the move from (11) to (12)) and (13). This observation is extremely general, and one can derive it directly from a model with a continuum of professions (though technical issues concerning existence must be dealt with). The argument that leads up to (23), then, can replicated under a very general set of conditions.

The second point is that (12) and (13) yield a family of solutions to the equilibrium wage function that are *ordered*, in the sense that if w and $\hat{w}$ are both solutions and $w(h) > \hat{w}(h)$, then the same inequality is true at all *other* professions. We derived this under the crude assumption that x was an exogenously given function. How generalizable is this observation?

The endogeneity of the x function arises from the possibility that it takes professionals to train professionals, so that x depends on w. One elementary formulation is a fixed coefficients "recursive" training technology: workers proceed incrementally over successive training levels, and to increase one's level of training from $h - dh$ to h requires a fixed proportion $\alpha(h) > 0$ of teachers with training level h: this costs $\alpha(h)w(h)$. This corresponds to the cost function

$$x(h) = \int_0^h \alpha(h')w(h')dh' \tag{26}$$

Combining this with (18), this implies that the wage profile in any limit steady state must belong to the family

$$w(h) = w(0)\exp[\int_0^h \frac{\alpha(h')}{\delta}dh'] \tag{27}$$

Smooth steady states are thus pinned down entirely, except for their level, which correspond to the wage $w(0)$ of workers with no training at all. Notice that this family of wage function satisfies the required monotonicity condition.

Or suppose, alternatively, that the training technology is Cobb-Douglas, with level-h training technology described by the function

$$\log s(h) = \int_0^h \alpha(h')\log t(h')dh' \tag{28}$$

where $s(h)$ is the number of type h students turned out by a process that uses $t(h')$ teachers of type $h' \in [0, h]$. Here training an h-type requires teachers of all levels up to level h, but there is scope for substitutability among teachers of different levels. Higher level teachers may be more effective, but also more expensive. Hence cost-effective training requires educational institutions to select an optimal teacher mix of different levels given their wage profile, to minimize the cost of turning out each student. This cost minimization exercise generates the training cost function

$$x(h) = \exp[\int_0^h \alpha(h') \log \frac{w(h')}{\alpha(h')} dh'] \tag{29}$$

implying from (18) that a limit steady state wage profile must satisfy the differential equation

$$w'(h) = \frac{1}{\delta} \alpha(h) \log \frac{w(h)}{\alpha(h)} \exp \left[\int_0^h \alpha(h') \log \frac{w(h')}{\alpha(h')} dh' \right] \tag{30}$$

Once again, it is evident that the family of wage functions (determined up to a constant of integration) satisfies the required monotonicity property.

In both the examples, it appears that what drives this monotonicity is the recursive nature of the technology. Professionals at level h or "below" are needed to generate a professional of level h. Suppose, then, that the unit cost of production of a type-h graduate depends only on the wage profile up to that type. Then, by constant returns to scale, we see that

$$x(h) = \psi \left(h; w(h')_{h' \leq h} \right) \tag{31}$$

for some family of functions $\psi(h, .)$. Combining this with (18), we see that

$$w(h) = \frac{1}{\delta} \psi \left(h; w(h')_{h' \leq h} \right) + [w(0) - \frac{1}{\delta} x(0)] \tag{32}$$

which can be seen to be a single-parameter family indexed by $w(0)$. Notice that for each h, ψ must be increasing in the sense that more expensive teachers cannot lower the cost of production. This observation, along with (32), proves the desired monotonicity property.[20]

The final ingredient in our argument is the observation that if there is a "monotonic" family of wage functions, *at most one* of them can be consistent with profit maximization in the production sector, (provided every type of worker is essential in the production sector). But the proof of this argument rested on the presumption that we have a single aggregate output. With multiple final outputs, it is conceivable that this final step of the argument might not hold. This will require further investigation. In particular, the jury remains open on the multiplicity question when demand-side effects (as in Baland and Ray [1991], Mani [1998] and Matsuyama [1999]) drive the story.

---

[20] We conjecture that this property holds also for a large range of non-recursive training technologies as well. Then training costs for any profession h depend on the entire wage profile, rather than just those below h, and equilibrium wage profiles must solve $w(h) = \frac{1}{\delta}\psi(h; \{w(.)\}) + w(0)$. As long as there is a unique solution to this integral condition for any given $w(0)$, it is clear that the uniqueness and monotonicity properties extend.

3.3 Discussion

Our model illuminates what — to our knowledge — no other theory in the literature discusses. It connects societal multiplicity to the richness of the professional structure, and uses intertemporal utility maximization to do so (witness the restrictions created by (21) and (22)). These features merit some discussion.

Models that exhibit societal multiplicity, such as Galor and Zeira [1993] or Banerjee and Newman [1993], typically exhibit one or more of the following two characteristics: (a) they assume a small number of professions, and (b) a 'warm glow' bequest motive. We argue now that the difference in results relies only on (a): the same uniqueness results extend in the case of a warm glow bequest motive when the occupational structure is rich enough. Consider the version of our model with warm glow bequests introduced in Section 2.5.2. There we showed that the present value utility of a child whose parent is in occupation h and is contemplating a shift to occupation p takes the form

$$V(p|h) \equiv u_1(b(w(h)) - x(p)) + v(w(p)), \tag{33}$$

which satisfies the single-crossing property. And consider also the limiting case of a continuum of professions $[0, 1]$, whence the first-order condition for a smooth steady state is that $p = h$ maximize $V(p|h)$ for every h. This gives rise to the differential equation

$$w'(h) = \frac{u_1'(b(w(h)) - x(h))}{\delta v'(w(h))} x'(h) \tag{34}$$

which again determines the wage function upto the constant $w(0)$. With $w(0)$ determined by the profit maximization condition, we obtain a unique (smooth) steady state with warm-glow bequests. This suggests that parallel results should emerge for economies with a finite but fine enough structure of occupations.

This history independence at the societal level is in surprising contrast to the history-dependence at the individual level, where individual families get locked into occupational choices of their ancestors. It suggests that while one-time policy interventions can permanently affect the fates of individual dynasties, such interventions applied to a positive measure of agents will inevitably be compensated for in the overall workings of the economy, leading to the same state of affairs (modulo some permutation of individual positions).

Note, moreover, that as we move to the continuum, the nonconvexities created at the individual level are entirely endogenous. With a discrete set of occupations, the jump from one profession to another is inherently nonconvex if there is a setup cost. No such presumption exists when there is a continuum of professions starting all the way from "unskilled labor". The return function $w(h)$ and the cost function $x(h)$ are both endogenous, in principle, and so — *a fortiori* — is their curvature. Yet Proposition 2 and its Corollary tell us that — provided more than one profession is necessary, *individual* outcomes are necessarily nonergodic. Let us combine these two observations carefully.

First, look at the steady state (unique with a single output and a fine enough grid) in which people settle at different professional choices, and in which there is utility inequality. The (limit) wage function is given in (18); we reproduce it here:

$$w(h) = (1/\delta)x(h) + w(0).$$

This is effectively a *linear* production function — at the level of an individual — that connects investments (x) to "outputs" w. There is no nonconvexity *at* the steady state, but inequality persists simply because different individuals are at different initial conditions at the steady state. This suggests that the endogenous nonconvexities alluded to in the previous paragraph must arise *out* of steady state.

To make this argument, assume that all dynamic competitive equilibria indeed converge to steady states. [In general, this is an open question which we solve for a special case, below.] Then, in particular, equal initial conditions must lead to final inequalities. Since all agents face exactly the same prices and technologies, this proves that a unique optimal path from the starting point of equality for all *cannot* exist. Because the utility function is strictly concave, this proves that *equilibrium prices must endogenously generate strict nonconvexities in individual feasible sets.*

Several applications may be considered; we mention one that reinterprets these results for a global economy with interaction. Suppose that the individuals of our model are countries, or more precisely the planning agencies of these countries. View 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.

Nevertheless, while individual fates can be altered, the world economy must exhibit a certain compositional balance. There will be high-tech exporters, but not too many of them. And not all developing countries need be primary commodity exporters, but there cannot be too few of them either. It may be hard to talk about economic policies that imitate a Korea or a Hong Kong in the world economy. Sequenced development that maintains global hierarchical compositions may the rule rather than the exception.

4 Dynamics

The discussion so far on the emergence of inequality (as opposed to its persistence) makes an important assumption. It is that starting from any initial configuration, an economy will indeed converge to a steady state. After all, Proposition 2 makes no claims regarding persistent inequality when the economy fails to converge to a steady state.

Moreover, a satisfactory theory of the long-run role of historical inequality should account for the dynamic process by which initial conditions determine longrun outcomes. For instance even if there are many possible steady states, it is conceivable that only a few of them are stable attractors, and others cannot be reached from a nontrivial set of initial conditions. In that case the steady state analysis overstates the multiplicity of long run outcomes. And even if convergence can be established, the precise map between initial conditions and eventual steady state reached, and the transitory process is of interest in its own right (e.g., does inequality tend to increase or decrease over time? how fast is the convergence? what are the transitory and long term effects of one-shot redistributions?)

To our knowledge, there is no general theorem that guarantees convergence in this class of models, and it is our belief that such a theorem would constitute an important step forward.[21] Competitive versions of the turnpike theorem are available (see Bewley [1982], Coles [1985] or Yano [1984]) but do not apply here, as those 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.

The purpose of this section is to report on a special case for which we have been able to establish convergence, and characterize the dynamics completely.[22] While the results reported here are of intrinsic interest, the reader who wishes to skip this section can do so without loss of continuity.

We focus on the two-skill model with exogenous training cost, described in Section 3.2.3 above. It is useful to recall the definition of perfect foresight competitive equilibrium in this context, and to note some of its characteristics. Imagine, therefore, that an infinite sequence of wages is given, one for each category of labor. We describe the maximization problem for each household as follows: there is a sequence of *values* $\{\bar{V}_t, \underline{V}_t\}$ describing the infinite-horizon payoffs to each generation at each date, conditional on starting skilled or unskilled. That is, for each t,

$$\bar{V}_t = \max_{c_t, x_t} [u(c_t) + \delta V_{t+1}(x_t)]$$

subject to the conditions that

$$c_t + x_t = v_t,$$

and

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

---

[21] Even in the simplistic warm-glow formulations of intergenerational behavior, general convergence arguments are hard to come by (see, e.g., Banerjee and Newman [1993] for a discussion).

[22] The results of this section are based on Ray [1990].

Likewise,$\underline{V}_t$ denotes the maximum value of the above problem for a currently unskilled household with income w_t instead of v_t. This describes how education decisions are made, given the sequence of wage rates. For given initial skill distribution $\lambda_0 \in (0, 1)$, a *competitive equilibrium* is therefore a sequence of wages and subsequent skill distributions $\{v_t, w_t, \lambda_t\}_{t=0}^{\infty}$ such that

[i] Given λ_0, the path of subsequent skill distributions $\{\lambda_t\}$ is generated by the maximization problems just described,

[ii] For each t, $v_t = v(\lambda_t)$ and $w_t = w(\lambda_t)$ if $\lambda_t < \lambda_1$, and $v_t = w_t = v(\lambda_1) = w(\lambda_1)$ if $\lambda_t \geq \lambda_1$.

Note that the definition of a competitive equilibrium uses the fact that if the marginal product of 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. In any case this is not very relevant because no competitive equilibrium will ever involve such a large fraction of the population being skilled at any date, i.e., in any equilibrium $0 < \lambda_t < \lambda_1$ for all $t \geq 1$. The proof of this is obvious. 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. On the other hand, λ_t cannot exceed λ_1 for any $t \geq 1$, for in that case skilled and unskilled wages are equalized. No one in the previous generation would have invested in such circumstances. In order to avoid qualifying statements for the initial value of λ, we make the inessential assumption that $\lambda_0 \in (0, \lambda_1)$.

Next observe that in any competitive equilibrium, there cannot be any date at which an unskilled household decides to educate its children, while a skilled household does not. This follows again from the logic underlying Proposition 1, which applies not only to steady states but also to any competitive equilibrium sequence. Since skilled households always earn higher wages, the 'single crossing property' is satisfied: the demand for education is higher among skilled households. To see this explicitly, note that if an unskilled household were to educate its children, then

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

or

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

By strict concavity and the fact that $\lambda_t < \lambda_1$ for all t, we may conclude that

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

But this means that a skilled household has a *strict* incentive to educate its children.

It follows from this observation that in any competitive equilibrium, if the proportion of skilled households increases from one generation t to the next, it must be the case that all skilled households at date t are educating their children, and some unskilled households as well. Moreover, some unskilled households must also be deciding to *not* educate their children — otherwise there would be no unskilled households at date $t+1$, which has already been ruled out. It must then be the case that when some unskilled households switch professions, they must be exactly indifferent between switching and not switching professions. Conversely, if the proportion of skilled households goes down from one generation to the next, it must be the case that some skilled households change professions while being exactly indifferent between switching and not. In particular, it is always the case that at every date, the lifetime utility of the skilled (and unskilled) must be equal to the utility they would have received were their descendants *never* to switch status: at every date t,

$$\bar{V}_t = \sum_{s=t}^{\infty} \delta^{s-t} u(v_s - x) \tag{35}$$

and

$$\underline{V}_t = \sum_{s=t}^{\infty} \delta^{s-t} u(w_s). \tag{36}$$

This implies that for a household that is skilled at date t:

$$\sum_{s=t}^{\infty} \delta^{s-t} u(v_s - x) \geq u(v_t) + \sum_{s=t+1}^{\infty} \delta^{s-t} u(w_s),$$

or equivalently,

$$u(v_t) - u(v_t - x) \leq \sum_{s=t+1}^{\infty} \delta^{s-t} [u(v_s - x) - u(w_s)], \tag{37}$$

with equality holding whenever a switch from "skilled" to "unskilled" occurs at date t. The left hand side of (37) is the current sacrifice $c_s(\lambda_t)$ for a skilled household in educating its children, which does not exceed the discounted present value benefit for the succeeding generation to be skilled (and by the above logic, of the discounted benefit of *all* succeeding generations being skilled rather than unskilled). Likewise, for the currently unskilled, the sacrifice involved in educating their children exceeds or just equals the benefit of all their descendants switching from the unskilled to the skilled profession:

$$u(w_t) - u(w_t - x) \geq \sum_{s=t+1}^{\infty} \delta^{s-t} [u(v_s - x) - u(w_s)], \tag{38}$$

with equality holding whenever a switch from "unskilled" to "skilled" does occur along the equilibrium path.

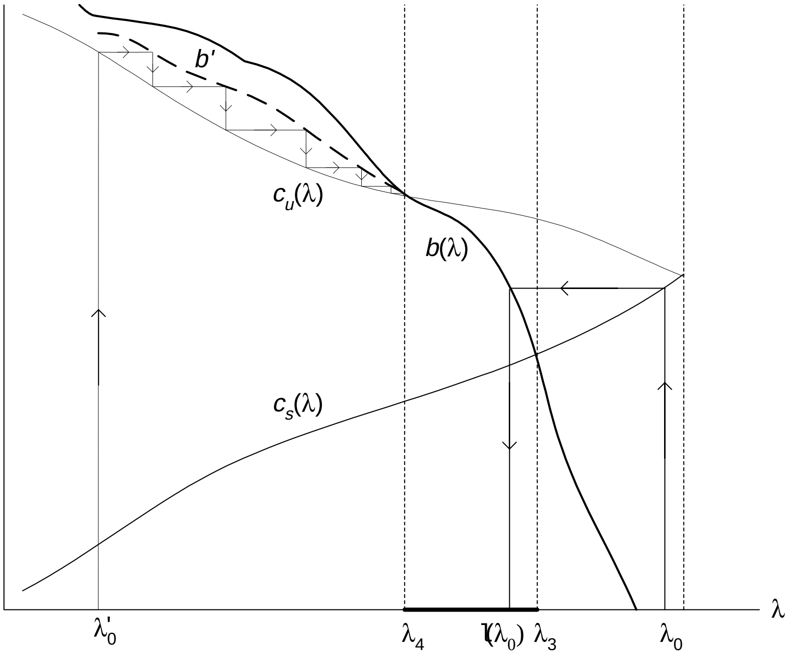

Figure 2: Dynamics in Two-Profession Model

To proceed further, it will be necessary to consider two possible zones in which λ might lie, when λ is *not* a steady state. We divide the non-steady state space into two complementary parts: in the first subset (denoted by A), the steady state condition fails owing to insufficient incentive of skilled households to educate their children. Equivalently, the first inequality in (14) fails. Recall from Section 3.2.3 that λ_3 is the highest steady state value of λ, where skilled households are just indifferent between educating their children and not. Then A is the range of 'high' skill ratios (λ_3, λ_1).

In the second subset (denoted by B), the steady state condition fails because unskilled families strictly prefer not to educate their children. Equivalently, the second inequality in (14) fails. B is the union of all skill ratios lower than λ_3 that do not constitute steady states. In Figure 2 in which the set of steady states is simply the interval (λ_4, λ_3), the set B is the set of skill ratios $(0, \lambda_4)$. In general, it is clear that A and B are disjoint owing to the strict concavity of u. In what follows, we relate the dynamics of λ to membership of the initial skill ratio in one of the sets A and B.

Proposition **5** *If $\lambda_0 \in A$, then there exists a unique competitive equilibrium from λ which goes to the steady state in one period: $\lambda = \lambda_0 > \lambda_1 = \lambda_t$ for all $t \geq 1$.*

If $\lambda_0 \in B$, then there exists a unique competitive equilibrium in which the proportion *of skilled people increases strictly in every period, and converges to some steady state:* $\lambda_t < \lambda_{t+1}$ *for all* $t \geq 0$.

If λ *is a steady state, there is a unique competitive equilibrium from* $\lambda_0 = \lambda$*, given by* $\lambda_t = \lambda$ *for all* t.

Hence from any initial condition, there is a unique competitive equilibrium which converges to a steady state. If the initial skill ratio λ_0 is a steady state, the equilibrium involves that ratio for ever thereafter. If it is a high ratio (in the set A) then the skill ratio falls in just one generation to a steady state, and stays there forever after. This is depicted in Figure 2 for the initial skill ratio depicted λ_0. Since this ratio is very high, the skill premium is too low to motivate educational investments that are consistent with a steady state. Accordingly, at such a date, some skilled households will not educate their children, and every unskilled household will behave likewise. This lowers the skill ratio for the succeeding generation. The eventual steady state $l(\lambda_0)$ is pinned down by the requirement that generation 0 skilled households are just indifferent between educating and not educating their children. Since from generation 1 onwards the economy will be in a steady state, the present value benefit of educating children for generation 0 households is given by the steady state benefit function $b(\lambda)$, which must equal the sacrifice for generation 0 skilled households:

$$c_s(\lambda_0) = b(l(\lambda_0)) \tag{39}$$

which determines the function $l(\lambda_0)$. The skill ratio $l(\lambda_0)$ must be a steady state because it is smaller than λ_0, so the sacrifice $c_s(l(\lambda_0))$ for skilled households must be smaller than the benefit $b(l(\lambda_0))$ — using equation (39), while the sacrifice for unskilled families must be larger than $b(l(\lambda_0))$.

In contrast, if the initial skill ratio is too low to constitute a steady state (i.e., is in the set B), then convergence to the eventual steady state will occur gradually rather than in one step. In Figure 2 this is represented by the initial skill ratio λ'_0. Then the skill ratio will subsequently increase over time, with unskilled households progressively switching to the high skill status. Then the dynamics is determined by the condition that they be indifferent between switching and not at every date. As more and more households become skilled in this fashion, the skill premium declines over time, reducing the benefit from switching. At the same time the cost of switching for unskilled households also falls, as the unskilled wage rises over time. Since the convergence does not occur in one step, the present value benefit of switching is not represented by the steady state benefit function $b(\lambda)$, but by the function depicted by b' in Figure 2 which is lower (reflecting the fact that the benefit of switching is falling over time). The dynamics is then pinned down by the equality of sacrifice for unskilled families and the present value benefit b', as depicted in Figure 2. It is evident from this that the eventual steady state will be the

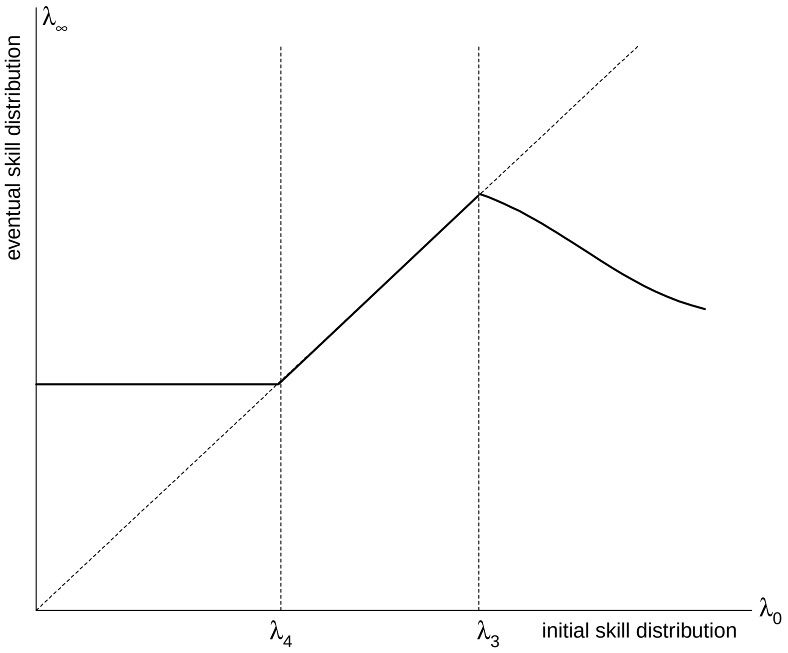

Figure 3: Map From Initial to Long-Run Skill Distribution

smallest steady state skill ratio lying above the initial skill ratio, in the case where the latter is in the low range B. In Figure 2 this is the skill ratio λ_4.

The corresponding map from initial to eventual skill ratios is depicted in Figure 3, and from the initial skill premium (a measure of inequality) to eventual long run per capita income in Figure 4.[23] These help predict the effect of initial conditions of the economy to its long run performance. These maps are nonmonotone, thus showing that orderings of countries by human capital, inequality and per capita income can get reversed over time. For instance countries with initial skill ratio higher than λ_3 will eventually end up with a lower skill ratio than countries starting with the skill ratio λ_3. Correspondingly countries that start with a high degree of equality (a low skill premium) end up more unequal and with a lower per capita income. Such reversals can only occur at the high end of the spectrum of initial skill ratios (i.e., when starting in the set A). When initial skill ratios lie below (either in the steady state set or in B), initial conditions and eventual outcomes are ordered in the same way.

Particularly interesting is the case where the economy starts in B, i.e., with sufficient inequality. Then inequality falls over time, accompanied by a process of progressive

---

[23] These figures correspond to the case depicted in Figure 2, where the set of steady states constitutes a single interval.

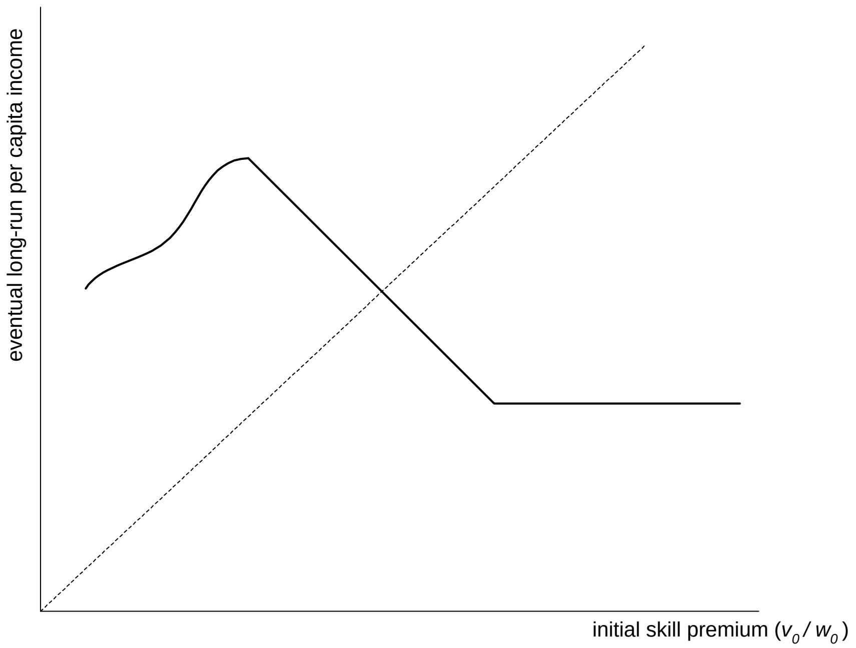

Figure 4: Map from Initial Skill Premium to Long-Run Per-Capita Income

increase in education and skill within the population, which serves to raise per capita income over time. However, the initial conditions do not (locally) affect the long run outcome, which is invariably the nearest steady state (λ_4 in these figures). Hence one shot redistributions in this case only have a transitory impact, that speed up the skill upgrading process. In contrast when the economy starts in a steady state (e.g., in the interior of (λ_4, λ_3)), one shot redistributions do have an immediate and permanent impact. At the other extreme, when the economy starts with a very high skill ratio (in the set A), one shot redistributions have an immediate and permanent effect, but *which perversely causes a move in the opposite direction.*

5 Other Forms of Capital

We now turn to the possibility that individuals may leave financial bequests in addition to — or perhaps in lieu of — educational or professional choices. This significantly qualifies our earlier results concerning the inevitability of persistent inequality. The main reason is that the possession of human capital is not the sole source of wealth. Nonhuman wealth could be distributed independently of the distribution of human capital. Then wealth disparities between unskilled and skilled families would be moderated. It is conceivable that unskilled families are just as well off as skilled families, or maybe even better off, if they are disproportionately favored in the distribution of nonhuman wealth. The distribution of nonhuman sources of wealth then adds one more dimension of multiplicity to income distribution, and indeed can permit the existence of equal steady states.

To confirm this intuition, we augment our model to accommodate forms of nonhuman wealth, such as physical capital. Since bequests could be made in the form of either nonhuman or human capital expenditures, we need to explicitly model the way in which households allocate bequests between these two forms. We continue to work with the case of two professions, and an exogenous training cost x.

We assume that the production function is represented by a CRS production function with three factors: physical capital, skilled and unskilled labor:

$$Y_t = F(K_t, S_t, U_t) = U_t f(k_t, \lambda_t) \tag{40}$$

where K_t is the aggregate amount of the physical capital good measured in the same units as output, S_t, U_t respectively denote inputs of skilled and unskilled labor, and $f(k, \lambda) \equiv F(k, \lambda, 1), k_t \equiv \frac{K_t}{U_t}, \lambda_t \equiv \frac{S_t}{U_t}$. We assume that f is strictly increasing, twice differentiable and strictly concave in both arguments, and satisfies Inada conditions. Capital and skilled labor are complementary: $f_{k\lambda} > 0$.

Moreover, for the sake of simplifying the steady state analysis, we shall additionally assume that skilled and unskilled labor can be aggregated into a single labor composite $H(S, U)$ which is a CRS concave function satisfying Inada conditions, such that the aggregate production function $F(K, S, U)$ can be written as a function $P(K, H(S, U))$ which is also concave, CRS and satisfies Inada conditions. This property is satisfied by the Cobb-Douglas production function, for instance.

Physical capital is owned by households, and can be accumulated via saving and investment. Each parent can then own and bequeath two forms of capital: physical, and human. Incomes derive from both labor and capital market earnings. The allocation of investment in children will depend on anticipated relative returns to the two forms of capital.

Let us first describe how a date t parent who inherited physical capital $\mathcal{K}_t$ allocates a given level of investment expenditure I between these two forms of capital:

$$I = \mathcal{K}_{t+1} - \mathcal{K}_t + e_t x \tag{41}$$

where $e_t \in \{0, 1\}$ denotes the decision whether or not to educate his child, and $\mathcal{K}_{t+1}$ is the capital bequeathed. The resulting income of the child will be

$$y_{t+1} = r_{t+1}\mathcal{K}_{t+1} + e_t(v_{t+1} - w_{t+1}) + w_{t+1} \tag{42}$$

where r_t, v_t, w_t denote factor prices given respectively by $f_k(k_t, \lambda_t)$, $f_\lambda(k_t, \lambda_t)$ and $f(k_t, \lambda_t) - k_t f_k(k_t, \lambda_t) - \lambda_t f_\lambda(k_t, \lambda_t)$. The date t parent will select e and $\mathcal{K}_{t+1}$ to maximize (42), subject to the constraint $\mathcal{K}_{t+1} \geq 0$, given the size of the investment I, inherited wealth $\mathcal{K}_t$ and anticipated factor prices at $t+1$.

Let $R_{t+1} \equiv \frac{v_{t+1}-w_{t+1}}{x}$ denote the rate of return to investment in human capital. Then note that the child's future income can be written as

$$y_{t+1} = r_{t+1}I + e_t x(R_{t+1} - r_{t+1}) + r_{t+1}\mathcal{K}_t + w_{t+1} \tag{43}$$

Unless the rate of return on human capital is at least as large as that on physical capital ($R_{t+1} \geq r_{t+1}$), no one will want to educate their children. Assuming that both kinds of labor are essential in production, an equilibrium must satisfy the condition $R_t \geq r_t$ at all dates t. And given this condition, a parent choosing an investment of at least $x - \mathcal{K}_t$, will find it optimal to educate its child; it will definitely select to educate if the rate of return on human capital strictly exceeds that on physical capital. Optimal education choices are given by

$$e(I; \mathcal{K}_t, R_{t+1}, r_{t+1}) = \begin{cases} 1 & \text{if } R_{t+1} > r_{t+1} \text{ and } I \geq x - \mathcal{K}_t \\ 0 & \text{if } R_{t+1} < r_{t+1} \text{ or } I < x - \mathcal{K}_t \\ \{0,1\} & \text{if } R_{t+1} = r_{t+1} \text{ and } I \geq x - \mathcal{K}_t \end{cases} \tag{44}$$

while the resulting income transferred to the child as a function of the investment is (assuming $R_{t+1} \geq r_{t+1}$):

$$Y_{t+1}(I) = \begin{cases} r_{t+1}I + [w_{t+1} + r_{t+1}\mathcal{K}_t] & \text{if } I < x - \mathcal{K}_t \\ r_{t+1}I + [v_{t+1} + r_{t+1}(\mathcal{K}_t - x)] & \text{if } I \geq x - \mathcal{K}_t \end{cases} \tag{45}$$

If the rate of return on human capital strictly exceeds that on physical capital, this results in a nonconvexity in investment returns, with a discrete upward jump at the point that the investment is just enough to afford tuition costs (after running down all the capital owned by the parent). This is depicted in Figure 5.

Turn now to the selection of the scale of investment by a date t parent with a current income of y

$$V_t(y) = \max_{I \geq -\mathcal{K}_t} \{u(y - I) + \delta V_{t+1}(Y_{t+1}(I))\} \tag{46}$$

where we suppress dependence of the value functions on factor prices and inherited capital stock. The income y_t of a date t parent equals $r_t\mathcal{K}_t + v_t$ if the parent is educated, and $r_t\mathcal{K}_t + w_t$ otherwise.

5.1 Steady States

We shall construct steady states where education status, capital stock and consumption of every family is maintained over time. Those currently educated will pay for their children's education, selecting an investment which pays for tuition $I = x$, while the

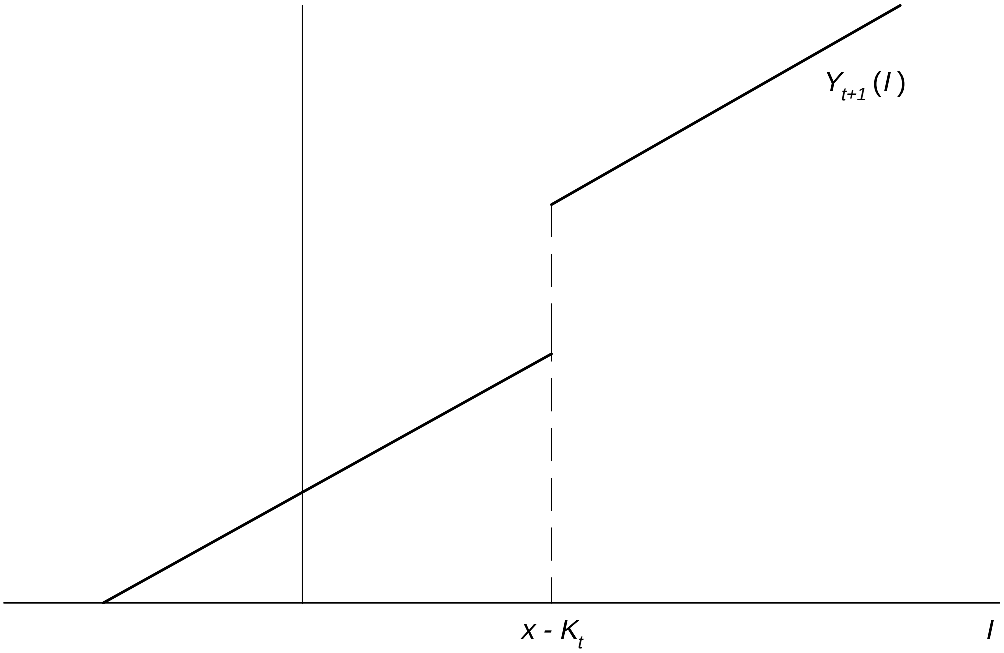

Figure 5: INVESTMENT RETURNS

uneducated do not invest in education for their children at all. Both kinds of families will hand down physical capital stocks in exactly the quantities they inherited. Consumption smoothing requires that

$$f_k(k,\lambda) = r = \frac{1}{\delta} \tag{47}$$

while education investment incentives require that

$$R \equiv \frac{v-w}{x} \geq r \tag{48}$$

It follows that the level of capital per (unskilled) worker $k(\lambda)$ — which is given by condition (47) — is an increasing function of λ. Investments in physical capital will therefore depend on the fraction of the workforce which is skilled, owing to the complementarity between physical and human capital.

Given conditions (47) and (48), it is evident that educated parents are selecting investment levels optimally. In the case where (48) holds as an equality, the nonconvexity in returns to investment does not exist, and the investment return function is given by $Y_{t+1}(I) = rI + [v + r(\mathcal{K}_t - x)]$ for all levels of investment (rather than only those exceeding $x - \mathcal{K}_t$). Then it is clear that consumption smoothing (involving $I = x$ and $\mathcal{K}_{t+1} = \mathcal{K}_t$) is optimal for an educated parent, as it satisfies the Euler and transversality conditions jointly necessary and sufficient for optimality of the consumption-saving policy. This is true irrespective of their inherited capital stock, which is handed down unmodified to their children. Since this policy never necessitates these families investing less than $x - \mathcal{K}_t$, this policy continues to be optimal even when (48) holds as a strict inequality, as the investment returns are adjusted downward in a region that the family never intended to select.

For uneducated parents, however, if there is a nonconvexity in investment returns, the optimality of not investing in their children's education requires the level of capital stock they inherit to be limited relative to inheritances received by skilled households. Otherwise it generates enough capital income to overcome the difference in earnings, and unskilled households too will invest in education. When the support of the income distribution of the unskilled and skilled overlap in this way, a steady state cannot therefore not be characterized by such a nonconvexity. The returns to human and nonhuman capital must be equal, rendering all families indifferent about how they allocate their bequests. This pins down the skill intensity of production and the labor allocation in the economy, given our assumption that the technology permits aggregation of the different types of labor into a composite labor input. Specifically, the steady state condition (47) implies that the level of capital stock is a function of the labor composite H, thus enabling us to express output in terms of inputs of skilled and unskilled labor input alone:

$$G(S, U) = F(K(H(S, U)), H(S, U)) \tag{49}$$

where the 'reduced form' production function G is also concave, CRS and satisfies Inada conditions. Then the marginal products of each type of labor, which incorporate the accompanying change in capital stock, behave exactly in the same way as in a model without any physical capital. In particular, the skill premium will be monotone in the skill intensity of the labor input, and will have the same limiting properties. We can therefore express the wages of the two types of labor as a function of λ, incorporating the implicit dependence of the capital stock on labor inputs.

It then follows that there exists a unique skill intensity λ^* in which the rates of return to the two forms of capital are equalized:

$$v(\lambda^*) - w(\lambda^*) = \frac{x}{\delta} \tag{50}$$

In such a steady state, the labor earnings of the educated net of the education costs they incur for their children $v(\lambda^*) - x$ strictly exceeds the earnings of the unskilled $w(\lambda^*)$. The educational choices are optimal for both types of families, irrespective of the capital they own, since the returns to investment do not have a nonconvexity. The steady state is thus compatible with arbitrary distribution of ownership of physical capital in the economy. In particular, it is compatible with a perfectly equal income distribution, e.g., if every unskilled family owns more physical capital than does every skilled family, just enough to offset their earning differentials and education costs:

$$\mathcal{K}_u^* = \mathcal{K}_s^* + v(\lambda^*) - x - w(\lambda^*) \tag{51}$$

This clarifies the key distinction between the economy with and without forms of non-human capital: alternative sources of wealth can offset labor market differentials to wipe out inequality entirely. The uneducated need not get trapped for ever in a low income-investment cycle: they can offset their lower labor market earnings with greater capital income. Indeed, there is nothing to prevent unskilled families from being better off relative to skilled families, if the distribution of wealth is sufficiently skewed in their favor.

Nevertheless 'unequal' steady states with a nonconvexity in investment returns also coexist in the same model. Turn now to such steady states where the return to human capital exceeds that for physical capital, which requires that:

$$v(\lambda) - w(\lambda) > \frac{x}{\delta} \tag{52}$$

Given our assumptions concerning the technology, this implies that skilled labor must be in short supply relative to that in the 'equal' steady state described above: the skill intensity in production λ must be less than λ^*, and hence so must the capital intensity ($k(\lambda) < k(\lambda^*)$). Educating one's child is optimal for everyone, provided parents are willing to invest on a scale large enough to cover training costs. It is necessary in such a steady state that unskilled parents do not have enough income to want to invest on such a scale. In particular the income distribution among the skilled and unskilled cannot overlap. Otherwise an unskilled family with the same income as a skilled family would choose to educate their children, since the latter do. Hence such steady states must be associated with income inequality of a stark form: the wealthiest unskilled family must be worse off compared with a skilled family with no physical capital at all. Compared with the 'equal' steady state, such a steady state is associated both with higher inequality, and lower investment in both forms of capital.

Recall that it is optimal for every educated family to educate their children and smooth consumption, irrespective of whether or not there is a nonconvexity in returns to investment. It remains only to check the condition that uneducated families will not want to educate their children. Since educational incentives cannot decrease if capital income rises, it is necessary for this that an uneducated family without any capital at all will not want to educate their children, i.e.,

$$u(w) \geq (1-\delta)u(w-x) + \delta u(v-x) \tag{53}$$

Conversely, once this condition is satisfied, it also defines an upper bound for the amount of capital that can be owned by an unskilled family (the value of $\mathcal{K}$ for which this incentive constraint would be satisfied as an equality, if the current income of the family were augmented by $r\mathcal{K} = \frac{\mathcal{K}}{\delta}$). Hence conditions (52) and (53) are necessary and sufficient to characterize such 'unequal' steady states.

If we confine attention to steady states where the unskilled own no physical capital at all, the model essentially reduces to that in the previous section.[24] The same is true when they own insufficient physical capital to afford training for their children. Hence there must be a continuum of unequal steady states varying in terms of inequality and per capita output, exactly as in the context without any physical capital, and these steady states will have similar qualitative properties. These 'unequal' steady states are marked by smaller levels of physical and human capital, higher inequality in labor market earnings, and lower output, compared to an 'equal' steady state described above.

We conclude that the range of steady states is wider in the presence of alternative forms of nonhuman wealth. While the persistence of inequality in the long run is not inevitable, it does necessitate a sufficiently 'progressive' distribution of physical capital which redresses the inequities resulting from differences in human capital. There is no inherent reason why the market mechanism will tend to generate such a distribution of capital ownership: indeed, in the model the distribution of nonhuman wealth is quite arbitrary. This gives rise to the question: if all families start equal with respect to ownership of all forms of capital, will the resulting longrun outcomes must necessarily involve inequality, and if so in what form? Exploring this question remains a task for future research.

Appendix 1. Details for Example in Section 3.2

We consider the extended example described in Section 3.2. We assume throughout that there exists ω and Ω such that for all n, $\infty > \Omega \geq w_0(n) \geq \omega > 0$. See footnote 18 for a justification.

Lemma 1 *There exists $\epsilon > 0$ such that for all n and h,*

$$w_n(h) - x(h) \geq \epsilon. \tag{54}$$

Proof. By the general conditions describing a steady state (see, e.g., (11)),

$$\begin{aligned} u\left(w_n(h) - x(h)\right) &\geq (1-\delta)u\left(w_n(h)\right) + \delta u\left(w_n(0) - x(0)\right) \\ &\geq (1-\delta)u\left(w_n(h)\right) \\ &\geq (1-\delta)u\left(\omega\right), \end{aligned}$$

---

[24] The only difference would be that the skilled would be endowed with additional sources of wealth, which will only enhance their incentive to educate their children.

where the second inequality uses the fact that $w_n(0) - x(0) \geq 0$ and the normalization that u is nonnegative, and the last inequality uses the fact that $w_n(h) \geq w_n(0) \geq \omega$. Now define $\epsilon \equiv u^{-1}\left([1-\delta]u\left(\omega\right)\right)$ to complete the proof. ■

Lemma **2** *There exists $K < \infty$ such that $w_n(h) \leq K$ for all K and all h.*

Proof. By the continuity (and therefore uniform continuity) of x and Lemma 1, there exists an index N such that for all $n \geq N$ and all available professions h,

$$w_n(h) - x(h^r_N) > 0, \tag{55}$$

where h^r_n denotes the profession right-adjacent to h in $\mathcal{H}(n)$. Hence all professions between h and h^r_N can be "accessed" by a household of profession h, for all $n \geq N$. For such pairs h and h', the comparison (11) applies, so that

$$u\left(w_n(h) - x(h)\right) \geq (1-\delta)u\left(w_n(h) - x(h')\right) + \delta u\left(w_n(h') - x(h')\right)$$

This implies, in particular, that

$$\delta u\left(w_n(h') - x(h')\right) \leq u\left(w_n(h)\right),$$

so that for all $h' \in (h, h^r_N)$ and all $n \geq N$,

$$w_n(h') \leq K(w_n(h)), \tag{56}$$

where

$$K(w) \equiv u^{-1}\left(\frac{1}{\delta}u(w)\right) + x(1). \tag{57}$$

Note that K is an increasing function. A simple iteration using (56) therefore establishes that for all $n \geq N$ and all h,

$$w_n(h) \leq K^{(N)}(\Omega),$$

where Ω, it will be recalled, is simply the upper bound on $w_n(0)$ and $K^{(N)}$ is simply the N-fold composition of K. This establishes boundedness (there are only finitely many values of w before the index N is reached). ■

In what follows, notice that for all $h > 0$ that are available professions for some n, its *lower n-adjacent profession h_n* — the adjacent profession to the "left" of h — is well-defined.

Lemma **3** *For each $h > 0$, $w_n(h) - w_n(h_n) \to 0$ as $n \to \infty$, and the convergence is uniform over h.*

Proof. We know that for each h and n, $w_n(h) - w_n(h_n) \geq 0$. So if the lemma is false, there is a subsequence of n (retain notation), a corresponding sequence of available professions h^n (together with their lower n-adjacent professions, denoted by h^n_n), and some $\eta > 0$ such that

$$w_n(h^n) - w_n(h^n_n) \geq \eta \tag{58}$$

for all n along this subsequence. By the same uniformity argument leading to (55) in the proof of Lemma 2, there is N such that for all $n \geq N$, the adjacent comparison (see (12) or (13)) between h^n and h^n_n applies, and

$$u\left(w_n(h^n_n) - x(h^n_n)\right) - u\left(w_n(h^n_n) - x(h^n)\right) \geq \frac{\delta}{1-\delta}\left[u\left(w_n(h^n) - x(h^n)\right) - u\left(w_n(h^n_n) - x(h^n_n)\right)\right] \tag{59}$$

By Lemma 2, all values in the domain of u (above) lie in some compact set, so all relevant continuous functions are uniformly continuous. Consequently, as $n \to \infty$ along our subsequence, the LHS of (59) converges to zero (by the uniform continuity of x and u), while the RHS stays bounded away from zero (by (58) and, again, the uniform continuity of x and u). This is a contradiction. ■

Recall the definition of $a_n(h)$ and $b_n(h)$ in the main text. We reproduce them here:

$$a_n \equiv \frac{u'\left(w_n(h_n) - x(h)\right)}{u'\left(w_n(h) - x(h)\right)}, \tag{60}$$

and

$$b_n \equiv \frac{u'\left(w_n(h) - x(h_n)\right)}{u'\left(w_n(h_n) - x(h_n)\right)}, \tag{61}$$

where h_n is the lower n-adjacent profession to h.

LEMMA **4** *For each $h > 0$, $a_n(h)$ converges to unity from above and $b_n(h)$ converges to unity from below as $n \to \infty$. The convergence is uniform in h.*

Proof. It is obvious that $a_n(h) > 1$ while $b_n(h) < 1$. Next, by Lemmas 1 and 2, all values in the domain of u' (in (60) and (61) above) are uniformly bounded below and above by strictly positive numbers.[25] Thus u' is uniformly continuous in the relevant subdomain, and Lemma 3 completes the proof. ■

By rearranging the inequalities in (21) and (22), we obtain for each n and available profession h,

$$\frac{(1-\delta)b_n(h) + \delta}{\delta}\left[x(h) - x(h_n)\right] \leq w_n(h) - w_n(h_n) \leq \frac{(1-\delta)a_n(h) + \delta}{\delta}\left[x(h) - x(h_n)\right],$$

---

[25] The only check that requires a little extra work is the boundedness below of $w_n(h_n) - x(h)$, but this is guaranteed by the same argument leading up to (55) in the proof of Lemma 2.

and summing these inequalities over all available professions h' in the grid (for n) such that $0 < h' \leq h$,

$$\sum_{h'>0;h'\leq h} \frac{(1-\delta)b_n(h')+\delta}{\delta}\left[x(h')-x(h'_n)\right] \leq w_n(h) - w_n(0) \leq \sum_{h'>0;h'\leq h} \frac{(1-\delta)a_n(h')+\delta}{\delta}\left[x(h')-x(h'_n)\right].$$

But the uniform convergence of $a_n(h)$ and $b_h(h)$ established in Lemma 4 shows that both the summation terms above must converge to $(1/\delta)x(h)$ as $n \to \infty$ (we use here the fact that $x(0) = 0$).

We have shown so far that for every limit point $w(0)$ of $w_n(0)$, $w_n(h) \to w(h)$ along the corresponding subsequence of n for every h that is available for some n onwards, where

$$w(h) = \frac{1}{\delta}x(h) + w(0). \tag{62}$$

The main text argues that there can be only one such value of $w(0)$ that is compatible with competitive equilibrium in the limit.

Appendix 2. Steps in the Proof of Proposition 5

Lemma **5** *If $\lambda_t > \lambda_{t+1}$, then $\lambda_t \in A$ and $\lambda_{t+1} = \lambda_{t+2}$.*
If $\lambda_t < \lambda_{t+1}$, then $\lambda_t \in B$ and $\lambda_{t+1} \leq \lambda_{t+2}$.

Warning. Note that the two statements in the lemma are *not* symmetric. The lack of symmetry will become even clearer later.

Proof of Lemma 5. *We begin by establishing the first part of the first statement.* Because $\lambda_t > \lambda_{t+1}$, (37) must hold with equality, and we have

$$u(v_t) - u(v_t - x) = \delta[u(v_{t+1} - x) - u(w_{t+1})] + \delta^2 M \tag{63}$$

where $M \equiv \sum_{s=t+2}^{\infty} \delta^{s-(t+2)}[u(v_s - x) - u(w_s)]$. Using (37) for period $t+1$, we see that

$$u(v_{t+1}) - u(v_{t+1} - x) \leq \delta M. \tag{64}$$

Combining (63) and (64), we see that

$$u(v_t) - u(v_t - x) \geq \delta[u(v_{t+1}) - u(w_{t+1})].$$

Because $\lambda_t > \lambda_{t+1}$, we see that $v_{t+1} > v_t$ and $w_{t+1} < w_t$. Therefore

$$u(v_t) - u(v_t - x) > \delta[u(v_t) - u(w_t)],$$

which shows that $\lambda_t \in A$.

The proof of the first part of the second statement is completely parallel, but because (as noted above) there is an asymmetry lurking here it will be useful to simply retrace these steps and convince ourselves that they indeed go through.

For this part, $\lambda_t < \lambda_{t+1}$, so that (38) must hold with equality, and we have

$$u(w_t) - u(w_t - x) = \delta[u(v_{t+1} - x) - u(w_{t+1})] + \delta^2 M \tag{65}$$

where M is defined just as before. Using (38) for period $t+1$, we see that

$$u(w_{t+1}) - u(w_{t+1} - x) \geq \delta M. \tag{66}$$

Combining (65) and (66), we see that

$$u(w_t) - u(w_t - x) \leq \delta[u(v_{t+1} - x) - u(w_{t+1} - x)].$$

Because $\lambda_t < \lambda_{t+1}$, we see that $v_{t+1} < v_t$ and $w_{t+1} > w_t$. Therefore

$$u(w_t) - u(w_t - x) < \delta[u(v_t - x) - u(w_t - x)],$$

which shows that $\lambda_t \in B$.

Next, we establish the second part of the first statement: that $\lambda_{t+1} = \lambda_{t+2}$. Suppose this is false. Then there are two cases to consider.

Case 1: $\lambda_{t+1} < \lambda_{t+2}$. Then at date $t+1$, (38) must hold with equality, so that

$$u(w_{t+1}) - u(w_{t+1} - x) = \delta M. \tag{67}$$

Combining (63) and (67), we see that

$$u(v_t) - u(v_t - x) = \delta[u(v_{t+1} - x) - u(w_{t+1} - x)]. \tag{68}$$

Because $\lambda_t > \lambda_{t+1}$, we have $v_t > w_t > w_{t+1}$. Consequently, by the strict concavity of the utility function,

$$u(v_t) - u(v_t - x) < u(w_t) - u(w_t - x) < u(w_{t+1}) - u(w_{t+1} - x). \tag{69}$$

Combining (68) and (69), we may conclude that

$$u(w_{t+1}) - u(w_{t+1} - x) > \delta[u(v_{t+1} - x) - u(w_{t+1} - x)].$$

But this means that $\lambda_{t+1} \notin B$. On the other hand, we have $\lambda_{t+1} < \lambda_{t+2}$, and this contradicts the first part of the second statement of the lemma, which we have already proved.

CASE 2: $\lambda_{t+1} > \lambda_{t+2}$. Then at date $t+1$, (37) must hold with equality, so that

$$u(v_{t+1}) - u(v_{t+1} - x) = \delta M. \tag{70}$$

Combining (63) and (70), we see that

$$u(v_t) - u(v_t - x) = \delta[u(v_{t+1}) - u(w_{t+1})]. \tag{71}$$

Because $\lambda_t > \lambda_{t+1}$, we have $v_t < v_{t+1}$. Consequently, by the strict concavity of the utility function,

$$u(v_t) - u(v_t - x) > u(v_{t+1}) - u(v_{t+1} - x). \tag{72}$$

Combining (71) and (72), we may conclude that

$$u(v_{t+1}) - u(v_{t+1} - x) < \delta[u(v_{t+1}) - u(w_{t+1})].$$

But this means that $\lambda_{t+1} \notin A$. On the other hand, we have $\lambda_{t+1} > \lambda_{t+2}$, and this contradicts the first part of the first statement of the lemma, which we have already proved.

Finally, we prove the second part of the second statement: that $\lambda_{t+1} \leq \lambda_{t+2}$. Suppose this is false. Then $\lambda_{t+1} > \lambda_{t+2}$. Thus at date $t+1$, (37) must hold with equality, so that

$$u(v_{t+1}) - u(v_{t+1} - x) = \delta M. \tag{73}$$

Combining (65) and (73), we see that

$$u(w_t) - u(w_t - x) = \delta[u(v_{t+1}) - u(w_{t+1})]. \tag{74}$$

Because $\lambda_t < \lambda_{t+1}$, we have $w_t < w_{t+1} \leq v_{t+1}$. Consequently, by the strict concavity of the utility function,

$$u(w_t) - u(w_t - x) > u(w_{t+1}) - u(w_{t+1} - x) \geq u(v_{t+1}) - u(v_{t+1} - x). \tag{75}$$

Combining (74) and (75), we may conclude that

$$u(v_{t+1}) - u(v_{t+1} - x) < \delta[u(v_{t+1}) - u(w_{t+1})].$$

But this means once again that λ_{t+1} satisfies the first inequality in (14), or equivalently, that $\lambda_{t+1} \notin A$. On the other hand, we have $\lambda_{t+1} > \lambda_{t+2}$, and this contradicts the first part of the first statement of the lemma, which we have already proved.

■

LEMMA **6** *If λ is a steady state, then there is a unique competitive equilibrium from $\lambda_0 = \lambda$, given by $\lambda_t = \lambda$ for all t.*

Proof. Immediate from Lemma 5. For if the competitive equilibrium is nonstationary, then it must be the case that either $\lambda \in A$ or $\lambda \in B$ (simply examine the first date that $\lambda_t \neq \lambda_{t+1}$ and apply Lemma 5). In either of these cases, λ cannot be a steady state. ■

A converse to this result is the subject of the next lemma.

Lemma **7** *If at any date t along a competitive equilibrium we have $\lambda_t = \lambda_{t+1}$, then $\lambda \equiv \lambda_t = \lambda_{t+1}$ is a steady state, and in particular $\lambda_s = \lambda_t$ for all $s \geq t$.*

Proof. Suppose not. Then by Lemma 6, it must be the case that either $\lambda \in A$ or $\lambda \in B$.

Case 1: $\lambda \in A$. In this case, renumbering time periods if necessary, we must have $\lambda_t = \lambda_{t+1} > \lambda_{t+2}$ (using Lemma 5). Thus (37) must hold with equality at date $t+1$, so that

$$u(v_{t+1}) - u(v_{t+1} - x) = \delta M, \tag{76}$$

while at date t

$$u(v_t) - u(v_t - x) \leq \delta[u(v_{t+1} - x) - u(w_{t+1})] + \delta^2 M \tag{77}$$

Combining (76) and (77), we see that

$$\begin{aligned} u(v_t) - u(v_t - x) &\leq \delta[u(v_{t+1}) - u(w_{t+1})] \\ &= \delta[u(v_t) - u(w_t)]. \end{aligned}$$

But this means that $\lambda \notin A$, which is a contradiction.

Case 2: $\lambda \in B$. In this case, renumbering time periods if necessary, we must have $\lambda_t = \lambda_{t+1} < \lambda_{t+2}$ (using Lemma 5). Thus (38) must hold with equality at date $t+1$, so that

$$u(w_{t+1}) - u(w_{t+1} - x) = \delta M, \tag{78}$$

while at date t

$$u(w_t) - u(w_t - x) \geq \delta[u(v_{t+1} - x) - u(w_{t+1})] + \delta^2 M \tag{79}$$

Combining (78) and (79), we see that

$$\begin{aligned} u(w_t) - u(w_t - x) &\geq \delta[u(v_{t+1} - x) - u(w_{t+1} - x)] \\ &= \delta[u(v_t - x) - u(w_t - x)]. \end{aligned}$$

But this means that $\lambda \notin B$, which is a contradiction.

So neither Case 1 nor Case 2 is possible. This means that λ is a steady state. Applying Lemma 6, we see that there is a unique stationary equilibrium, and we are done. ■

Proof of Proposition 5. To prove the first part of the proposition, note that *if* there is a competitive equilibrium, then by Lemmas 5 and 7, it must have the property discussed in the statement of the proposition. To check existence and uniqueness, define λ_1 by

$$u(v(\lambda)) - u(v(\lambda) - x) \equiv \delta(1-\delta)^{-1}[u(v(\lambda_1) - x) - u(w(\lambda_1))].$$

It is easy to see that λ_1 is well-defined and unique, and that $\lambda_1 < \lambda$. Now check that this gives us a competitive equilibrium, and that there is no other way of constructing an path that satisfies both (37) and (38).

To prove the second part of the proposition, we first need to strengthen the implication of Lemma 5 in this case. It will be enough to strengthen the second part of the statement of that lemma to: If $\lambda_t < \lambda_{t+1}$, then $\lambda_t \in B$ and $\lambda_{t+1} < \lambda_{t+2}$.

All of this is proved except for the stronger implication: $\lambda_{t+1} < \lambda_{t+2}$. To establish this, suppose that the assertion is false. Then, using Lemma 5, it must be the case that $\lambda_t < \lambda_{t+1} = \lambda_{t+2}$. By Lemma 7, we have $\lambda_{t+1} = \lambda_s$ for all $s \geq t+1$. Also, (38) must hold with equality at date t. Combining these two pieces of information, we see that

$$u(w_t) - u(w_t - x) = \delta(1-\delta)^{-1}[u(v_{t+1} - x) - u(w_{t+1})].$$

Now $\lambda_t < \lambda_{t+1}$, so that $w_t < w_{t+1}$. By the strict concavity of u and the equality above,

$$u(w_{t+1}) - u(w_{t+1} - x) < \delta(1-\delta)^{-1}[u(v_{t+1} - x) - u(w_{t+1})].$$

But this means that $\lambda_{t+1} \in B$ as well. But then by Lemma 7, it cannot be the case that $\lambda_{t+1} = \lambda_{t+2}$.

To prove existence and uniqueness from this initial condition, define recursively for each λ_t, the value of λ_{t+1} that solves the equation

$$u(w_t) - u(w_t - x) \equiv \delta[u(v_{t+1} - x) - u(w_{t+1} - x)], \tag{80}$$

where v_{t+1} and w_{t+1} are to be interpreted as the wages corresponding to λ_{t+1}.

To see that this is uniquely defined, note that

$$u(w_0) - u(w_0 - x) < \delta[u(v_0 - x) - u(w_0 - x)],$$

because $\lambda_0 \in B$. So there is a unique λ_1 that solves (80) for $t = 0$. Note that λ_1 must exceed λ_0. And this will be so whenever $\lambda_t \in B$. So it only remains to show that if $\lambda_t \in B$, then $\lambda_{t+1} \in B$ as well. To see thus simply use the fact that $\lambda_{t+1} > \lambda_t$, which implies that $w_{t+1} > w_t$. Using this information in (80) along with the strict concavity of u, we are done.

The trick to understanding uniqueness is that this is the *only* way to construct a competitive equilibrium from the given initial condition, because (38) will have to hold with equality.

Finally, part 3 of the proposition is already established. ■

References

Aghion, P. and P. Bolton (1997). "A Theory of Trickle-Down Growth and Development," *Review of Economic Studies* **64**, 151–172.

Baland, J-M. and D. Ray (1991). "Why Does Asset Inequality Affect Unemployment? A Study of the Demand Composition Problem," *Journal of Development Economics* **35**, 69–92.

Bandyopadhyay D. (1997). "Distribution of Human Capital and Economic Growth," Working Paper no. 173, Department of Economics, Auckland Business School.

Banerjee, A. and A. Newman (1993), "Occupational Choice and the Process of Development," *Journal of Political Economy* **101**, 274–298.

Barro, R. (1974), "Are Government Bonds Net Wealth?," *Journal of Political Economy* **82**, 1095-1117.

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

Champernowne, D. (1953), "A Model of Income Distribution," *Economic Journal* **63**, 318–351.

Clark, C. (1971). "Economically Optimal Policies for the Utilization of Biologically Renewable Resources," *Mathematical Biosciences* **12**, 245–260.

Dasgupta, P. and D. Ray (1986). "Inequality as a Determinant of Malnutrition and Unemployment: Theory," *Economic Journal* **96**, 1011–1034.

Dasgupta, P. and D. Ray (1987). "Inequality as a Determinant of Malnutrition and Unemployment: Policy," *Economic Journal* **97**, 177–188.

Dechert, R. and K. Nishimura (1983), "A Complete Characterization of Optimal Growth Paths in an Aggregated Model with a Non-Concave Production Function," *Journal of Economic Theory* **31**, 332–354.

Freeman, S.(1996), "Equilibrium Income Inequality among Identical Agents," *Journal of Political Economy* **104(5)**, 1047-1064.

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

LJUNGQVIST, L. (1993), "Economic Underdevelopment: The Case of Missing Market for Human Capital," *Journal of Development Economics* **40**, 219–239.

LOURY, G. C. (1981). "Intergenerational Transfers and the Distribution of Earnings," *Econometrica* **49**, 843–867.

MAJUMDAR, M. AND T. MITRA (1982), "Intertemporal Allocation with a Non-Convex Technology: The Aggregative Framework," *Journal of Economic Theory* **27**, 101–136.

MAJUMDAR, M. AND T. MITRA (1983), "Dynamic Optimization with a Non-Convex Technology: The Case of a Linear Objective Function," *Review of Economic Studies* **50**, 143–151.

MANI, A. (1998). "Income Distribution and the Demand Constraint." Mimeograph, Department of Economics, Vanderbilt University.

MATSUYAMA, K. (1999), "The Rise of Mass Consumption Societies," mimeo.

MAOZ, Y.D. AND O. MOAV (1999), "Intergenerational Mobility and the Process of Development," *Economic Journal*, 109 (October), 677-697.

MITRA, T. AND D. RAY (1984), "Dynamic Optimization on a Non-Convex Feasible Set: Some General Results for Non-Smooth Technologies," *Zeitschrift für Nationalökonomie* **44**, 151–175.

MULLIGAN, C. (1997), *Parental Priorities and Economic Inequality*, Chicago, IL: University of Chicago Press.

PIKETTY, T. (1997), "The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing," *Review of Economic Studies* **64**, 173–189.

RAY, D. (1990), "Income Distribution and Macroeconomic Behavior," mimeo.

RAY, D. AND P. STREUFERT (1993). "Dynamic Equilibria with Unemployment due to Undernourishment," *Economic Theory* **3**, 61–85.

SKIBA, A. (1978). "Optimal Growth with a Convex-Concave Production Function," *Econometrica* **46**, 527–540.

SOLOW, R. (1956). "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics* **70**, 65–94.