# Occupational Diversity and Endogenous Inequality

Dilip Mookherjee
Boston University

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

February 2006

Abstract

A traditional view of markets with intergenerational bequests within dynasties is that they equalize wealth across households. A more recent literature suggests that markets are inherently disequalizing. A third viewpoint argues that initial history is crucial in determining whether inequalities persist or not. By constructing a theory of equilibrium investment allocation between human capital and financial assets in the presence of borrowing constraints, we address these views in a unified way. Two attributes of occupational diversity turn out to be central to our understanding: *span*, the range of training costs across occupations, and *richness*, the variety of different training costs contained within the span. The former is used to generate a necessary and sufficient condition for markets to be disequalizing, while the latter is shown to be directly connected to the question of history-dependence. Among the several implications of the analysis is a greater proneness for markets to be disequalizing in poorer countries.

Mookherjee's research was supported by the MacArthur Foundation and by National Science Foundation Grant No. 0418434, and Ray's research by National Science Foundation Grant No. 0241070. Ray is also grateful to the London School of Economics for warm hospitality during a sabbatical year. We thank participants at several seminars for their helpful comments. The views expressed in this paper are the authors' own and do not represent the views of the MacArthur Foundation or the National Science Foundation. Some results contained in the present paper were informally presented in an AER *Papers and Proceedings* volume; see Mookherjee and Ray (2002). Address correspondence to `dilipm@bu.edu` and `debraj.ray@nyu.edu`.

## 1. INTRODUCTION

Do markets possess some intrinsic tendency to equalize or disequalize the fortunes of different households or societies? This is a fundamental question in the theory of income distribution. In order to pose it precisely, one needs a definition of what it means to "equalize" or "disequalize". In our view, the appropriate definition must be set in a context in which the dynamics of income distribution are generated *only* by the operation of the market, in the (hypothetical) absence of any heterogeneity or stochastic shocks in abilities, tastes or opportunities. This allows a clear conceptual separation between the role of the market *per se* and heterogeneity or random events in generating or propagating inequality.

For dynastic households, this allows us to frame the question as follows: consider two households with similar abilities and tastes but different wealth. Does the market by itself (in the absence of any stochastic shocks) induce the wealths of their descendants to draw ever closer together, and their difference to eventually vanish? If so, say that the market is *equalizing*. Conversely, the market is *disequalizing* if — even in the absence of heterogeneity or random events — it causes the wealths of equal or near-equal households to separate over subsequent generations, and such wealth differentials to persist in the long run. To be sure, one could ask the same question at different levels: for instance, is there a tendency for economies with similar characteristics (i.e., tastes and technology) but different historical conditions to converge or diverge?

Existing theories provide — explicitly or implicitly — answers to these questions, and on that basis we can classify such theories into three broad groups. First, there are theories that exemplify what might be called the *equalization view*, in which the intergenerational transmission of wealth makes for inter-agent convergence (Brock and Mirman (1976), Becker and Tomes (1979, 1986), Loury (1981)). In this view, long run inequality is just an ongoing tussle between the convergence and exogenous, ongoing stochastic shocks to abilities or opportunities ("luck", generally speaking), the disequalizing tendencies of which the market subsequently tends to iron out.

At the other extreme lies the *disequalization view* (Ray (1990), Ljungqvist (1993), Freeman (1996), Bandyopadhyay (1997), Mookherjee and Ray (2003)) which argues that markets *must* separate individual fortunes: despite the absence of stochastic shocks, *all* steady states must involve interhousehold inequality that persists across generations. Among other things, these models rely on "symmetry-breaking": the demand by the market for a diversity of occupations forces individuals in identical or near-identical situations to make distinct choices of profession, with implications for the subsequent emergence of inequality, not just in gross income, but in net lifetime *payoffs*.

In between these two approaches lies a third collection of models which may be described succintly as the *neutrality view* (Banerjee and Newman (1993), Galor and Zeira (1993), Ray and Streufert (1993), Aghion and Bolton (1997), Piketty (1997), Lloyd-Ellis and Bernhardt (2000), Matsuyama (2000, 2003), and Ghatak and Jiang (2002)).[1] These models permit both initial inequalities and initial equalities to persist. Typically, history determines where a society ends up in the long run.[2]

[1] Aghion and Bolton's model is characterized by an ergodic wealth distribution, owing partly to some strong parameter restrictions they impose. In the absence of such restrictions, their model would also belong to the general class of "neutral" models.

[2] There are also models in which the steady state distribution of wealth is fully indeterminate — see, e.g., Chatterjee (1994). Here the investment frontier that each household faces is linear (rather than strictly convex as in Loury (1981)), so with a dynastic bequest motive each family wants to maintain its wealth indefinitely, and a steady state is compatible with *any* interfamily wealth distribution. If the bequest motive is modified, as in Becker and Tomes (1979), to a paternalistic motive where parents care about the wealth of only their next generation, rather than the infinite succession

The purpose of this paper is to describe an inclusive framework which embeds all three approaches as special cases, thus allowing us to interpret the key differences in underlying assumptions. But the exercise is not just one of achieving generality: it yields new results and insights, which we describe below.

There are two common threads that run through our framework (as in all the literature cited here). The first emphasizes the *intergenerational* transmission of inequality via parental bequests. Such bequests have two broad components. There are *occupational bequests*, in which parents build up the inalienable capital of their children. Typically, such inalienable capital is embodied human capital: nutrition, education, societal positioning, the building-up of skills, the acquisition of professions. It may also involve the transfer of a business that must remain in the family for reasons of moral hazard. There are also *financial bequests*, which may be given to children in lieu of or to supplement occupational bequests.

The second thread is the assumption that parents cannot borrow to make these bequests. Once again, this theme is common to all the papers we cite. Loury (1981) summarizes it cleanly: "Legally, poor parents will not be able to constrain their children to honor debts incurred on their behalf."

This setup is completely standard. Yet it allows us to provide novel answers to two fundamental questions. First, we are able to provide a necessary and sufficient condition under which *every* steady state of the model must display persistent inequality, in line with the disequalization literature. [When the condition fails, at least *one* steady state must display perfect equality.] This condition, which we call the *span condition*, compares the "range" of the occupational structure (in the space of training costs for different occupations) with the *strength* of the parental bequest motive. The condition holds when occupational range or "span" is large relative to the intensity of the bequest motive.

Of course, we discuss the span condition in detail in the paper. Suffice it to mention here that suitable interpretations of the condition yield a variety of implications. A particularly important theme that also recurs in a later section of the paper is that poverty is correlated with the span condition. Therefore, poorer societies are more likely to exhibit disequalization. At the same time, the condition also suggests — somewhat in contrast to the previous observation — that quickly *growing* economies are more vulnerable to disequalization. Likewise, implications can be drawn for countries which are suddenly exposed to a globalized range of products and services. One can study whether increased reliance on physical capital in production is conducive to greater or less inequality. In short, the span condition represents not just an abstract construct but a full characterization that can be used to study a variety of applied issues.

Notice that the span condition provides a context in which inequality will be observed in all steady states, but it is entirely silent on the question of steady-state multiplicity. That multiple equilibria (depending on belief systems) and multiple steady states (depending on historical performance) may explain differences across ex-ante identical societies is a powerful idea in development economics, and can be traced back many decades in academic development thought, to the work of Paul Rosenstein-Rodan, Albert Hirschman and Allyn Young. In such models, history has long-run economic consequences, particularly for inequality, per-capita income, and the overall acquisition of human capital. Indeed, the multiplicity idea forms the basis of a recent literature in economic history that attempts to trace differences in the current functioning

---

of subsequent generations, then this arbitrariness disappears; steady state is consistent only with perfect interfamily wealth equality (under the assumptions made by Becker and Tomes on the strength of parental altruism). Alternatively, if investments are subject to diminishing returns at the househol level of each household, then again perfect long run equality is predicted, irrespective of the bequest motive. We focus on versions with a determinate theory of distribution.

of societies to events that happened in the distant past (see, e.g., Sokoloff and Engerman (2000), Acemoglu, Johnson and Robinson (2001), and Banerjee and Iyer (2005)).

A second objective of the paper is to study history-dependence more closely, and in so doing place the "neutrality" models in perspective. Our analysis throws light on the precise circumstances under which history-dependence can occur. These circumstances turn out to be related to a different feature of occupational structure: the *richness* of occupational choices. When there are just two occupations — say "skilled" and "unskilled" labor — history-dependence is endemic in the class of the models we consider: there is a wide variety of steady states. However, this finding contrasts sharply with that for a "perfectly rich" occupational structure in which the set of training costs constitutes an interval, and all occupations are indeed active in steady state.

We prove that — conditional on the interest rate — richness of occupational structure implies uniqueness of the steady state. That steady state will exhibit persistent inequality (if the span condition is met), but such inequality is now to be regarded as the *inevitable* outcome for the society. There is no other steady state.[3]

Once again, a detailed argument is relegated to the main body of the paper, but the implications of this finding are striking. It appears that the two-occupation model (very commonly used in this literature) paints a misleading picture of multiplicity, and that such multiplicity disappears as occupational richness increases.

To be sure, the uniqueness result is conditioned on the interest rate, and it also relies on the assumption that all occupations are active in steady state. When these assumptions are dropped — singly or jointly — history-dependence may emerge; we discuss these possibilities.

Notice that the richness condition (which drives uniqueness) is orthogonal to the span condition (which evaluates disequalization versus equalization). The two can vary quite independently. We return to these matters in some detail in the main body of the paper.

Our final set of results examines what happens when neither the span nor the richness condition is met. We know then that there exists a steady state with perfect equality. But there are other steady states with inequality. We are now squarely in the realm of the neutrality models. Markets may be equalizing or disequalizing, depending on initial conditions. The analysis of non-steady state dynamics is then necessary to identify the exact manner in which initial conditions affect the eventual outcome.

This is not the first paper to conduct a study of dynamics in the two-occupation model; see Section 7 for references. At the same time, as explained in that Section, our analysis is different both in content and implications. In particular, it reveals patterns intermediate to those stressed by the neutrality and disequalization literatures. For instance, initial equality of wealth does not always ensure the perpetuation of equality thereafter. For instance, if a country starts perfectly equal yet sufficiently poor, then the fortunes of different families will diverge and the economy converges to an unequal steady state. This echoes a similar theme in our analysis of the span condition. On the other hand if the country is sufficiently rich to start with, then equality is maintained thereafter. In intermediate cases, both the forces of disequalization and equalization co-exist. In summary, with a sparse set of occupations, both the average level of wealth and its inequality matter for the balance between the disequalization and equalization forces of the market.

[3]This observation has very different implications for policy. If there is a unique steady state (especially one involving inequality), then it paints a somewhat more depressing picture of the world: apart from direct and ongoing policies designed to shift the *parameters* of the model, certain inequalities will keep recurring. The view of policy as a shifter of initial conditions, which then drive the economy of its own accord into a salubrious steady state, is then no longer valid — unless occupational span and richness are themselves products of history, which, of course, they may well be.

The paper is organized as follows. Section 2 describes the core idea of the model by contrasting occupational versus financial bequests. Section 3 describes the model. Section 4 develops the span condition and its implications. Section 5 describes the uniqueness result when occupational structure is rich. Section 6 discusses other sources of multiplicity. Section 7 studies dynamics for the two-occupation model. Section 8 shows how the results can be extended to a a context where occupations are interpreted as firms of different scales and start-up costs. Section 9 concludes.

## 2. Occupational Structure and Disequalization

Economists have a simple shorthand for the study of occupational diversity, which is to reduce different qualifications and skills to aggregate quantities of "human capital". In other words, all human capital is — *even before we write down the definition of equilibrium for the society in question* — commonly expressible in some common efficiency unit.[4] This approach is summarized by Becker and Tomes in their 1986 paper:

> "Although human capital takes many forms, including skills and abilities, personality, appearance, reputation and appropriate credentials, we further *simplify* by assuming that it is homogeneous and the same "stuff" in different families." (Becker-Tomes (1986, p.56), emphasis ours)

This may be a useful starting point, but as we shall argue, the implications are quite drastic. The crucial assumption is that the relative returns to different occupations are exogenous, so that the reduction to efficiency units can be carried out separately from the behavioral decisions made in the population. We spell out the implications in the following example, taken from Mookherjee and Ray (2003).

2.1. **An Example.** Suppose that aggregate production is a function of just two inputs: skilled and unskilled tasks, satisfying the Inada conditions in each. There are just two occupations: skilled and unskilled labor. The latter can only do the unskilled tasks. To bring up a skilled offspring involves a fixed cost. Assume that this is the only way a parent can leave a "bequest"; i.e., there are no financial bequests.

Now observe that even if all parents in the economy have identical wealth and preferences, *they cannot all make the same bequests*. The reason is simple. If every parent leaves their child unskilled, the return to skilled labor will be enormously high, prompting investment in that category. At the same time, if all children are skilled, the return on *unskilled tasks* will become high, killing off the investment motive.[5] Hence even if all families start equal in generation 0, in the next generation their fortunes must separate. These two outcomes are utility-equivalent for generation 0, but they are not utility-equivalent for generation 1! Furthermore, in succeeding generations wealthier parents will have a greater incentive to train their children for $S$, so that the "primitive inequality" that sets in at the first generation will be reinforced: children of skilled parents will be more likely to acquire skills themselves. In any steady state, there is persistent inequality.

---

[4]So, for instance the investment choice in Loury's model is interpreted as a choice of "how much" education to acquire: there is no formal difference between human and physical capital. The so-called endogenous growth models (see, e.g., Lucas (1988)) continue, by and large, to retain this shorthand.

[5]If skilled labor cannot perform unskilled tasks, the unskilled wage will become very high by the Inada conditions. But even if they can perform unskilled tasks, this will equalize the two wage rates. Either interpretation has the same outcome.

2.2. **Discussion.** It is clear from the preceding argument that the divisibility or otherwise of investment opportunities is immaterial: the argument applies irrespective of how many occupations there are. What *is* needed is some minimal degree of occupational diversity (at least two essential occupations must entail distinct training costs). This is precisely the condition that fails when efficiency units are used to simplify the problem: every occupation can be "transformed" into every other with no change in relative prices.

Thus the endogeneity of relative prices is central to our argument. "Human capital" represents a *collection* of different occupations which supply distinct labor inputs to the production process, inputs that are *not* perfect substitutes for one another. Inter-occupational earning differences are then endogenously determined by relative supplies of people in different occupations, among other things. Such a formulation is supported by considerable empirical evidence.[6]

Why doesn't the same argument apply to machines as well as human occupations? The reason is simple: if the return from physical capital is *alienable* — so that shares in "capital" can be divisibly held — everyone can hold a "portfolio of machines" and derive the very same rate of return on it. In contrast, barring societies with slavery, the return on human occupations is not alienable.[7]

These remarks suggest, however, that the correct dividing line to be drawn is not between "physical" and "human" bequests, but bequests that result in endowments that are alienable (e.g. money) and endowments that are not (e.g. occupations). The latter may well include transfers of physical assets such as a family business which is not incorporated — perhaps for reasons of moral hazard or simply the lack of development of a deep stock market. These transfers are no different from human bequests in their implications for disequalization, and we include them in the category of "occupational bequests".[8] The remaining, alienable component we shall refer to as "financial bequests".

This discussion naturally provokes the following question: if financial bequests are *not* subject to the sort of reasoning in the example, does that example then fail when financial and occupational bequests are permitted to coexist? We assumed there that the only way for parents to transfer wealth to their children is through investment in their education. If financial bequests could supplement educational expenditures, they could be used to offset the inequality induced by differences in educational investments. For instance, if all parents started with equal wealth, they would have similar preferences over their children's future wealth *vis-a-vis* their own consumption. They could simultaneously sort into different occupations while ensuring their children all end up with the same wealth — those selecting less skilled occupations for their children could compensate them via higher financial bequests. The need to ensure occupational diversity would then no longer necessitate inequality in wealth.[9]

This is precisely the issue that motivates a more general exploration of the issues. To anticipate what is to come: occupational bequests will be associated with disequalization, financial bequests with equalization. The relative importance of these two components will then determine

---

[6]See e.g., Katz and Murphy (1992) for responsiveness of US skill premia to relative supply of skilled workers.

[7]It could still be divisible in principle, if an individual could hold a mix of "part time" occupations in any proportion she desired. To a certain extent we allow for this — after all, our space of occupations is abstract — but the point is that if every conceivable convex combination were allowed the minimal diversity condition must fail.

[8]In Section 8, we discuss these issues in more detail.

[9]Formally, an "occupation" is now to be interpreted as a pair: the first entry being financial holdings, the second being the occupation in the standard usage of that word. The diversity condition will now have to be imposed on these "occupations", and it may well fail.

which view of the market is to be the pertinent one. This is our conceptual point of entry into a wider discussion.

## 3. Model

**3.1. Technology.** There is a single aggregate output, one physical capital good,[10] and a compact set of occupations $\mathcal{H}$. Output $y$ *includes* undepreciated capital stock. The technology is described by an quasiconcave CRS production function $y = f(k, \boldsymbol{\lambda})$, where $k$ is physical capital and $\boldsymbol{\lambda}$ is an occupational distribution — some finite measure — on $\mathcal{H}$.

There is an exogenous, continuous training cost $x(h)$ for occupation $h$, denominated in units of the consumption good.[11] For expositional simplicity, assume that there exists an occupation with zero training cost.

**3.2. Remark on Capital Markets in Production.** As set up, the model accommodates two extreme cases concerning capital mobility (and the intermediate options as well). The first is that of a *closed economy*, in which the variable $k$ represents physical capital used in the production sector, and supplied entirely by the savings of households in the economy. Here the rate of return on capital will be endogenously determined.

Alternatively, there may be an international capital market (for production) with perfect mobility of capital, in which case the interest rate is exogenously given. A formally equivalent way of describing the production function would then be as $f(k, \lambda) = (1 + \bar{r})k + g(\boldsymbol{\lambda}, \bar{r})$, where $\bar{r}$ is the fixed interest rate, and $g$ is a reduced-form production function from which the capital input has been "maximized out".

**3.3. Factor Prices and Minimal Diversity.** We study a perfectly competitive economy. Normalize the price of final output to 1. Let $\mathbf{w} \equiv \{w(h)\}$ denote the wage function, $r$ the interest rate, and $\mathbf{p} \equiv (r, \mathbf{w})$ the factor price function. Say that $\mathbf{p}$ is a *supporting price* if there exists a strictly positive, profit-maximizing output associated with it. Equivalently, the unit cost of production, given $\mathbf{p}$, is precisely the price of output (unity).

A particular class of factor prices will play a key role in the paper. Say that $\mathbf{w}$ is *r-linear* if for every pair of occupations $h$ and $h'$, $w(h) - w(h') = (1 + r)[x(h) - x(h')]$. An $r$-linear wage function yields a rate of return on occupational investments exactly equal to $r$.

An important assumption of this paper is that not all occupations are perfect substitutes for one another. The strongest form of this statement is that all occupations are always in positive demand under every supporting price, but in the first part of the paper we assume something weaker. Specifically, say that the technology exhibits *minimal diversity* if for every $r$-linear supporting price the support of the profit-maximizing input combination contains at least two occupations with distinct training cost, one of which is an occupation with the maximal training cost.

The restriction to $r$-linear prices and the requirement that only two (or more) inputs be in demand make this assumption quite weak. Nevertheless, the requirement that at least one of them have the highest training cost is arguably more restrictive. In the appendix we show why matters are more complicated when this restriction is dropped, but the point is that it *can* be dropped: the appendix provides an extended characterization that parallels Proposition 1.

---

[10] As long as capital goods are alienable and shares in them can be divisibly held, having capital goods makes no difference to the analysis.

[11] As in Mookherjee and Ray (2003), this may be generalized to allow training costs to depend on the pattern of wages. We conjecture that the results will continue to hold.

If the set of occupations is finite and the production function satisfies Inada conditions in occupations then minimal diversity is trivially satisfied: *all* occupations are demanded at every supporting price. On the other hand the presumption that all occupational skills can be reduced to common efficiency units — being really an assumption about perfect substitutes — fails minimal diversity.

3.4. **Households.** There is a continuum of families in $[0,1]$, each represented by a single member in each generation $t = 0, 1, 2, \ldots$ At the start of generation $t$, a typical agent receives a financial bequest $b_t$ (to be augmented by the interest rate) and an occupation $h_t$ from her parent. At prices $\mathbf{p}_t = (r_t, \mathbf{w}_t)$, her implied wealth is $W_t \equiv (1 + r_t)b_t + w_t(h_t)$.[12]

The agent (correctly) anticipates factor prices $\mathbf{p}_{t+1} \equiv (r_{t+1}, \mathbf{w}_{t+1})$ for the next generation $t + 1$, and selects $b_{t+1}, h_{t+1}$ to maximize

$$U\left(W_t - x(h_{t+1}) - b_{t+1}\right) + V\left((1 + r_{t+1})b_{t+1} + w_{t+1}(h_{t+1})\right) \tag{1}$$

subject to the borrowing constraint $b_{t+1} \geq 0$. We assume that $U$ and $V$ are smooth, increasing and strictly concave.

The bequest motive implicit in (1) is more sophisticated than "warm-glow" (in that parents do not simply care about the bequest *per se* but rather what it does for the children), but is not "dynastic" (in that parents do not maximize some nonpaternalistic value function).[13] A special case that we occasionally invoke is one in which $U$ and $V$ both exhibit the same constant elasticity: $U(c) = (c^{1-\sigma} - 1)/(1 - \sigma)$ with $\sigma > 0$, and $V \equiv \delta U$, where $\delta \in (0, 1)$ represents the extent of parental altruism.

The condition $b \geq 0$ is a fundamental restriction stating that children cannot be held responsible for debts incurred by their parents. In this model, the capital market may be active but it only involves lending by households (those making financial bequests) to owners of firms producing the final output. Section 8 will show how the model can be extended to the context where households can also borrow (subject to credit limits).

3.5. **Equilibrium.** A *competitive equilibrium* given historical wealth distribution $\mathbf{W}_0 \equiv \{W_0(i)\}_{i \in [0,1]}$ is a sequence of factor prices $\mathbf{p}_t = (r_t, \mathbf{w}_t), t = 0, 1, 2, \ldots$ and associated $(k_t, \boldsymbol{\lambda}_t)$, $t = 0, 1, 2, \ldots$, such that:

(a) each family $i$ selects bequests and occupations optimally, given its inheritance and market prices;

(b) these decisions aggregate to $(k_t, \boldsymbol{\lambda}_t)$ at each $t$, and

(c) $p_t$ supports profit maximization using the inputs $(k_t, \boldsymbol{\lambda}_t)$, for every $t$.

---

[12] We use $r_t$ to augment the bequest for generation $t$, so consumption occurs at the end of the period. Any other consistent formulation works just as well.

[13] We could easily accommodate dynastic preferences with no essential difference to the main results.

Abstracting from uncertainty, this framework nests several models in the literature; in particular Mookherjee and Ray (2003) and Becker and Tomes (1979, 1986). Our model reduces to Becker-Tomes in the absence of occupational bequests, [14] and to a variant of our earlier model in the absence of financial bequests.[15]

3.6. **Steady States.** A *steady state* is a stationary competitive equilibrium: $(r_t, \mathbf{w}_t) = (r, \mathbf{w})$ and $(k_t, \boldsymbol{\lambda}_t) = (k, \boldsymbol{\lambda})$ for all $t$.

The following observation is useful (its proof, based on a single-crossing argument, is relegated to the appendix):

OBSERVATION **1.** *The wealth of (almost) every dynasty is stationary in any steady state.*

A cross-dynasty wealth distribution is therefore well-defined in each steady state. In what follows, an *equal steady state* will refer to a steady state with a degenerate wealth distribution; all other steady states will be called *unequal*.

## 4. THE SPAN CONDITION FOR DISEQUALIZATION

In this section, we develop the basic span condition that characterizes steady-state inequality. For necessary preliminaries, we need to retreat temporarily to a special case in which a nontrivial occupational structure is missing. [This case is not of intrinsic interest to us, but it serves as a useful intermediate step.]

4.1. **Financial Bequests Alone.** Suppose everyone earns an exogenously fixed wage $\omega$. There are only financial bequests at fixed interest rate $r$. Then a parent with wealth $W$ selects $b \geq 0$ to maximize $U(W - b) + V(\omega + (1+r)b)$. Let the resulting wealth of the child be denoted $W' \equiv \omega + (1+r)b$. It is evident that $W'(W; r, \omega)$ is increasing in $W$. We impose the following restriction on bequest behavior:

*Limited Persistence*. For any $r$ and $\omega$, if $W'(W; r, \omega) \leq W$ then $W'(\hat{W}; r, \omega) < \hat{W}$ for all $\hat{W} > W$.

This condition implies that future wealth as a function of current wealth crosses the $45^0$ line once and only once. Therefore limited persistence is a weaker form of a restriction imposed by Becker and Tomes (1979): $\frac{\partial W'}{\partial W} \in (0, 1)$, a condition which has widespread empirical support. Apart from empirical plausibility, we impose limited persistence to maintain comparability between different models.[16]

Under limited persistence, intergenerational wealth must converge from any starting point to a unique limit, which depends on $\omega$ and $r$: call it $\Omega(w, r)$. It is very easy to compute this function in specific models (see Section 4.4 for an example).

---

[14] This is not to say that Becker-Tomes don't have such bequests — they do — but the efficiency units assumption creates a mathematical equivalence of such bequests.

[15] The main difference from our earlier model is in the nature of the bequest motive: we assume here — in line with Becker-Tomes — that parents care only about the wealth of their children, apart from their own consumption. Our earlier formulation had nonpaternalistic "dynastic" preferences.

[16] The reason we use the weaker version of the condition is simply because the rate of interest may be endogenous, and it affects $\partial W'/\partial W$, so the stronger restriction would be an undesirable one on endogenous variables. For instance, in the isoelastic example mentioned in the text, it is easy to see that our limited persistence condition is always satisfied, whereas without exogenous restrictions on $r$ the stronger condition on derivatives will generally fail.

4.2. **Span.** Now return to the general model. As we've already remarked, $r$-linear wage functions will play a crucial role here. It is easy to see that for any $r$, there is at most one $r$-linear wage function such that $(r, \mathbf{w})$ is an $r$-linear supporting price. If there *is* one such function, say that $r$ is *allowable*.

To understand allowability, imagine that the production function is Inada on all factor inputs, including physical capital. Then *every* value of $r$ is allowable. If, on the other hand, there is global capital mobility at a given rate of return so that domestic production is effectively linear in capital, then — as is reasonable — there is only one allowable value of $r$: that given by the global rate of return.

Given some allowable $r$ and associated $r$-linear wage function, let $\bar{w}(r)$ and $\underline{w}(r)$ denote the highest and lowest values of that function on the set of occupations..

Finally, for any allowable $r$ (and its associated $r$-linear wage function), let $\mathcal{W}(r)$ denote the set of profit-maximizing national wealths, when the scale of output is such as to equate the total demand for labor across all types to the supply (normalized to one). Specifically, for any given scale of output, calculate the total demand for labor at the given factor prices, aggregating across all occupations. Then increase or decrease the scale by some multiple to ensure that the total labor demand equals one. Add all the labor incomes at these factor prices. Then add in the value of domestically held capital (at the price $1 + r$). This is an element of the set $\mathcal{W}(r)$.

If there is no capital mobility and the production function is strictly quasiconcave in all inputs, $\mathcal{W}(r)$ will contain a single element: total wealth is pinned down uniquely. On the other hand, if there is international capital mobility, $\mathcal{W}(r)$ consists of all numbers at least as great as labor income.[17]

It is important to note that these definitions have nothing to do with *equilibria*. Values such as $\bar{w}(r)$, $\underline{w}(r)$ and $\mathcal{W}(r)$, as indeed the various values of $r$ that are allowable, *are entirely technological concepts*. In principle, these can all be calculated simply by knowing the production and training cost functions.

PROPOSITION **1.** *Assume that the production function exhibits minimal diversity. Then the following statements are equivalent.*

1. *There exists a steady state with perfect equality.*

2. *There is an allowable r such that* $\bar{w}(r) \leq \Omega(\underline{w}(r), r) \in \mathcal{W}(r)$.

*In the special case in which perfect international capital mobility exists at rate* $\bar{r}$*, the second restriction in statement 2 can be dropped: an equal steady state exists if and only if* $\Omega(\underline{w}(\bar{r}), \bar{r}) \geq \bar{w}(\bar{r})$.

If we consider for a moment the special case of international capital mobility, and reverse the statements in the proposition, we see that *every* steady state must involve inequality — i.e., disequalization "wins" — if and only if

$$\Omega(\underline{w}(\bar{r}), \bar{r}) < \bar{w}(\bar{r}). \tag{2}$$

For reasons to become clear soon, we will refer to (2) as the *wide span condition*, where "span" refers to the range of training costs over different occupations. By the same token, the opposite inequality (which is equivalent to the existence of at least one equal steady state) will be referred to as the *narrow span* condition.

[17] Capital held domestically can vary from 0 to $\infty$ and bears no necessary relation to the amount of capital used in domestic production.

In the next section, we turn to a discussion (and proof) of Proposition 1. This part of the paper may be omitted on a first reading. However, the arguments presented here are not simply technical; they do aid in understanding the model.

4.3. **Analysis of the Span Condition.** Say that an occupation is *active* in a steady state if it lies in the support of occupations which experience positive demand. In any steady state, we must have

$$w(h) - w(h') \geq (1+r)[x(h) - x(h')] \tag{3}$$

for every $h$ and $h'$ such that $x(h) > x(h')$ and $h$ is active.

For if this were false, no parent would wish to "supply" the occupational slot $h$: they could do better by educating their children for $h'$ instead and supplementing the remainder with financial bequests.[18]

To prove that statement 1 in Proposition 1 implies statement 2, begin with the presumption that there exists a steady state with perfect equality.

OBSERVATION **2.** *If occupations $h$ and $h'$ are both active in a perfectly equal steady state,* (3) *must actually hold with equality for these occupations.*

*Proof.* Let $h$ and $h'$ both be active. Some pair of equally wealthy parents must split between the occupations for their children in order to supply each slot. Since their bequest preferences are identical, it must be the case that these parents are indifferent between $h$ and $h'$: those selecting the lower-cost occupation must compensate their children via higher financial bequests that exactly offset the earning differences between the two occupations. Hence the implicit rate of return across $h'$ and $h$ generated by the wage function must exactly equal $r$, the rate of return on financial bequests. ■

This shows why the $r$-linearity of the wage function may be related to equal steady states, though Observation 2 by no means settles this. [After all, all occupations aren't necessarily active at the steady state.] Yet the following assertion is generally true:

OBSERVATION **3.** *Consider any equal steady state $(r, \hat{\mathbf{w}})$. Then assuming minimal diversity, there is an equivalent steady state $(r, \mathbf{w})$ with exactly the same interest rate, occupational distribution, capital stock and payoffs (but possibly different wage function $\mathbf{w}$) such that*

$$w(h) - w(h') = (1+r)[x(h) - x(h')] \tag{4}$$

*for every $h$ and $h'$, and such that $w(h)$ coincides with $\hat{w}(h)$ at every active occupation $h$.*

*Proof.* By minimal diversity, an occupation $\bar{h}$ with the highest training cost ($x(\bar{h})$) must be occupied. Define $\underline{w}(r) \equiv \hat{w}(\bar{h}) - x(\bar{h})(1+r)$, and consider the unique $r$-linear wage function emanating from $\underline{w}$. [It "begins' at $\underline{w}$ and "ends" at $\hat{w}(\bar{h})$.]

For any occupation $h$, it must be that $w(h) \geq \hat{w}(h)$. Otherwise parents choosing the occupation with the highest training cost would do better to select occupation $h$ and then bequeath the remaining amount financially: a higher offspring wealth can the be achieved with the same total investment.

So the $r$-linear wage function $\mathbf{w}$ lies weakly above the actual wage function $\hat{\mathbf{w}}$. Moreover, by Observation 2, $\mathbf{w}$ coincides with $\hat{\mathbf{w}}$ for all *active* occupations. This shows that the profit-maximizing

[18]In this and all the arguments to follow in this section, entirely standard modifications need to be made to handle the cases in which there is a continuum of inputs. An "active occupation" need not have positive demand for it, but appropriate limiting arguments allow us to replicate all the assertions in the mai n text.

pattern of factor demands remains unaltered when $\hat{\mathbf{w}}$ is replaced by $\mathbf{w}$, and in particular that $r$ is allowable, with $(r, \mathbf{w})$ a supporting price.

To complete the argument, we show that the supply of occupations can be taken to be unchanged as a result of this wage adjustment. By minimal diversity, it follows that there is another active occupation, with training cost distinct from $x(\bar{h})$ — call it $h'$. Then $\hat{w}(h') = w(h')$. It follows that our "wage adjustment" makes no difference to feasible possibilities for all bequests between $x(h')$ and $x(\bar{h})$: this range could be "spanned" in any case by adopting occupation $h'$ and then using financial bequests. Therefore the only way in which the wage adjustment could make a difference to choices is if $\Omega(\underline{w}, r) < w(h')$. But then, by strict convexity of preferences, it is easy to see that a person choosing occupation $\bar{h}$ could not, in fact, be at a steady state relative to $\hat{\mathbf{w}}$: she would be running down descendant wealth relative to her own.[19] ■

The appendix contains a discussion of the role played by minimal diversity in establishing Observation 3.

The main argument can now be completed. The occupation $\bar{h}$ is demanded, so it has to be supplied, both under $\hat{\mathbf{w}}$ and (by Observation 3) under $\mathbf{w}$. This implies that limit wealth must be at least as large as $\bar{w}(r)$, which proves the first part of statement 2.

But more is needed. Domestically generated wealth will have to equal the "supply of wealth" generated through production, so in addition, we must have $\Omega(\underline{w}(r), r) \in \mathcal{W}(r)$.

In the special case in which perfect international capital mobility exists at rate $\bar{r}$, this last requirement is automatically met. By varying the amount of capital held domestically from 0 to $\infty$, $\mathcal{W}(\bar{r})$ can be seen to be the set of all values above national labor income. But the latter is certainly bounded above by $\bar{w}(\bar{r})$, so the first part of statement 2 automatically implies the second: $\Omega(\underline{w}(\bar{r}), \bar{r}) \in \mathcal{W}(\bar{r})$.

The sufficiency argument has no need for the critical intermediate step in Observation 3, and is consequently more straightfoward. Assume statement 2, and consider the supporting prices $(r, \mathbf{w})$ given there. Because $\mathbf{w}$ is $r$-linear, households facing these prices will have access to a linear investment frontier, just as in the simple Becker-Tomes world. All households will therefore converge to the steady state wealth $\Omega(\underline{w}(r), r)$. By the first part of statement 2, this quantity is at least as high as the wage earnings of the most expensive active occupation, so parents (at the limit wealth) are cetainly comfortable with choosing such an occupation for their children. But this means, in turn, that parents are comfortable with choosing *any* occupation for their children that is in positive demand at the price vector $(r, \mathbf{w})$. So all occupational demands, appropriately scaled so that aggregate labor demand equals unity, can be met willingly by supply.

To complete the construction of an equal steady state, we have to ensure that the capital market clears as well. This is the role of the second part of statement 2. We know that $\Omega(\underline{w}(r), r) \in \mathcal{W}(r)$, so we can always choose some value of national wealth (compatible with profit maximization) that equals desired limit wealth. Because the labor market has been cleared by construction, this takes care of the capital market. The proof of Proposition 1 is complete.

4.4. **Widespan: Interpretation and Applications.** It is easiest to interpret the span condition when there is perfect international mobility of capital. This fully pins down the rate of return to capital. In a later section we explore the implications of an endogenous rate of return.

To translate statement 2 in Proposition 1 into a more explicit assertion about span, consider the isoelastic example introduced earlier: $U(c) = (c^{1-\sigma} - 1)/(1 - \sigma)$ with $\sigma > 0$, and $V \equiv \delta U$, with

---

[19] This argument relies strongly on the existence of another active occupation, so that an individual with wealth greater than or equal to $\bar{w}$ can "locally" run down wealth along a line with rate of return $r$.

$\delta \in (0,1)$. Define $\rho \equiv [\delta(1+r)]^{1/\sigma}$. It is easy to calculate intergenerational wealth movements in the simple model with financial bequests alone:

$$W' = \frac{(1+r)\rho}{1+\rho+r}W + \frac{\rho}{1+\rho+r}\omega$$

if $W \geq \frac{\omega}{\rho}$, and $W' = \omega$ otherwise. This allows us to calculate limit wealth very quickly:

$$\Omega(\omega, r) = \begin{cases} \omega & \text{if } \rho \leq 1, \\ \frac{\rho}{1-r(\rho-1)}\omega & \text{if } \rho \in (1, 1+\frac{1}{r}), \\ \infty & \text{if } \rho \geq 1+\frac{1}{r}). \end{cases}$$

If $\rho \leq 1$, then there are no (limiting) financial bequests in steady state in the Becker-Tomes world. In particular, $\Omega(\underline{w}(r), r) = \underline{w}(r) < \bar{w}(r)$, and the widespan condition must hold. Otherwise, if we impose the Becker-Tomes limited persistence assumption on the constant rate of return $r$, we have $\rho \in (1, 1+\frac{1}{r})$. Letting $X$ stand for the highest training cost associated with any occupation, the widespan condition reduces to

$$\frac{X}{\underline{w}} > \frac{\rho - 1}{1 - r(\rho - 1)} \tag{5}$$

where to ease notation, we write $\underline{w} = \underline{w}(r)$ given that $r$ is exogenous.

Inequality (5) makes clear that Proposition 1 involves the span of training costs; it states that the range $X$ (relative to the lowest wage under the $r$-linear wage function associated with $r$) should not be "too large", where the term "too large" depends on parameters, notably the strength of bequest motive and the interest rate.

This interpretation makes it very easy to consider different economic situations and the associated likelihood of widespan. Here are some examples.

4.4.1. *Low Discounting.* The effect of a lower $\delta$ is rather obvious. If for some reason a society discounts the future by less, the bequest motive will be dulled and the widespan condition is more likely to apply. (5) confirms this: $\delta$ is positively related to $\rho$, and $\rho$ is positively related to the right hand side of (5).

4.4.2. *Poverty.* Suppose that two economies have different degrees of total factor productivity, but no other parametric difference. [So the productivity of all factors is scaled uniformly in one economy relative to the other.] Then widespan will hold once TFP drops below some threshold level, and such poor economies must be prey to disequalization. To see this in terms of (5), notice that the $r$-linear wage function associated with $r$ must shift downwards as TFP declines, leading to a corresponding decline in $\underline{w}$ and therefore increasing the likelihood of widespan.

It should be noted that this argument presumes that training costs are impervious to changes in TFP. There are two contrasting possibilities. If the training technology also varies across the two economies in the same way as TFP, then the observation of the previous paragraph is reinforced. On the other hand, lowered wages in the less productive country will also reduce the wages for educators, and this will bring down educational costs measured in terms of final output. This effect works in the opposite direction. The net effect can be analyzed using a broader notion of technology which includes education as well as final production, as in Mookherjee and Ray (2003). It is easy to see that if material inputs as well as human inputs enter the educational training function, the overall costs of education must fall by less than wages, so the first-order effect identified above is not nullified.

4.4.3. *Higher Growth.* While poverty is conducive to disequalization, higher *growth* may be positively related to it as well. For instance, if growth (from Hicks-neutral technical progress) causes all wages and costs to grow at a uniform rate, then — all other things being equal — the level of desired bequests will be dulled, raising the likelihood of widespan and disequalization. We omit a formal demonstration of this assertion, which proceeds by deriving an equivalent of the widespan condition (5) in the presence of neutral technical progress. Steady state wealth would then grow at the very same rate, but the *level* of that trajectory would be depressed by higher growth. It is this level that needs to be compared to the (unchanged) level coefficient of the highest training cost, which also grows at the same rate in steady state. In such a situation disequalization would become more likely with higher growth.

To the extent that poorer countries grow faster owing to a "catch-up" phenomenon in technology, the wide span condition is therefore more likely to hold in such countries on two counts: poverty and growth. Of course, the net result is ambiguous if growth is not positively correlated with initial poverty.

4.4.4. *Domestic Reliance on Physical Capital.* Now let us compare economies with differing degrees of reliance on physical (or alienable) capital in production. One simple way to do this is to suppose that final output is produced via a nested function

$$y = Ak^{\alpha}m^{1-\alpha},$$

where $m$ is a composite of the occupational inputs; or equivalently, an abstract intermediate good "produced" by occupational inputs. To capture the effect of changing reliance on physical capital, simply alter $\alpha$. To this end, set the marginal product of capital to the given world interest rate $r$ to obtain

$$\frac{k}{m} = \left(\frac{A}{r}\right)^{\frac{1}{1-\alpha}},$$

so that the indirect "reduced-form" production function is linear in $m$:

$$y = Bm,$$

where

$$B = A\left(\frac{A}{r}\right)^{\frac{\alpha}{1-\alpha}}.$$

Notice that $B$ essentially prices the composite in terms of the final output. If $B$ goes down for some reason, then occupational wages must suffer, in general, and if the wage function is restricted to be $r$-linear, every wage, and particularly $\underline{w}$, must decline. So a reduction in $B$, other things being equal, must contribute to a greater likelihood of disequalization.

Whether $B$ goes up or down with $\alpha$ depends on the ratio of $A$ (domestic productivity) to $r$.[20] In relatively "unproductive" economies, in which $A$ is small, an increase in physical capital intensity must lower $B$ and so — by (5) — is even more conducive to inequality. On the other hand, "productive" economies in which $A$ is large will benefit (in terms of equality) from a greater reliance on physical capital in production.

This sort of analysis throws novel light on an old question in development: what are the implications of greater mechanization in production for the evolution of inequality?

[20] The intuitive reason why this is the appropriate comparison is that $A$ really becomes the productivity of capital in domestic production in the limiting case when $\alpha = 1$. The issue is whether this domestic productity exceeds or falls short of the world rate of return on capital, proxied by $r$.

4.4.5. *The Global Rate of Return on Capital.* A change in the rate of return to capital has subtle effects as well. Look again at (5). When $r$ rises, $\rho$ also goes up. Both these effects work against the wide span condition, by raising the rate of return to financial bequests. So a first cut would suggest that an increase in the global rate of return to physical capital tends to be equality-enhancing. However, one should be careful of other effects, such as the possibility that $\underline{w}$ may be lowered by the increase in $r$. This effect runs in the opposite direction.

4.4.6. *Globalization.* Wide spans may be the outcome of a sudden move to globalization. To see this, consider an economy under autarky with a certain range of goods and services comprising its national output. This range will be naturally associated with a corresponding range of occupational slots. Now suppose that restrictions on trade and the activity of foreign firms are lifted, and the economy "globalizes". It is likely that occupational span will significantly widen. This will be true in the literal sense, as professional categories (that were previously relatively rare) emerge: software technicians, telecommunications experts, high-end service providers. But it is also true in the wider sense that we've defined occupations — as unincorporated businesses with high setup costs, such as medical groups, business consultants, advertising agents, begin to emerge.

While we haven't described yet what steady states must look like in the presence of the wide-span condition, our later analysis suggests that such inequalities will be felt most strongly in the highest end of the income or wealth distribution. There is no definite prediction for the bulk of the remaining income distribution that span the preexisting set of occupations. Notice, however, that such inequalities are no less significant as their on-the-street visibility may be extremely high.[21]

Our discussion has been necessarily brief and somewhat informal. The idea has been to show that the widespan condition can be readily applied to different economic situations to assess the chances of market-driven disequalization. A detailed exploration of widespan in each of these situations is beyond the scope of the current paper.

## 5. The Role of History

That multiple steady states, with selection within that set depending on historical conditions, may explain differences across ex-ante identical societies is a powerful and provocative idea. As already discussed, the idea has decades-old roots in development economics, and has found recent revival in formal models of multiplicity. Historians such as Sokoloff and Engerman have taken up this viewpoint (see also Acemoglu, Johnson and Robinson (2003), Banerjee and Iyer (2005) and Pandey (2004)). In the words of Sokoloff and Engerman, scholars "have begun to explore the possibility that initial conditions, or factor endowments broadly conceived, could have had profound and enduring impacts on long-run paths of institutional and economic development . . . "

Now, it is important to distinguish between two kinds of history-dependence. There is, first, history-dependence at the level of the individual unit (a dynasty, in our model). It should be obvious that our model chronically displays such history-dependence. Consider, for instance, the example in Section 2.1, in which two dynasties with identical wealth mutually separate (and stay separated thereafter) as a consequence of symmetry-breaking. The initial allocation of

[21] There is some evidence that this sort of change in "high-end" inequality has been characteristic of the globalizing Indian economy since 1991; see Banerjee and Piketty (2004).

ex-ante identical households to different occupations is essentially random, but the effects are persistent and felt over generations.

Second, there is the question of history-dependence "in the large", in the sense that initial history permanently affects the path taken *by the entire economy*. These might result in differences in aggregates, such as per-capita income or wealth holdings, and/or in distributions, such as the extent of wealth inequality. Depending on the context (and particularly on how "individual units" are defined), history-dependence in the large is typically what we are interested in.

5.1. **Market Forces and History-Dependence.** If we restrict ourselves to "market sources" of history-dependence in the large, one such source has received dominant emphasis in the literature.[22] This has to do with the possibility that *factor prices* — and in consequence the personal distribution of wealth — can systematically vary across different steady states. Often, such variation is used as a classification scheme; e.g., Galor and Zeira (1993, p. 45–46) "define an economy as 'developed' if the equilibrium wage of unskilled workers . . . is high," and as "less-developed" otherwise. Ljungqvist (1993) "establishes a continuum of steady states depending on the distribution of human and nonhuman wealth," and describes "an underdeveloped country [as] characterized by a high ratio of unskilled workers in the labor force, a small stock of physical capital, a low gross national product, a high rate of return on human capital, and a corresponding large wage differential between skilled and unskilled workers." Banerjee and Newman (1993, Section 4.4) use the same kind of multiplicity to define "prosperity and stagnation".

The multiplicity of factor prices is easy enough to obtain in a two-factor model. Imagine that there are only two labor categories: a highly skilled "elite" and a low-skilled "proletariat". Then, if the elite base is very narrow to start with, then the returns to the proletariat will be low and the elite return will be very high. At the same time, the enormously high costs of transit to the elite class will prevent the members of the proletariat from assisting their progeny to achieve elite occupational status.[23] The very high elite income, in turn, permits the elite to maintain their priviled position: the costs of "training" for their elite progeny to remain elite can be readily borne by their parents. such an outcome displays all the properties described by Lungqvist and others.

But there are *other* steady states that are also self-reinforcing. Some of these involve a high proletarian wage and a relatively low elite income, though the latter must still exceed the former. The wage gap cannot be too small otherwise the elite must abandon their elite status. If elite professions are necessary for economic activity, then this cannot happen. in short, a diversity of steady states can be sustained by the system. *Which* steady state will come to be depends on initial inequalities.

So much for factor prices. One can also conceive of multiplicity as the emergence of diverse "modes of production". In one mode a particular technique of production is dominant, employing a particular subset of factors. In another a different technique is used, along with a different set of factors — perhaps overlapping to some degree in the former. Once again, imaginative terminology can be used for the purposes of interpretation and exposition. Banerjee and Newman (1993) call this the "endogeneity of economic institutions as part of the process of development." They show that an "economy might converge to a steady state in which there is (almost) only self-employment in small-scale production" (what they call "the cottage"), "alternatively, it may end up in a situation in which an active labor market and both large- and small-scale production

[22] There could be "non-market" sources of history dependence, such as the distribution of political power. These lie outside the purview of the present model, but we shall comment on them below.

[23] The assumption of missing capital markets for education is used here.

prevail. Which of the two types of production organization eventually predominates . . . depends on the initial distribution of wealth." To our knowledge, Banerjee and Newman (1993) are the only authors to study this aspect of the multiplicity question.

In what follows, we examine both these sources of history-dependence in the large.

5.2. **A Uniqueness Theorem.** In this section we introduce a new feature of equilibrium occupational structure. Say that a steady state is *rich* if the support of the training costs of employed occupations is an interval.

Several remarks are in order. First, richness is orthogonal to the span condition; it is perfectly compatible with either wide or narrow spans. [So the presence or absence of inequality is not at issue here.] Second, richness is an abstraction, necessitating in particular that there be a continuum of occupations. This is not to be taken literally, of course, but comments on approximation are needed and will be made later. Finally, richness implies there are no indivisibilities in the set of investment opportunities, but of course such a set may still be nonconvex: that is a matter for equilibrium prices to settle.

PROPOSITION **2.** *Fix the interest rate $r$. Then there cannot be two rich steady states with the same support of inputs and distinct wage functions.*

*Given the support, a steady state wage function can be fully described. The rate of return on human capital equals $r$ for occupations up to training cost $\theta \equiv [\Omega(\underline{w};r) - \underline{w}]/(1+r)$; families located in such occupations all have the same wealth $\Omega(\underline{w};r)$ and leave financial bequests provided $\Omega(\underline{w};r) > \underline{w}$.*

*If the wide span condition holds then there are occupations with training costs that exceed $\theta$. Over this latter range families make/receive no financial bequests, receive (unequal) wages satisfying the differential equation*

$$w'(x) = \frac{U'(w(x) - x)}{V'(w(x))}, \tag{6}$$

*and the marginal rate of return on human capital exceeds $r$ (for almost all $x$ in this range).*

*If $f$ is strictly quasiconcave in occupational inputs, the uniqueness of wage functions also implies the uniqueness of occupational structure.*

This proposition seriously challenges one branch of received wisdom on history-dependence. It suggests that such multiplicity may be an artifact of something that is usually thought of as an inncocuous simplifying assumption: that there are two (or a small number of) occupations. It is hardly necessary to point out that a rich array of occupations and training costs is a far better approximation to the truth. If a result is entirely overturned in that case, it is time to question that result.

That *individual* dynasties may be locked into outcomes that depend fundamentally on historical accident is something we do not question. What Proposition 2 shows is that the *overall distribution* of such outcomes may be pinned down uniquely.

There are several important implications of this result. First, it suggests the existence of policy externalities: an intervention for one or more individual "units" (people, or perhaps regions or countries) will be neutralized by opposite effects on other "units". This may be all right if the unit in question is all that the policymaker cares about (such as a country). It may be more worrisome in other cases.

A second implication is that with occupational richness, the span condition fully determines whether inequality is inevitable. With wide spans we already know this. With narrow spans there is the possibility that equal and unequal steady states coexist, but such multiplicity will be ruled out by richness. In brief, the endogenous inequality literature is vindicated in the wide-span case, and the exogenous inequality literature in the narrow-span case.

Other implications can be drawn from the predicted shape of the wage function. One that deserves particular emphasis is that inequality stems entirely from unequal occupational earnings. The marginal rate of return on occupational investments above $\theta$ is strictly higher than the interest rate, and these occupations drive all the observed inequality. Financial bequests arise (if at all) only at the bottom of the wealth distribution. When they do arise they serve an *equalizing* role: there is a wealth mass point at the very bottom, at $\Omega(\underline{w}; r)$.[24] The relative positioning of human and financial capital bequests in the Becker-Tomes (1986) is turned on its head: the latter is used only for small bequests, and human capital is used for large bequests.[25]

Is this blatantly counterfactual? We think not. First, it is well known from decomposition studies that earnings inequality accounts for most of overall income inequality. For instance, Fields (2004) summarizes observations from several studies, writing that that "labor income inequality is as important or more important than *all other income sources combined* in explaining total income inequality". Second, there is evidence that *within* the class of financial bequests, *intentional* bequests may not be that important. For instance, Gokhale *et al* (2001) argue that most financial bequests in the US economy are unintentional, the result of premature death and imperfect annuitization. In the iso-elastic example, this would correspond to the case with $\rho$ falls below unity. Our theory then predicts that there are no *intentional* bequests anywhere in the wealth distribution, so human capital differences entirely account for all inequality, perhaps supplemented by unintended financial bequests (which we do not formally model) at the high end.[26] Finally, it is possible to view large financial bequests observed at the top end of the distribution as a form of occupational investment by parents. Presumably the wealthiest sections of the population do not invest much in bank accounts or Treasury bonds, but instead manage these funds across various high-risk high-return investment opportunities that require large setup investments and access to private information networks. These high end investment activities may themselves be viewed as an occupation, to enter which an agent requires a large upfront investment.

5.3. **Why Uniqueness.** The detailed argument for Proposition 2 is provided in the Appendix. Here we provide a broad outline. The first step is to note that every family must have steady state wealth at least $\Omega(\underline{w})$, since all families always have at least the same investment opportunities as in the Becker-Tomes world. The second step is to note that if $\Omega(\underline{w}) > \underline{w}$ then the rate of return on active occupations with training cost upto $\theta$ must be $r$. This is an easy consequence of the first step.

The third step is to check that the wage function as described in the Proposition is indeed a steady state. The steady state is constructed on the assumption that those selecting occupations with training cost above $\theta$ will select no financial bequests at all, so the wage function must be such as to induce a parent with a given occupation to select the same occupation for its child. The requirement that the chosen occupation be optimal within a local neighborhood of occupations ties down the slope of the wage function as given in (6). Along with the boundary condition

[24]If the narrow span condition applies, this wealth is large enough to encompass the entire earnings distribution, and we obtain the equal steady state.

[25]In the Becker-Tomes model, small bequests are made entirely in the form of human capital, over a range where the rate of return on human capital exceeds that on financial capital. Owing to diminishing returns to human capital, the rate of return eventually falls to the rate of return on financial capital. Thereafter all additional bequests take the form of financial capital.

[26]For a model of unintended bequests arising from uncertain life span, see, e.g., Fuster (2000).

provided by the wage corresponding to training cost $\theta$, this differential equation pins down the entire wage function.[27]

It remains to check there cannot be any other steady state with the same interest rate. This argument is broken down into two steps. First, no other steady state starting with the same wage $\underline{w}$ for the least skilled occupation can arise. By the arguments above, any such steady state must coincide exactly upto $\theta$. Moreover if over some range of occupations above $\theta$ there are no financial bequests then the slope of the wage function must satisfy the same differential equation (6) over that range. And over any range where there are financial bequests the marginal rate of return has to equal $r$. Combining these two facts it is easy to infer that any such wage function — if distinct from the one constructed above — must lie below the latter for some occupations, and coincide for all others. This contradicts the requirement that the wage functions be consistent with profit maximization. Indeed, this argument establishes that all steady states must be of the form described in Proposition 2, with financial bequests upto $\theta$, and none above, and wage functions respectively $r$-linear upto $\theta$, and governed by the differential equation (6) thereafter.

The final step is to verify that there cannot be any other steady state wage function of this form starting with a *different* wage for the least skilled occupation. Profit maximization requires these functions must cross, and they cannot cross at any occupation where there are no financial bequests in both steady states (because they must follow the same differential equation (6) at any such occupation). Neither can they cross at an occupation at which the wage functions are both locally $r$-linear, since they are then locally parallel. So they must cross at an occupation in which the wage function is locally $r$-linear in one steady state, and not in the other. But this implies that the poorest families in the former steady state are investing more in total than families with the same level of wealth in the other steady state, despite obtaining a lower marginal rate of return.

This final step relies on the assumption in the proposition that all steady states have the same support of inputs. When this condition is dropped, we enter a world which may be interpreted as displaying various "modes of production."

5.4. **Multiple Modes of Production.** We now examine the possibility that multiple steady states may arise from the coexistence of different "modes of production". At one level it is hard to understand what such a concept may mean. After all, different methods or modes of production may simply be subsumed under the formal heading of a single production function, which may be viewed as the outer envelope of different modes. Proposition 2 applies equally to this case, as long as each "mode of production" uses, however infinitesimally, every one of the available inputs. Then two steady states must have the same support, and if they are rich, Proposition 2 applies to yield uniqueness.

It is, however, possible that different "modes of production" use entirely different or at best partially overlapping occupational inputs. For instance, Banerjee and Newman (1993), in interpreting their model, "identify self-employment with self-sufficient peasants and cottage industries, and entrepreurial production with large-scale capitalist agriculture and factory production." Then the multiplicity of steady states is once again a possibility. Let us discuss the conditions that are needed.

---

[27] It can subsequently be checked that the function as constructed provides a marginal rate of return on educational investments strictly above $r$ for (almost) all such occupations. In turn this implies that those families will not want to supplement their educational investments with financial bequests, so the wage function as constructed does constitute a full-blown steady state, even when financial bequests are permitted.

To fix ideas, suppose that occupations lie in $[0,1]$, with training cost given by $x(h) = \psi h$. Suppose further that the production function $f$ is "built from two modes of production," as follows. Mode $A$ exhibits a log-linear production function which uses only labor from the subinterval $[0,a]$, given by

$$\ln y = \ln A + \int_0^a \alpha(h) \ln \lambda(h) dh,$$

where $\int \alpha(h) dh = 1$ on $[0,a]$ and $A$ is a TFP parameter.

Mode $b$ is another log-linear production function which uses only labor from the subinterval $[b,1]$, given by

$$\ln y = \ln B + \int_b^1 \beta(h) \ln \lambda(h) dh,$$

where $\int \beta(h) dh = 1$ on $[b,1]$ and $B$ is a TFP parameter.

We suppose that $b < a$ to allow for overlapping occupations, though as will be clear from the exposition below this is not at all necessary for the argument.

In this example, physical capital has been "netted out" from production and there is a fixed rate of return $r$ on such capital.

Our production function $f$ is constructed by using any combination of these two technologies. It is very easy to check that $f$ satisfies all the assumptions of our model. Yet if $f$ is "unpacked" into its two modes, the way we have here, we see that barring nongeneric cases, *only one* of the two modes will actually be operative, the mode with the lower unit cost of production.

The calculation of unit costs for each mode, given some wage function $\mathbf{w}$ defined on $[0,1]$ is a completely standard exercise. For mode $A$ unit cost is given by

$$c_A(\mathbf{w}) \equiv \frac{1}{A} \exp\{\int_0^a \alpha(h) \ln \frac{w(h)}{\alpha(h)} dh\}, \tag{7}$$

while the corresponding expression for mode $B$ is

$$c_B(\mathbf{w}) \equiv \frac{1}{B} \exp\{\int_b^1 \beta(h) \ln \frac{w(h)}{\beta(h)} dh\}. \tag{8}$$

Now consider wage functions that satisfy the steady state conditions. Figure 1 will accompany the description here.

For ease of exposition, choose invidividual utility functions so that in the Becker-Tomes world, $\Omega(\omega, r) = \omega$ for all $\omega$. As we've seen, this can be guaranteed by taking, for instance, the isoelastic case with $\rho \leq 1$.

Now consider a wage function — call it $\mathbf{w}_A$ — that will support mode $A$ in steady state. Because all occupations along the interval $[0,a]$ must be willingly supplied, and because financial bequests have been shut down by construction, the same arguments used to establish (6) in Proposition 2 may be employed to see that

$$w_A'(h) = \frac{\psi U'(w_A(h) - \psi h)}{V'(w_A(h))} \text{ for } h \in [0,a], \tag{9}$$

where $\psi$ is the coefficient of training cost on occupational label. Choose the "intercept" $\underline{w}$ of this wage function so that the unit cost of production, $c_A(\mathbf{w}_A)$, is precisely equal to the normalized output price of unity.

This defines $\mathbf{w}_A$ on $[0,a]$. To define it on $[a,1]$ we have to be sure that no dynasty will want to acquire those occupations in the mode $A$ steady state. We also want to make sure that mode $B$

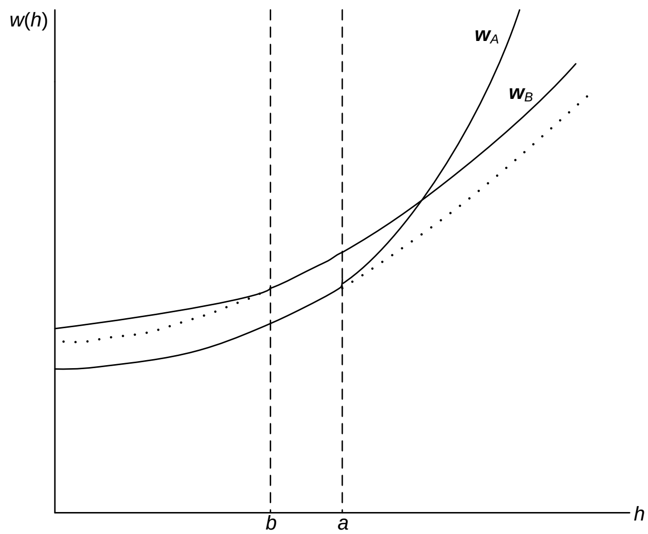


FIGURE 1. Two Steady-State Wage Functions

is relatively unprofitable, so let's set these wages as high as we can consistent with with the first requirement. This is done by defining $\mathbf{w}$ on $[a, 1]$ implicitly according to the equation

$$U(w_A(a) - \psi h) + V(w_A(h)) = U(w_A(a) - \psi a) + V(w_A(a)), \tag{10}$$

so that even the highest-wealth individual in mode $A$ will not want to "deviate" to occupational choices more expensive than $a$. Now $\mathbf{w}_A$ is fully defined on *all* of $[0, 1]$.

Figure 1 shows that these two different segments of the wage function have markedly different slopes. Beyond the point $a$, the function $\mathbf{w}_A$ lies above the what it would be had the differential equation in (9) been "continued" (the dotted line in the figure). This equation maintains individuals at their going occupations, while (10) is constructed to prevent an individual *at a* from moving to higher occupations. Given the presence of wealth effects, the wage function can afford to rise steeply beyond $a$ without pulling away individuals from their steady-state occupations between 0 and $a$. As we shall see, this feature is necessary for potential multiplicity.

To make sure that $\mathbf{w}_A$ is a steady state a critical condition needs to be satisfied:

$$1 = c_A(\mathbf{w}_A) \leq c_B(\mathbf{w}_A), \tag{11}$$

where $c_A$ and $c_B$ are given by (7) and (8).

Now we will attempt to find another wage function — call it $\mathbf{w}_B$ — that will support mode $B$ in steady state. Because all occupations between $b$ and 1 must be supplied, such a wage function *must* follow the differential equation analogous to (9) on $[b, 1]$:

$$w_B'(h) = \frac{\psi U'(w_B(h) - \psi h)}{V'(w_B(h))} \text{ for } h \in [b, 1]. \tag{12}$$

Choose the "intercept" $w(b)$ of this wage function so that the unit cost of production, $c_B(\mathbf{w}_B)$, is equal to the normalized output price of unity.

To define $\mathbf{w}_B$ on $[0, b]$ we have to be sure that no dynasty will want to acquire those occupations. Analogously to the argument above, we will want mode $A$ to be relatively unprofitable, so once again set these wages as high as possible, but with with the first requirement. This is done by defining $\mathbf{w}$ on $[0, b]$ implicitly according to the equation

$$U(w_B(h) - \psi h) + V(w_B(h)) = U(w_B(b) - \psi b) + V(w_B(b)). \tag{13}$$

Now even the lowest-wealth individual in mode $B$ will not want to "deviate" to such occupations. This defines $\mathbf{w}_B$ on all of $[0, 1]$.

Now comes the critical condition: for $\mathbf{w}_B$ to be a steady state, mode $B$ must outperform mode $A$ at these factor prices. We therefore require that

$$1 = c_B(\mathbf{w}_B) \leq c_A(\mathbf{w}_B). \tag{14}$$

It is possible to write down a specific example for which both (11) and (14) are satisfied.

## 6. Multiplicity via the Interest Rate

Another qualification to the statement of uniqueness is worth noting: it is conditional on the interest rate.

To better understand this phenomenon, assume that the set of training costs forms an interval $[0, X]$, and that every $x$ corresponds to an active occupation in steady state. Let $w(x, r)$ stand for the steady state wage function when the interest rate is exogenously given by $r$. With the richness assumption in place, we know that this function is uniquely defined. In the same vein, let $\lambda(x, r)$ be the density of labor demands generated by the production sector once profit is maximized at prices $(r, \{w(x, r)\})$ and *total* labor demand is restricted to add up to the population. And let $K(r)$ represent the demand for physical capital under the very same conditions.

Equilibrium requires (among other things) that the demand for physical capital equals the willingness to supply it, but the latter is just the total flow of financial bequests to the next generation. Recall that such bequests occur for occupations with wages that do not exceed $\Omega^*(r)$, the Becker-Tomes wealth generated from the lowest wage under $w(x, r)$ with return $r$. For any such occupation with training cost $x$, the financial bequest must be precisely

$$\frac{\Omega^*(r) - w(x)}{1 + r},$$

which, if added over all such occupations, yields a grand total of

$$\int_{x: w(x,r) \leq \Omega^*(r)} \left[ \frac{\Omega^*(r) - w(x, r)}{1 + r} \right] \lambda(x, r) dx.$$

The equality of this expression with $K(r)$ clears the capital market.

Using the fact that $w(x)$ must be $r$-linear in the relevant interval, we can simplify this requirement a bit. Define

$$M(r) \equiv \frac{\Omega^*(r) - \underline{w}(r)}{1 + r}.$$

Then it is easy to see that capital-market clearing reduces to the condition that

$$K(r) = \int_0^{\min\{X, M(r)\}} [M(r) - x] \lambda(x, r) dx. \tag{15}$$

Given that the other equilibrium conditions are already subsumed in the definitions of $w(x, r)$, $\lambda(x, r)$, and $K(r)$, (15) is a necessary *and* sufficient description of a steady state. It is an equation in a single variable, $r$.

There may well be multiple solutions to this equation. But this is a "typical" general equilibrium multiplicity that arises when the "indirect" effects of a change in $r$ outweigh the "direct" effects.

For instance, we expect that for a generic set of parameters the set of steady state interest rates will be finite, and each steady state interest rate and wage function will be locally isolated, so there will be no scope for the extreme hysteresis properties of two occupation models.[28] In this respect the potential multiplicity arising from the interest rate is qualitatively similar to the potential multiplicity of Walrasian equilibria.

On the other hand, multiple interest rates could affect the extent of borrowing constraints (when agents are allowed to borrow to some extent rather than not at all), with corresponding effects on efficiency and the wealth distribution. Piketty (1997) and Banerjee and Duflo (2004) emphasize this particular avenue to history dependence. In turn this gives rise to the question of how such multiplicities can survive when capital is free to flow across countries. We leave this as an interesting question for future research.

## 7. Occupational Sparseness and Dynamics

When occupational structure is not perfectly rich, multiple steady states can appear. To be sure, all the multiple steady states must lie in a very narrow region when the structure, while not perfectly rich, is sufficiently fine. However, the region expands as the structure becomes sparser. The case of two occupations represents a natural extreme point for the study of dynamics, which we now consider.

We are particularly interested in cases in which the narrow span condition is satisfied. We know then that an equal steady state always exists. But there are other — unequal — steady states as well. The objective of this section is to understand what sort of initial conditions lead us into the different steady states.

This is not the first exercise on dynamics in the occupational choice model; see, e.g., Ghatak and Jiang (2002) and Matsuyama (2000, 2003). But the setting is different on two central counts. First, we explictly consider two kinds of bequests, financial and occupational, and because we dispense with a warm-glow formulation, this distinction matters for the analysis. Second, we use end-point conditions on the production function that necessitate the active use of both occupations in equilibrium. As we shall see, this formulation allows for a variety of cases and yields new results. In particular, we shall show that our model yields disequalization for poor economies, *even if such economies start out perfectly equal*.

Consider, then, just two occupations, one skilled (S) which requires a training cost of $x$, and the other unskilled (U) which requires no training. We shall also assume that the interest rate is exogenously fixed at $r$. Denote the skill ratio by $\lambda$, and wages in the two occupations by $w_s(\lambda)$ and $w_u(\lambda)$. Under the usual Inada conditions, these are respectively decreasing and increasing functions, with $w_s(\lambda) \to \infty(0)$ and $w_u(\lambda) \to 0(\infty)$ as $\lambda \to 0(\infty)$. [Sometimes, when there is no confusion, we will simply drop the argument $\lambda$ to ease the notation.]

In this context multiple steady states arise quite naturally. For instance, if $x$ is small enough so that the narrow span condition is satisfied, then an equal steady state exists and is uniquely defined: it has a skill ratio $\lambda^*$, determined by the condition that $w_s(\lambda^*) - w_u(\lambda^*) = (1+r)x$. And

[28] We have not formally developed such genericity results. But note that small perturbations of the extent of parental altruism will shift the $\Omega^*(r)$ function, i.e, the supply of savings. For a given technology, the demand for capital from the production sector will be unaffected. Hence such perturbations will eliminate any instance of a continuum of steady state interest rates.

unequal steady states also exist, under the mild additional assumption that consumption is constrained to be nonnegative.[29] Our model is therefore capable of generating different parameter regions in some of which there is a unique equal steady state, others where only unequal steady states exist, and where both types co-exist. We focus only on the latter case in this section.

Define $\tilde{\lambda}$ by the condition that $\Omega(w_u(\tilde{\lambda})) = x$. Then every $\lambda \in (0, \tilde{\lambda})$ constitutes an unequal steady state.[30] Hence a continuum of unequal steady states can co-exist with the equal steady state. Moreover, within the set of unequal steady states, those with a lower skill ratio are associated with lower per capita income and higher inequality. Occupational sparseness can thus precipitate severe history-dependence.

We now explore the nature of such history dependence in more detail, by describing dynamics from arbitrary initial conditions. Particularly interesting is the case in which both equal and unequal steady states co-exist, a condition that we impose from now on. What determines whether the market is equalizing or disequalizing: does the economy converge (if at all) to the equal or an unequal steady state?

We begin by characterizing unequal steady states. Use $Z(W;\omega)$ to denote the indirect utility of a parent in a Becker-Tomes world with current wealth $W$ and flow earnings $\omega$ of the child (at the given interest rate, which we are suppressing in the notation).

PROPOSITION **3.** *$\lambda$ is an unequal steady state skill ratio if and only if*

$$\lambda < \lambda^* \tag{16}$$

*and in addition*

$$\max_{b \geq 0}[U(\Omega(w_u) - x - b) + V(w_s + b(1+r))] \leq Z(\Omega(w_u), w_u). \tag{17}$$

The argument is straightforward: if $\lambda = \lambda^*$ then we are effectively in a Becker-Tomes world with a linear investment frontier and a constant rate of return $r$ on all investments, where there cannot be any long run wealth inequality. So $\lambda < \lambda^*$ is necessary for wealth inequality. Unequal steady states must therefore involve a nonconvexity in investment returns: upto $x$ only financial bequests are possible, while at $x$ there is a discontinuous upward jump in investment returns (the size of which depends on the gap between the rate of return on human and financial capital). Since the return on human capital is higher, any parent wishing to invest $x$ or more must first invest in education, and make up the remainder with financial assets. Those investing less than $x$ will only invest in financial assets.

If $\lambda < \lambda^*$ then skilled families have wealth $\Omega(w_s) > \Omega(w_s^*)$. They will want to invest in their children's education because they want to invest at least $x$ if $\lambda = \lambda^*$; now they are even richer and the returns to investing in education are even higher.

Hence one only needs to check the incentives of the unskilled: they must not want to invest $x$ or more. This is the role of condition (17). In steady state unskilled households must have a wealth of $\Omega(w_u)$. The left side is the maximum payoff of a parent with this wealth conditional on investing at least $x$. The right side is what they attain with financial investments alone. This

[29] If, however, the Inada conditions on the production function do not hold, there can be a lower bound to unskilled wages, in which case unequal steady states may not exist if the training cost required to enter the skilled occupation is sufficiently low.

[30] In any such state, the unskilled wage is low enough that the corresponding Becker-Tomes steady state wealth falls below $x$, i.e., unskilled families cannot afford education at all. On the other hand, skilled families will want to invest in education: they are willing to do so even at the equal steady state skill ratio, and their incentive is still larger when the skill ratio is lower.

condition implies that the wealth of the unskilled falls below the skilled wage: $\Omega(w_u) < w_s$.[31] So there must be wealth inequality in these steady states.

Let $\Gamma$ denote the set of unequal steady state skill ratios. In general $\Gamma$ consists of a continuum of skill ratios, because it is characterized by a set of inequalities.[32] Turning finally to the dynamics:

PROPOSITION **4.** *Suppose there are two occupations, a fixed interest rate, and both the equal steady state with skill ratio $\lambda^*$ and a continuum of unequal steady states $\Gamma$ exist. Then from an arbitrary initial wealth distribution at date 0, the economy converges to a steady state. If the equilibrium skill ratio in the first generation exceeds $\bar{\lambda}$ (the highest skill ratio across all unequal steady states), the economy converges to the equal steady state. Otherwise it converges to an unequal steady state.*[33]

This proposition says that the dynamics depend on the historical wealth distribution, which determines the equilibrium skill ratio at the very beginning. Consider for instance the case where all families start with equal initial wealth.

**Corollary.** *Suppose all families start with the same wealth $W_0$. Then there exists a threshold $\bar{W}$ such that if $W_0 \leq \bar{W}$, the economy converges to an unequal steady state, while if $W_0 > \bar{W}$ it converges to the equal steady state.*

This result distinguishes our model from the previous "neutral to inequality" literature, in which initial equality always implies equality for ever thereafter. It also distinguishes it from previous "endogenous inequality" literature in which convergence (if it occurs) from any initial condition must be to an unequal steady state. We obtain a more nuanced theory which combines elements of both literatures. Societies that start perfectly equal converge to an unequal steady state if they are sufficiently poor initially; otherwise they converge to an equal steady state. Initial and eventual per capita wealth are then positively related, and poor countries do not eventually catch up with rich countries. The market plays a disequalizing role in poor countries, and an equalizing role in non-poor countries.

Initial wealth matters because it affects the (initial) incentive to invest in human capital, given the presence of borrowing constraints. Sufficiently poor countries cannot make the required initial investments in skill to boost the unskilled wage to a level that can initiate a virtuous upward spiral for unskilled families. Analogous to the endogenous inequality literature, in the initial generation all families have the same investment preferences. To ensure the supply of some skilled people in the following generation, symmetry must be broken: some families choose to invest and others don't, while they are all indifferent between investing and not. From the next generation onwards, this creates inequality which cannot be reversed: skilled families have higher wealth than unskilled families. Earnings inequality emerges and remains stationary. Over succeeding generations the wealth of the unskilled fall, while those of the skilled rise, so the operation of financial transfers exacerbates the perpetuation of inequality. Here the "endogenous inequality" forces prevail and are accentuated by the existence of financial capital.

---

[31] Otherwise unskilled parents would be investing at least $x$ in their children, in which case they would be better off educating them rather than provide only financial bequests.

[32] If $\lambda$ is an unequal steady state where the constraint (17) holds as a strict inequality, then an open neighborhood of it also satisfies (17) as a strict inequality.

[33] In the former case the skill proportion rises. In the latter case, the dynamics are as follows. If the date 0 skill ratio is an unequal steady state skill ratio then it converges to that steady state and the equilibrium skill ratio is stationary. If the date 0 skill ratio is not an unequal steady state skill ratio, then it rises over time and converges to the smallest steady state skill ratio lying above it.

On the other hand if an economy starts perfectly equal and sufficiently rich, then $\lambda = \lambda^*$ at the very first generation itself. Symmetry is broken as usual, with some households investing in human capital and others not, while all are indifferent. Those not investing in human capital invest in financial capital instead. In contrast to the economies which start poor, here $\lambda = \lambda^*$ in the first generation, ensuring that the rates of return on both kinds of capital are equalized. This implies that the composition of investments between the two forms of capital does not matter, and perfect wealth equalization must obtain in the next generation as well. The same logic repeats itself thereafter, with perfect wealth equality in every succeeding generation, with financial transfers perfectly offsetting differences in educational investments throughout. This is exactly the logic of perpetuation of equality in the "neutral to inequality" literature: because financial transfers are adequately available, the forces of "endogenous inequality" never have an opportunity to appear.

There is a third case intermediate between the previous two, where both forces of "equalization" and "disequalization" appear, and the former prevail. If the economy starts rich but not so rich that $\lambda$ in the first generation is less than $\lambda^*$, wealth inequality emerges initially. But the inequality is "moderate": unskilled wages are high enough so that parents leave financial bequests to their children, causing the wealth of the unskilled to grow. In turn this causes the demand for education (and hence $\lambda$) to grow, raising unskilled wages even further, and lowering the wage gap between skilled and unskilled. The wealth of the unskilled rises faster than that of the skilled, resulting in convergence. Here financial bequests induce "trickle down" and the market is equalizing.

What about the more general case where the initial wealth distribution is non-degenerate? Then initial inequality also matters: even for a country with high initial per capita wealth. If this wealth is distributed sufficiently unequally the equilibrium skill ratio at the beginning can fall below $\bar{\lambda}$, causing the economy to converge to an unequal steady state. This, of course, is familiar from preexisting literature.

In sum, we obtain a novel connection between initial poverty and the prevalence of disequalization forces in the market. But more generally the results that emerge with investment sparseness is that historical conditions matter for the long run, and the market can be equalizing or disequalizing depending on favorable or unfavorable initial conditions. This is very much along the lines of the "neutral to inequality" literature. What is different is the particular set of initial conditions for inequality to emerge and persist, which can now include cases of perfect equality as well.

## 8. Capital Markets with Frictions and the Size Distribution of Firms

So far, we've used the term "occupations" to describe a variety of economic activities that are inalienable. Chief among them, of course, are occupations — jobs or professions — in the traditional sense of the term. The remarks in this section (and the accompanying appendix) explicitly address an alternative interpretation, in which "occupations" are firms of different sizes, the ownership of which cannot be fully diversified in a stock market. Thus our framework can incorporate models of distributional dynamics through the capital market rather than human capital (as in Aghion and Bolton (1997), Piketty (1997), Lloyd-Ellis and Bernhardt (2000), Matsuyama (2000, 2003) or Banerjee and Duflo (2004)).

As before, an individual starts life with wealth inherited from her parents, but now starts up one of several firms with varying fixed costs (interpret labor as a firm with zero fixed costs). The individual faces borrowing constraints depending on the size of inherited wealth, and so is subject to an upper bound on the size of her firm. Firms of different sizes produce intermediate goods that are generally imperfect substitutes for one another, with the case of perfect substitutes a limiting case. Intermediate goods combine to produce a single consumption good according to a production function. An assumption on the "essentiality" of different intermediate goods in the production of the final good is then a convenient way of ensuring diversity in firm size.

The model developed in preceding sections represented the capital market as a frictionless transaction between lender-households and borrower-firms producing final consumption goods. We now need to focus on frictions in the capital market. To do so, we simplify by assuming that production of the final good from intermediates does not require any physical capital at all. Physical capital is required only in the production of intermediate goods, and we focus on the capital market for these sectors.

The CRS production function for the final consumption good is thus given by $y = f(\boldsymbol{\lambda})$, where $\boldsymbol{\lambda}$ is a firm size distribution over the set of intermediate good sectors $\mathcal{H}$. We shall refer to a sector by its capital requirement $k$, so the "training cost" function is just $x(k) \equiv k$. Perfect competition in the final good sector determines returns $w(k)$ to the production of different intermediate goods: $w(k)$ depends on the firm size distribution $\lambda$ via marginal products.

An individual with inheritance $I$ is subject to a borrowing constraint: it can invest in a firm of size $k$ provided $k \leq \Gamma(I; r)$, where for each $r$, $G$ is increasing in $I$ with $\Gamma(I; r) \geq I$.[34] Because parents care for child wealth, one can suppose without loss of generality that a parent allocates her overall bequest $I$ between physical capital $k$ and financial assets $b$. [The child would make the same choice, given $I$.] Within the region specified by the borrowing constraint, $b$ can indeed be negative: for given $I$, the parent can bequeath any combination $(b, k)$ such that $b = I - k$ and $k \leq \Gamma(I; r)$.

With some minor modifications, our previous analysis applies largely unchanged. The Appendix contains more detail.

## 9. Conclusion

The aim of this paper is to provide insight into three seemingly different views of market-driven inequality: *equalization*, embodied in traditional theories of income distribution, *disequalization*, central to the recent endogenous inequality literature, and *neutrality*, that either can happen depending on historical conditions. The results draw attention to an aspect of technology that has received little attention in the literature: the span and richness of occupational structure, which we jointly refer to as *occupational diversity*.

We present three sets of results. In the first, we make precise the *span condition* that distinguishes equalization from disequalization in steady state. Roughly speaking, the condition compares the importance of financial bequests to the range of occupational structure (as measured by the diversity in training costs). To our knowledge, this is a new condition that not only unifies opposing views of market-driven inequality, but also provides insight into the endogenous emergence of unequal distributions, and its links to other economic variables such as poverty.

Second, we study a fundamental issue in development models — that of systemic history-dependence. In our model, the long-run fate of an individual dynasty is highly dependent on history. But what of the system as a whole, the overall *distribution* of occupations and wealth? The component of occupational diversity that matters here is occupational richness: whether "in between" any two occupations with distinct training cost there lies a third with intermediate training costs. This richness criterion is quite different from the span condition, and indeed is

[34] This borrowing limit is the only capital market imperfection; the borrowing and lending rates of interest are assumed to be the same. Several microfoundations for this kind of imperfection are available in the literature.

orthogonal to it. We prove that if occupational structure is rich in this sense and all occupations are active, there can only be one steady state. Systemic history-dependence is ruled out.

This is not to say that history-dependence is an impossibility. It is certainly possible if there is a sparse set of occupations, or if different sets of occupations are active in different steady states. We discuss these possibilities.

Indeed, the possibility of history-dependence with a sparse set of occupations, especially the case in which equal and unequal steady states coexist, motivates our third exercise: the study of equilibrium dynamics. We show that such dynamics do indeed converge, and reveal an interplay between the market forces of equalization and disequalization that depends both on initial level of per capita wealth as well as how unequally it is distributed.

Concerning the issue of history dependence at the macro-level, and related issues of convergence or divergence, the results here suggest that such history dependence requires *either* the presence of significant indivisibilities in human capital investment, *or* very different sets of occupations ("modes of production") in different steady states. The latter is only a possibility if certain sets of occupations fully shut down over steady states. As for the former, casual empiricism suggests that there is little evidence of major gaps in sets of investment opportunities: between most unskilled and skilled occupations one can think of many intermediate occupations. Nevertheless one may argue that there are some key nonconvexities in returns to education that resemble the effect of indivisibilities — e.g. the relative lack of a premium for partial completion of a college degree. In most developed countries with relatively little public subsidization of higher education, the choice of whether or not to go to college or graduate school seems to resemble a significant sparseness. Further thought needs to be devoted to the question whether models with indivisibilities are empirically relevant. And if the evidence suggests that they are not, then stories of macroeconomic history dependence will have to be based on political economy channels rather than pure market-based occupational choice mechanisms.

## REFERENCES


**Acemoglu, D., Robinson J.A. and S. Johnson** (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review* **91**, 1369–1401.

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

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

**Banerjee, A. and Newman, A.** (1993). "Occupational Choice and the Process of Development." *Journal of Political Economy* April 1993, 101(2), pp. 274–298.

**Banerjee A. and Duflo E.** (2004), "Growth Theory Through the Lens of Development Economics," mimeo, Department of Economics, MIT.

**Banerjee, A. and L. Iyer** (2005), "History, Institutions and Economic Performance: The Legacy of Colonial Land Tenure Systems in India," forthcoming, *American Economic Review*.

**Becker G. and Tomes, N.** (1979), "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87(6), 1153–89.

**Becker G. and Tomes, N.** (1986), "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, 4(3), S1–S39.

**Chatterjee, S.** (1994), "Transitional Dynamics and the Distribution of Wealth in a Neoclassical Growth Model," *Journal of Public Economics,* 54, 97–119.

**Fields, G.** (2004), "A Guide to Multisector Labor Market Models," mimeo., Cornell University.

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

**Fuster, L.** (2000), "Capital Accumulation in an Economy with Dynasties and Uncertain Lifetimes," *Review of Economic Dynamics* **3**, 650–674.

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

**Ghatak, M. and N. Jiang** (2002), "A Simple Model of Inequality, Occupational Choice and Development" *Journal of Development Economics* **69**, 205–226.

**Gokhale, J., Kotlikoff, L., Sefton, J. and Weale M.** (2001) "Simulating the transmission of wealth inequality via bequests", *Journal of Public Economics*, 79(1), 93-128.

**Katz, L. and Murphy, K.** (1992), "Change in Relative Wages, 1963-1987: Supply and Demand Factors", *The Quarterly Journal of Economics*, 107, 35-78.

**Lloyd-Ellis H. and Bernhardt D.** (2000), "Enterprise, Inequality and Economic Development," *Review of Economic Studies*, 67(1), 147–168.

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

**Loury, G.** (1981), "Intergenerational Transfers and the Distribution of Earnings." *Econometrica* July 1981, 49(4), pp. 843–867.

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

**Matsuyama, K.** (2000), "Endogenous Inequality", *Review of Economic Studies*, 67, 743–759.

**Matsuyama, K.** (2003), "On the Rise and Fall of Class Societies," mimeo, Department of Economics, Nnorthwestern University.

**Mookherjee, D. and D. Ray** (2002), "Is Equality Stable?," *American Economic Review* **92** (Papers and Proceedings), 253–259.

**Mookherjee, D. and D. Ray** (2003), "Persistent Inequality," *Review of Economic Studies*, 70(2), April 2003, 369-394.

**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, Boston University.

**Sokoloff, K. and S. Engerman** (2000). "History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World," *Journal of Economic Perspectives* **14**, 217–232.

# APPENDIX

**Proof of Observation 1.** Let $W_t$ denote the wealth of a typical member of generation $t$. We claim that if $W_{t+1} < W_t$, $W_{s+1} \leq W_s$ for all $s > t$. Suppose not; then there exist dates $\tau$ and $s$ with $s > \tau$ such that (a)$W_{\tau+1} < W_\tau$, (b) $W_{s+1} > W_s$, and (c) $W_m = W_{\tau+1} = W_s$ for all intermediate dates $\tau + 1 \leq m \leq s$ (this last requirement is, of course, vacuous in case $s = \tau = 1$). But then a *strictly* higher wealth ($W_\tau$ compared to $W_s$) generates a strictly lower descendant wealth ($W_{\tau+1}$ compared to $W_{s+1}$), which contradicts a familiar "single-crossing" argument based on the strict concavity of the utility function. Therefore dynastic wealth converges in this case.

Next suppose that $W_{t+1} > W_t$. Then reversal of the same logic as above implies that family wealth is nondecreasing across generations. It cannot grow indefinitely for any positive measure of families since average wealth in the population is bounded. So it must converge for all but a measure zero of families. ■

**A Counterexample When Minimal Diversity fails.** The minimal diversity condition states that for all $r$-linear supporting prices, the support of the occupational input distribution must contain at least two occupations with distinct training costs, one of which has the highest training cost. If the condition fails in the sense that only one occupation is required, it is easy to construct counterexamples to Proposition 1, and we omit such a construction here.

More subtly, the same result might fail when no active occupation has the highest training cost. Here is an example.

Suppose that there are four occupational inputs,[35] and the production function is given by

$$f(\boldsymbol{\lambda}) = A\sqrt{\lambda_1\lambda_2} + B\sqrt{\lambda_3\lambda_4},$$

where $A < B$. Capital has been "netted out"; the interest rate is constant at some $r > 0$.

Think of occupations 1 and 2 as having similar (low) training cost, and occupations 3 and 4 as having similar (but high) training costs. Also think of $B$ as a large number, so the unique $r$-linear supporting wage vector has only occupations 3 and 4 demanded, while occupations 1 and 2 shut down. This will happen if $\mathbf{w}$, $A$ and $B$ have the property that

$$1 = \frac{B}{\sqrt{w_3 w_4}} > \frac{A}{\sqrt{w_1 w_2}}. \tag{18}$$

(The LHS is the reciprocal of the unit cost if occupations 3 and 4 alone are used, while the RHS is the reciprocal of the unit cost if occupations 1 and 2 alone are used.)

At these wages, take the generational utility function to be such that limit wealth is smaller than even $w_3$. [However, assume that it is bigger than $w_2$: this is used later.] So the widespan condition holds.

Yet we describe an equal steady state. Continue to lower the intercept of the wage function a bit more — to $\mathbf{w}'$ — so that in (18), we now have

$$\frac{B}{\sqrt{w_3' w_4'}} > \frac{A}{\sqrt{w_1' w_2'}} = 1. \tag{19}$$

Now the zero-profit condition can be (just) met using occupations 1 and 2, but of course this is not a supporting price because infinite profits can be made using occupations 3 and 4. But now artificially raise the wages of 3 and 4 even more — to $w_3''$ and $w_4''$ — so that

$$\frac{B}{\sqrt{w_3'' w_4''}} < \frac{A}{\sqrt{w_1' w_2'}} = 1. \tag{20}$$

Now only occupations 1 and 2 will be demanded at this wage vector, which isn't $r$-linear of course. [Notice that these wages are "locally $r$-linear" over occupations 1 and 2, though.]

Now in (18), if the inequality is not by much, the final wages $w_1'$, $w_2'$, $w_3''$ and $w_4''$ will not be that different from $\mathbf{w}$. So the limit wealth under these wages will still be bigger than $w_2'$ but smaller than $w_3''$. We therefore have a steady state with perfect equality, even though the span condition fails.

[35] We believe that as long as strict concavity in each input is assumed, at least four occupations are needed for a counterexample.

**A Generalization of Proposition 1.** In light of the example above, here is a somewhat more involved theorem when minimal diversity is weakened to

*Weak Minimal Diversity* (WMD). At least two occupations with distinct training costs are demanded at any supporting price vector with the property that $w(h) - w(h') = (1+r)[x(h) - x(h')]$ among all occupations demanded at this price vector.

WMD is strictly weaker than minimal diversity. The example above satisfies WMD but not minimal diversity.

Let $\mu(\omega, r)$ be the steady state payoff to an agent in the Becker-Tomes world with interest rate $r$ and flow earnings $\omega$. Say that $\mathbf{w}$ is a *quasilinear wage function associated with* $r$ if for some $\underline{w} > 0$,

$$w(h) = \begin{cases} \underline{w} + (1+r)x(h) & \text{if } \underline{w} + (1+r)x(h) \leq \Omega(\underline{w}, r), \\ V^{-1}\left(\mu(\underline{w}, r) - U(\Omega(\underline{w}, r) - x(h))\right) & \text{otherwise.} \end{cases} \tag{21}$$

This wage function is linear (with slope $1+r$) upto the Becker-Tomes limit wealth $\Omega(\underline{w}, r)$, and coincides with the indifference curve of the parent with that limit wealth for larger bequests. So for any occupation $h$ with training cost exceeding the Becker-Tomes limit bequest, it is the wage which just deters agents in the Becker-Tomes steady state from deviating to occupation $h$.

Now say that $r$ is *allowable* if $\mathbf{p} = (r, \mathbf{w})$ is a supporting price and $\mathbf{w}$ is a quasilinear wage function associated with $r$. Indeed, there will be a unique such associated function (compatible with supportability); call this *the* quasilinear wage function associated with $r$. For any allowable $r$, define $\mathcal{W}(r)$, the set of profit-maximizing national wealths under the supporting price system $(r, \mathbf{w})$ where $\mathbf{w}$ is the associated wage function. Define $\underline{w}(r)$ to be the lowest wage under $\mathbf{w}$, the quasilinear wage function associated with $r$. Finally, define $\bar{w}(\mathbf{p})$ to be the highest wage paid out to occupations in the profit-maximizing set at the supporting price $\mathbf{p}$.[36]

PROPOSITION **5.** *Assume that the production function exhibits WMD. Then the following statements are equivalent.*

1. *There exists a steady state with perfect equality.*

2. *There is an allowable $r$ such that $\bar{w}(\mathbf{w}, r) \leq \Omega(\underline{w}(r), r) \in \mathcal{W}(r)$, where $\mathbf{w}$ is the quasilinear wage function associated with $r$.*

*Proof.* First show that [2] implies [1]. Pick allowable $r$ such that [2] is satisfied. Pick the associated wage function $\mathbf{w}$. Allocate households across occupations according to the pattern demanded by the production sector, and assign them a common wealth of $\Omega(\underline{w}, r)$. This is a steady state, since by construction no household nor firm would seek an occupation with training cost exceeding $b(\underline{w}, r)$, and households are selecting their investments optimally on the segment of the linear Becker-Tomes investment frontier with investments not exceeding $b(\underline{w}, r)$. The condition $\Omega(\underline{w}(r), r) \in \mathcal{W}(r)$ guarantees capital market clearance.

Now we show that [1] implies [2]. Suppose there is an equal steady state with wage function $\hat{\mathbf{w}}$ and rate of interest $r$. Let $h'$ denote the "highest" occupation demanded in steady state.[37] Define $\underline{w} \equiv \hat{w}(h') - (1+r)x(h')$. Consider the new wage function $\mathbf{w}'$ given by

$$w'(h) = \begin{cases} \underline{w} + (1+r)x(h) & \text{if } h \text{ is lower than } h', \\ \hat{w}(h) & \text{otherwise.} \end{cases} \tag{22}$$

Using exactly the same arguments as in the proof of Proposition 1, we may conclude that $\hat{w}(h) \geq w'(h)$ for all occupations $h$, with equality holding for active occupations. So $\mathbf{w}'$ along with $r$ also constitutes a supporting price vector, and gives rise to the same demand for occupations from the production sector.[38] Condition WMD then implies that there must be an occupation $h''$ lower than $h'$ which also has positive demand.

Given [1], we know that everyone has the same steady state wealth under $(r, \hat{\mathbf{w}})$. By exactly the same arguments used in the proof of Proposition 1, we can show that this implies $\Omega(\underline{w}, r) \geq w'(h')$.

To complete the proof, consider a quasilinear wage function associated with $r$ — call it $\mathbf{w}$ — starting from the same initial wage $\underline{w}$ under $\mathbf{w}'$. Because $\Omega(\underline{w}, r) \geq w'(h')$, it must be that $w(h) = w'(h)$ for every $h$ lower than $h'$. It must also

[36] The set of profit-maximizing occupations at any support price $\mathbf{p}$ will be generically unique. However, if the production function is not strictly quasiconcave in occupational inputs uniqueness may be lost for some supporting prices. In that case simply take a support with the lowest (infimum) upper bound on active occupations, and define $\bar{w}(\mathbf{p})$ to be the wage associated with that bound.

[37] Say occupation $h'$ is *higher* than $h''$ if $x(h') > x(h'')$.

[38] In the light of footnote 36, $h'$ may need to be redefined downwards in case $\mathbf{w}'$ yields more than one set of input demands; this makes no difference to the analysis.

be true that $w(h) \geq w'(h)$ for every $h$ higher than $h'$, otherwise (by construction of quasilinearity) households would prefer to switch to such occupations rather than leave their children with wealth $\Omega(\underline{w}, r)$. Therefore $\mathbf{w}$ along with $r$ must also constitute a set of supporting prices, so $\mathbf{w}$ is *the* quasilinear wage function associated with $r$ and $r$ is allowable.

Because demands and supplies are unchanged, the wage $w'(h')$ must also be the highest active wage under $\mathbf{w}$. So we have shown that $\Omega(\underline{w}, r) \geq \bar{w}(r)$. That $\Omega(\underline{w}, r)$ must also belong to $\mathcal{W}(r)$ follows from exactly the same argument as in the proof of Proposition 1. ■

**Proof of Proposition 2.** We divide the proof into a number of steps. Throughout we fix the interest rate at $r$ and suppress it in the notation. Also we shall refer to a BT investment locus as a linear investment frontier with constant rate of return of $r$, $\underline{w}$ as the wage of the least skilled profession, and $\Omega(\underline{w})$ as the corresponding steady state wealth.

*Step 1: given $\underline{w}$ every steady state wage function must be r-linear upto an investment of $\theta \equiv \frac{\Omega(\underline{w})-\underline{w}}{1+r}$.*

To prove this note first that every family must have steady state wealth at least $\Omega(\underline{w})$, since they all have access to the linear bequest technology. Hence all those families selecting occupations with training costs below $\theta$ must be indifferent between investing in human capital and financial assets. This implies that the rate of return on human capital must be $r$ upto $\theta$.

*Step 2: There exists a wage function of the form described in Proposition 2 which constitutes a steady state.*

This is proven through a sequence of Lemmas.

*Lemma 1.* Suppose that the wage function over the range $x \geq \theta$ is described by the differential equation (6), with $w(\theta) = \Omega(\underline{w})$. Then for any $x \in (\theta, \mathbf{x}]$ consider a parent with wage $w(x)$ who is choosing a level of investment $x' \in (\theta, \mathbf{x}]$ for its child, and financial bequests are not allowed. It is optimal for such a parent to select $x' = x$.

Let $M(x, x') \equiv U(w(x) - x') + V(w(x'))$ denote the expected utility of such a parent who selects $x'$. And define $N(x) \equiv M(x, x)$. Then $N$ is differentiable by construction and (6) implies that $N'(x) = U'(w(x) - x)w'(x)$, so

$$N(x) = \int_{\theta}^{x} U'(w(a) - a)w'(a)da. \tag{23}$$

Now take any $x' \in (\theta, x)$: we shall show that $N(x) \equiv M(x, x) \geq M(x, x')$. Then

$$\begin{aligned}
M(x,x) &= M(x',x') + \int_{x'}^{x} U'(w(a) - a)w'(a)da \\
&\geq M(x',x') + \int_{x'}^{x} U'(w(a) - x')w'(a)da \\
&= M(x',x') + U(w(x) - x') - U(w(x') - x') \\
&= M(x,x'),
\end{aligned}$$

where the step involving the inequality uses the concavity of $U$ and the fact that $x' < x$. A similar argument in the reverse direction shows the same result with respect to any $x' \in (x, \mathbf{x}]$. This proves Lemma 1.

*Lemma 2.* $w'(x) > 1 + r$ for (almost) all $x \in (\theta, \mathbf{x}]$.

To prove this, note first there cannot be any $\epsilon > 0$ such that $w(x) - w(\theta) \leq (1 + r)(x - \theta)$ for all $x \in (\theta, \theta + \epsilon)$. The reason is that $w(\theta) = \Omega(\underline{w})$, the unique steady state wealth on the BT investment frontier corresponding to $\underline{w}$. If a parent with wealth $w(x)$ where $x \in (\theta, \theta + \epsilon)$ were confronted with this investment frontier, the child would attain a wealth strictly less than $w(x)$. Given a rate of return for investments beyond $\theta$ which is $r$ or worse, the child would be left with even less, i.e., the parent would select $x' < x$. This contradicts Lemma 1.

Next note that there cannot exist any interval $(x_1, x_2)$ with $x_2 > x_1 > \theta$ such that $w(x)$ is locally $r$-linear over this interval. Corresponding to the BT investment frontier that coincides with this section of $w(.)$, there can be at most one steady state wealth. But Lemma 1 shows that for every $x$ in this interval, a parent with wealth $w(x)$ leaves his child with the same wealth, and we obtain a contradiction.

Finally we show there cannot be any interval $(x_1, x_2)$ with $x_2 > x_1 > \theta$ such that $w'(x) < 1 + r$ for all $x$ in this interval. Suppose otherwise. By the argument two paragraphs above, there must exist points $x$ in $(\theta, x_1)$ where $w'(x) > 1 + r$. Since $w'(x)$ is continuous by construction, there must exist $x_3 < x_4 < x_5$ such that $w'(x)$ equals $1 + r$ at $x = x_4$, exceeds $1 + r$ over $(x_3, x_4)$, and is less than $1 + r$ over $(x_4, x_5)$. Consider the BT investment frontier passing through $(x_4, w(x_4)) \in \Re^2$. Since a parent with wealth $w(x_4)$ selects $x' = x_4$ in the problem described in Lemma 1, the marginal rate of substitution of that parent equals $1 + r$ at a bequest of $x_4$. Hence the wealth $w(x_4)$ must be the unique steady state wealth associated with this BT investment frontier. Now consider any $x'' \in (x_4, x_5)$. Confronted with the BT investment frontier passing through $(x_4, w(x_4))$, such a parent would have invested less than $x''$. With the marginal rate of return for investments beyond $x_4$ strictly less than $1 + r$, such a parent would invest even less. This contradicts the result of Lemma 1.

Hence there cannot exist any non-degenerate interval above $\theta$ where $w'(x) \leq 1 + r$, which establishes Lemma 2.

*Lemma 3.* Given the opportunity to supplement educational investments with a financial bequest, no parent with $x > \theta$ would prefer to do so.

This follows from the result of Lemma 2, which implies that the marginal rate of substitution for every parent in occupation $x > \theta$ is at least $1 + r$.

*Lemma 4.* Given the opportunity to invest less than $\theta$, no parent in occupation $x > \theta$ would prefer to do that instead of selecting the same occupation for its child.

This follows from the fact that any such parent is wealthier than those with wealth $w(\theta)$, and the latter prefer $\theta$ to any lower investment.

Combining the above set of Lemmas, it follows that the wage function constitutes a steady state, in the sense that it is globally optimal for every parent to choose the same occupation for its child, with only those in occupations below $\theta$ investing in financial assets, just enough to leave them with the same level of wealth as themselves. This completes Step 2.

*Step 3. There cannot be any other steady state wage function starting from the same $\underline{w}$ for the least skilled occupation.*

Suppose otherwise. By Step 1 such a wage function must coincide with the constructed steady state upto $\theta$. Next note that over any range of occupations $(x_1, x_2)$ above $\theta$ where parents select positive financial bequests, the slope of $w(x)$ must equal $1 + r$. And if these parents select zero financial bequests, the first-order condition for their occupational choice implies $w'(x)$ must satisfy (6) throughout this range. Hence if a steady state involves no financial bequests above $\theta$, it must coincide with the constructed steady state. A distinct steady state wage function must therefore involve positive bequests over some range above $\theta$, in which case the slope of the wage function over that range lies below that of the constructed steady state. That wage function must lie uniformly below the constructed steady state beyond some occupation. For if they ever crossed they must cross at an occupation that leaves no financial bequest, and at such a point the same differential equation (6) applies. We then obtain a contradiction of the property that two distinct wage functions must cross, in order to be consistent with profit maximization.

*Step 4. There cannot be any other steady state wage function.*

From the previous steps it follows that a steady state wage function must be of the type constructed: $r$-linear upto $\Omega(\underline{w})$, and following (6) thereafter. So two distinct steady state wage functions must be of this form, with differing wages for the least skilled occupation ($\underline{w}_1, \underline{w}_2$ say, with $\underline{w}_2 > \underline{w}_1$). Let these wage functions be denoted by $w_1(x), w_2(x)$ respectively, and the corresponding investments of the poorest households by $\theta_1, \theta_2$, where $\theta_2 > \theta_1$.

These two wage functions must cross. Clearly they cannot cross at an occupation $x$ at or above $\theta_2$, since such an occupation makes no financial bequest in either steady state, so the slope of both wage functions must be the same at that point. They also cannot cross at an occupation $x \leq \theta_1$ since both wage functions have slope $1 + r$ there. So they must cross at an occupation $x \in (\theta_1, \theta_2)$.

We can now find $x_1$ in a right neighborhood of $x$ with $w_1'(x_1) > 1 + r$, such that the wealth $W = w_1(x_1)$ is attained in the other steady state with an investment $x_2 \leq \theta_2$. Since we are to the right of the crossing $x$, the same wealth is attained with a lower investment: $x_1 < x_2$. This is inconsistent with utility maximization: if the household with wealth $W$ is investing optimally in the $w_2(.)$ steady state with a marginal rate of return equal to $r$, it is underinvesting in the latter given that it enjoys a higher marginal rate of return. This concludes the proof of Step 4 and hence of Proposition 2.

**Proof of Proposition 4.**

We begin by noting properties of competitive equilibrium at any given date.

LEMMA **1.** *Let the wealth distribution at the beginning of any date t be described by the cdf $F_t$. Then there is a unique competitive equilibrium at date t giving rise to skill ratio at $t+1$: $\lambda_{t+1} \leq \lambda^*$. If $\lambda_{t+1} < \lambda^*$ the equilibrium is characterized by a wealth threshold $W_t$ (satisfying $\lambda_{t+1} = 1 - F_t(W_t)$) such that an unskilled family with wealth at this threshold is indifferent between educating his child and not at date t:*

$$U(W_t - x) + V(w_s(\lambda_{t+1})) = Z(W_t, w_u(\lambda_{t+1})). \tag{24}$$

*Moreover, the competitive equilibrium has the following properties:*

(a) *$\lambda_{t+1} = \lambda^*$ if and only if $1 - F_t(W^*) \geq \lambda^*$ (where $W^*$ is defined by $I(W^*, w_u^*) = x$, with $I(W, \underline{w})$ denoting the optimal bequest of a parent with wealth W facing a BT investment frontier corresponding to $\underline{w}$).*

(b) *If $F_{t+1}(W_t) < F_t(W_t) = 1 - \lambda_{t+1}$, then $\lambda_{t+2} > \lambda_{t+1}$.*

(c) *If $F_{t+1}(W_t) \geq F_t(W_t) = 1 - \lambda_{t+1}$, then $\lambda_{t+2} = \lambda_{t+1}$.*

*Proof of Lemma 1:* Existence and uniqueness of competitive equilibrium skill ratio at any date with a given wealth distribution follows from the fact that we have a continuum economy, and the demand for education (i.e., for investment of at least $x$) is decreasing in the skill ratio anticipated for the following date. Specifically, if there are two competitive equilibrium skill ratios $\lambda, \lambda' > \lambda$, then there must exist a positive measure of households who invest when $\lambda_{t+1} = \lambda'$ but not when $\lambda_{t+1} = \lambda$. But the incentive to invest must be lower in the former case as the skill premium is lower — if these households weakly prefer to invest in education when $\lambda_{t+1} = \lambda'$, they must strictly prefer to do so when $\lambda_{t+1} = \lambda$.

At this equilibrium, investment incentives are ordered by wealth. Since both occupations are essential, there must be some households investing and others not investing. Since households differ only by wealth, there must exist a wealth threshold where a family is indifferent between investing and not. Such a household must be indifferent between just investing $x$ and investing some amount less than $x$ — otherwise it prefers to invest more than $x$ to investing exactly $x$, and such a household must strictly prefer to invest at least $x$ to any amount less than $x$. Hence the wealth threshold is characterized by (24).

Part (a) follows from the following argument. Households with wealth at least $W^*$ are willing to invest at least $x$ when $\lambda = \lambda^*$. If $1 - F_t(W^*) \geq \lambda^*$ then $\lambda^*$ is a competitive equilibrium skill ratio. For if we set $\lambda_{t+1} = \lambda^*$, the investment frontier coincides with a BT linear investment frontier corresponding to a constant flow earning of $w_u^*$. Then all those with wealth at least $W^*$ are willing to invest at least $x$, and are indifferent between education and financial bequests. So we can select $\lambda^*$ households from this group, and require them to invest in education, while requiring that none of the remaining households in the economy invest in education. Then each household will be choosing optimally and we have a competitive equilibrium. The converse is obvious.

To prove (b), suppose on the contrary $F_{t+1}(W_t) < F_t(W_t)$ and $\lambda_{t+2} \leq \lambda_{t+1}$. Then any household with wealth above $W_t$ strictly prefers to invest in education at date $t$, and must continue to do so at $t+1$. But there are more households with wealth above $W_t$ at $t+1$, so $\lambda_{t+2} > \lambda_{t+1}$, a contradiction.

Finally, (c) is proven as follows. Suppose $F_{t+1}(W_t) \geq F_t(W_t)$. Then we cannot have $\lambda_{t+2} > \lambda_{t+1}$, as this would imply that a household with wealth $W_t$ would strictly prefer not to invest in education at $t+1$, and there are more households poorer than $W_t$ at $t+1$ than $t$, which implies that $\lambda_{t+2} \leq \lambda_{t+1}$. On the other hand if $\lambda_{t+2} < \lambda_{t+1}$ then every skilled family will want to invest in education at $t+1$ (since they want to do so even at $\lambda^*$), and there are $\lambda_{t+1}$ skilled families at $t+1$, implying that $\lambda_{t+2} \geq \lambda_{t+1}$, a contradiction. This concludes the proof of Lemma 1.

Note that there is no monotone structure on the set of unequal steady states, because increases in $\lambda$ lower the cost of investing in education for the unskilled (as they become richer), and also the benefit of education. In general, therefore, $\Gamma$ is the union of intervals $[\lambda^i, \lambda^{i+1}]$, with $i = 0, 2, 4, ....$ Condition (17) is satisfied as an equality at each $\lambda^i$, as a strict inequality in every $\lambda$ in $(\lambda^i, \lambda^{i+1})$ with $i$ even, and is violated in every $\lambda$ in $(\lambda^i, \lambda^{i+1})$ with $i$ odd.

Before we proceed to the dynamics, we need the following notation. Let $W^i$ denote $\Omega(w_u(\lambda^i))$, the steady state wealth of the unskilled at the boundary unequal steady state $\lambda^i$.

LEMMA **2.** (a) *$W_t = W^i$ implies $\lambda_{t+1} = \lambda^i$.*

(b) *$W_t \in (W^i, W^{i+1})$ with i even implies $\lambda_t \in (\lambda^i, \lambda^{i+1})$ and $W_t \geq \Omega(w_u(\lambda_{t+1}))$.*

(c) *$W_t \in (W^i, W^{i+1})$ with i odd implies $\lambda_t \in (\lambda^i, \lambda^{i+1})$ and $W_t < \Omega(w_u(\lambda_{t+1}))$.*

*Proof of Lemma 2.* Recall condition (24) relating the threshold wealth $W_t$ with the competitive equilibrium skill ratio $\lambda_{t+1}$. Compare this with the relation between steady state wealth $W^i$ and skill ratio at any boundary (unequal) steady state skill ratio $\lambda^i$:

$$Z(W^i, w_u(\lambda^i)) = U(W^i - x) + V(w_s(\lambda^i)). \tag{25}$$

Part (a) follows from comparing these two conditions. For part (b), note that $W_t > W^i$ implies that anticipating the skill ratio $\lambda^i$ at $t+1$, the threshold wealth type $W_t$ would strictly prefer to invest in education, so $\lambda_{t+1} > \lambda^i$. Conversely this type would prefer not to invest in education anticipating a skill ratio of $\lambda^{i+1}$, so $\lambda_{t+1} < \lambda^{i+1}$. Hence $\lambda_{t+1}$ is an unequal steady state ratio, with

$$\begin{aligned} Z(\Omega(w_u(\lambda_{t+1})), w_u(\lambda_{t+1})) &\geq \max_{b \geq 0} [U(\Omega(w_u(\lambda_{t+1})) - x - b) + V(w_s(\lambda_{t+1}) + b(1+r))] \\ &\geq U(\Omega(w_u(\lambda_{t+1})) - x) + V(w_s(\lambda_{t+1}))]. \end{aligned}$$

Now compare with (24) to infer that $W_t \geq \Omega(w_u(\lambda_{t+1}))$.

Next turn to part (c). The same argument as for part (b) shows that $\lambda_{t+1} \in (\lambda^i, \lambda^{i+1})$. Since $i$ is now odd, $\lambda_{t+1}$ is not a (unequal) steady state skill ratio. Moreover $\lambda_{t+1} < \lambda^{i+1} < \lambda^*$. Proposition 3 now implies

$$Z(\Omega(w_u(\lambda_{t+1})), w_u(\lambda_{t+1})) < \max_{b \geq 0} [U(\Omega(w_u(\lambda_{t+1})) - x - b) + V(w_s(\lambda_{t+1}) + b(1+r))]. \tag{26}$$

We claim that this implies

$$Z(\Omega(w_u(\lambda_{t+1})), w_u(\lambda_{t+1})) < [U(\Omega(w_u(\lambda_{t+1})) - x) + V(w_s(\lambda_{t+1}))]. \tag{27}$$

Otherwise the maximum on the right hand side of (26) is attained at some positive $b$: a household with wealth $\Omega(w_u(\lambda_{t+1}))$ prefers to supplement educational investment with a positive bequest. The convexity of preferences then implies that a pure educational investment would in turn dominate any financial bequest less than $x$, contradicting the hypothesis. Finally the result that $W_t < \Omega(w_u(\lambda_{t+1}))$ follows upon comparing (24) with (27). This concludes the proof of Lemma 2.

To complete the proof of Proposition 4, consider first case (b), in which $\lambda_{t+1} \in (\lambda^i, \lambda^{i+1})$. Lemma 2 shows in this case that $W_t \geq \Omega(w_u(\lambda_{t+1}))$. Then all unskilled households at $t$ (i.e., whose parents had wealth below $W_t$ at $t$) will also have wealth below $W_t$ at $t+1$, The reason is that those with wealth between $W_t$ and $\Omega(w_u(\lambda_{t+1}))$ will leave less to their children. And those at or below $\Omega(w_u(\lambda_{t+1}))$ will leave more than they themselves inherited, yet their children's wealth cannot exceed $\Omega(w_u(\lambda_{t+1}))$. So the mass of the wealth distribution below $W_t$ is not smaller at $t+1$ than at $t$. Part (c) of Lemma 1 now implies that $\lambda_{t+2} = \lambda_{t+1}$. In turn this implies that $W_{t+2} = W_{t+1}$. So the same story applies at $t+1$ as at $t$. The equilibrium skill ratio will remain stationary at $\lambda_{t+1}$ for all $T \geq t+1$. Along this process, the wealth of the unskilled will converge to $\Omega(w_u(\lambda_{t+1}))$, while those of the skilled will converge to $\Omega(w_s(\lambda_{t+1}))$, so the economy converges to the unequal steady state associated with skill ratio $\lambda_{t+1}$.

Next consider case (c), in which $\lambda_t \in (\lambda^i, \lambda^{i+1})$ with $i$ odd and $W_t < \Omega(w_u(\lambda_{t+1}))$. Now the wealth of all unskilled households rises towards $\Omega(w_u(\lambda_{t+1}))$. It is still possible that $W_{t+1} = W_t$ and thus $\lambda_{t+1} = \lambda_t$, but if so their wealths will move even closer to $\Omega(w_u(\lambda_{t+1}))$. Eventually at some date $t+T$, we must have $F_{t+T}(W_t) < F_t(W_t)$. Then Lemma 1 implies that $\lambda_{t+T} > \lambda_t$. But comparing with condition (25) applied to $i+1$, it follows that $\lambda_{t+T} < \lambda^{i+1}$. Applying the same argument from $t+T$ onwards, it follows that the equilibrium skill ratio is a nondecreasing sequence bounded above by $\lambda^{i+1}$. So it must converge. It can only converge to a steady state skill ratio, which must therefore be $\lambda^{i+1}$.

Finally consider the case where $\lambda_{t+1} > \bar{\lambda}$, in which case $W_t > \Omega(w_u(\bar{\lambda}))$. Then also Lemma 2 implies $W_t < \Omega(w_u(\lambda_{t+1}))$. The same logic as in case (c) now implies the equilibrium skill ratio is a monotone sequence, bounded above by the equal steady state ratio $\lambda^*$. So it must converge, and to a steady state skill ratio. Since the only steady state skill ratio above $\lambda_{t+1}$ is $\lambda^*$, this is the skill ratio it must converge to. This completes the proof of Proposition 4.

**Occupations as Firms.** A parent of wealth $W$ that anticipates an interest rate $r$ and firm size distribution $\boldsymbol{\lambda}$ for the next generation then faces the following choice: select an asset portfolio $(b, k)$ for its child to maximize

$$U(W - b - k) + V((1+r)b + w(k; \lambda)) \tag{28}$$

subject to the constraint

$$b \geq \Gamma^{-1}(k; r) - k, \tag{29}$$

which is a restatement of the household optimization problem (1), with a more general version of the borrowing constraint.

The definition of a competitive equilibrium is slightly different, owing to the requirement of clearing of the capital market. It is a sequence $(r_t, \boldsymbol{\lambda}_t, w_t(k))$ such that: (i) a parent $i$ with wealth $W_t(i) \equiv (1+r_t)b_t(i) + w_t(k_t(i))$ derived from its asset portfolio $(b_t(i), k_t(i))$ selects a portfolio $(b_{t+1}(i), k_{t+1}(i))$ for its child to maximize (28) with $r = r_{t+1}$, $w = w_{t+1}$; (ii) $\boldsymbol{\lambda}_t$ is generated by distribution of $k_t(i)$, and $\int b_t(i)di = 0$; (iii) $w_t(k) = w(k; \boldsymbol{\lambda}_t)$.

It is straightforward to check that Proposition 1 then extends. So turn now to the conditions for an equal steady state. We shall continue to suppose *minimal diversity*, i.e., that the production function is such that at least two distinct intermediate goods are essential to produce the final good.

In what follows we use the following notation. Let $\boldsymbol{\lambda}^*(r; w(0))$ denote the cost-minimizing demand for intermediate inputs in the production of the final good, when the price of the $k$-sector-input is $w(k) = w(0) + (1+r)k$. Next, $\bar{k}(\boldsymbol{\lambda})$ denotes mean physical capital corresponding to distribution $\boldsymbol{\lambda}$: $\bar{k}(\boldsymbol{\lambda}) \equiv E[k|\boldsymbol{\lambda}] \equiv \int k d\lambda(k)$. And $k^*(\boldsymbol{\lambda})$ denotes the highest capital size active in the distribution $\boldsymbol{\lambda}$. $C^*$ will represent the minimum value of $E[k|\boldsymbol{\lambda}]$ across all distributions satisfying the production constraint $f(\boldsymbol{\lambda}) \geq 1$. The production viability condition PV will continue to apply as before; given PV an interest rate $r$ will be allowable if $r \leq \frac{1}{C^*} - 1$, and the corresponding return on the intermediate good sector with zero start-up cost will be $\underline{w}(r) = 1 - (1+r)C^*$. Let $w^*(k;r) \equiv \underline{w}(r) + (1+r)k$ be the associated $r$-linear return function, and $\boldsymbol{\lambda}^*(r) \equiv \boldsymbol{\lambda}^*(r, \underline{w}(r))$ the corresponding demand for intermediate goods. As before, $\Omega^*(r)$ denotes the steady state Becker-Tomes wealth corresponding to a linear investment frontier with interest rate $r$ and base earnings $\underline{w}(r)$, and $Y(r)$ the set of profit-maximizing outputs of the final good given full employment of all households in the economy.

PROPOSITION **6.** *An equal steady state exists if and only if there is an interest rate $r \leq \frac{1}{C^*} - 1$ such that: (i) $\Omega^*(r) \geq [1 - (1+r)\bar{k}(\boldsymbol{\lambda}^*(r))] + (1+r)\Gamma^{-1}(k^*(\boldsymbol{\lambda}^*(r);r))$; (ii) $\Omega^*(r) \in Y(r)$; (iii) $\Gamma(\bar{k}(\boldsymbol{\lambda}^*(r))) \geq k^*(\boldsymbol{\lambda}^*(r);r)$.*

Condition (i) is the new version of the narrow occupational span NOS condition. The form of the capital market clearing condition (ii) is unchanged. The new condition (iii) states that the borrowing constraint should not be too severe, relative to the span of capital sizes — it should permit households to invest in the most capital intensive sector. So it can be interpreted as a supplementary span condition, pertaining to the distribution of firm size relative to the capital market imperfection.

The proof of the Proposition is straightforward. In an equal steady state our previous arguments establish that $w^*(.;r)$ is a equilibrium price function for intermediate goods: profit maximization in the final good sector holds, and households are indifferent across different capital sizes, so they can be arrayed according to $\boldsymbol{\lambda}^*(r)$. Since they face a linear investment frontier, the common wealth of households must be $\Omega^*(r)$, and a household selecting capital size $k$ must hold financial wealth $b^*(k)$ where $(1+r)b^*(k) + w^*(k;r) = \Omega^*(r)$, i.e., $b^*(k) = \frac{\Omega^*(r) - w^*(k;r)}{1+r}$. The requirement that the capital market clears is that $\int b^*(k) d\lambda^*(k;r) = 0$, or $\Omega^*(r) = \int w^*(k;r) d\lambda^*(k;r) = y^*$, the output of the final good. This establishes (ii).

To obtain condition (i), the new version of NOS, note that those selecting larger firm scales will be borrowing from those selecting smaller scales, so the gross returns $w^*(k;r)$ must be adjusted for interest payments to yield net returns: $n^*(k,r) \equiv w^*(k,r) + (1+r)[\Gamma^{-1}(k;r) - k] = \underline{w}(r) + (1+r)\Gamma^{-1}(k;r)$. NOS requires that $\Omega^*(r) \geq n^*(k;r)$ for all active intermediate sectors. Since $\Gamma^{-1}(k;r)$ is increasing in $k$ for any $r$, NOS reduces to the condition imposed on the sector with the largest scale: $\Omega^*(r) \geq \underline{w}(r) + (1+r)\Gamma^{-1}(k^*(\boldsymbol{\lambda}^*(r);r))$. Using the fact that $C^* = \bar{k}(\boldsymbol{\lambda}^*(r))$, we get $\underline{w}(r) = 1 - (1+r)\bar{k}(\boldsymbol{\lambda}^*(r))$. Inserting this in the equation of the previous sentence, we obtain (i).

Finally condition (iii) follows from the borrowing constraint applied to each active capital size $k$: $b^*(k) \geq \Gamma^{-1}(k;r) - k$ or $\Gamma(b^*(k) + k; r) \geq k$. The left hand side is $\Gamma(I^*;r)$ where $I^* = b^*(k) + k$ is the constant wealth inherited by each agent, which in turn equals the mean capital size $\bar{k}(\lambda^*(r))$ (using the capital market clearing condition again). So (iii) obtains for all active capital sizes, and hence for the largest one.

The uniqueness result Proposition 2 also extends analogously when the firm size distribution is rich in the sense that the support of any technically feasible firm size distribution constitutes an interval. We discuss this informally. Conditional on the interest rate $r$, the steady state must take the following form. Upto capital size $\theta \equiv \frac{\Omega(w(0),r) - w(0)}{1+r}$ the rate of return on physical capital will equal $r$, the poorest households will have wealth $\Omega(w(0), r)$, and will hold nonnegative financial assets (i.e., lend) according to $b(k) = \frac{\Omega(w(0),r) - w(k)}{1+r}$. Richer households will hold no financial assets, will operate firms of size exceeding $\theta$, partly financed by borrowing. The rate of return over this range of capital size will exceed $r$. Every rich household will borrow the maximum amount allowed. The investment frontier will be $W(I;r)$ which will be linear at rate $r$ for investment or inheritance $I$ upto $\theta$, and above $\theta$ will be given by the differential equation $W'(I) = \frac{U'(W(I)-I)}{V'(W(I))}$.

The corresponding returns to physical capital $w(k;r)$ can be derived as follows. Note that $W(I;r)$ is the maximized value of $(1+r)I + \{w(k;r) - (1+r)k\}$ subject to $k \leq \Gamma(I;r)$. So conjecturing that the rate of return to capital in firms exceeds $(1+r)$ everywhere above $\theta$ (and equals $(1+r)$ below $\theta$), every non-poor household inheriting $I > \theta$ will select $k$ at the upper endpoint $\Gamma(I;r)$. This implies that above $\theta$:

$$W(I;r) = (1+r)I + [w(\Gamma(I;r)) - (1+r)\Gamma(I;r)]$$

or

$$w(\Gamma(I;r)) = W(I;r) - (1+r)[I - \Gamma(I;r)].$$

Hence for $k > \theta$:

$$w(k;r) = W(\Gamma^{-1}(k;r);r) - (1+r)[\Gamma^{-1}(k;r) - k]$$

from which we can indeed verify that the return to physical capital exceeds $r$:

$$w'(k;r) = (1+r) + [W'(\Gamma^{-1}(k;r);r) - (1+r)]\Gamma^{-1'}(k;r) > 1+r.$$

Inequality in this context stems from the higher returns to firms above size $\theta$, which in turn owes to the inability of poor households to enter those sectors owing to the nature of credit rationing. We obtain a more natural interpretation of inequality at the top end of the distribution, compared to the case of pure human capital. All bequests are 'financial' in this world, and the largest financial bequests are observed at the top rather than bottom end of the distribution. The right notion of 'occupation' thus seems to involve more than just human capital or acquisition of skill: it embraces also the wherewithal to enter sectors of the economy with large start-up costs.