Queen's Economics Department Working Paper No. 893

Enterprise, Inequality and Economic Development

Dan Bernhardt Huw Lloyd-Ellis

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

12-1993

Discussion Paper #893

Enterprise, Inequality and Economic Development

by

Dan Bernhardt
Queen's University

and

Huw Lloyd-Ellis
University of Toronto

December 1993

Enterprise, Inequality and Economic Development

by

Huw Lloyd-Ellis
Department of Economics
University of Toronto
Toronto, Ontario, Canada
M5S 1A1

Dan Bernhardt
Department of Economics
Queen's University
Kingston, Ontario, Canada
K7L 3N6

December 1993

Abstract

We develop a dynamic general equilibrium model of economic development with altruism in which the evolution of the extent of entrepreneurship, the rate of rural-urban migration, the scale and structure of production and the degree of income and wealth inequality are endogenously determined. The model generates a development process that has distributional characteristics consistent with those of the Kuznets hypothesis. In early stages of development, agents face binding financing constraints so that production is only carried out on a small scale. Few agents earn entrepreneurial profits and the lack of competition for labor keeps the wage low. The scale of production gradually expands as the descendants of entrepreneurs face less stringent financing constraints. Income and wealth inequality become increasingly acute as the entrepreneurial rich get richer and the poor remain so. Eventually, competition for workers drives up wages. As the labor share of income rises, the quality of entrepreneurs improves and the scale of production and the profits of an entrepreneur, controlling for wealth and ability, fall. Consequently, as the development process continues, income and wealth disparities eventually decline. With no engine for technological advancement, the economy converges to an invariant, non–degenerate wealth and income distribution. The economy may even be asymptotically efficient, as the 'right' entrepreneurs take on investment projects.

We derive the time path of the optimal redistribution policy and detail the impact of aggregate shocks at different stages. Finally, we provide insight into why the 'Kuznets curve' may appear in some economies but not in others, and how long–run fluctuations in economic activity can arise endogenously.

Keywords : Size-distribution, heterogeneity, altruism, wealth constraints.
JEL Classification Numbers: E0, O1, O4

This paper has benefited from comments by Jim Bergin, Charles Beach, Mick Devereux and Burton Hollifield and seminar participants at UBC, Toronto and Queen's. All remaining errors and omissions are our own. Funding from SSHRCC is gratefully acknowledged.

1. Introduction

What is the relationship between economic development and income inequality? How are the extent and scale of productive entrepreneurial activities determined over the development cycle? These and other questions are addressed here in the context of a dynamic general equilibrium dual economic development model with altruism which permits an analysis of the links between wealth and income inequality, entrepreneurship and economic development. The term 'economic development' distinguishes the analysis from a model of sustained growth and is used here to refer to the transition from a traditional rural-agrarian based economy to an advanced urban-manufacturing based economy.

In his classic analysis of the long-run growth patterns of contemporary developed countries, Kuznets (1955) suggests that in the early stages of economic development the extent of income inequality tends to worsen, while in later stages it improves. Paukert (1973), Ahluwalia (1976), Lydall (1979) and Williamson (1985) document these empirical regularities. Summers, Kravis and Heston (1984) offer similar evidence for a cross-section of countries between 1950 and 1980: most low income countries experienced rising inequality; middle income countries, little movement in income inequality; and high-income countries, steadily declining inequality. Figure 1 suggests an inverted U-shaped Kuznets' relationship between contemporaneous GNP per capita and income inequality.

Figure 1 goes here

Classical economists argued that high income inequality is a necessary condition for rapid economic growth (Lindert and Williamson 1985). If the rich invest much of their wealth, while the poor consume most of theirs, then an economy characterized by highly unequal distributions of wealth would save more and grow faster than one with a more equitable distribution. However, as the World Development Report (1991) indicates, there is little evidence of this relationship. The classical view implicitly assumes that the rich invest their wealth in the most productive activities. However, in a world of imperfect capital markets (e.g. Townsend 1979 or Newman 1991), informational asymmetries may prevent potentially efficient entrepreneurs from undertaking productive investments due to their lack of wealth. Greenwood and Jovanovic (1991) show how the development of a financial credit market can be an important source of growth and have a significant impact on the distributions of income and wealth. Blanchflower and Oswald (1990) find that the distribution of inherited wealth is a major factor in determining the level of entrepreneurial activity [1] . Evans

[1] Those who inherited £5000, for example, were approximately twice as likely to set up in business.

and Jovanovic (1989) document that entrepreneurs in the U.S. are capital constrained and Levy (1993) finds similar results for Tanzania and Sri Lanka. Since enterprise is generally recognized to be a key factor in the development of an economy (e.g. Schumpeter 1934), the extent of inequality will have important macroeconomic implications if capital markets fail to allocate resources optimally.

In this paper, we provide a theoretical basis for Kuznets' inverted U-shaped hypothesis in the context of a competitive economy with financial constraints and derive conditions which determine whether such a relationship arises. We determine how the distribution of wealth affects the extent and structure of economic activity which, in turn, is a determining factor in the distribution of income and, hence, the future distribution of wealth.

The framework for our analysis is related to that of Banerjee and Newman (1993). Agents live for a single period and optimally bequest a portion of their lifetime wealth to their children. Given his inheritance, each agent chooses between several occupations so as to maximize lifetime income. An agent can become an entrepreneur, in which case he must use his inherited wealth to pay a fixed project set-up cost and to pay for as much productive capital as he can afford or desires. Once in business, the entrepreneur hires as much labor as desired at the equilibrium wage rate which he pays out of earned revenues, receiving all remaining revenues as pure profit. Individuals who do not become entrepreneurs either remain in subsistence agriculture or enter the manufacturing labor force.

Agents vary in their abilities in that their personal costs of undertaking an investment project differ. Two factors determine whether an agent undertakes his project. If an agent's inherited wealth is too low relative to the cost and investment requirement to undertake a profitable project, or if the equilibrium wage rate is so high that he is better off becoming a laborer, then the agent will not become an entrepreneur. These occupational differences give rise to a non-degenerate distribution of incomes in the economy, even among those agents who are endowed with the same wealth, and result in a non-degenerate distribution of inherited wealth among their descendants. The exact evolution of the economy depends on the distribution of project costs, but we show that under certain reasonable assumptions it will exhibit the characteristics described below.

In the initial stages of development, few individuals can afford to become entrepreneurs and those who do, operate on a small scale because of their wealth constraints. Thus, there is little competition for labor so that the equilibrium wage rate is driven down to the lowest possible wage that will entice workers away from subsistence. The resulting distribution of income is somewhat unequal with most agents remaining in subsistence or earning the low equilibrium wage and a handful obtaining relatively large entrepreneurial profits because of this low wage.

This early stage of dual economic development is characterized by increasing entrepreneurial activity on an increasing scale, as descendants of entrepreneurs accumulate wealth and become less constrained. As a result, there is persistent migration of agents from rural to urban areas. However, as long as only a small proportion of agents become entrepreneurs, there is still relatively little competition for workers, so the equilibrium wage rate remains low. Income inequality therefore becomes increasingly acute because, while those at the lower end of the distribution experience a small increase in incomes as they move from subsistence farming to manufacturing, those entrepreneurs at the upper end of the distribution earn rapidly increasing profits. The distribution of wealth becomes more and more unequal because wealth persists along family lineages. An agent whose parent was a wealthy entrepreneur is more likely to become an entrepreneur even though project start-up costs are uncorrelated. This is because even relatively costly projects are worth undertaking when wage rates are low, especially if sufficient capital can be employed. Thus, a primary factor determining whether an agent becomes an entrepreneur is his inherited wealth. The child of a poor, non-entrepreneur remains as such unless he is especially efficient.

Eventually, the economy develops to the point that there are enough entrepreneurs operating at a sufficient scale that competition bids up the market clearing wage. As the equilibrium wage rate rises and the profits earned from operating any project at a given scale fall, less able agents begin to prefer to work as laborers and entrepreneurial quality rises. Provided production exhibits sufficient decreasing returns to scale, aggregate output still increases, reflecting the fact that the marginal product of poorer entrepreneurs exceeds that of wealthier ones. The optimal scale of production declines and there is a shift toward more capital-intensive production. The share of labor income in aggregate production rises. Income inequality does not generally fall when wages first increase because, while an entrepreneur's wage bill increases, those who inherit more may operate at a less constrained, more profitable, level. Still, as wages rise, operating profits eventually fall and income inequality declines. In turn, this results in a more equitable distribution of inherited wealth as the inheritances received by descendants of workers rise toward the levels received by descendants of entrepreneurs. Production becomes increasingly efficient and gradually, efficiency replaces wealth as the primary factor in determining the occupation of an agent, so that wealth does not persist as much along family lineages. The growth rate is non–monotonic during the development process. In general, growth rates depend on the interaction between the rising equilibrium wage, the associated redistribution of wealth and the relaxation of financing constraints.

Section 2 sets out the economy and we derive the optimal actions of agents in Section 3. Section 4 characterizes the macroeconomic equilibrium and the resulting distributions of income, wealth and bequests. Section 5 provides the terminology used to order distributions. Section 6 details the process of economic development from an initial period in which all agents have the same low level of wealth, to a long run level of development. The equilibrium evolution of relative income inequality resulting from this development process is analyzed in more detail in Section 7, in terms of the evolution of the associated Lorenz curves. In Section 8, we consider in more detail how the degree of social mobility evolves through time. Section 9 details how the optimal wealth redistributional policy evolves throughout the development process, and how it depends on the structure of production. In Section 10 we show that, because of the harsher consequences of imperfect capital markets in developing economies, proportionate negative wealth shocks have greater effects on them than on established economies. Finally, Section 11 alters the technological assumptions to see how the process of development is affected. In particular, it is shown how the endogenous evolution of wages can interact with the absence of functioning capital markets to generate endogenous cycles in output levels. All proofs are in the appendix.

2. The Environment

There are countably many time periods, $t = 0, 1, 2, ...$, and a continuum of family lineages such that in every period t, the economy is populated by a unit measure of agents. Each agent is active for one period, then reproduces one agent, so that the economy's population is constant over time. An agent's endowment consists of a bequest inherited from his parent. The agent consumes some of his lifetime wealth and the remainder is bequeathed to his offspring.

Agents have identical preferences represented by the utility function

$$u(C_t, B_{t+1}),$$

where C_t is an agent's consumption and B_{t+1} is his bequest. The utility function is twice differentiable, strictly increasing in both arguments and strictly quasi-concave. Both goods are assumed strictly normal, so that as wealth increases both consumption and bequests rise. The assumption that the preference is for the bequest itself, not for the offspring's utility, reduces the complexity of the analysis and is intended to capture the idea that agents adhere to a tradition for bequest-giving (see Adreoni 1989 or Bernheim 1991).

At time t=0 a new production technology comes into existence. Entrepreneurs who pay the necessary fixed project start-up cost (discussed below), produce a single consumption good according to the common production function

$$f(k_t, l_t),$$

where l_t is the number of workers employed on a project and k_t is the capital employed. The production function is twice differentiable, strictly increasing in both arguments and strictly concave. Capital and labor are complements, $f_{kl} > 0$, and the production function is quasi-homothetic. Third-order derivatives are assumed to be of negligible magnitude relative to first- and second-order ones, so the production function can be well approximated by a second-order Taylor series expansion. There is 100% capital depreciation between periods and the consumption good is perfectly convertible to capital.

Agents are distinguished by two characteristics: their initial wealth inheritances, b, and their personal costs of undertaking a project, x. Thus, an agent's "type" is given by the pair (b, x). Project set-up costs are drawn from a time-invariant distribution $H(x)$ with support $[\underline{x}, \bar{x}]$ and strictly positive density $h(x)$. The distribution of start-up costs reflects innate differences in entrepreneurial ability. Project quality is uncorrelated with inherited wealth:

$$H(x|b) = H(x).$$

The distribution of inherited wealth is determined endogenously as described below.

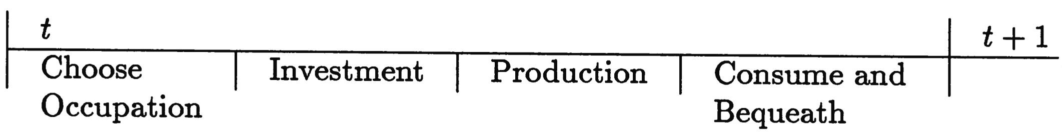


Figure 2

There is no individual or aggregate uncertainty. Agents make decisions taking the equilibrium wage rate and their own potential profits as given. The timing of decisions in any period t is depicted in Figure 2. At the beginning of each period an agent observes his own inheritance, b, and his cost draw, x. Since there is no financial market through which agents can pool their resources, the scale of the project is limited by an agent's inheritance and set-up cost [2] . Each agent chooses among three occupations. If his inheritance is large enough, he can invest all or part of it in his project

[2] We could introduce a capital market of the type considered by Banerjee and Newman (1993), or Galor and Zeira (1993), but, since this would merely result in agents being able to borrow up to a fixed proportion of their inherited wealth, it would be qualitatively superfluous.

and become an entrepreneur in the manufacturing sector. Investment must be financed out of inherited wealth and includes both the level of productive capital employed and the fixed project set up cost. Alternatively, agents can work in the manufacturing sector at the prevailing equilibrium wage rate, w_t, and save their inheritance until the end of the period. Saving is in the form of "hoarding" and no interest is paid. Finally, agents may prefer to remain in subsistence and receive the fixed marginal product of land, γ, in addition to their inheritance. The manufacturing sector is located in urban areas, where agents incur a cost of living equal to $\nu \geq 0$, that is not paid by agents in subsistence[3] .

The occupational choice of an agent of type (b, x) determines his lifetime income given the equilibrium obtaining that period, $y(b, x, w_t)$. His total lifetime wealth equals

$$W_t = W(b, x, \iota_t, w_t) = y(b, x, w_t) - \iota_t \nu + b,$$

where ι_t is a location indicator that takes on the value 1 if the agent lives in an urban area and is zero otherwise. At the period's end, agents allocate their total final wealth between their own consumption and bequests to their descendants so as to maximize utility.

Example

To illustrate our results, we employ an example featuring Cobb-Douglas preferences,

$$u(C_t, B_t) = C_t^{1-\omega} B_t^{\omega} \qquad 0 < \omega < 1,$$

a linear-quadratic production function,

$$f(k_t, l_t) = \alpha k_t - \frac{\beta}{2} k_t^2 + \xi l_t - \frac{\rho}{2} l_t^2 + \sigma l_t k_t,$$

and a distribution of start-up costs with support $[0, 1]$ and a linear density,

$$h(x) = 2mx + (1 - m) \qquad -1 \leq m \leq 1.$$

Note that the distribution of costs is uniform if $m = 0$.

[3] Other costs could be included, such as relocation costs, without affecting the qualitative nature of the results. Allowing for a utility cost of working in the manufacturing sector increases the complexity of the analysis unless, as in Banerjee and Newman (1993), preferences exhibit homogeneity of degree one.

3. Optimal behavior

An agent with final wealth W_t, chooses consumption C_t and bequest B_{t+1} to solve

$$\max\ u(C_t, B_{t+1})$$
$$s.t.\quad C_t + B_{t+1} = W_t,$$

yielding the optimal consumption and bequest policies, $C(W_t)$ and $B(W_t)$. Since consumption and bequests are strictly normal goods, i.e. since $0 < B'(W) < 1$, there is a one-to-one correspondence between final and bequeathed wealth. We assume that the pre-industrial economy has reached a steady state with constant bequests defined by

$$b^0 = B(\gamma + b^0).$$

Those agents who draw a sufficiently low cost technology and hence are willing and able to become entrepreneurs, do so. Wages are paid out of end–of–period revenues. After paying set-up cost, x, an entrepreneur chooses capital to maximize profits subject to the constraint that capital is financed out of his remaining wealth, $b - x$. Thus, the profits earned by an agent of type (b, x) who undertakes a project equal

$$\pi(b, x, w_t) = \max_{k_t, l_t}\ f(k_t, l_t) - w_t l_t - k_t - x$$
$$s.t.\quad 0 \leq k_t \leq b - x$$

This problem yields the capital demand function

$$k(b, x, w_t) = \min[k^c(b, x), k^u(w_t)],$$

where $k^c(b_t, x_t) = \max[b - x, 0]$ and $k^u(w_t)$ is the unconstrained optimal level of capital, and the associated labor demand function

$$l(b, x, w_t) = \min[l^c(b, x, w_t), l^u(w_t)],$$

where $l^c(b, x, w_t)$ and $l^u(w_t)$ are, respectively, the constrained and unconstrained demands for labor. The assumption of quasi-homotheticity implies that the marginal effect of an increase in wealth on the demand for labor by a constrained entrepreneur, $l_k(w_t)$, is solely a function of the wage rate and not firm size.

Denote the associated profits of the entrepreneur by

$$\pi(b, x, w_t) = \min[\pi^c(b, x, w_t), \pi^u(x, w_t)],$$

where again u indicates the unconstrained profit level, and c the constrained level. The profit function $\pi(b, x, w_t)$ gives the income y of an entrepreneur with cost x given his inheritance and the wage at time t. The 'dual' to this is a cost function $\tilde{x}(y, b, w_t)$ that identifies the project set-up cost required for an entrepreneur with inheritance b at time t to earn the income level y. Such a cost function is given by

$$\tilde{x}(y, b, w_t) = \min[\tilde{x}^c(y, b, w_t), \tilde{x}^u(y, w_t)],$$

where $\pi^c(b, x^c, w_t) = y$ and $\pi^u(x^u, w_t) = y$. The set-up cost associated with a given profit level y, $\tilde{x}(y, b, w_t)$, is strictly decreasing in y, increasing and concave in inherited wealth, b, and decreasing in the wage, w_t. For any bequest b and wage rate w_t, there is a one-to-one relationship between $\tilde{x}$ and y.

Agents can shift costlessly between sectors. The reservation wage in the manufacturing sector, below which potential workers prefer to remain in subsistence is given by

$$\underline{w} = \gamma + \nu.$$

When the wage is $\underline{w}$, non-entrepreneurs are indifferent between manufacturing and agriculture. We focus on the equilibrium in which agents only relocate in response to excess demand in the manufacturing sector. This assumption precludes the possibility of homogeneous labor relocating in opposite directions in equilibrium. This accords with casual observation, leaves the qualitative results unaffected, and can be motivated by arbitrarily small transportation costs. If the equilibrium wage exceeds $\underline{w}$, all non-entrepreneurs prefer to work in the manufacturing sector and there is no subsistence. In this case, agents just choose between undertaking a project or working for someone who has undertaken one.

For an agent with inherited wealth b to undertake a project, he must draw a start-up cost that is less than b. Even if such an agent can afford to become an entrepreneur there may still exist a marginal set-up cost level, $x^m(b, w_t)$, defined implicitly by

$$\pi(b, x^m, w_t) = w_t,$$

at which the agent would be indifferent between working and becoming an entrepreneur. We assume that $\pi(b^0, \underline{x}, \underline{w}) > \underline{w}$ so that some agents always wish to undertake projects.

The start-up cost of the agent type with inherited wealth b who is just willing and able to undertake a project at time t is therefore given by

$$z(b, w_t) = \min[b, x^m(b, w_t)].$$

Agents with $x > z(b, w_t)$ always become workers rather than undertake a project, even if they can afford to operate one profitably. Agents with $x \leq z(b, w_t)$ become entrepreneurs.

Figure 3 goes here

Figure 3 illustrates this function, which separates the population (represented by all possible (b, x) combinations) into agents who become constrained or unconstrained entrepreneurs and those who become workers or remain in subsistence[4] . Observe that, because of decreasing returns to scale in production, an increase in inherited wealth adds more, for a given start-up cost, to the profits of poor entrepreneurs than it does to those of more wealthy ones, so that $z(b, w_t)$ is weakly concave. A rise in the wage increases the relative attractiveness of wage laboring for all types so that the start-up cost of the marginal investor, $z(b, w_t)$, falls.

4. Macroeconomic Equilibrium

(a) Characterization

The endowment stream of any particular lineage evolves according to

$$b_{t+1} = B(y(b_t, x_t, w_t) - \iota_t \nu + b_t),$$

where x_t is the cost draw of the t^{th} member of the lineage. An agent's inheritance depends not only on his parent's inheritance, location and cost realization, but also on the past distribution of wealth via its effect on the equilibrium wage. Therefore, it is not possible to analyze the behavior of the economy by studying a single lineage because the transition function varies as the equilibrium wage rate evolves. The distribution of wealth evolves according to an endogenous non–stationary probability transition function, $P_t(b'|b)$, so that

$$G_{t+1}(b') = \int P_t(b'|b) G_t(db),$$

where $G_0(b) = 0$ for all $b < b^0$, and $G_0(b) = 1$ for all $b \geq b^0$.

To characterize the macroeconomic equilibrium in each period, we aggregate the optimal decisions of agents of different types, (b, x). The fraction of agents with inheritance b that become entrepreneurs when the wage is w_t is given by

$$\eta(b, w_t) = H(z(b, w_t)).$$

4 In the terms used by the credit rationing literature, there can exist, simultaneously, firms which are rationed on the intensive and extensive margins (i.e. would operate even without capital if they could just pay the set-up costs). We could rule out the latter by assuming $f(0,\cdot)=0$.

We also define, respectively, the total start-up costs incurred, labor employed, capital invested and output produced by these entrepreneurs:

$$\chi(b, w_t) = \int_{\underline{x}}^{z(b,w_t)} x h(x)\, dx,$$

$$\lambda(b, w_t) = \int_{\underline{x}}^{z(b,w_t)} l(b, x, w_t) h(x)\, dx,$$

$$\kappa(b, w_t) = \int_{\underline{x}}^{z(b,w_t)} k(b, x, w_t) h(x)\, dx,$$

$$q(b, w_t) = \int_{\underline{x}}^{z(b,w_t)} f(k(b, x, w_t), l(b, x, w_t)) h(x)\, dx.$$

The total income accruing to all agents with inheritance b is given by

$$\psi(b, w_t) = \int_{\underline{x}}^{z(b,w_t)} \pi(b, x, w_t) h(x)\, dx + [1 - H(z(b, w_t))] w_t.$$

All of these variables increase with inherited wealth and decrease with the wage rate. Integrating over inheritances yields, respectively, the aggregate fraction of agents who become entrepreneurs, $E_t(w_t) = \int \eta(b, w_t) G_t(db)$; the aggregate set-up cost expenditures, $X_t(w_t) = \int \chi(b, w_t) G_t(db)$; the aggregate labor demand, $L_t(w_t) = \int \lambda(b, w_t) G_t(db)$; the aggregate capital demand, $K_t(w_t) = \int \kappa(b, w_t) G_t(db)$; the aggregate manufacturing output, $Q_t(w_t) = \int q(b, w_t) G_t(db)$; and the aggregate net income in the advanced sector, $Y_t(w_t) = \int \psi(b, w_t) G_t(db)$. All of these aggregates are time-varying functions of the wage because they also depend on the distribution of inherited wealth.

A **competitive equilibrium** for an economy with an initial distribution of wealth $G_t(\cdot)$ is a tuple $\{w_t^e, E_t^e, L_t^e, S_t^e\}$ such that:

1. Given the wage w_t^e, an agent of type (b, x) selects his occupation to maximize utility.

2. Agents of type (b, x) who become entrepreneurs maximize profits by choosing optimal levels of labor and capital subject to the constraint that $k \leq b - x$.

3. Markets clear:

$$E_t^e(w_t^e) + L_t^e(w_t^e) + S_t^e(w_t^e) = 1,$$

where $S_t^e(w_t^e) = 0$ if $w_t^e > \underline{w}$.

We distinguish between two types of equilibrium that may arise.

A **dual economy equilibrium** obtains at time t if and only if the distribution of inheritances, $G_t(\cdot)$, is such that for all $w \geq \underline{w}$,

$$E_t(w) + L_t(w) < 1.$$

In such an equilibrium, many agents are either fully constrained from undertaking projects, or can only invest small amounts in productive capital. Thus, both the extent and the scale of productive activities is so limited that the demand for labor generated is not great enough to absorb the remainder of the population at the reservation wage rate. Surplus labor remains in existence in the rural-agricultural sector, so the wage remains at $\underline{w}$.

Throughout we implicitly assume that migration occurs in the rural-to-urban direction only. A sufficient condition for this to be true is that the increase in the urban population between time t=0 and t=1 is sufficient to absorb at least those agents with inherited wealth b^0 that wish to become entrepreneurs:

$$E_1(\underline{w}) + L_1(\underline{w}) - E_0(\underline{w}) - L_0(\underline{w}) > E_0(\underline{w})(1 - E_0(\underline{w})).$$

Under the development process that we describe this condition replicates itself every period and, hence, if it is true in period $t = 1$, it is true in every period (see Proposition 2). An implication is that all agents in subsistence have inherited wealth b^0, because their ancestors must have always been in subsistence. This assumption does not affect the evolution of the distribution of wealth or the underlying development process, but is used to characterize the distribution of income.

A **non-dual economy equilibrium** obtains at time t, if the distribution of inheritances $G_t(\cdot)$ is such that there exists a wage $w > \underline{w}$ for which

$$E_t(w) + L_t(w) = 1.$$

In such an equilibrium, agents are wealthy enough for there to be competition amongst entrepreneurs for workers. The equilibrium wage rate is bid up past $\underline{w}$ and there is no surplus labor in the backward sector. For such an equilibrium to exist, it must be feasible for the economy to support the entire population in the urban sector. To ensure this, it is assumed that the distribution of set-up costs and $\underline{w}$ are such that if all agents were unconstrained then there would be no subsistence:

$$[1 + l^u(\underline{w})]H(x^m(\underline{w})) > 1.$$

In other words, the maximum potential efficiency of the economy is not achieved before the rural labor surplus is exhausted.

(b) Equilibrium distributions

At time t, a proportion $1 - H(z(b, w_t))$ of agents with inheritance b realize a set-up cost so high as to make undertaking a project either infeasible or undesirable. If $b = b^0$, then the ancestors of such agents have never been entrepreneurs. $G_t(b^0)$ is the measure of such agents and, hence, the measure of agents whose ancestors have never been entrepreneurs at the beginning of the next period is $G_{t+1}(b^0) = [1 - H(z(b^0, w_t))]G_t(b^0)$. The fraction of non-entrepreneurs that inherit wealth b^0 and remain in subsistence, rather than work as wage laborers, is

$$p_t = \frac{S_t}{[1 - H(z(b^0, w_t))]G_t(b^0)}.$$

The assumption which guarantees that migration is in the rural-to-urban direction only, also ensures that $0 \leq p_t < 1$[5] . If $b > b^0$, then the parents of such agents resided in the urban sector. All non-entrepreneurs at these inheritance levels enter the industrial work force and receive income w_t. The remaining fraction of agents with inheritance b, $H(z(b, w_t))$, receive a low enough cost realization, $x < z(b, w_t)$, to undertake their projects. Since $\tilde{x}(y^*, b, w_t)$ is the value of x such that $\pi(b, x, w_t) = y^*$, the measure of entrepreneurs with profits less than some income level y^*, is given by $H(z(b, w_t)) - H(\tilde{x}(y^*, b, w_t))$.

The distribution of income conditional on inherited wealth and the equilibrium obtaining at time t can be represented by the cumulative distribution function,

$$\Phi(y|b, w_t) = \begin{cases} 0 & \text{if } y < \gamma \text{ or } y < w_t, b > b^0 \\ S_t/G_t(b^0) & \text{if } \gamma \leq y < w_t, b = b^0 \\ 1 - H(z(b, w_t)) & \text{if } w_t \leq y \leq \underline{\pi}(b, w_t) \\ 1 - H(\tilde{x}(y, b, w_t)) & \text{if } \underline{\pi}(b, w_t) < y \leq \bar{\pi}(b, w_t) \\ 1 & \text{otherwise} \end{cases}$$

5 By assumption $S_t < S_{t-1} - H(z(b^0, w_t))G_t(b^0)$. Thus, $S_t < [1 - H(z(b^0, w_t))]G_t(b^0)$ if and only if $S_{t-1} < G_t(b^0) = [1 - H(z(b^0, w_t))]G_{t-1}(b^0)$. Since $S_0 < 1 - H(z(b^0, \underline{w}))$, the result follows by induction.

where $\underline{\pi}(b, w_t) = \max[w_t, \pi(b, b, w_t)]$ is the lower support on equilibrium profits, $\bar{\pi}(b, w_t) = \min[\pi^c(b, \underline{x}, w_t), \pi^u(\underline{x}, w_t)]$ is the upper support. The distribution of income conditional on b is a mixed distribution, consisting of a point mass at $y = w_t$ (and at $y = \gamma$ if $b = b^0$), and a density of entrepreneurial incomes, $\phi(y|b, w_t)$, on the interval $[\underline{\pi}(b, w_t), \bar{\pi}(b, w_t)]$.

The unconditional distribution of incomes is therefore given by

$$\Phi_t(y) = \int \Phi(y|b, w_t) G_t(db).$$

The final wealth of an agent of type (b, x) in period t, $W_t(b, x)$, is simply the sum of his inheritance and lifetime income, net of any urban living costs. Hence, the distribution of final wealth is given by

$$F_t(W) = \int F(W|b, w_t) G_t(db),$$

where $F(W|b, w_t) = \Phi(W + \iota_t \nu - b|b, w_t)$ is the distribution of income conditional on inherited wealth.

Similarly, the unconditional distribution of bequests is given by

$$G_{t+1}(b') = \int P(b'|b, w_t) G_t(db),$$

where $P(b'|b, w_t) = F(B^{-1}(b')|b, w_t) = \Phi(B^{-1}(b') + \iota_t \nu - b|b, w_t)$ is the distribution of bequests by agents with wealth b[6] . Thus, $P(b'|b, w_t)$ defines the probability transition function for inheritances.

We can also characterize the distribution of firm sizes. A firm's "size" is defined here as the amount of capital employed by it. If $k < \hat{k}(w_t)$, then the fraction of agents with inherited wealth b that hire k units of capital is given by $h(b - k)$ since such agents are constrained. Hence the measure of entrepreneurs with wealth b employing less than $k^* \in [b - z(b, w_t), \hat{k}(w_t)]$ units of capital is given by $H(z(b, w_t)) - H(b - k^*)$. The fraction of entrepreneurs with wealth b employing up to k units of capital is then given by

$$J(k|b, w_t) = \begin{cases} 0 & \text{if } b > \hat{b}(w_t), k < \hat{k}(w_t) \\ \frac{H(z(b,w_t)) - H(b-k)}{H(z(b,w_t))} & \text{if } b < \hat{b}(w_t), k < \hat{k}(w_t) \\ 1 & \text{otherwise,} \end{cases}$$

so that the time t distribution of firm sizes is

$$J_t(k) = \int J(k|b, w_t) G_t(db).$$

6 Strict normality of bequests ensures the inverse bequest function, $B^{-1}(\cdot)$, exists.

5. X-Dispersion

Since our aim is to characterize and explain the extent of inequality over time, we must first define what we mean by 'changes in inequality'. There are many measures of inequality, most of which are variants of the area under a Lorenz curve. Section 7 analyzes the evolution of inequality over time in terms of Lorenz dominance. However, because their means evolve over time, it is more useful to characterize the evolution of the distributions of income and wealth in another way:

Definition: Distribution $\Phi_1(\cdot)$ is more **X-Dispersed** than distribution $\Phi_2(\cdot)$, if there exists a y^* such that $0 < \Phi_1(y^*) < 1$ and $0 < \Phi_2(y^*) < 1$ and

$$\Phi_1(y) \geq \Phi_2(y) \quad \forall\, y < y^*$$

$$\Phi_1(y) \leq \Phi_2(y) \quad \forall\, y \geq y^*$$

where the inequality is strict on an interval of positive measure.

Figure 4 goes here

The two distributions are illustrated in Figure 4. X-Dispersion is a demanding criterion because it does not order distributions with either multiple or zero intersections on the interval $(0, 1)$. However, X-Dispersion permits a comparison of distributions even when their means differ, and, it turns out, is the notion of inequality that allows us to characterize the equilibrium dynamics of the economy. If the means of distributions 1 and 2 are equal, then X-Dispersion is equivalent to a mean preserving spread in the Rothschild-Stiglitz sense. More generally, if distribution 1 is more X-Dispersed than distribution 2, then the second moment of distribution 1 about y^* exceeds that of distribution 2. Clearly, if y^* is close to the means of the distributions, relative X-Dispersion reflects the relative variance of the two distributions.

6. The dynamics of economic development

Stage 1 : Dual economic development

We first characterize the dual economy stage of the development process in which the manufacturing sector has not advanced sufficiently to absorb all agents. The evolution of aggregate variables is characterized by:

Proposition 1 : *During the dual stage of the development process the rate of enterprise, $E_t(\underline{w})$, the rate of labor force participation, $L_t(\underline{w})$, the level of investment, $K_t(\underline{w})$, aggregate set-up costs, $X_t(\underline{w})$, aggregate output, $Q_t(\underline{w})$,* *aggregate net income, $Y_t(\underline{w})$, and final and bequeathed aggregate wealth, all rise monotonically.*

Proposition 2 : *Migration is strictly positive and occurs in the rural–to–urban direction only.*

To understand these propositions, consider the effects of the $t = 0$ introduction of the new production technology process to the economy in which agents have been in subsistence forever, receiving the subsistence income, γ, and bequeathing the stationary bequest, b^0. Provided that the equilibrium obtaining is a dual economy one, a small fraction $\eta(b^0, \underline{w})$ of agents, those with very low project set up costs, migrate to urban areas to become entrepreneurs. The final wealth of workers, after paying the cost of urban living ν, just equals that of subsisters so that they continue to bequeath b^0. The profits of entrepreneurs are at least as great as the wage, so that their final wealth exceeds $\gamma + b^0$. Since $B'(W) > 0$, it follows that entrepreneurs bequeath more than b^0. Hence, the distribution of inheritances at $t = 1$, $G_1(\cdot)$, dominates that at $t = 0$, $G_0(\cdot)$, in the first-order stochastic sense.

At $t = 1$, the same fraction $\eta(b^0, \underline{w})$ of agents inheriting b^0 become entrepreneurs. However, now there is a group of agents whose parents were entrepreneurs and who inherit more than b^0. These agents can both afford higher cost projects and can undertake low cost ones at a larger scale than those inheriting b^0. As a result, both the fraction of agents becoming entrepreneurs and their aggregate output increase. The associated demands for labor and capital rise, as does the aggregate level of start-up costs.

Since there is migration away from rural areas, the proportion of the population receiving subsistence income γ decreases. The measure of agents with inherited wealth b earning less than any given income $y \geq \underline{w}$, $\Phi(y|b, \underline{w})$, decreases with b. Since the distribution of inheritances at $t = 1$ dominates that at $t = 0$ in the first-order stochastic sense, it follows that the distribution of income at $t = 1$, $\Phi_1(\cdot)$, dominates that at $t = 0$, $\Phi_0(\cdot)$, in the first-order stochastic sense. The distribution of final wealths evolves in a similar manner so that the distribution of inheritances $G_2(\cdot)$ dominates $G_1(\cdot)$ in the first-order stochastic sense. Proposition 2 ensures that no lineage moves back to subsistence after having been in the urban sector so that all agents in subsistence must inherit b^0.

With first-order stochastic growth in inheritances, the fraction of agents that become entrepreneurs again increases in period $t = 2$ and the process continues. Output rises, the demand schedule for labor shifts further to the right and the distributions of income, wealth and bequests all grow in the first-order stochastic sense. The development process continues in this fashion as long as there is surplus labor in the economy and the wage remains at $\underline{w}$ (see Figure 5). Thus, we can think of this stage of development as a formalization of the Lewis (1954), and Fei and Ranis (1966) models of dual economic development with unlimited labor supplies.

Figure 5 goes here

First-order stochastic dominance does not, by itself, say anything about the dispersion of incomes and wealths in the economy. However, in this stage of development, there is always a strictly positive measure of agents who receive γ. As Figure 6 illustrates, this is a special case of X-Dispersion where $\Phi_t(\cdot)$ is more X-Dispersed than $\Phi_{t-1}(\cdot)$. The distributions of final wealth and inheritances evolve in a similar fashion with a positive, but declining, mass at the lower bounds, $W^0 = \gamma + b^0$ and b^0 respectively:

Proposition 3 : *Stage 1 of the development process is characterized by increasing X-Dispersion in the distributions of income, wealth and bequests.*

Figure 6 goes here

As long as the wage remains at $\underline{w}$, the gains from enterprise accrue to entrepreneurs. The relatively few entrepreneurs receive profits that are substantially higher than the wage, while the majority of the population continue to receive $\underline{w}$ or less. The former group expands and their profits increase in the first-order stochastic sense. Hence, the distribution of income exhibits greater absolute dispersion as long as the wage remains at $\underline{w}$. This income inequality results in a distribution of inheritances that exhibits similar properties. An agent whose parent had high wealth last period is less likely to be wealth constrained and, therefore, is more likely than less wealthy agents to become an entrepreneur and earn a high income this period. As a consequence, lineage wealth tends to persist and wealth inequality becomes progressively more acute. In the numerical example illustrated in Figure 7, dual economic development persists until period $t = 10$.

Figure 7 goes here

Finally, consider the evolution of the distribution of firm sizes during the early stages of development. Since, for any cost draw, the wealth of agents stochastically increases, so must the level of capital employed by them if they are constrained entrepreneurs. Note, however, that the number of entrepreneurs also increases (because fewer agents are constrained) and there is no increase in investment by unconstrained entrepreneurs, so that while the total capital stock increases unambiguously, first-order stochastic dominance in firm sizes holds only as long as there are not too many relatively inefficient agents becoming entrepreneurs on a small scale. A sufficient condition for this is condition (A), that the distribution of costs is not too skewed towards high cost projects.

Proposition 4 : *There exists a* $\delta > 0$ *such that if*
(A) No excess skewness: $\frac{h(x)-h(\underline{x})}{x-\underline{x}} < \delta \ \forall \ x \in [\underline{x}, x^m(\underline{w})]$
then, the distribution of firm sizes, $J_t(\cdot)$*, grows in the first-order stochastic sense and exhibits increasing X-Dispersion throughout the stage of dual economic development.*

Stage 2 : The Transitional Phase

During the dual development stage, since the population is constant, and both the demand for labor and the supply of entrepreneurs are increasing, the supply of surplus labor in the traditional sector must eventually be exhausted. Hence there exists a date τ^1 at which the equilibrium wage rate, w_{τ^1}, rises above $\underline{w}$. In Figure 8(a) the area below the upper envelope generated by the supply and demand curves for labor, represents aggregate income, Y_t. The rectangle $OwLC$ represents the aggregate wage bill. The remainder represents the profits earned by entrepreneurs. The area wAB, is the aggregate surplus earned in the labor market. The area $LBCD$, represents the economic profit earned from having sufficiently low start-up costs.

Figure 8 goes here

The effects of the rising wage on the equilibrium distributions of income, wealth and bequests are complex. The rising wage implies that the profits of an entrepreneur of any given type (b, x) falls, but this may be offset by the rise in profits due to the continued relaxation of financing constraints. For the latter effect to dominate so that aggregate income rises, it is sufficient that the supply and demand curves both shift up (see Figure 8b). At τ_1, this follows immediately from the fact that the distribution of inheritances dominates that at $\tau_1 - 1$ in the first-order stochastic sense. More generally,

Proposition 5 : *There exists a* $\delta > 0$ *such that if*
(A) No excess skewness: $\frac{h(x)-h(\underline{x})}{x-\underline{x}} < \delta \ \forall \ x \in [\underline{x}, x^m(\underline{w})]$
then, in the transitional phase of economic development the equilibrium wage rate, w_t*, aggregate net income,* $Y_t(w_t)$*, and final and inherited aggregate wealth rise monotonically.*

To understand why condition (A) is important, suppose that it does not hold, so that the distribution of start-up costs is highly skewed towards high-cost projects. Then the increase in the equilibrium wage may sharply reduce the fraction of entrepreneurs at low profit (high cost) levels as they switch into the labor force. If this offsets the effect of the stochastic increase in inherited wealth, the density of incomes at these profit levels may be less at time τ_1 than at time $\tau_1 - 1$. Condition (A) ensures first-order stochastic growth in incomes and hence bequests when $\bar{\pi}_{\tau_1} > \bar{\pi}_{\tau_1 - 1}$ (see Figure 9).

Figure 9 goes here

Until wealth is redistributed away from the richest agents in the economy, the supply of entrepreneurs, $E_t(w)$, and the aggregate demand for labor $L_t(w)$, must continue to rise. However, this transitional phase cannot continue forever. Strict concavity in production implies that eventually the richest and most efficient entrepreneur, $(\bar{b}_t, \underline{x})$, can produce at the optimal scale. At this point, the income of such an agent declines with the wage. Eventually, this declining income offsets the rising lineage wealth, so that at some time $\tau^2 \geq \tau^1$, the bequest of such an agent, $\bar{b}_{\tau^2+1}$, is less than his inheritance, $\bar{b}_{\tau^2}$. Thus, the distribution of inheritances in the next period, $G_{\tau^2+1}(\cdot)$, cannot dominate $G_{\tau^2}(\cdot)$ in the first-order stochastic sense. It is possible that $\tau^2 = \tau^1$ so that the transitional phase only lasts for one period. More typically, the transitional phase lasts for several periods until time $\tau^2 > \tau^1$, after which the next stage of development begins.

Stage 3 : Advanced Economic Development

In period τ_2, inherited wealth starts to be transferred from the top tail of the distribution of inheritances downwards. The consequences for aggregate variables in the economy, as well as for the distributions of income and wealth, are more subtle.

Proposition 6 : *There exists a $\delta > 0$ and a $\sigma > 0$ such that if*
(A) No excess skewness: $\frac{h(x)-h(\underline{x})}{x-\underline{x}} < \delta \ \forall\, x \in [\underline{x}, x^m(\underline{w})]$, *and*
(B) Weak classical effect: $\frac{B'(W)-B'(W^0)}{W-W^0} < \sigma \ \forall\, W$,
then, during the stage of advanced economic development, the equilibrium wage rate, w_t, aggregate net income $Y_t(w_t)$ and aggregate final and inherited wealth all rise monotonically.

Were the utility function thrice-differentiable, then $B''(W) < \sigma$ implies condition (B).

Essentially, conditions (A) and (B) ensure that past increases in aggregate production translate into greater wages and average bequests. Although aggregate final wealth at $\tau_2 - 1$ exceeds that at $\tau_2 - 2$, it need not follow that aggregate inherited wealth at the beginning of period τ_2

exceeds that in period $\tau_2 - 1$. If bequests are too much of a luxury good, then the redistribution of wealth away from the rich may decrease total bequeathed wealth by more than the increase caused by the rise in aggregate final wealth. Condition (B) ensures that this "classical effect" is sufficiently weak.

Condition (A) ensures that the supply and demand curves for labor both shift up so that the equilibrium wage and aggregate income must increase. To see this, suppose the distribution of inheritances at some time t, $G_t(\cdot)$, dominates $G_{t-1}(\cdot)$ in the second-order stochastic sense. Since the production function is strictly concave in capital and quasi-homothetic, a unit transfer of wealth from rich to poor entrepreneurs leads to a rise in the demand for labor because the marginal product of capital is greater for smaller scale projects. However, a transfer of wealth from the rich to the poor may reduce the measure of entrepreneurs, if, as a result, a large number of high cost projects become undesirable or infeasible. Condition (A) ensures that the former "scale effect" is not offset by the latter "occupational effect", so that the demand curve shifts outwards. Since Condition (A) ensures that the supply of entrepreneurs rises and the population is constant, the supply schedule for labor must shift inwards. Note that the more concave is the production function, the less binding condition (A) needs to be. This is because the greater is the degree of decreasing returns to scale, the stronger is the scale effect resulting from the redistribution of inheritances.

As in the transitional phase, the slope of the c.d.f. for income, $\Phi_t(\cdot)$, decreases over time on the interval $(w_t, \bar{\pi}_t]$. Now however, the maximum profit level falls, so that $\Phi_{t-1}(\cdot)$ and $\Phi_t(\cdot)$ intersect once and only once at some income level y_t^*. The distribution of income evolves in this way because relatively wealthy agents are more likely to be entrepreneurs and receive lower profits, while poorer agents are predominantly workers and experience an increase in income.

Given that aggregate net income increases and the evolution of the distribution is as described above, it follows that $\Phi_t(\cdot)$ dominates $\Phi_{t-1}(\cdot)$ in the second-order stochastic sense. In turn, since total inherited wealth and income increase, total final wealth must also increase and, hence, $F_t(\cdot)$ dominates $F_{t-1}(\cdot)$ in the second-order stochastic sense. This implies that, given condition (B), the distribution of bequests $G_{t+1}(\cdot)$ dominates $G_t(\cdot)$ in the second-order stochastic sense. By induction from period τ_2, the demand and supply schedules for labor continue to shift up and the economy continues to grow.

In general, the growth rate during the process of development is non–monotonic. Inspection of Figure 7 reveals that for the example economy, the growth rate declines gradually in the early stages of development, rises again when the wage starts to increase and then declines as the economy develops further. During the dual stage of development, the growth rate declines because more agents become unconstrained and the marginal return to their wealth falls. The increase in growth when wages initially rise is due to the effective redistribution of wealth towards poorer agents who are more productive on the margin. As the economy develops into the advanced stages, because there is no engine for long–run growth, wage increases decline and an increasing number of agents become unconstrained, so that the growth rate falls.

Moreover, as illustrated in Figure 10:

Proposition 7 : *There exists a $\delta > 0$ and a $\sigma > 0$ such that if*
(A) No excess skewness: $\frac{h(x)-h(\underline{x})}{x-\underline{x}} < \delta \; \forall \; x \in [\underline{x}, x^m(\underline{w})]$, *and*
(B) Weak classical effect: $\frac{B'(W)-B'(W^0)}{W-W^0} < \sigma \; \forall \; W$,
then the stage of advanced development is characterized by declining X–Dispersion in the distributions of income, wealth and inheritances.

Figure 10 goes here

The upward shifts in the supply and demands for labor can be consistent with either a decrease or an increase in the equilibrium number of laborers, depending on the respective wage elasticities of supply and demand. If, ceteris paribus, labor demand is sufficiently inelastic, then equilibrium wage laboring rises. If, labor demand is sufficiently elastic, as in Figure 11, the equilibrium labor force declines:

Proposition 8 : *There exists $\zeta < 0$ such that if, in addition to conditions (A) and (B), the wage elasticity of demand for labor is sufficiently high,*

$$\left(\frac{\underline{w}}{w - \underline{w}}\right) \int_{\underline{w}}^{w} \lambda_w(b, \hat{w}) d\hat{w} > \zeta \quad \forall \; b, w,$$

then, during the stage of advanced economic development, the rate of enterprise increases and wage laboring declines.

Figure 11 goes here

Characterization of the evolution of aggregate productive investment, $K_t(w_t)$, and aggregate start-up costs, $X_t(w_t)$, is more complex. Were the number of entrepreneurs to decline, then, because this is the net effect of an increase in relatively low wealth entrepreneurs and a decrease in higher wealth entrepreneurs, both $K_t(w_t)$ and $X_t(w_t)$ would decline unambiguously. If, as in the example, the equilibrium number of entrepreneurs were to rise, then the paths of these variables depend on the net effect of the negative impact of the wage increase and the positive impact of the redistribution of inheritances. Under condition (A), a second-order stochastic increase in inheritances causes the capital demand schedule to shift right. The responsiveness of the aggregate capital stock to wage changes is given by:

$$\frac{dK_t}{dw} = \int \left[\int_{b-\hat{k}(w)}^{z(b,w)} k_w(b,x,w)h(x)\,dx + k(b,z(b,w),w)\eta_w(b,w) \right] G_t(db).$$

For a given wage increase, the equilibrium level of capital falls if the wage elasticity of demand for capital is sufficiently high. If labor and capital are strong complements in production, if the supply of entrepreneurs is highly wage elastic and if the fraction of constrained entrepreneurs is small, then the demand for capital is quite wage elastic.

Similarly, the responsiveness of aggregate start–up costs, $X_t(\cdot)$, to changes in the wage depends on the wage elasticity of the supply of entrepreneurs:

$$\frac{dX_t}{dw} = \int z(b,w)\eta_w(b,w)G_t(db).$$

A second-order stochastic redistribution of inheritances may cause the schedule for $X_t(w)$ to shift right or left. However, again, the equilibrium level of aggregate start-up costs declines if the wage elasticity of the supply of entrepreneurs is sufficiently high.

In the example (see Figure 12) both $X_t(w_t)$ and $K_t(w_t)$ rise in period τ^1 because the wage does not rise much above $\underline{w}$. Subsequently, aggregate start-up costs fall as the rising wage causes less efficient agents to become workers rather than entrepreneurs. The capital stock also falls initially because the relatively large wage changes have effects which offset the effects of increased investment due to redistribution. The capital stock then rises slightly again as more and more agents continue to become unconstrained but the the wage increases are relatively small. What is unambiguous, is that output increases more rapidly than the sum of capital and start-up costs. This must be the case if net aggregate income increases.

Figure 12 goes here

If the aggregate capital stock falls and the number of entrepreneurs rises, then average firm size must decline. However, the aggregate capital stock need not fall. Figure 13a illustrates the time path of the average firm size for the example. Recall that the rate of enterprise, E_t, increases throughout the process, but that the capital stock falls at first, but then rises during the advanced stages. The evolution of the average firm size reflects this.

Figure 13 goes here

Although, due to rising wages, average firm size does not generally rise, the dispersion of firm sizes evolves in a way similar to that of the distributions of income and wealth:

Proposition 9 : *There exists a $\delta > 0$ and a $\sigma > 0$ such that if*
(A) No excess skewness: $\frac{h(x)-h(\underline{x})}{x-\underline{x}} < \delta \ \forall \ x \in [\underline{x}, x^m(\underline{w})]$, *and*
(B) Weak classical effect: $\frac{B'(W)-B'(W^0)}{W-W^0} < \sigma \ \forall \ W$,
then the distribution of firm sizes, $J_t(\cdot)$, exhibits declining X-Dispersion throughout the advanced stage of economic development.

Figure 13b depicts the time path of the variance of firm sizes for the example. It can be seen that the variance rises and reaches a peak during the dual stage of development and then declines throughout the advanced stage. If the stage of maximum efficiency were attained, the variance would be zero because all entrepreneurs would be producing at the optimal scale $k^u(\bar{w})$. In general, however, the limiting distribution of firm sizes is non-degenerate so that the variance converges to a positive value.

Stage 4 : Long run economic development

The economy cannot grow without bound. In particular, there exists a state of maximum efficiency in which all agents who wish to become entrepreneurs can do so at the unconstrained optimum scale given the equilibrium wage rate. An upper bound for the equilibrium wage is implicitly defined by the efficient market clearing condition,

$$[1 + l^u(\bar{w})]H(x^m(\bar{w})) = 1.$$

In this state, all firms produce the unconstrained maximum level of output. Hence, the aggregate levels of labor, capital, output and set-up costs are time invariant. Were this not the case then a change in the distribution of inheritances would affect the wage.

There is no guarantee that the economy will reach such a state of development. However, since it increase monotonically with time and is bounded, the wage must converge to some long run value $w^* \in [\underline{w}, \bar{w}]$. This implies that

Proposition 10 : *For all $\epsilon > 0$ there exists a T, such that for all $t > T$, the following distributions converge to time invariant distributions:*
(a) The conditional distribution of income:
$\Phi(y|b, w^*) - \Phi(y|b, w_t) < \epsilon \ \forall \ y, b,$

(b) The conditional distribution of final wealth:
$F(W|b, w^*) - F(W|b, w_t) < \epsilon \; \forall \; W, b,$
(c) The conditional distribution of bequests:
$P(b'|b, w^*) - P(b'|b, w_t) < \epsilon \; \forall \; b', b.$

Note that the support of the distribution of inherited wealth must converge to $[\underline{b}, \bar{b}]$, where $\underline{b} = B(w^* - \nu + \underline{b})$ and $\bar{b} = B(\pi^u(\underline{x}, w^*) - \nu + \bar{b})$. The stationary process governing the evolution of the distribution of inheritances is monotone since the distribution of inheritances amongst the offspring of a rich agent dominates that amongst the offspring of a poor agent in the first-order stochastic sense. Moreover, the process $P(\cdot|b, w^*)$ satisfies the Monotone Mixing Condition (see Hopenhayn and Prescott, 1992). That is, for any $b^* \in (\underline{b}, \bar{b})$, after a sufficient number of generations, there is a positive probability that the descendant of one of the richest agents in the economy receives an inheritance below b^* and a positive probability that the descendant of one of the poorest types eventually receives an inheritance above b^*. To see this, consider the descendants of the richest agent at time t. With positive probability, his children will realize a cost draw in any neighborhood of $\bar{x}$. Similarly, with positive probability his grandchildren will realize such a cost draw, and so on. Hence, with positive probability after a sufficient number of generations, his descendant will be a worker who inherits wealth in any neighborhood of $\underline{b}$. Similarly, there is positive probability that the child of the poorest agent realizes a cost draw in any neighborhood of $\underline{x}$ and becomes an entrepreneur. With positive probability, his descendant, after a sufficient number of generations, will inherit wealth in any neighborhood of $\bar{b}$. Hence, after a sufficient length of time, say N generations, the distribution of inheritances amongst the N^{th} descendants of the richest and poorest agents living at time t converge to the same distribution. By monotonicity, the distribution of inheritances amongst the N^{th} descendants of all agents living at time t must converge to this distribution.

Proposition 11 : *There exists a $\delta > 0$ and a $\sigma > 0$ such that if*
(A) No excess skewness: $\frac{h(x)-h(\underline{x})}{x-\underline{x}} < \delta \; \forall \; x \in [\underline{x}, x^m(\underline{w})]$, *and*
(B) Weak classical effect: $\frac{B'(W)-B'(W^0)}{W-W^0} < \sigma \; \forall \; W$,
then the distribution of inheritances converges to a unique time-invariant limiting distribution, $G^*(\cdot)$, which is independent of the initial distribution of inherited wealth.

Figure 14 goes here

Since the distribution of inheritances converges, so must the distribution of income:

$$\Phi^*(\cdot) = \int \Phi(\cdot|b, w^*) G^*(db).$$

The long-run distributions for both income and wealth are illustrated in Figure 14. Note that the limiting distributions of both inheritances and incomes are non-degenerate. In the long-run, wealth disparities continue to exist, but do not persist over time. The limiting distribution of inheritances, and hence that of income, is independent of the initial distribution of inheritances, $G_0(\cdot)$. Thus, economies which start out with more unequal distributions, while they may not follow the same cycle of development, will end up with the same long-run distribution. This provides a theoretical rationale for the observation that the relationship between inequality and per capita income seems to be more variable across low income economies than across higher income ones (see Figure 1).

The two extreme cases of the long-run development path of the economy are worth noting. The economy may reach the state of maximum efficiency in finite time. In this case, all production is undertaken at the efficient scale and no agent is constrained from undertaking a project if he so desires. Thus, the equilibrium of the economy and the occupational choices of agents are independent of the distribution of inheritances. Consequently, the conditional distribution of income is independent of inherited wealth : $\Phi(y|b,\bar{w}) = \Phi(y|\bar{w}) \; \forall \; b$. In this economy, inequality is very low because the wage is at its highest possible level.

Alternatively, the economy may never achieve enough momentum to leave the dual economic development stage. Because the wage remains at $\underline{w}$, wealth and income inequality remain very high (see Figure 15a). Moreover, production is very inefficient because most entrepreneurs are constrained from operating at the efficient scale. Even so, the economy still satisfies the mixing condition, although the persistence of wealth along lineages is very high. Figure 15b illustrates the long-run equilibrium wage for the simulated example, for different values of ω, the share of final wealth bequeathed to the next generation. As can be seen, if ω is too low the economy does not leave the dual development stage, whereas if it is sufficiently high the state of maximum efficiency is attained in finite time.

Figure 15 goes here

7. The evolution of inequality: A Lorenz curve analysis

We now characterize the behavior of the entire Lorenz curve of the income distribution during the development process we have described. The Lorenz curve of an income distribution details the fraction of total income received by the lowest fraction p of the population as p varies from 0 to 1. Let $\hat{y}_t(p)$ represent the income level of the poorest fraction p of the population at time t. This is defined implicitly by $\Phi_t(\hat{y}_t) = p$. The income share of the poorest fraction p is then given by the Lorenz function,

$$Z_t(p) = \frac{1}{Y_t} \int_0^{\hat{y}_t(p)} y\Phi(dy),$$

where, since the population has measure 1, Y_t is the mean income for the economy.

Definition : Income distribution $\Phi_1(\cdot)$ **Lorenz dominates** income distribution $\Phi_2(\cdot)$ if

$$Z_1(p) \geq Z_2(p) \quad \forall\, p.$$

That is, the entire Lorenz curve of distribution $\Phi_1(\cdot)$ lies below that of $\Phi_2(\cdot)$. Lorenz dominance is a sufficient condition for a reduction in inequality as measured by most commonly used measures of inequality. However, Lorenz dominance is a demanding criterion and does not allow inequality comparisons amongst income distributions that have Lorenz curves which intersect. Still, one can show that

Proposition 12 : *There exists a $\delta > 0$ such that if*
(A) No excess skewness: $\frac{h(x)-h(\underline{x})}{x-\underline{x}} < \delta \;\forall\, x \in [\underline{x}, x^m(\underline{w})]$
then there exists a time period $t^* < \tau_2$, such that for all $t \leq t^*$, income inequality increases in the Lorenz dominance sense.

Figure 16 shows a typical Lorenz curve of the income distribution during the dual stage of development. The linear segment OA corresponds to the traditional sector, the linear segment AB corresponds to wage laborers and the convex segment BC corresponds to entrepreneurs.

Figure 16 goes here

Since there is strictly positive migration away from the traditional sector, its share of aggregate income must decline. This alone implies that inequality cannot decline unambiguously during the dual stage of development. Two conditions are required to ensure that the Lorenz curve shifts from $OABC$ to $OA'B'C$ as illustrated and hence, that inequality increases in the Lorenz dominance sense. First, the growth rate in per capita income must exceed the growth rate in the income of agents migrating from the rural to the urban sector. This is trivially true in the first period and holds in subsequent periods provided that the growth rate in the average incomes of agents in the advanced sector is sufficiently high (i.e. the returns to wealth are sufficiently high and $\underline{w}$ is sufficiently low). This always holds if the cost of urban living, ν, is zero. Second, the profit of the wealthiest and most efficient entrepreneur must increase faster than the mean income in the economy. Again this is certainly the case in the first period and is true during the early stages of development as long as the returns to wealth are sufficiently high and $\underline{w}$ is sufficiently low. Eventually, however, the maximum income in the economy ceases to grow more rapidly than the mean. In particular, we know that prior to the advanced stages of development, the maximum income stops growing, whereas the mean continues to rise.

Figure 17 goes here

During the advanced stages of economic development, the Lorenz curve no longer includes the segment corresponding to the traditional sector. The maximum income in the economy falls and the mean rises, so the slope at C must decline. A necessary and sufficient condition to rule out intersecting Lorenz curves so that inequality declines unambiguously, is that the wage grows faster than per capita income (see Figure 17a):

$$\frac{w_t}{Y_t} \geq \frac{w_{t-1}}{Y_{t-1}}.$$

This holds if the sum of the elasticities of the demand for, and supply of, labor is sufficiently large. Observe in Figure 17b that for a given increase in the equilibrium wage, the associated increase in net aggregate income (represented by the shaded area) is less, the more elastic are the supply of, and demand for, labor. If the sum of these elasticities is sufficiently high, then an increase in the wage has a relatively large negative impact on the number of entrepreneurs and on their labor demands. Then, the negative impact of an increase in the wage on aggregate profits will be relatively large, so that the increase in per capita net income will be small relative to the wage increase. Provided that individual firm labor demand curves are sufficiently elastic and that the marginal product of capital is sufficiently low for all wealth levels, the wage rises more rapidly than average income:

Proposition 13 : *There exists $\theta < 0$ such that if, in addition to conditions (A) and (B), the sum of the wage elasticities of supply of and demand for, labor is sufficiently high:*

$$\left(\frac{\underline{w}}{w-\underline{w}}\right)\int_{\underline{w}}^{w}\left[\eta_w(b,\hat{w})+\lambda_w(b,\hat{w})\right]d\hat{w} > \theta \qquad \forall\, b, w,$$

then, for all $t > \tau^2$, income inequality declines in the Lorenz dominance sense.

Lorenz dominance is a demanding criterion. A less demanding measure of inequality is the Gini coefficient, which is twice the area between the Lorenz curve and the 45^o line[7] . Lorenz dominance is sufficient, but

[7] Hence, it ranges from 0 to 1, where 1 corresponds to maximum inequality.

not necessary, for declining inequality as measured by the Gini coefficient. Clearly, from Proposition 13, the Gini coefficient increases until period t^* and, under the conditions of Proposition 14, declines after period τ_2. We consider what happens to it between t^* and τ^2 in the context of our example. Figure 18(a) depicts the evolution of the Gini coefficient. The richest agents become unconstrained during the dual development stage, so that their incomes stop growing. Meanwhile, the fraction of agents who have left subsistence and receive at least the wage $\underline{w}$ continues to grow. As a result, the turning point occurs during the dual stage of development, before the wage actually starts to rise. For comparison with Figure 1, Figure 18(b) illustrates the relationship between the ratio of the income share of the richest 20% to that of the poorest 20% and the level of per capita income.

Figure 18 goes here

These results can be contrasted with those of Anand and Kanbur (1993), who analyze the path of inequality during the process of intersectoral migration. Stage 1 of the process described here is an endogenous generalization of the 'Kuznets process' that they discuss. At the beginning of the development process inequality rises unambiguously, but at the point at which migration stops, the direction of change in inequality is ambiguous. However, by analyzing the endogenous process which occurs after the migration period ends, we derived weak conditions that ensure inequality falls.

A final result related to income inequality relates to the notion of Generalized Lorenz Dominance introduced by Shorrocks (1983).

Income distribution $\Phi_1(\cdot)$ is said to **generalized Lorenz dominate** distribution $\Phi_2(\cdot)$ if

$$Y_1 Z_1(p) \geq Y_2 Z_2(p) \quad \forall\, p.$$

If welfare is measured as the sum of individual utility functions then generalized Lorenz dominance is a sufficient condition for increased welfare.

Proposition 14 : *There exists a $\delta > 0$ and a $\sigma > 0$ such that if*
(A) No excess skewness: $\frac{h(x)-h(\underline{x})}{x-\underline{x}} < \delta \;\forall\, x \in [\underline{x}, x^m(\underline{w})]$, *and*
(B) Weak classical effect: $\frac{B'(W)-B'(W^0)}{W-W^0} < \sigma \;\forall\, W$,
then the distribution of income in each period dominates that in the previous one in the Generalized Lorenz sense. Hence, social welfare as measured by a utilitarian social welfare function increases throughout the development process.

This follows directly from the fact that the income distribution in each period dominates that in the previous one in the first- or second-order stochastic sense. Thus, even though inequality increases during the dual development stage, the aggregate welfare of each generation is always greater than that of the last. Moreover, while many lineages become less well off in the advanced stages, aggregate welfare continues to grow.

8. Upward mobility

We now characterize the evolution of the ease with which poor lineages can climb the "social ladder". We capture the notion of upward mobility by tracking the probability that a child of one of the poorest agents in the economy makes it to the p^{th} percentile of the wealth distribution. At time t this probability is given by

$$\Pi_t(p) = 1 - P(\hat{b}_{t+1}(p)|\underline{b}_t, w_t),$$

where $\hat{b}_{t+1}(p)$ is the lowest wealth in the p^{th} percentile at time t.

Proposition 15 : *The probability with which the child of the poorest agent reaches the p^{th} percentile of the wealth distribution, $\Pi_t(p)$, is non-increasing for all p during Stage 1. There exists a percentile $p^* \in (0, 1)$ such that for all $p > p^*$, there exists a date $t(p)$, during the advanced stage of development, such that $\Pi_t(p)$ increases monotonically for all $t > t(p)$. Also, if $p_2 > p_1$ then $t(p_2) \leq t(p_1)$.*

In the dual development stage, as an increasing number of agents migrate into the urban sector, the distribution of wealth spreads out over time and the minimum wealth needed to reach any percentile gradually increases. Since the inheritance level of the poorest agent remains constant at b^0, it becomes less and less likely that such lineages can climb into the upper percentile groups.

Once the advanced stage of development is reached, for sufficiently high p, the wealth associated with the p^{th} percentile, $\hat{b}_{t+1}(p)$, declines (see Figure 19a). The rising wage makes it more likely that the children of the poorest agents will attain higher wealth levels. However, the increasing wage also makes wage laboring more desirable. This offsets the impact of the increasing inherited wealth, reducing the likelihood that the child of one of the poorest agents earns more than a given profit. As long as the distribution of wealth amongst the children of the least wealthy agents grows in the first-order stochastic sense, $\Pi_t(p)$ declines monotonically. Of course, it is possible that the bequest of the poorest, most efficient entrepreneur falls. But even then, the proportion of the children of poorer agents that reach the $p-th$ percentile increases *relative* to that of wealthier agents. This is because they are more constrained on average, so that a higher wage has a smaller effect on their expected entrepreneurial profits.

Finally, consider any two percentiles p_1 and p_2, such that $p_2 > p_1$. If a lineage reaches percentile p_1 it must also have reached percentile p_2. It follows that if the probability of reaching percentile p_1 is rising, so must the probability of reaching percentile p_2. As a result, upward mobility relative to percentile p_2 must start increasing at least as early as upward mobility relative to percentile p_1: $t(p_1) \geq t(p_2)$.

Figure 19(b) plots the time path of $\Pi_t(p)$ for the example with $p = 0.8$ and $p = 0.9$. When a sufficient proportion of lineages have been entrepreneurs at least once, upward mobility falls to zero rapidly and remains there for the rest of the dual development stage. The smaller is p, the longer it takes for 'upward mobility' to fall. In the advanced stage, once the economy has evolved sufficiently far along its development path, $\Pi_t(p)$ rises gradually again to its long-run value. The smaller is p, the greater the time taken for upward mobility to rise again in the advanced stage.

Figure 19 goes here

9. The optimal distribution of wealth

Due to the capital market imperfections, the economy always operates within its production possibilities frontier, unless the stage of maximum efficiency is reached. Therefore total income in the economy could be increased by redistributing inherited wealth. We now consider the nature and effects of such redistributions across agents in the economy in a single time period. We assume that the social planner maximizes aggregate output by re-allocating wealth prior to agents learning the quality of their projects:

$$\max_{\mu_t(b)\geq 0} \int \psi(b, w_t)\mu_t(b)db$$

$$\text{s.t.} \quad \int b\mu_t(b)db \leq \int bG_t(db); \qquad \int \mu_t(b)db = 1.$$

Proposition 16 : *Under the conditions stated in Proposition 6, there exists a time period s, where $0 \leq s < \tau_1$, such that*

(a) if $t < s$, the optimal distribution is a two-point distribution with some agents receiving nothing and the rest receiving a positive inheritance level b^*,

(b) if $t > s$, the optimal distribution is such that all agents receive the same inheritance,

(c) if the stage of maximum efficiency is attained, then no redistribution of inheritances raises aggregate output.

The social planner never allocates different positive levels of wealth to agents, but may give some agents no wealth at all. The social planner chooses the wealth level that maximizes expected output per unit of wealth. Note that s may be 0. By assumption, the distribution of set-up costs is not too skewed to the right, so that in the advanced stage of development, a second-order redistribution of inheritances (of which a mean-preserving squeeze is a special case) always increases output. However, in earlier stages of development, the marginal product of capital may be high, even for the richest entrepreneurs. At these stages of development, the gains in aggregate income arising from redistribution come from two sources which act in opposite directions. The concavity of the production function implies that the marginal product of capital is greater for small entrepreneurs than large. However, an increase in the wealth of an agent increases the likelihood of his becoming an entrepreneur, perhaps raising output, so that inegalitarian distributions of wealth may still be optimal.

Let b^* be the wealth level that maximizes output per unit of wealth (see Figure 20a). Ceteris paribus, the more skewed the distribution of set-up costs toward high cost projects or the less concave the production technology, the greater is b^* and hence the more inegalitarian is the optimal distribution of wealth. If per capital wealth is less than b^*, then it is always possible to increase net income via a redistribution which allocates b^* to some agents and nothing to the rest. Once per capita wealth exceeds b^*, then a perfectly egalitarian distribution is optimal. Were this not the case then the optimal distribution would be at least a two-point distribution with some agents receiving more than b^*. But then it would be possible to increase net income by giving all agents the same wealth.

Figure 20 goes here

Figure 20b illustrates the optimal path for aggregate net income for the example economy and compares it to the actual path. In this example, where $m = 0.2$, an inegalitarian policy is optimal until period 8, after which the optimal policy switches to an egalitarian one. Note that if $m \leq 0$, the optimal distribution would be egalitarian throughout the development cycle.

10. Aggregate Shocks

In this section, we consider the impact of aggregate exogenous shocks (e.g. an 'oil price shock') at different stages of development. In particular, we consider the response of aggregate variables to a one-time equiproportional reduction in the real wealth of all lineages. That the economy eventually returns its "no-shock" development path follows from the fact that the limiting distribution of wealth is independent of the economy's history. We characterize the results numerically assuming a uniform distribution of start–up costs, although the qualitative nature of the results are robust to the functional forms.

Consider a one-time 30% reduction in the inherited wealth of all agents. The absolute reduction in wealth is therefore greater in a more developed economy. Figures 21 (a)-(d) illustrate the effects on aggregate income and wealth of shocks occurring in representative periods $t = 4$, $t = 8$, $t = 12$ and $t = 16$, that correspond to the different stages of the development process. They illustrate the general finding that:

The immediate impact of an equiproportional negative wealth shock on per capita income increases during the dual development stage, then reaches a peak and declines during the advanced stage of development.

This is true in both absolute terms and as a percentage of the "no-shock" level of per capita income. As the economy develops, both the fraction of constrained entrepreneurs and the marginal product of their wealth declines. In the advanced stages of development this implies that the immediate effect of a reduction in inherited wealth on per capita income declines with the level of development. During the dual development stage, the income of non-entrepreneurs is unaffected by the reduction in inherited wealth. As the number of entrepreneurs increases, the negative effect of the shock on aggregate income also increases, offsetting the declining effect of the reduction in the fraction of constrained entrepreneurs.

Figure 21 goes here

Figures 21 (a) to (d) also illustrate that

The time taken to recover from an equiproportional wealth shock declines with the level of economic development.

This result obtains even though the wealth shock is bigger in absolute terms for a more developed economy. The earlier the stage of development, the greater the effect on both the decision to become an entrepreneur and the scale of production. Consider an unconstrained entrepreneur. Losing a sufficiently small portion of wealth has no effect on his decisions. In contrast, every unit that a constrained entrepreneur loses affects both his scale and his occupational decision. Earlier in development, more agents are constrained and, further, they are constrained more severely. As a result, the consequences of the shock are worse, it takes longer for them to recover and the aggregate impact of the shock persists for longer. In fact, the gap between actual aggregate income and the level which would have occurred in the absence of a shock, widens for a few periods before narrowing again.

The later the stage of development at which the shock occurs, the smaller is the resulting total loss in output. However, this does not account for the fact that the absolute size of the shock increases with the level of development. A better measure is the multiplier which is equal to the ratio of the overall loss in output to the initial reduction in wealth:

The multiplier effect of a negative wealth shock on aggregate income declines exponentially with the level of economic development.

Again, Figure 21b illustrates this result quite clearly. The multiplier effect on income is represented by the area above the time paths.

11. Convex technology and endogenous cycles

Throughout our analysis we have assumed sufficient structure for the equilibrium wage to rise with time. This, in turn, is sufficient to generate the development cycle that we characterize. A natural question to ask is what happens if sufficient condition (A), that the distribution of start–up costs is not too skewed towards high cost projects, is not met. During the dual-economy and transitional stages of development, the qualitative results with respect to both aggregate variables and distributions are independent of the assumed technology. In particular, until some time τ^2, the distribution of inheritances again grows in the first-order stochastic sense.

However, if the distribution of start–up costs is sufficiently skewed towards high cost projects, then eventually the rising wage has adverse effects on aggregate production. In particular, the rising wage reduces the incomes and inheritances of the richest types. If, for most projects, the efficient scale of production requires large inheritances to finance the set up costs and capital, then the consequence of a higher wage in the previous period may be to reduce the level of efficient production in the economy, and even to reduce the number of entrepreneurs, lowering output and welfare. In turn, the equilibrium wage falls. As the wage falls both the increase in the wealth of poorer agents and the decrease in the wealth of the richest agent slows. Eventually, the distribution of inheritances declines in the second-order stochastic sense. With convexity in the functions relating the demand for labor and the supply of entrepreneurs to inheritances, this in turn, results in an increase in the demand for labor and a decrease in the supply. Then, the equilibrium wage and aggregate income again rise and the economy's decline is reversed. As long as average inherited wealth does not fall too far, the wage at which this reversal occurs exceeds $\underline{w}$. However, it is even possible that the economy reverts to the dual stage before recovering.

If condition (A) is violated to the extent that wages do not rise monotonically, then the economy exhibits endogenous cycles in aggregate production, and wages evolve pro–cyclically.

Figure 22 shows the path followed by aggregate income in the example for different values of m (the slope of the density function for start-up costs). The actual long-run path followed by per capita income — the periodicity and magnitude of these cycles — is sensitive to changes in parameter values. For example, as the slope of the density of start-up costs is increased, the long run economy goes from exhibiting no cycles to two-period cycles, then to four-period cycles and back to no cycles before exhibiting three-period cycles and two-period cycles again.

Figure 22 goes here

In this cyclical economy, utilitarian welfare does not increase monotonically and the distributions of income and wealth exhibit cyclical levels of inequality. Moreover, because the economy always operates within its production possibilities boundary, government intervention to redistribute income can always raise the equilibrium wage, output and aggregate income. With such intervention, the evolution of the economy would follow the development cycle that occurs when sufficient condition (A) holds.

12. Concluding Remarks

This paper demonstrates how enterprise in the face of financing constraints can explain empirical regularities associated with the size distribution of income during the process of economic development. In particular, we provide a theoretical rationale for the Kuznets' curve and explain the associated movements in the Lorenz curve. Moreover, we chart the evolution of the distribution of firm sizes, the degree of social mobility, the optimal distribution of wealth and the impact of aggregate shocks at different stages of development. We offer insight into why the Kuznets curve may appear in some economies but not others, and show how long-run fluctuations in economic activity may arise endogenously.

We should note that in our model, agricultural and manufactured products are not traded across sectors. If instead, they could be traded in a perfect goods market, real–wage equalization across sectors would occur. This would break the duality arising in Stage 1, so that only a non–dual equilibrium would be possible. In this case, the economy would evolve in the way described in the later stages of the development process.

Finally, in the model described here, the distribution of agent qualities is time invariant so the only source of economic growth comes from the relaxation of borrowing constraints as agents accumulate wealth. The optimal scale of production is bounded, so the economy either reaches a steady state output with zero growth or evolves in a regular cyclical manner. In a related paper, Lloyd-Ellis (1993) looks at the relationship between income inequality and sustained growth. He develops an endogenous growth model with human capital accumulation in which the interactions between long–run growth, inequality and government policy are detailed.

References

Adreoni, J. (1989), 'Giving with impure altruism: Applications to charity and Ricardian equivalence', *Journal of Political Economy*, vol. 97, pp. 1447-1458.

Ahluwalia, M.S. (1974), 'Income distribution and development: some stylized facts', *American Economic Review*, vol. 66, pp.128-135

Anand, S. and S. Kanbur, (1993), 'The Kuznets process and the inequality-development relationship', *Journal of Development Economics*, vol. 40, pp. 25-52.

Banerjee, A.V. and A.F. Newman (1993), 'Occupational choice and the process of development', *Journal of Political Economy*, vol. 101, (2), pp. 274-298.

Bernheim, B.D. (1991), 'How strong are bequest motives? Evidence based on estimates of the demand for life insurance and annuities', *Journal of Political Economy*, vol. 99, (5), pp.899-927

Blanchflower, D.G. and A.J. Oswald (1990), 'What makes a young entrepreneur?', NBER Working Paper No. 3252.

Bourguignon, F. (1990), 'Growth and inequality in the dual model of development: The role of demand factors', *Review of Economic Studies*, vol. 57, pp. 215-228.

Evans, D. and B. Jovanovic (1989), 'An estimated model of entrepreneurial choice under liquidity constraints', Journal of Political Economy, vol. 97, (4), pp.807-827.

Fei, J.C.H. and G. Ranis (1966), 'Agrarianism, dualism and economic development', in I. Adelman and E. Thorbecke (eds.), *The Theory and Design of Economic Development*, Baltimore, John Hopkins Press, pp.3-44.

Galor, O. and J. Zeira (1993), 'Income distribution and macroeconomics', *Review of Economic Studies*, vol. 60, (1), pp. 35-52.

Greenwood, J. and B. Jovanovic (1990) 'Financial development, growth and the distribution of income', *Journal of Political Economy*, vol. 98, (5), pp. 1076-1107.

Hopenhayn, H. and E.C. Prescott (1992), 'Stochastic monotonicity for dynamic economies', *Econometrica*, vol. 60, (6), pp.1387-1406.

Kuznets, S. (1955), 'Economic growth and income inequality', *American Economic Review*, vol. 45, (1), pp.1-28.

Lloyd-Ellis, H. (1993), 'Enterprise, education and the distribution of gains from growth', unpublished manuscript, University of Toronto.

Lewis, A.W. (1954), 'Economic development with unlimited supplies of labor', *Manchester School of Economics and Social Studies*, vol. 22, pp.139-151.

Lindert, P.H. and J.G. Williamson (1985) 'Growth, equality and history', *Explorations in Economic History*, vol. 22, pp. 341-77.

Lydall, H. (1979), *A Theory of Income Distribution*, Oxford, Clarendon Press.

Newman, A.F. (1991), 'The capital market, the wealth distribution and the employment relation', mimeo, Northwestern University.

Paukert, F. (1973), 'Income Distribution at different levels of development: a survey of evidence', *International Labor Review*, vol. 108.

Romer, P.M., (1986), 'Increasing returns and long-run growth', *Journal of Political Economy*, Vol. 94, pp.1002-37.
Schumpeter, J.A. (1934) *The Theory of Economic Development*, Cambridge, Mass., Harvard University Press.
Shorrocks , A.F. (1983), 'Ranking income distributions', *Economica*, vol. 50, pp. 3-17.
Summers, R., I.B. Kravis and A. Heston (1984), 'Changes in the world income distribution', *Journal of Policy Modeling*, Vol. 6, pp.237-269
Townsend, R. M. (1979), 'Optimal contracts and competitive markets with costly state verification', *Journal of Economic Theory*, vol. 21, pp. 265-293.
Williamson, J.G. (1985), *Did British Captalism Breed Inequality*, Boston : Allen & Unwin.

Technical Appendix

Lemma 1 : $l(b, x, w_t)$ and $k(b, x, w_t)$ are weakly increasing in b, weakly decreasing in x and weakly decreasing in w_t.

Proof : The Lagrangian for the firm's constrained optimization problem is

$$\ell(b, x, \varphi) = \max_{k,l} f(k, l) - wl - k + \varphi[b - x - k]$$

where φ is the associated Lagrange multiplier. The associated first-order conditions are

$$\frac{d\ell}{dk} = f_k(k, l) - 1 - \varphi = 0; \quad \frac{d\ell}{dl} = f_l(k, l) - w = 0; \quad \frac{d\ell}{d\varphi} = b - x - k = 0.$$

Total differentiation in the constrained case yields the following:

$$k_b(b, x, w) = -k_x(b, x, w) = 1, \qquad l_b(b, x, w) = -l_x(b, x, w) = -\frac{f_{kl}}{f_{ll}}$$

$$k_w(b, x, w) = 0, \qquad l_w(b, x, w) = \frac{1}{f_{ll}}$$

In the unconstrained case (i.e. $\varphi = 0$), b and x do not affect demands:

$$k_b(b, x, w) = k_x(b, x, w) = l_b(b, x, w) = l_x(b, x, w) = 0,$$

$$k_w(b, x, w) = \frac{-f_{kl}}{\Delta}, \qquad l_w(b, x, w) = \frac{f_{kk}}{\Delta},$$

where, given strict concavity, $\Delta = f_{kk}f_{ll} - f_{kl}^2 > 0$, $f_{kk} < 0$, $f_{ll} < 0$, so the results hold. ■

Lemma 2 : $\pi(b, x, w)$ is weakly increasing in b, strictly decreasing in x and strictly decreasing in w.

Proof: Differentiating and applying the relevant envelope conditions yields :

$$\pi_b(\cdot) = f_k - 1 \geq 0; \quad \pi_x(\cdot) = -f_k \leq -1; \quad \pi_w(\cdot) = -l(b, x, w) < 0.$$

If the wealth constraint binds then $f_k > 1$. If not then $f_k = 1$. ■

Lemma 3 : $\tilde{x}(y, b, w_t)$ is strictly decreasing and convex in y, weakly increasing and concave in b and strictly decreasing in w_t. Also, $\tilde{x}_{yb} < 0$, $\tilde{x}_{yw} < 0$ and $\tilde{x}_{ybb} > 0$.

Proof : $\tilde{x}$ is implicitly defined by $\pi(b, \tilde{x}, w) = y$. Total differentiation and substitution using Lemma 2 yields the following :

$$\tilde{x}_y = \frac{1}{\pi_x} = -\frac{1}{f_k} < -1,$$

$$\tilde{x}_b = -\frac{\pi_b}{\pi_x} = \frac{f_k - 1}{f_k} \geq 0,$$

$$\tilde{x}_w = -\frac{\pi_w}{\pi_x} = -\frac{l(b, \tilde{x}(y, b, w), w)}{f_k} < 0.$$

Further differentiation yields the second-order differentials

$$\tilde{x}_{yy} = \tilde{x}_{bb} = \tilde{x}_{yb} = \frac{f_{kk} f_{ll} - f_{kl}^2}{f_k^3 f_{ll}} < 0,$$

$$\tilde{x}_{bw} = \tilde{x}_{yw} = \frac{f_k f_{lk} + l(b, \tilde{x}, w)(f_{kk} f_{ll} - f_{lk}^2)}{f_k^3 f_{ll}} < 0,$$

where the negativity follows from the strict concavity of the production function. Finally, differentiating $\tilde{x}_{yb}$ w.r.t. b yields

$$\tilde{x}_{ybb} = -\frac{3[f_{kk} f_{ll} - f_{lk}^2]^2}{f_k^4 f_{ll}} + o(\Delta) > 0,$$

because third-order derivatives of $f(\cdot)$ are assumed to be of negligible magnitude. ■

Lemma 4 : $z(b, w)$ is increasing and concave in b, decreasing in w and $\lim_{b \to \infty} z_b(b, w) = 0$.

Proof : Let $\tilde{b}(w_t)$ denote the inheritance below which the marginal entrepreneur is constrained on the extensive margin, so that $x^m(\tilde{b}, w_t) = \tilde{b}$. Let $\hat{b}(w_t)$ denote the inheritance level above which the marginal entrepreneurs are unconstrained, so that $x^m(\hat{b}, w_t) = \hat{x}^m(w_t)$. In general, $x^m(b, w)$ is implicitly defined by $\pi(b, x^m, w) = w$. Totally differentiating and applying the relevant envelope conditions yields

$$z_w(b, w_t) = \begin{cases} 0 & \text{if } b < \tilde{b}(w_t) \\ -\frac{1 + l^c(b, z(b, w_t), w_t)}{f_k} & \text{if } \tilde{b}(w_t) < b < \hat{b}(w_t) \\ -(1 + l^u(w_t)) & \text{if } b > \hat{b}(w_t) \end{cases}$$

$$z_b(b, w_t) = \begin{cases} 1 & \text{if } b < \tilde{b}(w_t) \\ \frac{f_k - 1}{f_k} < 1 & \text{if } \tilde{b}(w_t) < b < \hat{b}(w_t) \\ 0 & \text{if } b > \hat{b}(w_t) \end{cases}$$

$$z_{bb} = \begin{cases} \frac{f_{kk} f_{ll} - f_{kl}^2}{f_{ll} f_k^3} < 0 & \text{if } \tilde{b}(w_t) < b < \hat{b}(w_t) \\ 0 & \text{otherwise} \end{cases}.$$

Since $z_b(b, w) = 0$ for all $b > \hat{b}(w)$, it follows that $\lim_{b \to \infty} z_b(b, w) = 0$.

Definition 1 : Distribution $F_1(\cdot)$ dominates $F_2(\cdot)$ in the first-order stochastic sense if

$$\int_{x\leq y} F_1(dx) \leq \int_{x\leq y} dF_2(dx) \quad \forall\, y$$

This is written $F_1(\cdot)$ FSD $F_2(\cdot)$.

Theorem 1 : Consider any variable which takes the form $M_t = \int \mu(b)G_t(db)$, where $\mu_b(b) \geq 0 \;\forall\, b$; $\mu_b(b) > 0$ on some interval. If $G_t(b)$ FSD $G_{t-1}(b)$ then $M_t > M_{t-1}$.

Proof : $M_t - M_{t-1} = \int \mu(b)G_t(db) - \int \mu(b)G_{t-1}(db)$. Partially integrating,

$$M_t - M_{t-1} = \int \mu_b(b)[G_{t-1}(b) - G_t(b)]\, db > 0. \qquad \blacksquare$$

Lemma 5: If $G_t(\cdot)$ FSD $G_{t-1}(\cdot)$ and $w_t = \underline{w}$ then
(a) $\Phi_t(\cdot)$ FSD $\Phi_{t-1}(\cdot)$,
(b) $F_t(\cdot)$ FSD $F_{t-1}(\cdot)$, and
(c) $G_{t+1}(\cdot)$ FSD $G_t(\cdot)$.

Proof : The change in the distribution of income for any y can be expressed as

$$\Phi_t(y) - \Phi_{t-1}(y) = \int \Phi_b(y|b, \underline{w})[G_{t-1}(b) - G_t(b)]\, db.$$

But $\Phi_b(y|b,\underline{w}) = -h(\tilde{x}(y,b,\underline{w})\tilde{x}_b$, and from Lemma 3, $\tilde{x}_b \geq 0$. Thus, $\Phi_b(y|b,\underline{w}) \leq 0$ and, hence, part (a) follows from Theorem 1. Parts (b) and (c) are analogous. ■

Proof of Proposition 1 : $E_t(\underline{w})$, $X_t(\underline{w})$, $L_t(\underline{w})$, $K_t(\underline{w})$, $Q_t(\underline{w})$ and $Y_t(\underline{w})$ are all expected values of increasing functions of b and, hence, are all of the same form as M_t in Theorem 1. Since, from Lemma 5, $G_t(\cdot)$ FSD $G_{t-1}(\cdot)$ implies that $G_{t+1}(\cdot)$ FSD $G_t(\cdot)$ and $G_1(\cdot)$ FSD $G_0(\cdot)$, the assertions of Proposition 1 follow by induction. ■

Proof of Proposition 2 : A necessary condition for migration in the rural-urban direction only is that the increase in the urban population exceed the measure of agents whose parents were in subsistence that become entrepreneurs :

$$E_t(\underline{w}) + L_t(\underline{w}) - E_{t-1}(\underline{w}) - L_{t-1}(\underline{w}) \geq \eta_{t-1}(b^0, \underline{w})(1 - E_{t-1}(\underline{w}) - L_{t-1}(\underline{w})).$$

We can decompose the urban population into two groups: those who inherit wealth b^0 at time $t-1$ and those who inherit more. Proposition 1 found that the increase in the urban population from $t-1$ to t resulting from enterprise by lineages that had inheritance b at $t-1$, is non-negative because borrowing constraints are relaxed. Thus, a sufficient condition for rural-to-urban migration is that the increase in the urban population resulting from enterprise by agents with inheritance b^0 exceeds the measure of agents with inheritance b^0 undertaking projects (since this includes those whose parents were in subsistence):

$$\Big[\int [\eta(b,\underline{w}) + \lambda(b,\underline{w})]P(db|b^0) - [\eta(b^0,\underline{w}) + \lambda(b^0,\underline{w})]\Big] G_{t-1}(b^0) \geq \eta(b^0,\underline{w})G_t(b^0)$$

But $G_t(b^0) = (1-\eta(b^0,\underline{w}))G_{t-1}(b^0)$ and so, canceling on both sides leaves

$$E_1(\underline{w}) + L_1(\underline{w}) - E_0(\underline{w}) - L_0(\underline{w}) \geq E_0(\underline{w})(1 - E_0(\underline{w})),$$

which is true by assumption. Since, from Lemma 6, $G_t(\cdot)$ **FSD** $G_{t-1}(\cdot)$ implies that $G_{t+1}(\cdot)$ **FSD** $G_t(\cdot)$ and $G_1(\cdot)$ **FSD** $G_0(\cdot)$, Proposition 2 follows by induction. ∎

Lemma 6 : There exists a $\delta > 0$ such that if
(A) No excess skewness: $\frac{h(x)-h(\underline{x})}{x-\underline{x}} < \delta \ \forall \ x \in [\underline{x}, x^m(\underline{w})]$
then, there exists a $b^{**}(w)$ such that

$$\psi_{bb}(b,w) > 0, \ \forall \ b < b^{**}(w); \quad \psi_{bb}(b,w) < 0, \ \forall \ b > b^{**}(w).$$

Proof : The expected income of an agent with inheritance b is given by $\psi(b,w)$. Differentiating with respect to b yields

$$\psi_b(b,w) = \int_0^{z(b,w)} [f_k(b,x,w) - 1]h(x)\,dx.$$

Differentiating again with respect to b gives

$$\psi_{bb}(b,w) = \int_{\underline{x}}^{z(b,w)} [f_{kk} + f_{kl}l_k(w)]h(x)\,dx + [f_k(\cdot) - 1]h(z(b,w))z_b(b,w).$$

Suppose the distribution of start-up costs were uniform. Then differentiating a third time, assuming zero third-order derivatives of the production function, yields

$$\begin{aligned}\psi_{bbb}(b,w) =& [f_{kk} + f_{kl}l_k(w)]z_b(b,w) + [f_k(b,z(b,w),w) - 1]z_{bb}(b,w) \\ &+ [f_{kk} + f_{kl}l_k(w)][1 - z_b(b,w)]z_b(b,w) \leq 0.\end{aligned}$$

Since $\psi_{bbb}(b,w) \leq 0$, the result follows. Continuity implies this result holds for distributions of start-up costs with densities $h(\cdot)$, that are not too skewed toward high costs. ∎

Proof of Proposition 3 : This is a special case of X-Dispersion. In the case of the distribution of income, just let $y^* = \gamma$ and we have $1 > \Phi_{t-1}(\gamma) > \Phi_t(\gamma) > 0$ and

$$\Phi_t(y) \geq \Phi_{t-1}(y) \;\forall\; y < \gamma \quad \Phi_t(y) \leq \Phi_{t-1}(y) \;\forall\; y \geq \gamma.$$

A similar analysis holds for final wealth and bequests. Just set $W^* = \gamma + b^0$ and $b^* = b^0$. ■

Proof of Proposition 4 : The change in the distribution of firm sizes is

$$J_t(k) - J_{t-1}(k) = \int J_b(k|b)[G_{t-1}(b) - G_t(b)]\, db,$$

where

$$J_b(k|b, w_t) = \frac{H(b-k)}{H(z(b,w_t))}\left[\frac{h(z(b,w_t))z_b(b,w_t)}{H(z(b,w_t)} - \frac{h(b-k)}{H(b-k)}\right].$$

Were start-up costs uniformly distributed, then, since $z_b(b,w) \leq 1$ and $z(b,w) > b - k \;\forall\; k < k^u(w_t)$, this expression would be negative. Thus, by Theorem 1, $G_t(\cdot)$ FSD $G_{t-1}(\cdot)$ implies that $J_t(\cdot)$ FSD $J_{t-1}(\cdot)$. By continuity, this argument generalizes to cost distributions that are not too skewed. Since the lower support on firm sizes is fixed at $\underline{k}(b^0, \underline{w}) = b^0 - z(b^0, \underline{w})$ and a strictly positive measure of firms are of this size, the distribution of firm sizes must exhibit increasing X-Dispersion throughout the dual stage of development. ■

Lemma 7 : The area under the upper envelope of the supply and demand curves for labor is equal to aggregate net income.

Proof : Aggregate income is the sum of total profits and total wages,

$$Y_t(w_t) = \int \int_{\underline{x}}^{z(b,w_t)} \pi(b, x, w_t) h(x)\, dx G_t(db) + w_t L_t(w_t)$$

Using differentiability in the wage rate and the fact that $\pi(b, z(b,w), w) = w$,

$$Y_t(w_t) = \int_b \int_{w_t}^{\infty} \left[- \int_{\underline{x}}^{z(b,w)} \pi_w(b,x,w) h(x)\, dx - w h(z(b,w)) z_w \right] dw G_t(db)$$
$$+ w_t L_t(w_t).$$

From Lemma 2, $\pi_w(b,x,w) = -l(b,x,w)$, so that this can be reduced to

$$Y_t(w_t) = \int_b \int_{w_t}^{\infty} \left[\lambda(b,w) - w\eta_w(b,w)\right] dw G_t(db) + w_t L_t(w_t).$$

But $\frac{d(w\eta)}{dw} = w\eta_w + \eta$, and so

$$Y_t(w_t) = \int_b \int_{w_t}^{\infty} [\lambda(b,w) + \eta(b,w) - \frac{d(w\eta)}{dw}]\, dw G_t(db) + w_t E_t(w_t) + w_t L_t(w_t)$$

However, $-\int_{w_t}^{\infty} \left(\frac{d(w\eta)}{dw}\right) dw = w_t \eta(b, w_t)$, and so,

$$Y_t(w_t) = \int_b \int_{w_t}^{\infty} [\lambda(b,w) + \eta(b,w)]\, dw G_t(db) + w_t E_t(w_t) + w_t L_t(w_t),$$

which is the area under the upper envelope created by the supply and demand curves. ∎

Lemma 8: There exists a $\delta > 0$ such that if
(A) No excess skewness: $\frac{h(x)-h(\underline{x})}{x-\underline{x}} < \delta \ \forall\ x \in [\underline{x}, x^m(\underline{w})]$
then if $G_t(\cdot)$ FSD $G_{t-1}(\cdot)$ and $w_t > w_{t-1}$,
(a) $\Phi_t(\cdot)$ FSD $\Phi_{t-1}(\cdot)$ as long as $\bar{\pi}_t > \bar{\pi}_{t-1}$,
(b) $F_t(\cdot)$ FSD $F_{t-1}(\cdot)$ as long as $\bar{W}_t > \bar{W}_{t-1}$, and
(c) $G_{t+1}(\cdot)$ FSD $G_t(\cdot)$ as long as $\bar{b}_{t+1} \geq \bar{b}_t$.

Proof :

(a) Since $\Phi_t(y) = 0$ and $\Phi_{t-1}(y) \geq 0 \forall\ y < w_t$ it must be that $\Phi_t(y) \leq \Phi_{t-1}(y)\ \forall\ y < w_t$. For incomes $y > w_t$, decompose the change in the distribution of income as follows:

$$\begin{aligned}
\Phi_t(y) - \Phi_{t-1}(y) =& \int_b \Phi(y|b,w_t)G_t(db) - \int_b \Phi(y|b,w_{t-1})G_{t-1}(db)\\
=& \int_b \Phi(y|b,w_t)G_t(db) - \int_b \Phi(y|b,w_t)G_{t-1}(db)\\
&+ \int_b \Phi(y|b,w_t)G_{t-1}(db) - \int_b \Phi(y|b,w_{t-1})G_{t-1}(db)\\
=& \int_b [G_{t-1}(b) - G_t(b)]\Phi_b(y|b,w_t)\, db\\
&+ \int_b [\Phi(y|b,w_t) - \Phi(y|b,w_{t-1})]G_{t-1}(db)\\
=& \int_b [G_{t-1}(b) - G_t(b)]\Phi_b(y|b,w_t)\, db\\
&+ \int_b \int_{w_{t-1}}^{w_t} \Phi_w(y|b,w)\, dw G_{t-1}(db).
\end{aligned}$$

Since $\Phi_b(y|b,\underline{w}) = h(\tilde{x})\tilde{x}_b \leq 0$ and $G_t(\cdot)$ FSD $G_{t-1}(\cdot)$, the first term must be negative. Since $w_t > w_{t-1}$ and $\Phi_w(y|b,w) = -h(\tilde{x})\tilde{x}_w > 0$ from Lemma 3, the second term is positive. Hence, the sign of the expression is, in general, ambiguous. However, differentiating with respect to y, for $y \in (w_t, \bar{\pi}_t]$ the change in the slope of the distribution function is:

$$\phi_t(y) - \phi_{t-1}(y) = \int [G_{t-1}(b) - G_t(b)]\Phi_{yb}(y|b, w_t)\, db + \int \int_{w_{t-1}}^{w_t} \Phi_{yw}(y|b, w)\, dw G_{t-1}(db).$$

Were start-up costs uniformly distributed on $[0, 1]$, then from Lemma 3, $\Phi_{yb}(y|b, w_t) = -\tilde{x}_{yb} > 0$ and $\Phi_{yw}(y|b, w) = -\tilde{x}_{yw} > 0$. Hence, using Theorem 1, $\Phi_t(y) - \Phi_{t-1}(y)$ must increase in y on the interval $(w_t, \bar{\pi}_t]$. If $\bar{\pi}_t > \bar{\pi}_{t-1}$, this implies that the cumulative distribution functions do not intersect. It follows that $\Phi_t(\cdot)$ FSD $\Phi_{t-1}(\cdot)$.

Were the distribution of start-up costs skewed toward low-cost projects, then this result would still hold. However, were the distribution highly skewed toward high-cost projects then the change in the density of incomes would be ambiguous. The wage increase would sharply reduce the fraction of agents at low profit (high cost) levels. If this offsets the effect of the stochastic increase in inherited wealth, the density of incomes at these profit levels may fall. Continuity ensures that there exists a positive bound $\delta > 0$ on skewness such that the effect of the increase in inherited wealth dominates.

(b) For $W < w_t - \nu + \underline{b}_t$, clearly $F_t(W) \leq F_{t-1}(W)$. Consider wealths $W > w_t - \nu + \underline{b}_t$. Decompose the change in the distribution of final wealth as in (a):

$$F_t(W) - F_{t-1}(W) = \int [G_{t-1}(b) - G_t(b)]F_b(W|b, w_t)db + \int \int_{w_{t-1}}^{w_t} F_w(W|b, w)\, dw G_{t-1}(db).$$

Were start-up costs uniformly distributed, then $F_{bW} = 0$ and $F_{wW} > 0$ imply that $F_t(W) - F_{t-1}(W)$ increases in W. Hence, $\bar{W}_t > \bar{W}_{t-1}$ implies $F_t(W)$ FSD $F_{t-1}(W)$. By continuity, this result generalizes to cost distributions that are not too skewed.

(c) Decompose the change in the distribution of bequests as follows:

$$G_{t+1}(b') - G_t(b') = \int [G_{t-1}(b) - G_t(b)]P_b(b'|b, w_t)\, db + \int \int_{w_{t-1}}^{w_t} P_w(b'|b, w)\, dw G_{t-1}(db).$$

Were start-up costs uniformly distributed, then $P_{bb'} > 0$ and $P_{wb'} > 0$ imply that $G_{t+1}(b') - G_t(b')$ increases with b'. Hence, $\bar{b}_{t+1} > \bar{b}_t$ implies $G_{t+1}(b')$ FSD $G_t(b')$. By continuity, this result generalizes to cost distributions that are not too skewed. ■

Proof of Proposition 5 : It is sufficient to show that if $G_t(\cdot)$ FSD $G_{t-1}(\cdot)$ then the supply and demand schedules for labor shift up. This follows from Proposition 1 and the fact that there is no surplus labor in subsistence. Since $G_t(\cdot)$ FSD $G_{t-1}(\cdot)$, it follows that $w_t > w_{t-1}$ and $Y_t > Y_{t-1}$ for all $t \in [\tau^1, \tau^2]$. The change in aggregate final wealth is the sum of the changes in aggregate net income and aggregate inherited wealth, both of which are positive. ■

Definition 2 : Distribution $F_1(\cdot)$ dominates $F_2(\cdot)$ in the second-order stochastic sense if

$$\int_{-\infty}^{y} F_1(x)\,dx \leq \int_{-\infty}^{y} F_2(x)\,dx \quad \forall y.$$

This is written $F_1(\cdot)$ SSD $F_2(\cdot)$.

Theorem 2 : Consider any variable which takes the form $M_t = \int \mu(b)G_t(db)$, where $\mu_{bb}(b) \leq 0$, $\mu_{bb}(b) < 0$ for some b and $\lim_{b\to\infty} \mu_b(b) = 0$. If $G_t(\cdot)$ SSD $G_{t-1}(\cdot)$ then it follows that $M_t > M_{t-1}$.

Proof : $M_t - M_{t-1} = \int \mu(b)G_t(db) - \int \mu(b))G_{t-1}(db)$. Partially integrating twice yields

$$M_t - M_{t-1} = \Big[\mu_b(b) \int [G_{t-1}(\hat{b}) - G_t(\hat{b})]\,d\hat{b}\Big]_0^{\infty} + \int \mu_{bb}(b)\Big(\int_0^b [G_t(\hat{b}) - G_{t-1}(\hat{b})]d\hat{b}\Big)\,db.$$

Since $\lim_{b\to\infty} \mu_b(b) = 0$, the first term vanishes and hence the result follows. ■

Lemma 9 : Let $G_t(\cdot)$ SSD $G_{t-1}(\cdot)$ and $w_t > w_{t-1}$. There exists a $\delta > 0$ such that if

(A) No excess skewness: $\frac{h(x)-h(\underline{x})}{x-\underline{x}} < \delta \; \forall \; x \in [\underline{x}, x^m(\underline{w})]$

then

(a) if $Y_t > Y_{t-1}$ then $\Phi_t(\cdot)$ SSD $\Phi_{t-1}(\cdot)$,

(b) if total final wealth increases then $F_t(\cdot)$ SSD $F_{t-1}(\cdot)$, and

(c) if total bequeathed wealth increases then $G_{t+1}(\cdot)$ SSD $G_t(\cdot)$.

Proof : Consider again the change in the slope of the income distribution function:

$$\phi_t(y) - \phi_{t-1}(y) = \int [G_{t-1}(b) - G_t(b)]\Phi_{yb}(y|b, w_t)\, db + \int \int_{w_{t-1}}^{w_t} \Phi_{yw}(y|b, w)\, dw G_{t-1}(db).$$

Were start-up costs uniformly distributed, then if $w_t > w_{t-1}$, the second term is positive. Since $G_t(\cdot)$ SSD $G_{t-1}(\cdot)$ and, from Lemma 3, $\Phi_{ybb} = -\tilde{x}_{ybb} < 0$, Theorem 2 implies that the first term is also positive. Hence, $\Phi_t(y) - \Phi_{t-1}(y)$ increases with $y \in (w_t, \bar{\pi}_t]$, so $\bar{\pi}_t < \bar{\pi}_{t-1}$ implies the c.d.f.'s intersect only once:

$$\Phi_t(y) \leq \Phi_{t-1}(y) \text{ if } y < y_t^*,$$
$$\Phi_t(y) \geq \Phi_{t-1}(y) \text{ if } y > y_t^*.$$

If $Y_t > Y_{t-1}$, part (a), that $\Phi_t(\cdot)$ SSD $\Phi_{t-1}(\cdot)$ follows.

Analogous arguments establish parts (b) and (c). By continuity, these results hold for distributions of start-up costs that are not too skewed. ■

Proof of Proposition 6 : We first show that the supply and demand curves shift up if $G_t(b)$ SSD $G_{t-1}(b)$ so that both the wage and aggregate income must rise. Since, $\eta_b(b, w) = h(z(b, w))z_b$, it follows from Lemma 4 that $\lim_{b\to\infty} \eta_b(b, w) = 0$. Hence, as in Theorem 2, the change in the supply of entrepreneurs, $E_t(w_t)$, equals

$$E_t(w) - E_{t-1}(w) = \int \eta_{bb}(b, \underline{w}) \int_0^b [G_t(\hat{b}) - G_{t-1}(\hat{b})]\, d\hat{b}\, db.$$

If the distribution of set-up cost were uniform on [0,1] then

$$\eta_{bb}(b, w_t) = z_{bb}(b, w_t) \leq 0,$$

where the inequality is strict on a set of positive measure. By Theorem 2, the supply of entrepreneurs would increase for a given wage. By continuity, the supply of entrepreneurs would still increase provided that the distribution of start-up costs is not too skewed. In turn, the increase in the supply of entrepreneurs implies that the supply of labor must fall.

Since $\lim_{b\to\infty} \lambda_b(b, w) = 0$, the shift in the demand schedule for labor equals

$$L_t(w) - L_{t-1}(w) = \int \lambda_{bb}(b, \underline{w}) \int_0^b [G_t(\hat{b}) - G_{t-1}(\hat{b})]\, d\hat{b}\, db.$$

Were the distribution of set-up costs uniform, then

$$\lambda_{bb}(b,w_t) = -(1-z_b(b,w_t))^2 l_k(w_t) + z_{bb}(b,w_t) l(b,z(b,w_t),w_t) \leq 0$$

and, by Theorem 2, the demand schedule for labor would shift out. By continuity, this result generalizes to distributions of start-up costs that are not too skewed.

Although net income and therefore final wealth increase, this need not imply that aggregate bequeathed wealth rise. The change in total bequeathed wealth is

$$\int B(W)F_t(dW) - \int B(W)F_{t-1}(dW) = \int B'(W)[F_{t-1}(W) - F_t(W)]dW$$

By Theorem 2 this expression is positive so long as $B'(W)$ does not increase too rapidly at high wealth levels. By induction, using Lemma 9, the results of Proposition 6 follow. ■

Proof of Proposition 7 : Follows directly from Lemma 9. ■

Proof of Proposition 8 : We can decompose the overall change in the labor force as

$$L_t - L_{t-1} = \int \lambda_b(b,w_t)[G_{t-1}(b) - G_t(b)]db + \int\int_{w_{t-1}}^{w_t} \lambda_w(b,w)\,dw G_{t-1}(db)$$

where

$$\lambda_w(b,w) = \int_{\underline{x}}^{z(b,w)} l_w(b,x,w)h(x)\,dx + l(b,z(b,w),w)h(z(b,w))z_w(b,w).$$

The slope of the individual firm's labor demand curve is given by

$$l_w = \begin{cases} \frac{1}{f_{ll}} & \text{if } k < k^u(w_t) \\ \frac{f_{kk}}{f_{kk}f_{ll} - f_{kl}{}^2} & \text{otherwise.} \end{cases}$$

$|l_w|$, decreases with f_{ll} and so the slope of the demand curve for labor increases with f_{ll}, ceteris paribus. Hence, for sufficiently small f_{ll}, the equilibrium labor force declines. ■

Proof of Proposition 9 : The upper support on firm size during Stage 3 is the unconstrained optimal level $k^u(w_t)$. Since the wage increases over time, the upper support on firm size must fall. The lower support on firm size is $\underline{k}_t = \underline{b}_t - z(\underline{b}_t, w_t)$. Since both the lower support on inheritances, $\underline{b}_t$, and the wage rate rise, $\underline{k}_t$ must also rise over time. The question then is what happens to the distribution $J_t(\cdot)$ between these bounds.

Suppose the distribution of start-up costs were uniform. Then the c.d.f. for the distribution of firm sizes for $k \in [\underline{k}_t, k^u(w_t)]$ is given by

$$J_t(k) = \int \left[1 - \frac{b-k}{z(b, w_t)}\right] G_t(db),$$

which is linear in k. It follows immediately that the distributions $J_t(\cdot)$ and $J_{t-1}(\cdot)$ intersect only once on the interval $[\underline{k}_t, k^u(w_t)]$. By continuity, this argument generalizes to sufficiently uniform cost distributions. ■

Proof of Proposition 10 : The wage increases monotonically and is bounded above by $\bar{w}$. Hence, it must converge to some $w^* \in [\underline{w}, \bar{w}]$. In particular, for all $\epsilon > 0$ there must exist a T such that

$$w^* - w_t < \frac{\epsilon}{l^u(w_t)}, \; \forall\, t > T.$$

Since the l.h.s. decreases over time and the r.h.s. increases, if this inequality holds for $t = T$ it must hold for all $t > T$. Using the differentiability of $\Phi(y|b, w)$ in w, we have:

$$\Phi(y|b, w^*) - \Phi(y|b, w_t) = \int_{w_t}^{w^*} \Phi_w(y|b, w)\, dw, \quad \forall\, y, b.$$

Were the distribution of start-up costs uniform on $[0, 1]$, then from Lemma 5,

$$\Phi(y|b, w^*) - \Phi(y|b, w_t) = \int_{w_t}^{w^*} \frac{l(b, \tilde{x}(y, b, w), w)}{f_k(b, \tilde{x}(y, b, w), w)}\, dw.$$

But $l(b, \tilde{x}(y, b, w), w) \leq l^u(w_t)$, $\forall\, w > w_t$ and $f_k(b, \tilde{x}(y, b, w), w) \geq 1$, and so

$$\Phi(y|b, w^*) - \Phi(y|b, w_t) \leq (w^* - w_t) l^u(w_t).$$

Hence, for all $t \geq T$

$$\Phi(y|b, w^*) - \Phi(y|b, w_t) < \epsilon, \; \forall\, y, \; b.$$

The same analysis holds for parts (b) and (c). Continuity implies the argument extends to distributions of start-up costs that are not too skewed. ■

Proof of Proposition 11 : From proposition 10, the evolution of the distribution of inheritances converges to a stationary, monotone Markov process, $P(\cdot|b, w^*)$. Hopenhayn and Prescott (1992) detail conditions for this class of Markov processes which ensure that the limiting distribution is unique and invariant.

Define the probability that the n^{th} descendent of an agent with inheritance b receives an inheritance in the interval $[b_1, b_2]$ as $P^n([b_1, b_2]|b)$. The Monotone Mixing Condition requires that for each inheritance level $b^* \in [\underline{b}, \bar{b}]$ there exists $\varepsilon > 0$ and an N such that

$$P^N([\underline{b}, b^*]|\bar{b}) > \varepsilon \quad \text{and} \quad P^N([b^*, \bar{b}]|\underline{b}) > \varepsilon.$$

Consider the inheritance $b^* = \frac{\underline{b}+\bar{b}}{2}$. For sufficiently large N, there is a positive probability that the N^{th} descendant of an agent with $\bar{b}_t$ at time t receives an inheritance in any neighborhood of $\underline{b}$. To verify this, observe that for any $\varpi > 0$, there is positive probability that the child of such an agent realizes a start-up cost $x \in [\bar{x} - \varpi, \bar{x}]$. Similarly, with positive probability his grandchild will realize a start-up cost $x \in [\bar{x} - \varpi, \bar{x}]$ and so on. For ϖ sufficiently small, and after a sufficient number of generations, N, there is positive probability that his descendant will inherit a wealth level in any neighborhood of $\underline{b}$. Thus, for sufficiently large N, the probability that a lineage reaches the interval $[\underline{b}, b^*]$ is positive. By a similar argument, the probability that the N^{th} descendant of an agent with $\underline{b}_t$ receives an inheritance in the interval $[b^*, \bar{b}]$ is also positive, for N sufficiently large.

Thus, given any initial distribution at time $t \geq T$, $G_t(\cdot)$, the associated sequence converges:

$$\lim_{n \to \infty} G_{t+n}(\cdot) = \int \lim_{n \to \infty} P^n(\cdot|b) G_t(\,db) = P^*(\cdot) \qquad \blacksquare$$

Proof of Proposition 12 : The slope of the Lorenz curve at a given point is equal to the ratio of the income of the corresponding percentile of the distribution to the mean income of the distribution. The segment OA, in Figure 18, therefore has a constant slope $\frac{\gamma}{Y_t}$, the segment AB has a constant slope, $\frac{w}{Y_t}$, and the segment BC has a slope that is strictly increasing with p.

Consider the shift in the Lorenz curve during the dual stage of development from $OABC$ to $OA'B'C$ (see figure 18). Since there is strictly positive migration away from the traditional sector, its share of aggregate income must decline. Hence, the segment corresponding to the traditional sector must shift to OA', where A' lies above and to the left of A. This alone implies that inequality cannot decline unambiguously during the dual stage of development. Two observations can be made about the shift in the linear segment AB. First, its horizontal component increases because of the increase in wage laboring. Secondly, its slope must decrease since the wage does not change, but mean income rises. A necessary condition for

Lorenz dominance is that the slope of the line joining $A'A$ exceeds the slope of $A'B'$. This condition can be written as

$$\frac{Y_t - Y_{t-1}}{Y_{t-1}} > \frac{\nu}{\gamma}\left(\frac{S_{t-1} - S_t}{S_{t-1}}\right) \tag{C1}$$

A sufficient condition for (C1) to hold is that the growth rate in per capita income exceeds the growth rate in the income of agents migrating from the rural to the urban sector. Note that (C1) is always satisfied if the cost of urban living, ν, is zero. Since the proportion of agents who become entrepreneurs increases, the horizontal component of the segment BC also increases. Hence, if (C1) holds, the fact that the slope of AB exceeds that of $A'B'$ implies that the point B' must lie below and to the left of B.

Now consider the change in the convexity of the Lorenz curve segment BC:

$$Z_t''(p) - Z_{t-1}''(p) = \frac{\hat{y}_t'(p)}{Y_t} - \frac{\hat{y}_{t-1}'(p)}{Y_{t-1}} = \frac{1}{\phi_t(\hat{y}_t(p))Y_t} - \frac{1}{\phi_{t-1}(\hat{y}_{t-1}(p))Y_{t-1}}.$$

Observe that during the dual economy stage of development, if $\Phi_t(\cdot)$ FSD $\Phi_{t-1}(\cdot)$ then $\hat{y}_t(p) \geq \hat{y}_{t-1}(p)$. Now

$$\phi_t(\hat{y}_t) - \phi_{t-1}(\hat{y}_{t-1}) = \int \phi(\hat{y}_t|b,\underline{w})G_t(db) - \int \phi(\hat{y}_{t-1}|b,\underline{w})G_{t-1}(db)$$

$$= \int \phi_b(\hat{y}_t|b,\underline{w})[G_{t-1}(b) - G_t(b)]\,db + \int [\phi(\hat{y}_t|b,\underline{w}) - \phi(\hat{y}_{t-1}|b,\underline{w})G_{t-1}(db).$$

Were the distribution of start-up costs uniform, then $\phi_b(y|b) = -\tilde{x}_{yb}(y,b,\underline{w}) \geq 0$ and so, since $G_t(\cdot)$ FSD $G_{t-1}(\cdot)$, the first term is positive. Also, since $\hat{y}_t(p) > \hat{y}_{t-1}(p)$ and $\tilde{x}_{yy} < 0$,

$$\phi(\hat{y}_t|b) - \phi(\hat{y}_{t-1}|b) = \tilde{x}_y(\hat{y}_{t-1},b,\underline{w}) - \tilde{x}_y(\hat{y}_t,b,\underline{w}) \geq 0.$$

Hence, the second term is positive, so that $\phi_t(\hat{y}_t) > \phi_{t-1}(\hat{y}_t)$. Since $Y_t > Y_{t-1}$,

$$Z_t''(p) < Z_{t-1}''(p) \;\forall\; p \geq 1 - E_{t-1}.$$

By continuity, this result extends to cost distributions that are not too skewed.

It follows that if the slope at C is increasing, the Lorenz curves cannot intersect on this segment. This is the case if the profit of the wealthiest and most efficient entrepreneur increases faster than the mean income in the economy:

$$\frac{\pi(\bar{b}_t,\underline{x},\underline{w})}{Y_t} \geq \frac{\pi(\bar{b}_{t-1},\underline{x},\underline{w})}{Y_{t-1}} \tag{C2}$$

If (C1) holds throughout the dual stage of development, then t^* is the period in which the profit of the richest entrepreneur ceases to grow faster than the mean income. In all periods after t^*, the slope of the Lorenz curve at C must decline. As a result inequality cannot increase unambiguously after time t^*. The decreasing convexity of the Lorenz curves between time $t-1$ and time t implies that they intersect once and only once. If, in addition, (C1) ceases to hold, then the Lorenz curve intersects twice. ∎

Proof of Proposition 13 : During the advanced stages of economic development, the Lorenz curve no longer includes the segment corresponding to the traditional sector, OA and the slope at C must decline. Again it can be shown that the Lorenz curves intersect at most once for all $p \in [1-E_{t-1}, 1]$.

Consider the percentile p^* at which $\Phi_t(\cdot)$ and $\Phi_{t-1}(\cdot)$ intersect:

$$p^* = \Phi_t(y^*) = \Phi_{t-1}(y^*).$$

Since $\Phi_t(y) > \Phi_{t-1}(y)$ for all $y > y^*$, it must be true that for any $p > p^*$, $\hat{y}_t(p) < \hat{y}_{t-1}(p)$. Given that $Y_t(w_t) > Y_{t-1}(w_{t-1})$, this implies that

$$\frac{\hat{y}_t(p)}{Y_t} < \frac{\hat{y}_{t-1}(p)}{Y_{t-1}} \ \forall\, p > p^*.$$

But this says that the slope of the Lorenz curve must be less at time t than at time $t-1$ for all $p > p^*$. Since the Lorenz curves meet at $p = 1$, it follows that

$$Z_t(p) > Z_{t-1}(p) \ \forall\, p > p^*.$$

For $p < p^*$ we know that $\hat{y}_t(p) > \hat{y}_{t-1}(p)$, so the above argument cannot hold. Consider, now, the change in the density function for any $p < p^*$. This can be decomposed into:

$$\begin{aligned}\phi_t(\hat{y}_t) - \phi_{t-1}(\hat{y}_{t-1}) &= \int \phi_b(\hat{y}_t(p)|b, w_t)[G_{t-1}(b) - G_t(b)]\, db \\ &+ \int\int_{w_{t-1}}^{w_t} \phi_w(\hat{y}_t(p)|b, w)\, dw G_{t-1}(db) \\ &+ \int [\phi(\hat{y}_t|b, w_t) - \phi(\hat{y}_{t-1}|b, w_t)] G_{t-1}(db).\end{aligned}$$

Were the distribution of start-up costs uniform, then, as in Lemma 9, the first two terms must be positive and, as in Proposition 13, the third term must also be positive. Hence, the Lorenz curve at t must be less convex than at $t-1$:

$$Z_t''(p) < Z_{t-1}''(p) \ \forall\, p \in [1 - E_{t-1}, p^*].$$

Thus, the Lorenz curves can intersect at most once. If the supply and demand curves for labor are sufficiently wage elastic that $\frac{w_t}{Y_t} > \frac{w_{t-1}}{Y_{t-1}}$, then the Lorenz curves cannot intersect at all. By continuity, this argument generalizes to cost distributions that are not too skewed. ∎

Proof of Proposition 14 : Second-order stochastic dominance is a sufficient condition for Generalized Lorenz Dominance (see Shorrocks 1983). ∎

Proof of Proposition 15 : Part (a) follows from the first-order stochastic growth of the distribution of wealth and the fact that the lower support on bequests remains fixed. Part (b) is less trivial. Let p_t^* be the percentile at which the distribution of wealths at t and $t-1$ intersect. Then for $p > p_t^*$ it follows that $\hat{b}_{t+1}(p) < \hat{b}_t(p)$. Then, a sufficient condition for $\Pi_t(p) > \Pi_{t-1}(p)$ is that $P(b'|\underline{b}_t, w_t) \leq P(b'|\underline{b}_{t-1}, w_{t-1})$. Since $P_{b'b} > 0$ and $P_{b'w} > 0$ it follows that the slope of $P(b'|\underline{b}_t, w_t)$ increases with t. Since the fraction of the poorest agents that become workers declines with time, it follows that the c.d.f.'s do not intersect if and only if the upper support on their wealths increases. This is the case if the most efficient of these poorest agents remains constrained.

It is possible that some of the poorest agents become unconstrained and, eventually, that the upper support on their childrens' wealths falls. Even in this case, the result holds unambiguously. Observe first that $P(\underline{b}_t|\underline{b}_t, w_t) = G_{t+1}(\underline{b}_t) = 0$ and $P(\underline{b}_t|\underline{b}_{t-1}, w_{t-1}) > G_t(\underline{b}_t)$ since the fraction of agents remaining on the lower support is greater among the poorest agents than among the entire population. It follows that

$$P(\underline{b}_t|\underline{b}_t, w_t) - P(\underline{b}_t|\underline{b}_{t-1}, w_{t-1}) < G_{t+1}(\underline{b}_t) - G_t(\underline{b}_t). \qquad (*)$$

Next observe that, if sufficient conditions (A) and (B) hold, then from Lemma 3, $P_{b'b} \geq 0$. Hence, for all b',

$$P_{b'}(b'|\underline{b}_t, w_t) - P_{b'}(b'|\underline{b}_t, w_{t-1}) \leq P_{b'}(b'|b, w_t) - P_{b'}(b'|b, w_{t-1}) \quad \forall\, b > \underline{b}_t.$$

Multiplying by $g_t(b)$ and integrating,

$$\begin{aligned} & P_{b'}(b'|\underline{b}_t, w_t) - P_{b'}(b'|\underline{b}_t, w_{t-1}) \\ & \leq \int P_{b'}(b'|b, w_t)G_t(db) - \int P_{b'}(b'|b, w_{t-1})G_t(db) \\ & \leq \int P_{b'}(b'|b, w_t)G_t(db) - \int P_{b'}(b'|b, w_{t-1})G_{t-1}(db) \\ & \quad - \int P_{b'b}[G_{t-1}(b) - G_t(b)]db. \end{aligned}$$

Since $P_{b'b} \geq 0$ it follows that for all b',

$$P_{b'}(b'|\underline{b}_t, w_t) - P_{b'}(b'|\underline{b}_t, w_{t-1}) \leq g_{t+1}(b') - g_t(b') \quad \forall\, b'. \qquad (**)$$

Integrating over $b' \geq \underline{b}_t$ using $(*)$ and $(**)$, and noting $P(b'|\underline{b}_t, w_t) < P(b'|\underline{b}_{t-1}, w_t)$ yields

$$P(b'|\underline{b}_t, w_t) - P(b'|\underline{b}_{t-1}, w_{t-1}) < G_{t+1}(b') - G_t(b') \quad \forall b'.$$

That is, the probability that the poorest agent receives more than any given wealth increases relative to the probability that any agent earns more than that wealth. Since both sides of this inequality monotonically decrease in b', this condition also implies

$$P(\hat{b}_{t+1}(p)|\underline{b}_t, w_t) - P(\hat{b}_t(p)|\underline{b}_{t-1}, w_{t-1}) < G_{t+1}(\hat{b}_{t+1}(p)) - G_t(\hat{b}_t(p)) = 0.$$

Finally, consider any two percentiles p_1 and p_2, $p_2 > p_1$. If at some date t $\hat{b}_t(p_1) < \hat{b}_{t-1}(p_1)$ then, due to decreasing X-dispersion, $\hat{b}_t(p_2) \leq \hat{b}_{t-1}(p_2)$. Hence, $\Pi_t(p_1) > \Pi_{t-1}(p_1)$ implies $\Pi_t(p_2) \geq \Pi_{t-1}(p_2)$. Equivalently, if $t > t(p_1)$ then $t \geq t(p_2)$, so that $t(p_1) \geq t(p_2)$. ■

Proof of Proposition 16 : Under conditions (A) and (B), any second-order stochastic redistribution of inheritances that raises the wage causes aggregate income to rise. In this case, the optimal distribution policy puts all the mass of the distribution at the mean. We therefore only need consider the optimal policy when the wage is $\underline{w}$.

The function relating an agent's expected income to his wealth, $\psi(b, w)$, was characterized in Lemma 7 and is illustrated in Figure 20a. Consider any two inheritance levels b^l and b^h and those agents receiving these inheritances. Let p be the fraction of these agents receiving b^l. Then, their average wealth is $b^a = pb^l + (1-p)b^h$ and their average income is $p\psi(b^l, \underline{w}) + (1-p)\psi(b^h, \underline{w})$. If $b^a < b^*$, then their average income per unit of inherited wealth is less than if they all received b^*:

$$\frac{\psi(b^*, \underline{w})}{b^*} > \frac{p\psi(b^l, \underline{w}) + (1-p)\psi(b^h, \underline{w})}{b^a}.$$

This can be re-written as

$$(\frac{b^a}{b^*})\psi(b^*, \underline{w}) > p\psi(b^l, \underline{w}) + (1-p)\psi(b^h, \underline{w}).$$

But the left hand side of this inequality is just the total income attained by giving a fraction $\frac{b^a}{b^*}$ of these agents inheritance b^* and the rest zero. This is true even if one of the inheritance levels exceeds b^*, as long as the average, b^a, is lower. It follows that if the per capita wealth for the entire population is less than b^*, then the net income maximizing redistribution gives a positive fraction of the population b^* and the rest 0. Otherwise, there would be at least one other inheritance level to which mass would be assigned. But then the average wealth of agents at this inheritance level and at zero must be less than b^* and so it would then be possible to increase net income by giving some these agents b^* and the rest zero.

Now suppose $b^a > b^*$. In this case, even if each agent received wealth b^* there would be some left over. How should this additional wealth be redistributed to maximize income? Observe that for wealth levels b^l, b^h, such that their mean b^a exceeds b^*,

$$\psi(b^a, \underline{w}) > p\psi(b^l, \underline{w}) + (1-p)\psi(b^h \underline{w}).$$

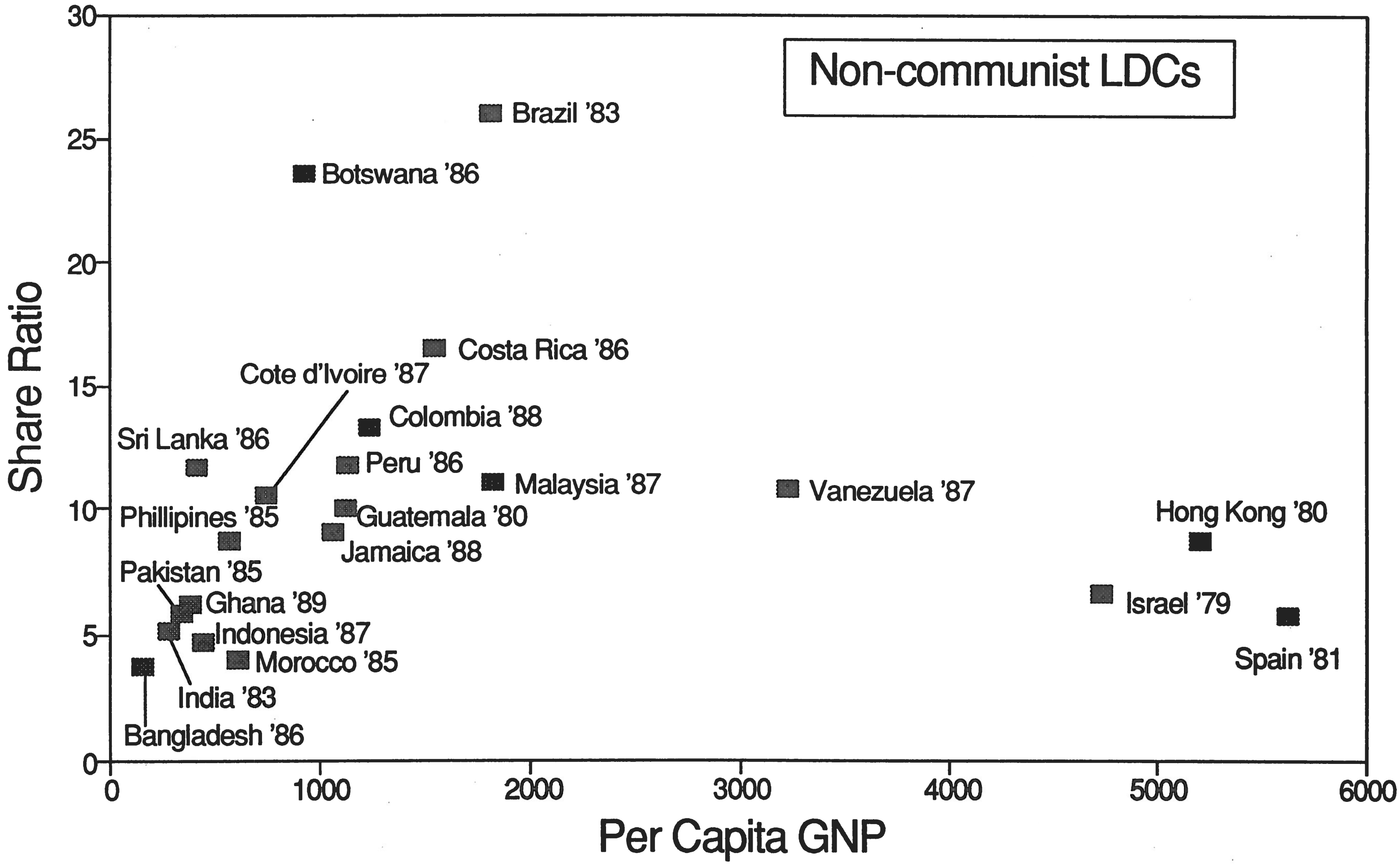
Thus, their expected income can be increased by giving them all the same inheritance, b^a. This is true even if one of the inheritance levels is below b^*, so long as the mean is greater. Hence, as long as we can find two inheritance levels such that the mean wealth of the associated agents exceeds b^*, then it is always possible to increase total income via a redistribution which gives all agents the same inheritance. It follows that if per capita wealth for the entire population exceeds b^*, then the net income maximizing redistribution would give all agents the same inheritance. Otherwise, the optimal distribution would be at least a two-point distribution. But then it would then be possible to increase net income by giving all agents the same wealth. ∎

Notation Table

b	inherited wealth
b^0	pre-industrialization steady-state inheritance
$f(\cdot)$	production function
$h(\cdot)$	density function for start-up costs
k	capital level
k^c	wealth constrained capital investment
k^u	unconstrained optimal capital investment
l	labor level
l^c	constrained demand for labor
l^u	unconstrained optimum level of labor
m	slope of example set-up cost density function
p	proportion of population
$q(w,b)$	output of entrepreneurs with inherited wealth b
$u(\cdot)$	agent utility function
w	wage
$\underline{w}$	reservation wage
$\bar{w}$	efficient market clearing wage
x	start-up cost
$\underline{x}$	lower bound on distribution of start-up costs
$\bar{x}$	upper bound on distribution of start-up costs
$x^m(b,w)$	start-up cost of indifferent agent
$\tilde{x}(y,b,w)$	start-up cost that yields profit level y
y	income of individual agent
$z(b,w)$	start-up cost of marginal agent with inheritance b
$B(\cdot)$	agent bequest function
$C(\cdot)$	agent consumption function
$E_t(w)$	aggregate rate of enterprise for given wage
$F_t(\cdot)$	c.d.f. for final wealth
$G_t(\cdot)$	c.d.f. for inherited wealth
$H(\cdot)$	c.d.f. for start-up costs
$J(\cdot)$	c.d.f. for firm sizes
K_t	aggregate capital stock
L_t	aggregate labor force
$P(\cdot\|\cdot)$	probability transition function for inheritances
$Q_t(w)$	aggregate output
$S_t(w)$	fraction of population in subsistence
W	final wealth of individual agent
$X_t(w)$	aggregate start-up costs
$Y_t(w)$	aggregate net income
$Z_t(\cdot)$	Lorenz function

$\alpha, \beta, \xi, \rho, \sigma$	parameters of example production function
γ	subsistence income
$\eta(b, w)$	measure of entrepreneurs with inheritance b
ι_t	location indicator
$\kappa(b, w)$	capital of entrepreneurs with inheritance b
$\lambda(b, w)$	labor demanded by entrepreneurs with inheritance b
$\pi(b, x, w)$	profit of individual entrepreneur
τ_1	date when wage begins to increase
τ_2	date when maximum income level begins to decrease
ν	cost of urban living
$\phi_t(\cdot)$	density of incomes
$\chi(b, w)$	set-up costs of entrepreneurs with inheritance b
$\psi_t(b, w)$	expected income conditional on own inheritance
ω	share of wealth bequeathed in example
$\Pi_t(p)$	occupational mobility
$\Phi_t(\cdot)$	c.d.f. for income

Figure 1
Inequality vs. per capita income



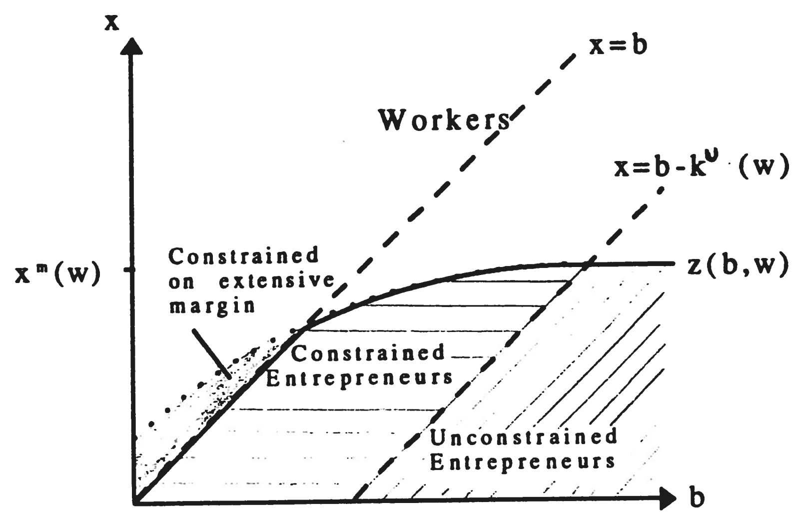


Figure 3

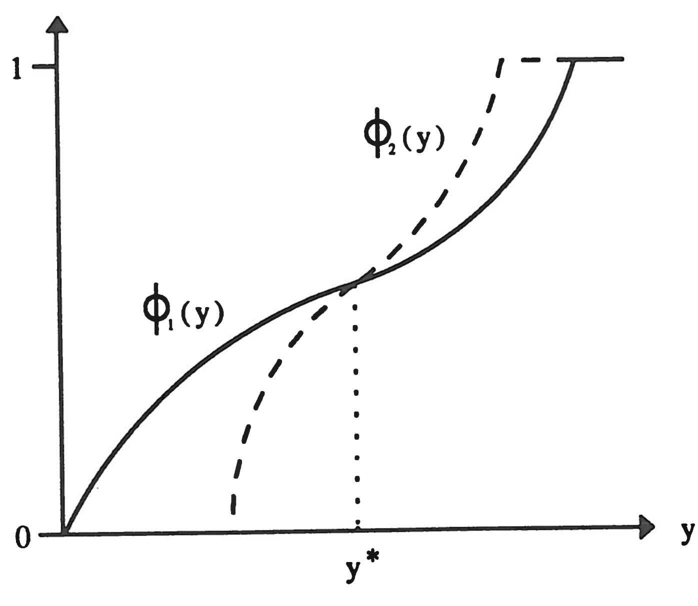


Figure 4

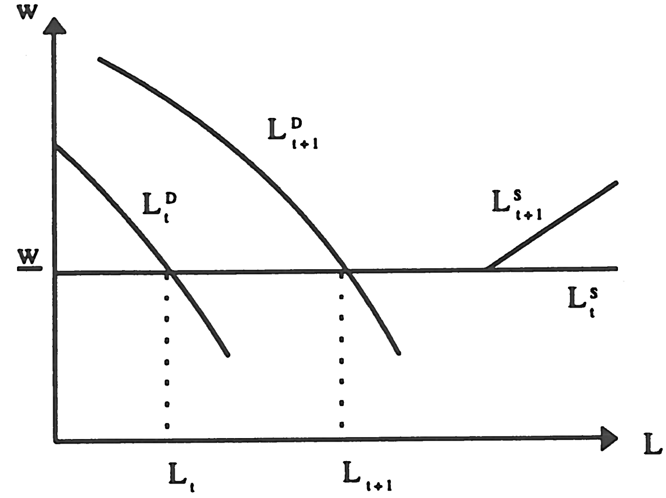


Figure 5

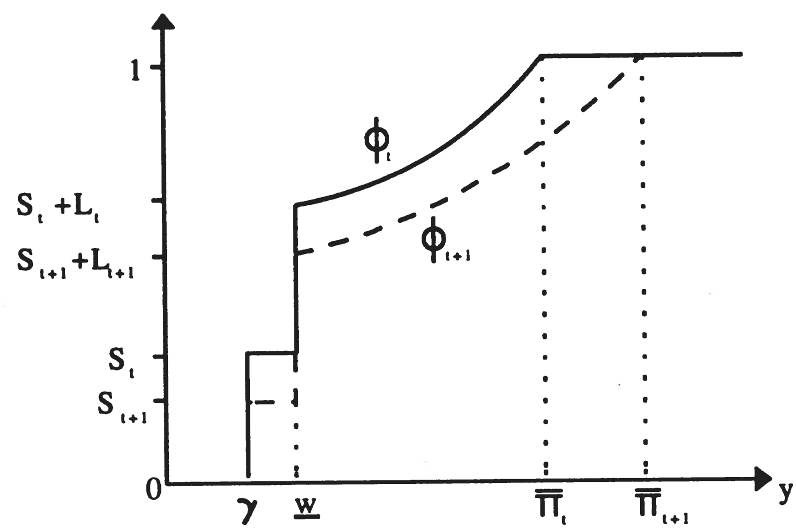


Figure 6

Figure 7(a)

Output, Income and Wages

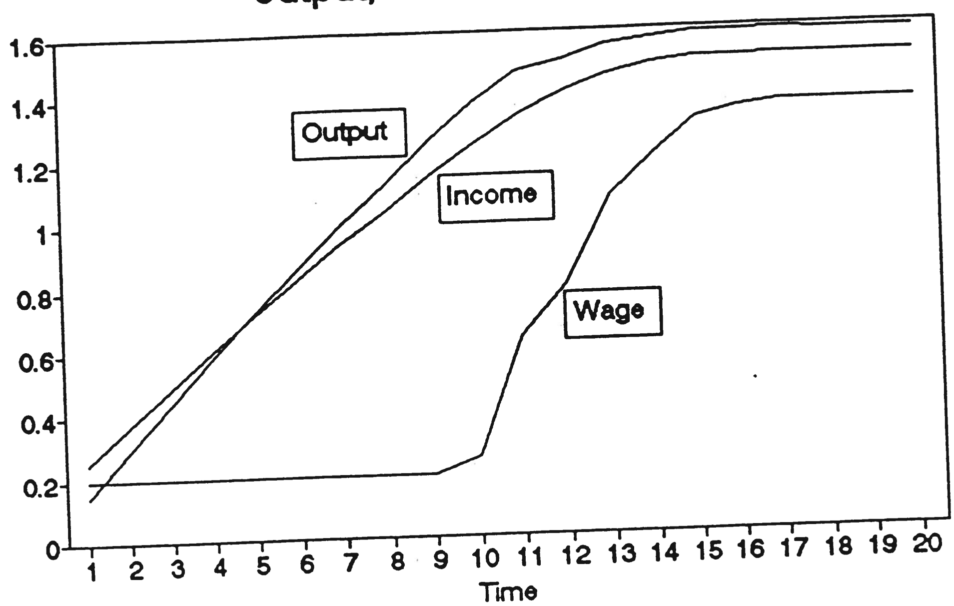
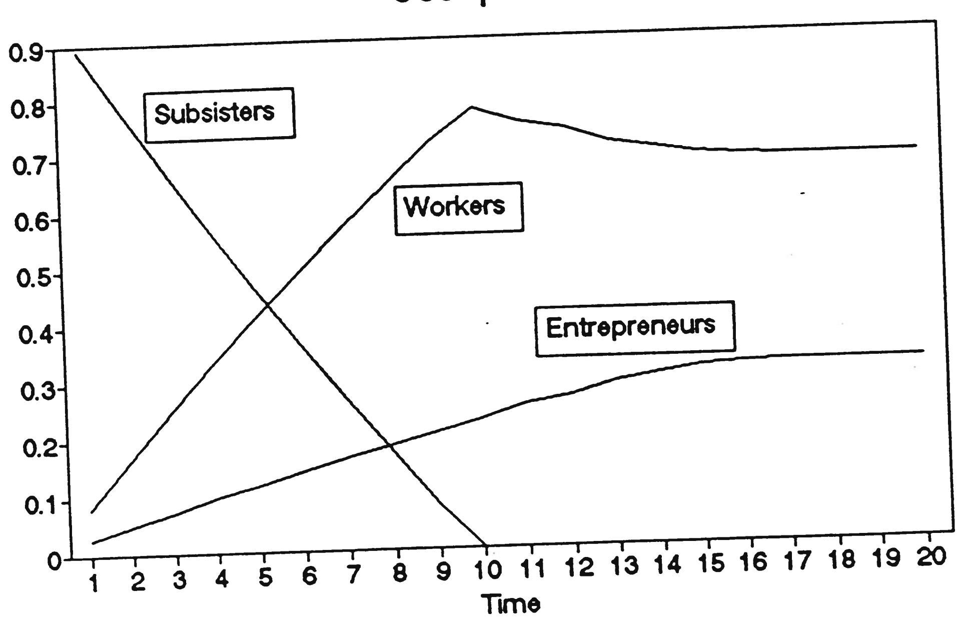
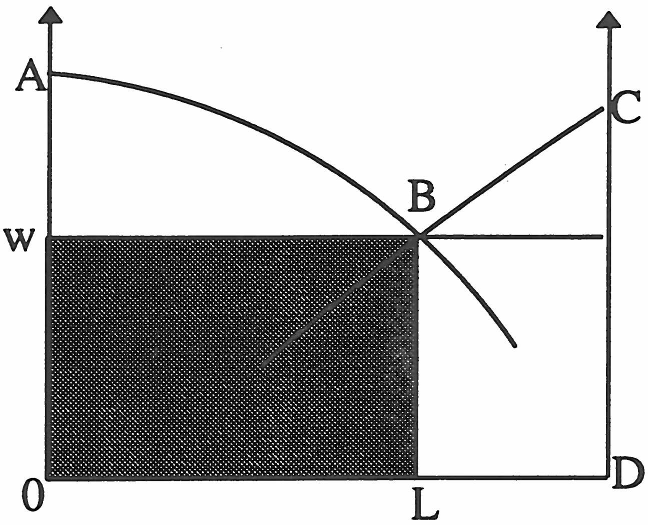


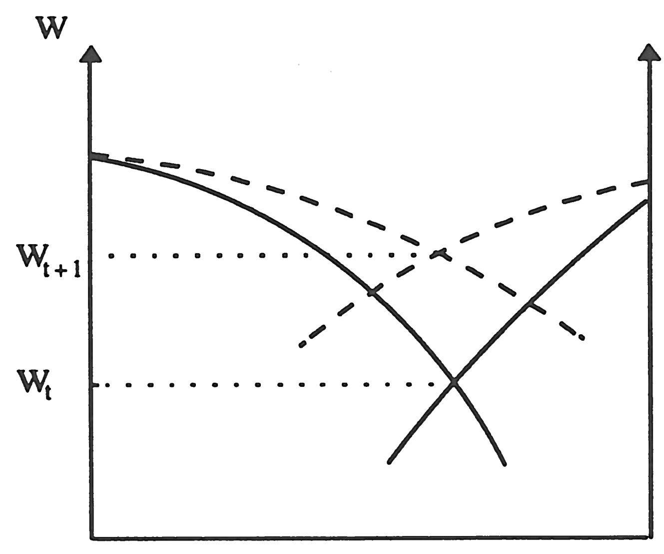
Figure 7(b)

Occupations





(a)



(b)

Figure 8

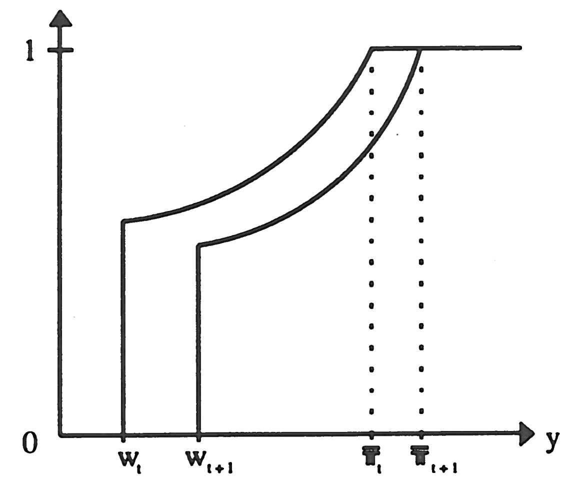


Figure 9

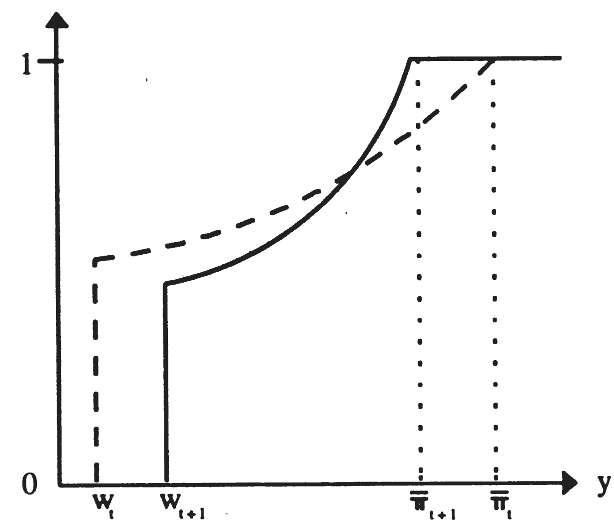


Figure 10

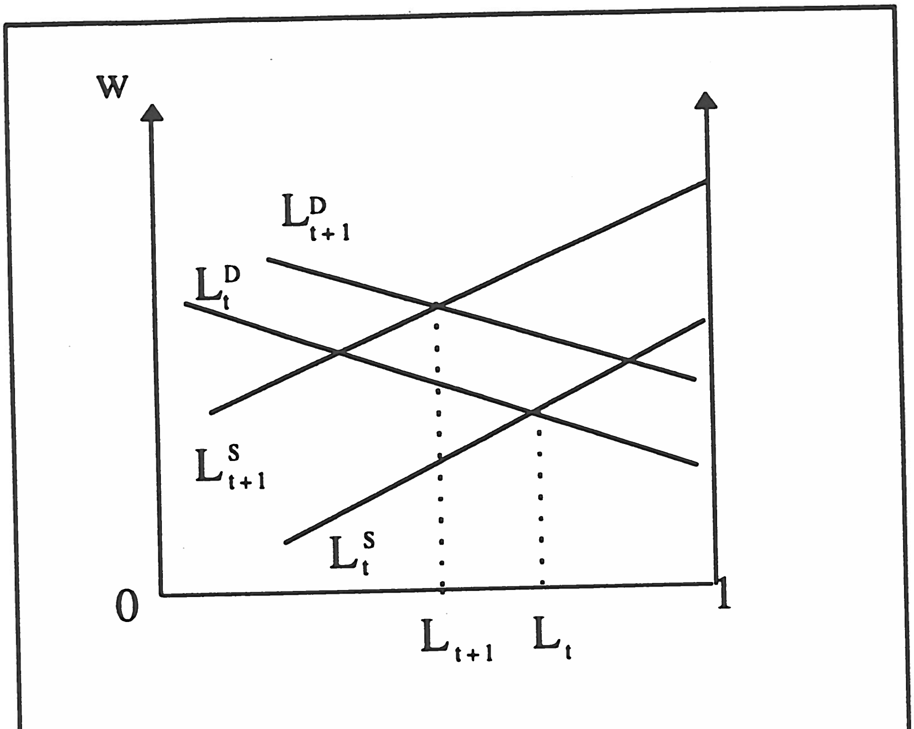


Figure 11

Figure 12

Investment

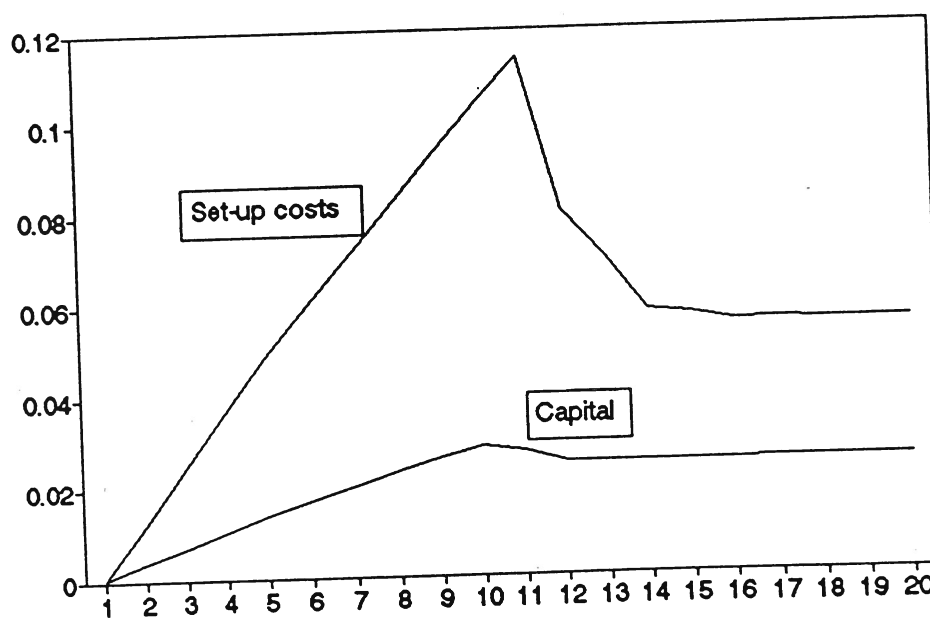


Figure 13(a)

Average and Optimal Firm Size

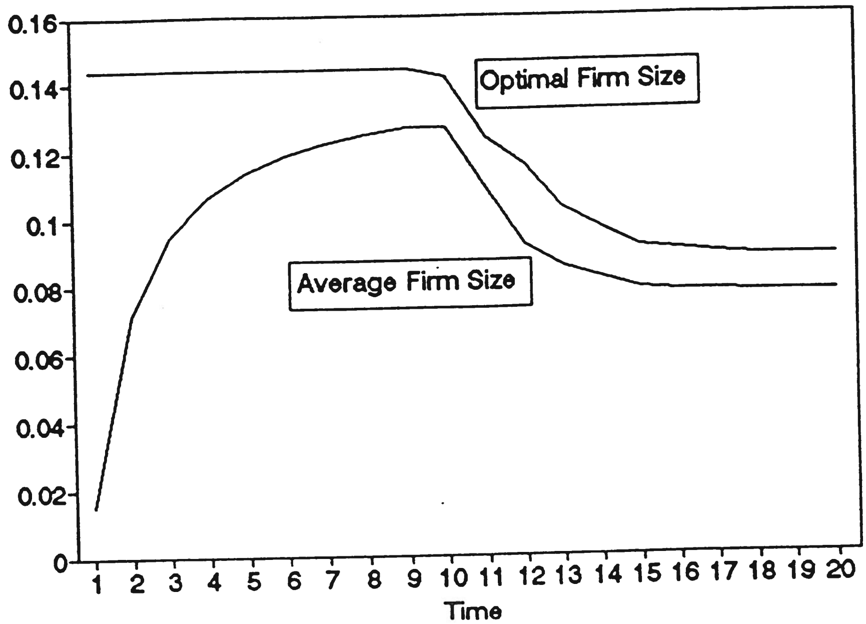


Figure 13(b)

Variance of Firm Sizes

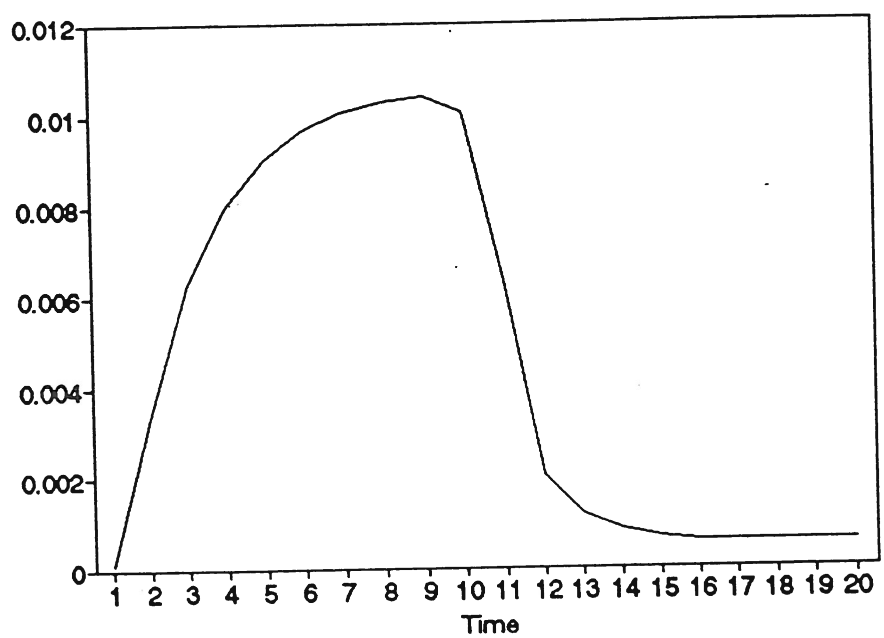


Figure 14(a)

Long-run wealth distribution

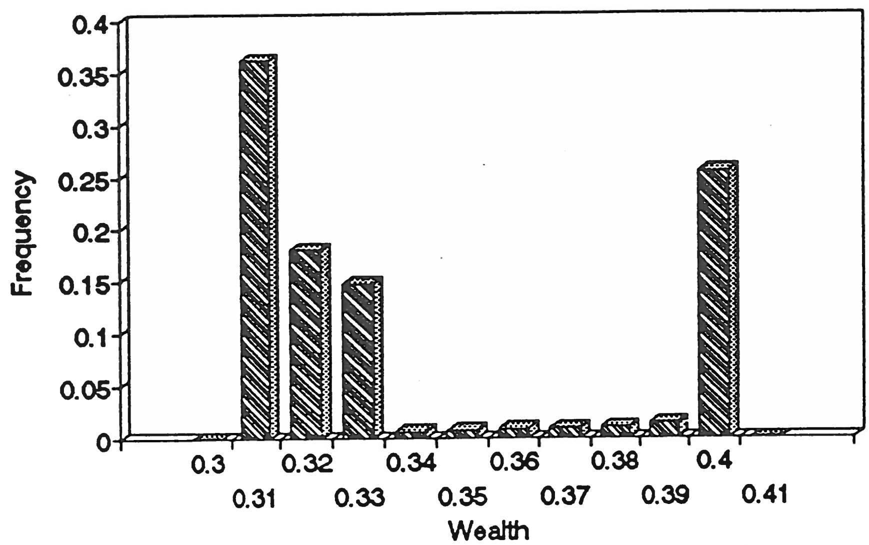


Figure 14(b)

Long-run income distribution

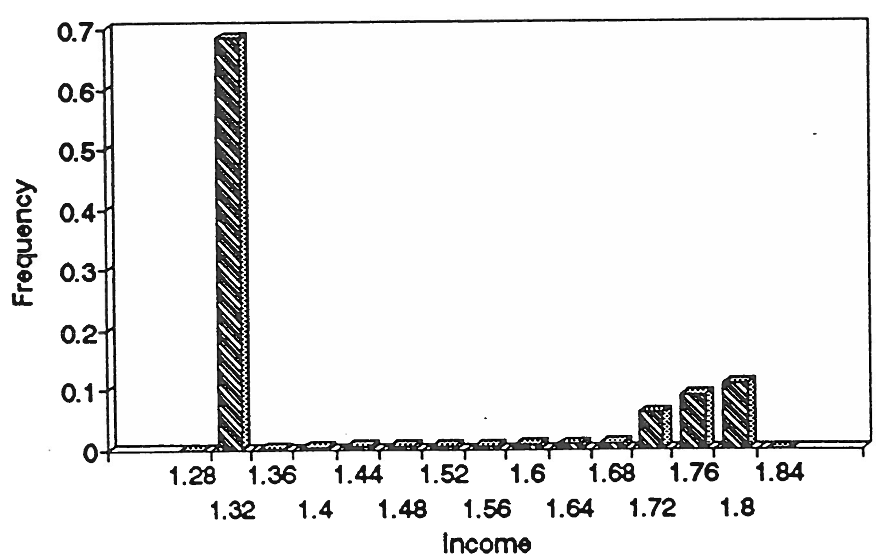


Figure 15(a)

Low development income distribution

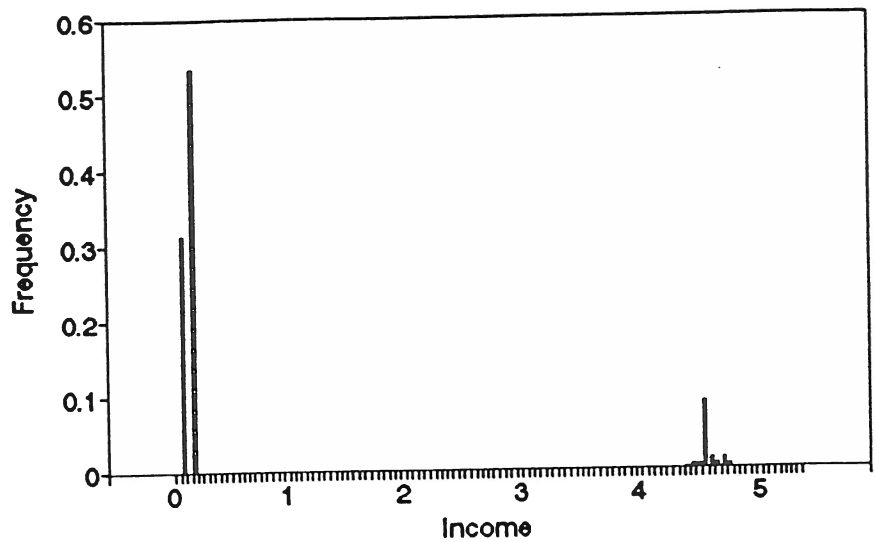
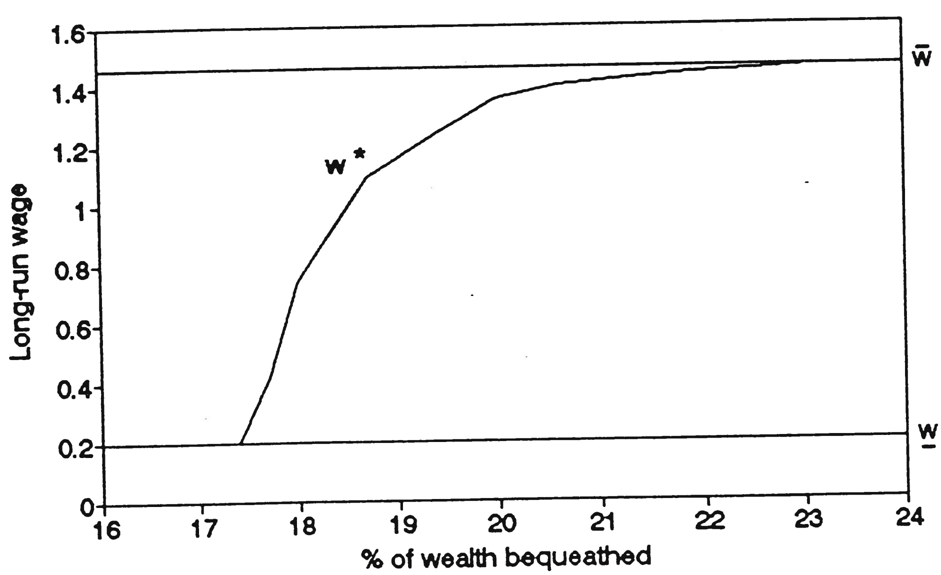
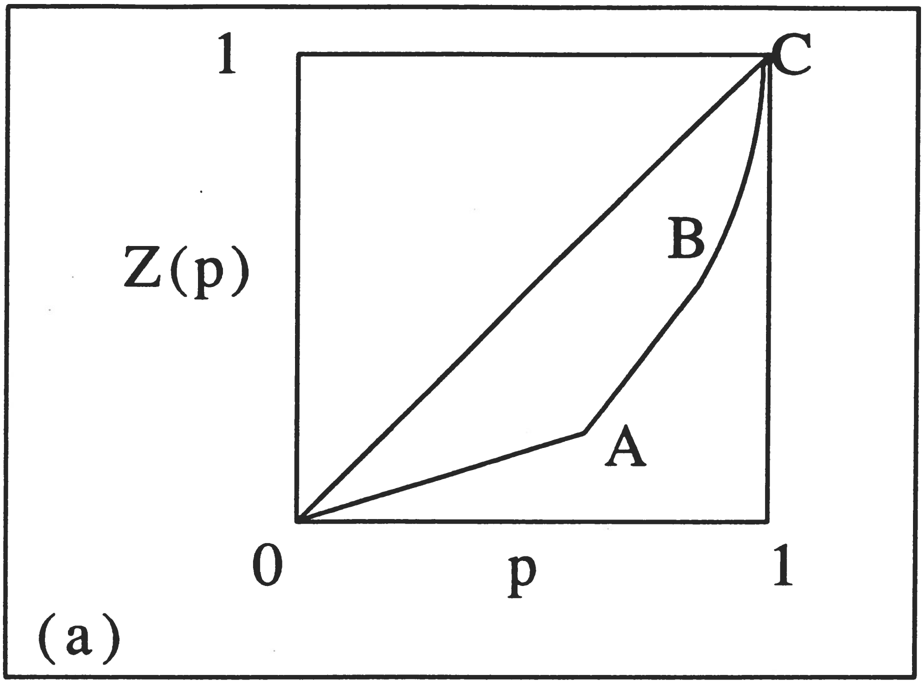


Figure 15(b)

Long-run wages





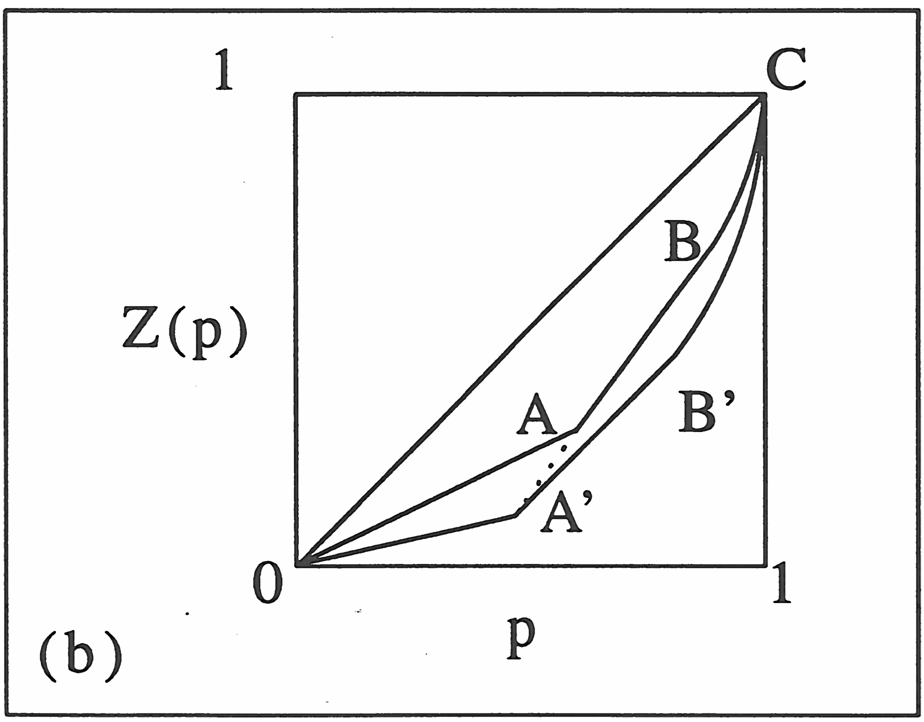
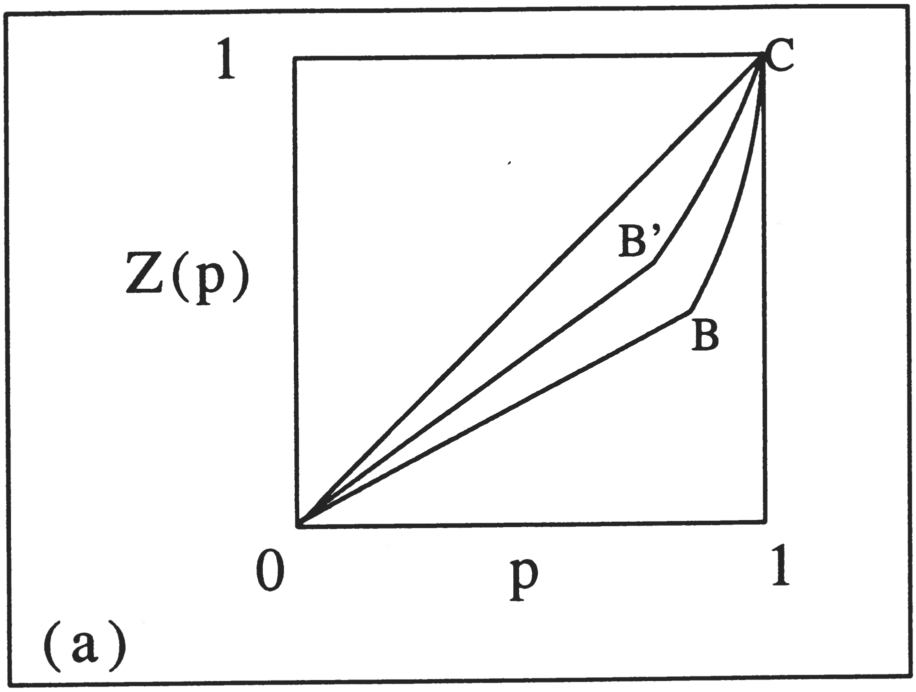


Figure 16



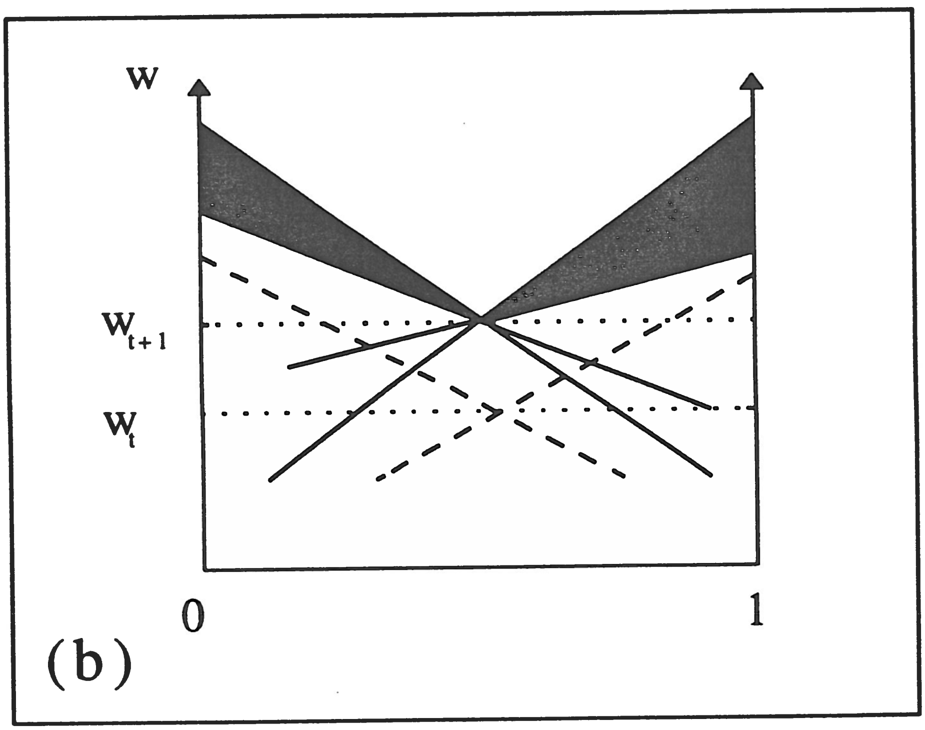


Figure 17

Figure 18(a)

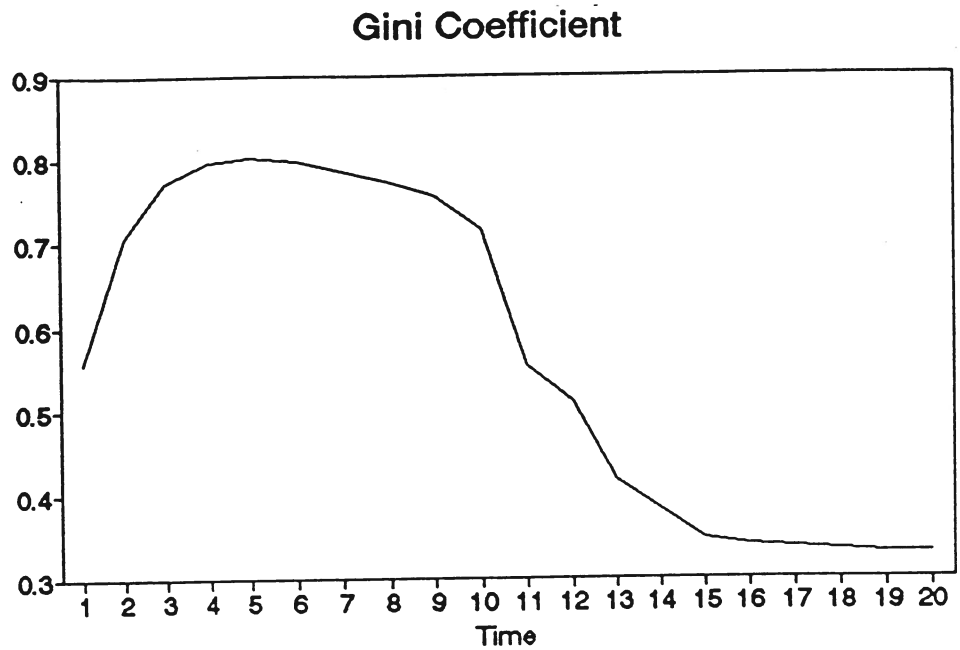
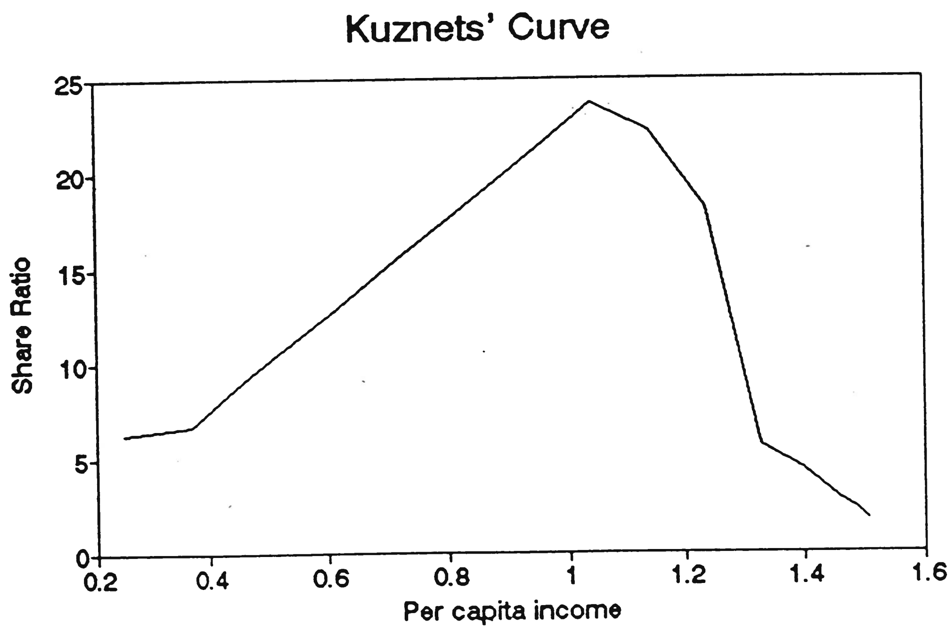


Figure 18(b)



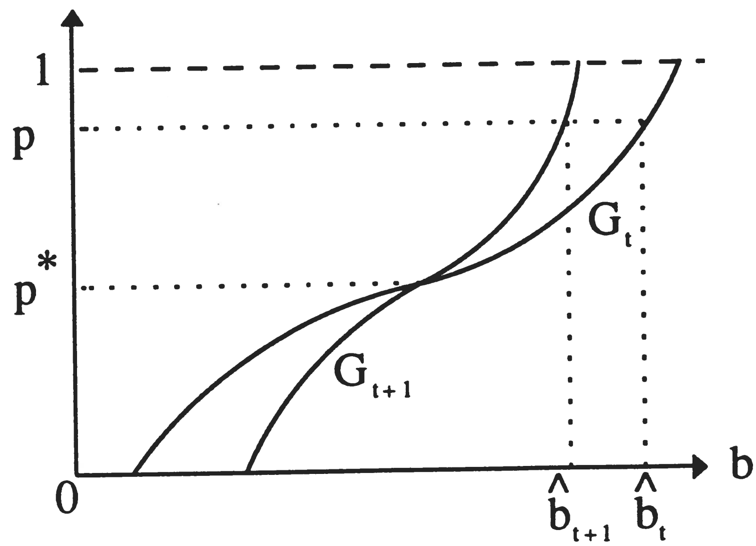
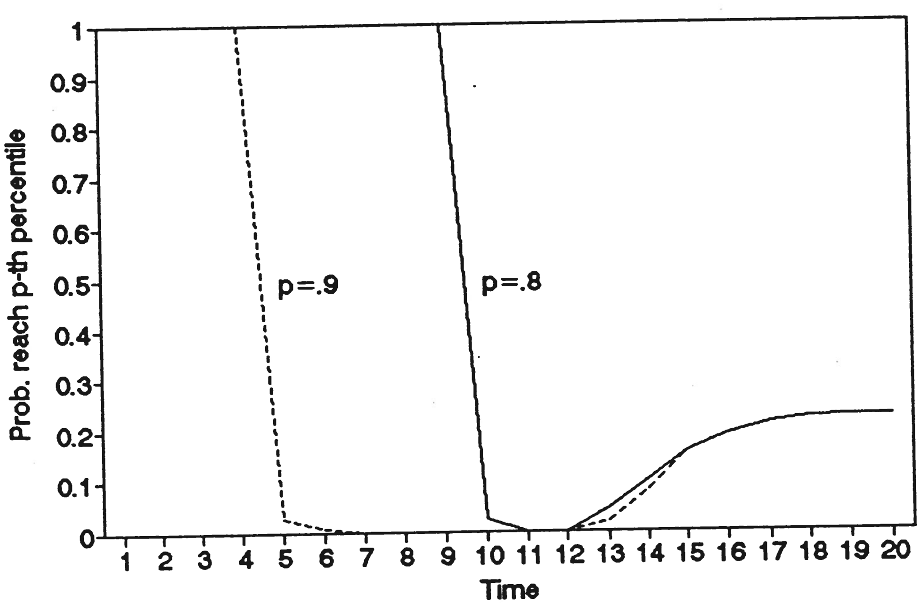


Figure 19(a)

Figure 19(b)

Upward Mobility : p= .8, .9



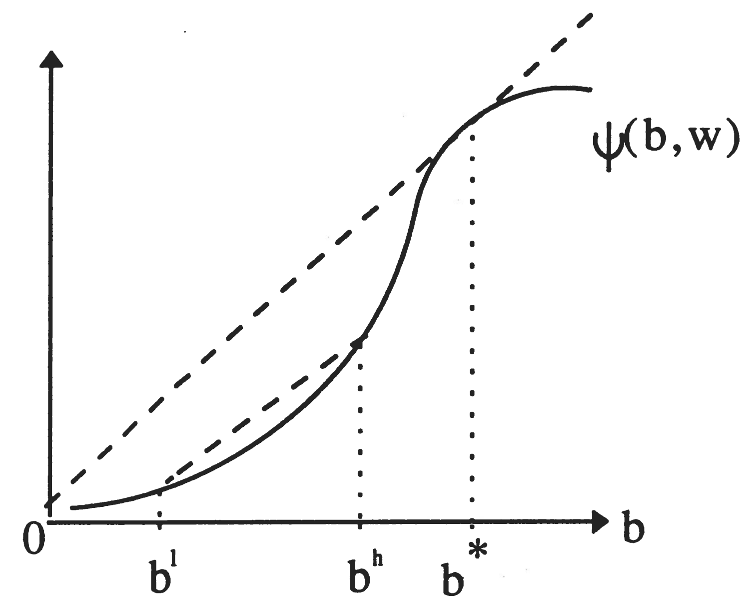


Figure 20(a)

Figure 20(b)

Actual vs. Optimal Income

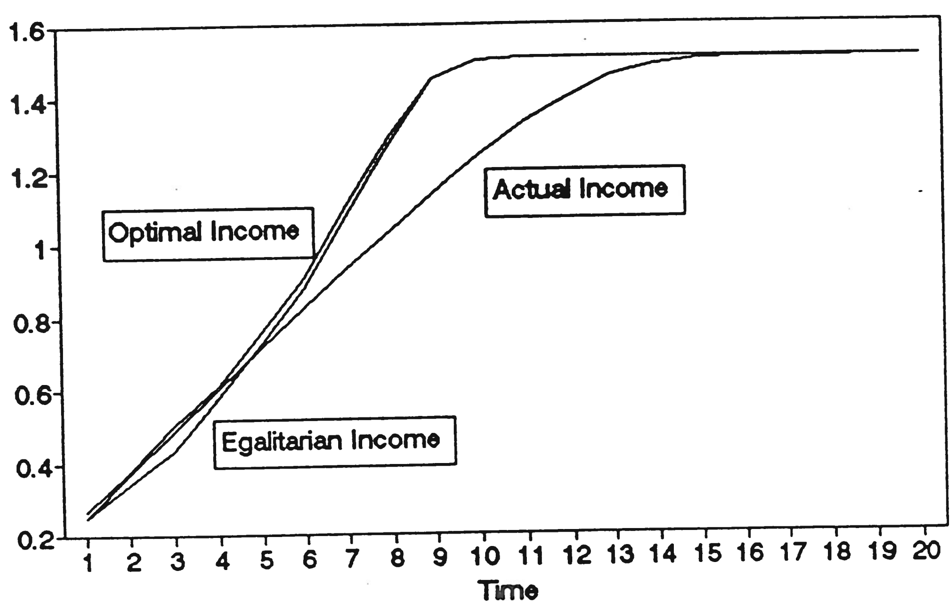


Figure 21(a)

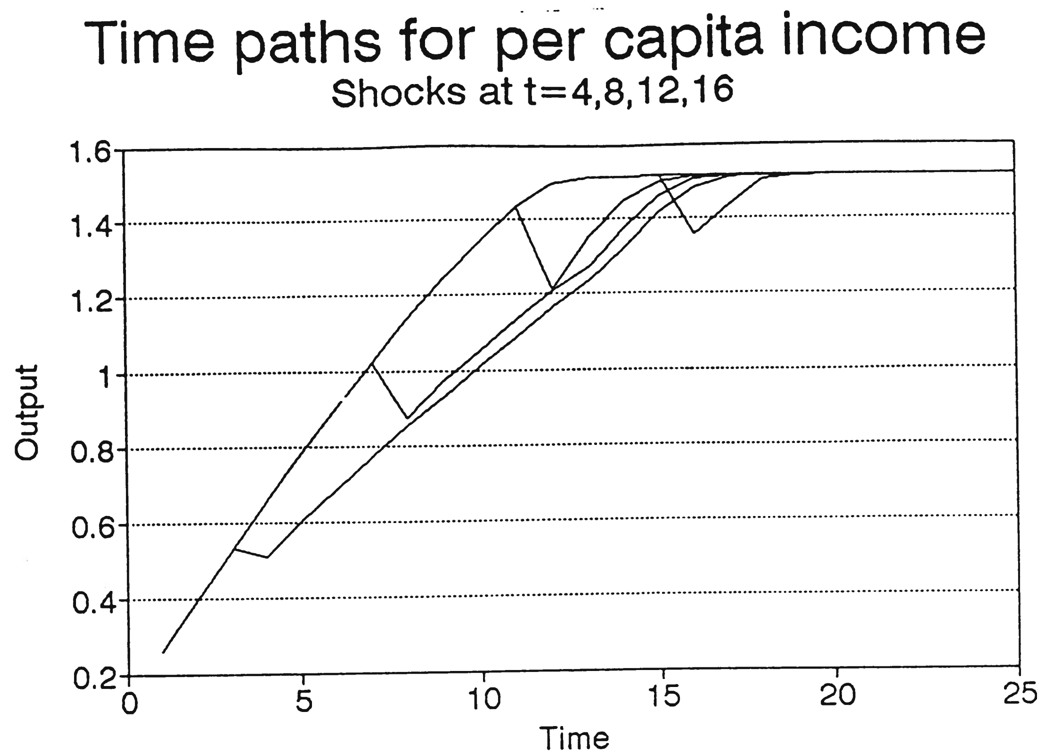


Figure 21(b)

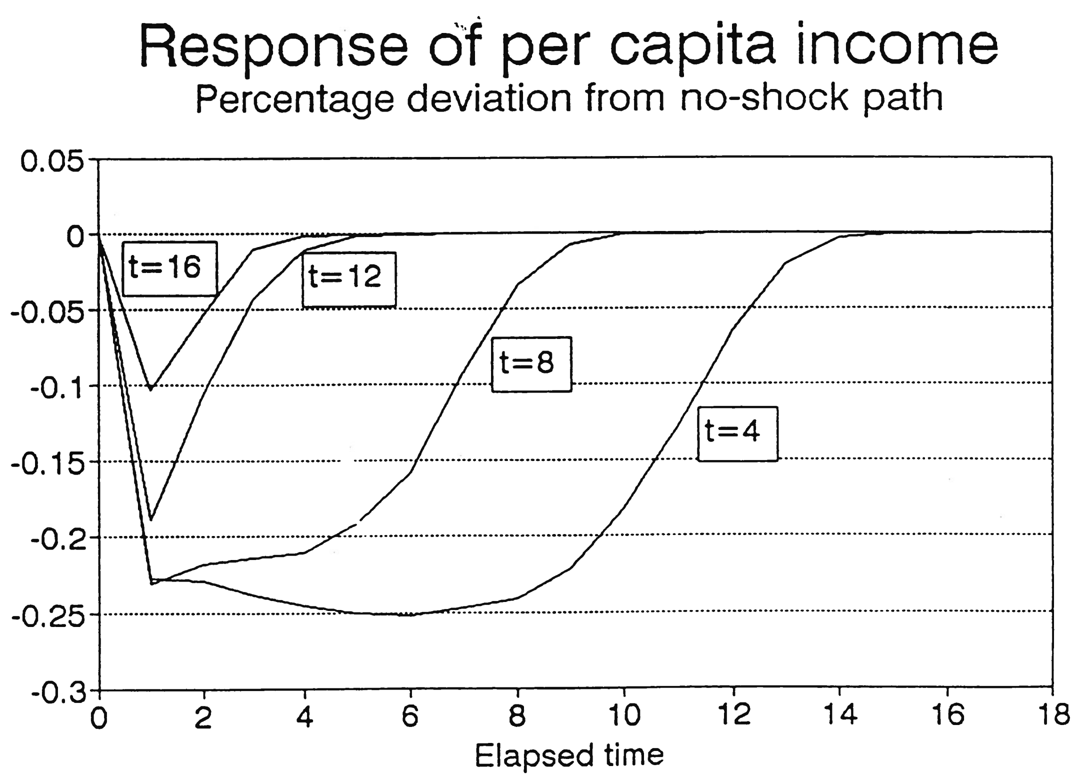


Figure 21(c)

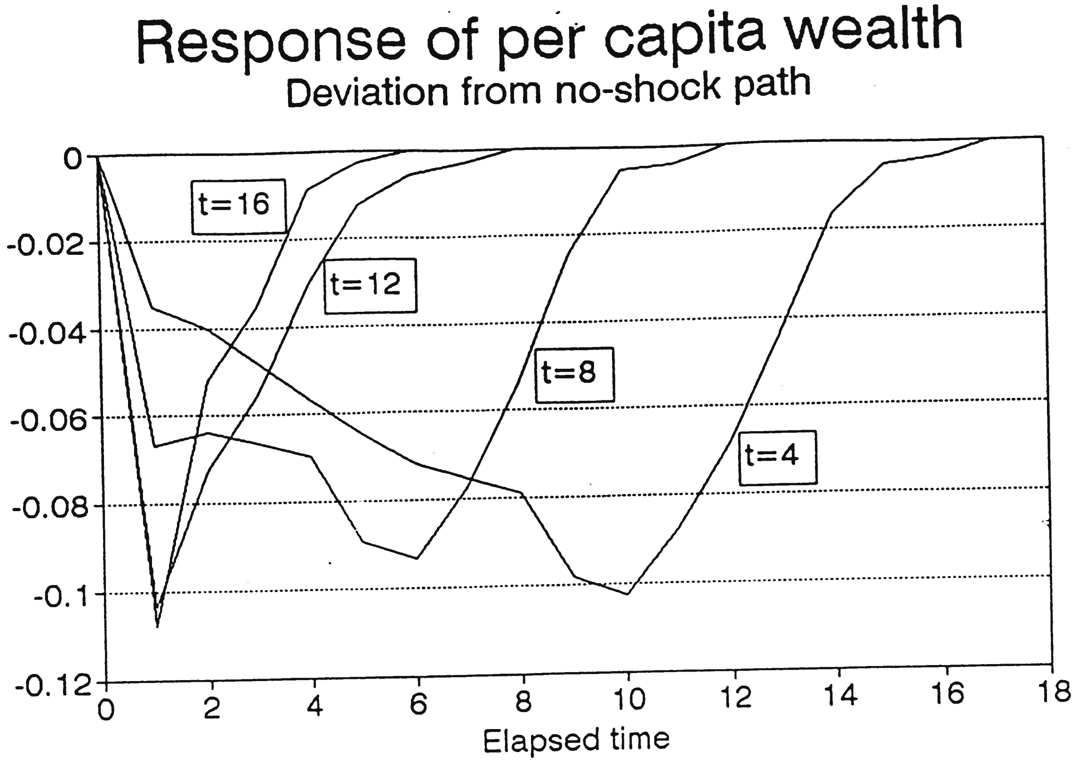


Figure 21(d)

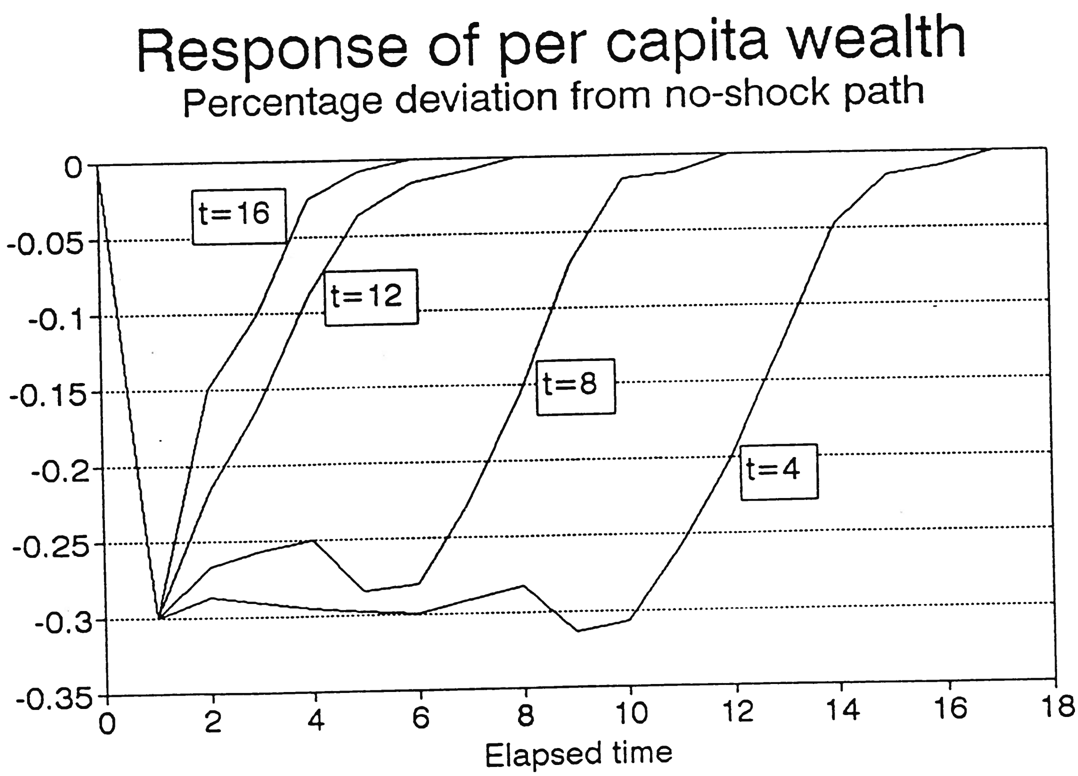
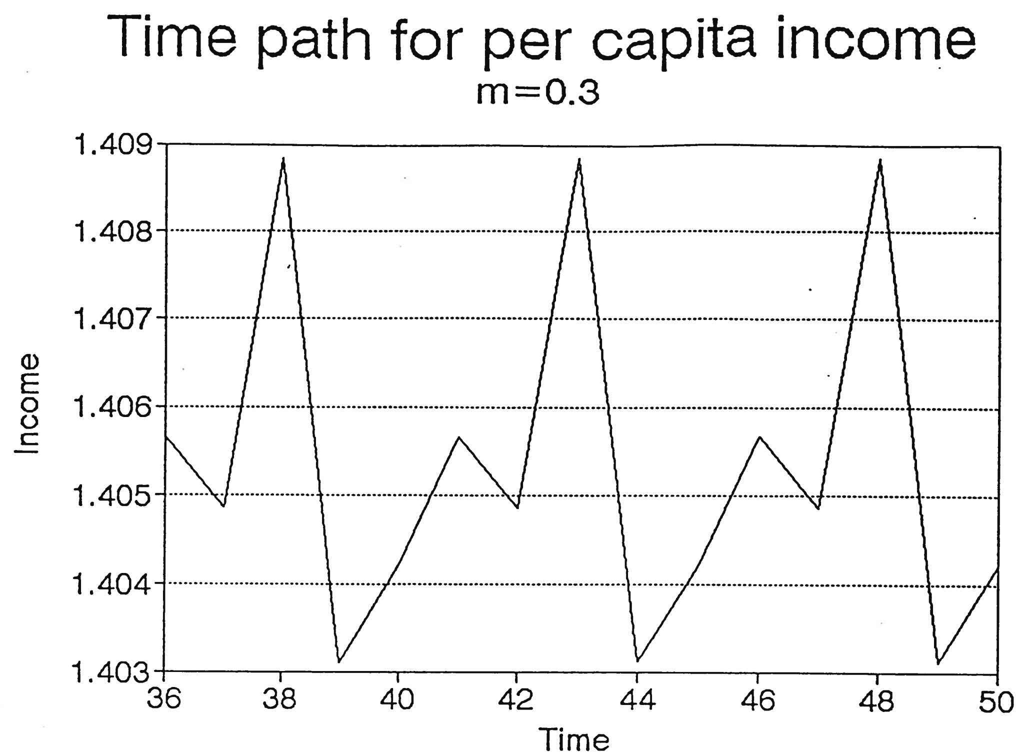
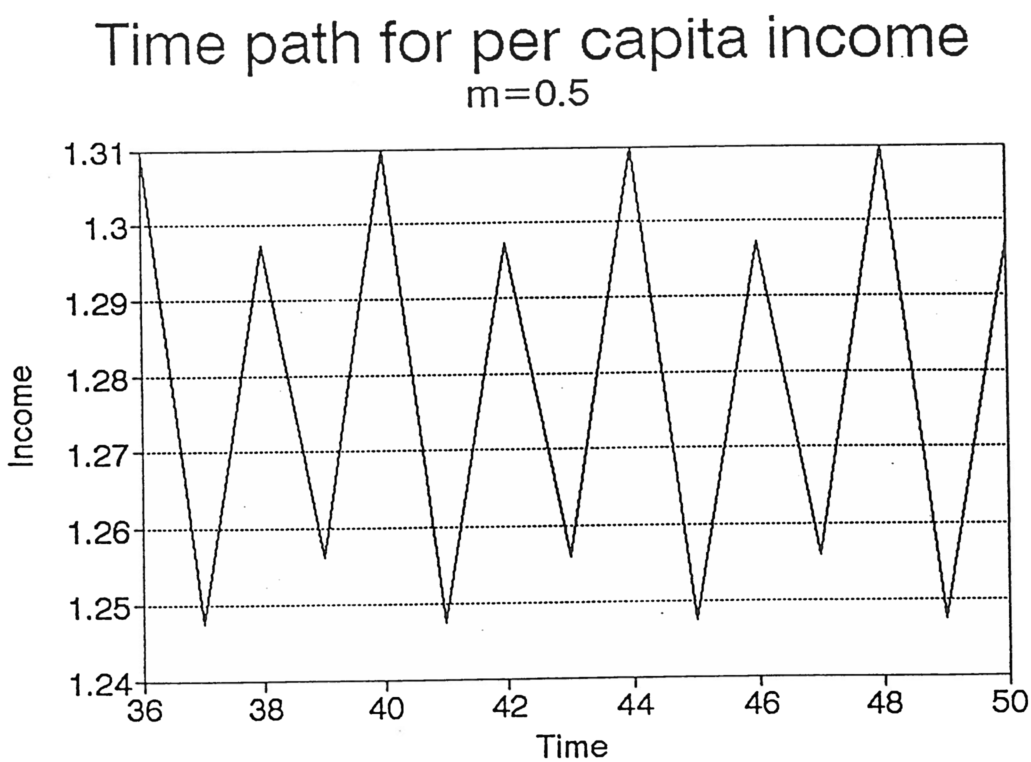


Figure 22



Time path for per capita income
m=0.3
Income
1.409
1.408
1.407
1.406
1.405
1.404
1.403
36
38
40
42
44
46
48
50
Time



Time path for per capita income
m=0.5
Income
1.31
1.3
1.29
1.28
1.27
1.26
1.25
1.24
36
38
40
42
44
46
48
50
Time