DISCUSSION PAPER NO. 226

INTERGENERATIONAL TRANSFERS AND THE EQUILIBRIUM DISTRIBUTION OF EARNINGS

by

Glenn C. Loury

June 1976

Intergenerational Transfers and the Equilibrium Distribution of Earnings

Introduction:

The study of the distribution of income is as old, it seems, as economics itself. Classical writers--Smith, Ricardo, Marx--were mainly concerned about the distribution of national product (income) among the various factors of production. Much recent research has focused on that question as well.[1] However, the distribution of income among persons, in particular the distribution of labor's share, has also been a topic of great interest.[2] It is an especially relevant issue when, as now, there exist substantial class differentiations among those who do not derive a significant part of their income from property.

Modern theories of the distribution of earnings do not explicitly consider these class distinctions. Individuals are typically viewed as "earning" their economic reward by accepting a drawing from a (partially controlled) lottery.[3] The distribution of rewards for a given economic agent may depend upon his characteristics and even his choices, but not on the actions of those who share his environment.[4] These essentially independent random drawings are then aggregated to yield the distribution of earnings in the population.[5]

The most recent work in this area, especially that growing out of the human capital school, seeks a grounding of the theory of income distribution in the rational behavior of individual agents.[6] Earnings are assumed to depend on the innate talents and acquired skills of the worker, and decisions to acquire skills are taken so as to maximize welfare. This approach has been successful in explaining many of the documented empirical regularities of earnings distributions. While considerations of social class have not played a prominent role in this recent literature, there has been some discussion of the role of accumulated wealth and inheritance in determining the distribution of earnings.[7] Moreover, it is now generally acknowledged that the ability of an individual to convert his natural talents into earnings is affected in an important way by the economic success of his parents.[8]

Thus, it seems natural at this stage to seek a theory of the distribution of earnings which explicitly acknowledges the role of social origin in the determination of economic achievement.[9] Such a theory must necessarily look across generations, as the distribution of income among today's workers is closely related to the distribution of family background among those who will constitute the labor force of the next generation. We should also require of such a theory that the effects of family origin on earnings find their explanation in the rational behavior of the family members themselves.

This essay develops a simple model of the distribution of earnings which attempts to satisfy these criteria. Individuals begin with a random endowment of innate capacities, but their acquisition of other skills is affected by the allocation of family resources to their training. It is shown that, given an exogenous distribution of innate capacities, the economy always tends to a unique equilibrium distribution of earnings. The concept of equilibrium is quite natural in that, should this distribution occur among a given generation of workers, their optimal allocations of family resources will lead to the reproduction of that same distribution of earnings among their offspring in the subsequent generation.[10]

Some elements of the effects of social stratification are incorporated into the model by positing that capital markets are balkanized. That is, individuals cannot borrow funds for their traning from a perfectly competitive capital market, but rather must depend wholly on funds generated within their family to finance the acquisition of skills. This characterization is designed to rigorously illustrate how the effects of social origin can combine with the natural abilities of individuals to produce an earnings distribution which is difficult to rationalize as the "natural" or meritocratic outcome of the economic process.

In the following sections we set out the basic structure of the model and discuss the decision making process within the family. This family allocation rule may then be used to consider the evolution of the aggregate distribution of income from generation to generation. Our concept of equilibrium is precisely defined and some of its properties are demonstrated. We then consider a specific example in which the equilibrium distribution of earnings may be calculated under certain assumptions on family behavior, technology, and the distribution of abilities.

Basic Structure of the Model:

We imagine an economy composed of a large number of individuals, each of whom lives for exactly three units of time. Each unit of time (periods) may be thought of as a generation. Assume for now that the population is stationary so that an equal number of individuals enter and leave the economy in any period. The life of an individual is divided into three stages, corresponding to his age. In the first period of life (youth) each individual is engaged in the acquisition of skills. The second period of the life cycle (maturity) is devoted to productive activity, and the individual earns his lifetime income. The final stage is one of retirement in which the agent may contemplate the meaning of life.

The social organization of this economy is structured around the "family." That is, while each person has an individual existence, they all see themselves as elements of a larger entity and behave accordingly. A family is the common household formed by three generations of individuals, the younger of whom are descendent from the older. One may imagine that at the beginning of the second period of life each individual produces an offspring. In this way, all new entrants to the economy are attached to a mature individual, and are considered the descendant of that individual. A family then consists of a retired individual, a mature individual descendent from him, and young individual descendent from the mature agent. Thus, every individual in the economy belongs to a family at each stage of his life.

For reasons that will become clear presently, consumption is considered a common family activity. Individuals derive their own personal consumption from the aggregate family consumption in each period. We do not examine the distribution of family consumption among the members, taking this to be determined by custom and need. We will, however, examine in the next section how the aggregate level of family consumption is decided upon. For now, we assume that all family decisions are made by the productive member. That is, the family head is the mature individual; at the end of his second period of life he turns over the leadership responsibility to his newly matured offspring. In this way continuity in the family existence is maintained across generations.

Production of the one perishable commodity in this economy is an individualized activity. That is, the output of each mature agent is purely a function of his individual productivity, unaugmented by the use of any other factors of production. An individual's earnings may thus be identified with his output. There is no durable capital and no store of value, so that saving in the conventional sense is not possible. Thus, non-productive individuals are dependent upon the productive family member to provide consumption.[11]

The productivity (output) of a mature individual depends on two distinct factors. The first of these is his innate endowment of natural economic ability. It is assumed that each individual begins life with a random endowment of natural productive capacity. The second element in determining productivity is the level of acquired skills attained by the individual in the first period of his life. Let α denote the innate endowment and e represent the level of training during youth. Then an individual's output, x, is given by

$$(1) \qquad x = h(\alpha, e),$$

where $h(\cdot, \cdot)$ is a function whose properties will be specified later. In each period the mature individuals earn their income according to (1), and must decide how to divide it between consumption for the family and investment in the training of their offspring. This investment process provides a means of transferring purchasing power between periods.

The Family Decision Making Process:

In this section we will treat the problem of determining how a family decision maker allocates his earnings between consumption and investment in his offspring. While it is possible to simply specify a consumption function which would apply to all individuals, it seems preferable to deduce this behavior from first principles. Below we consider alternative approaches.

Because there is no possibility of storage of output in this model, mature individuals will be dependent upon their offspring to provide family consumption during their retirement. Thus, they have a direct individualistic incentive to expend some of their earnings on training for their offspring, as this increases the income and hence consumption of the family in the subsequent period. It is possible to use this observation to develop a theory of internal family transfers.[12]

One might imagine mature individuals allocating their income so as to maximize their utility, which would depend on family consumption during their maturity and their retirement. This latter consumption could be taken as some known function of offspring's income. Yet, consistency would require that the planned family consumption of a mature individual undertaking this utility maximization bear the same functional relation to his income as that presumed of consumption in the subsequent period as a function of offspring's income. Moreover, the natural ability to produce would most reasonably be assumed unknown to parents when they decide on investment in their children. This leads one to pose the following problem:

Find the function $C^*(y)$ such that

$$\text{(2)} \qquad \underset{0 \leqq c \leqq y}{\text{Max}}\ E_\alpha U(c, C^*(h(\alpha, y-c))) \rightarrow c = C^*(y), \quad \forall y \geqq 0$$

where $u(\cdot, \cdot)$ is the common utility function of mature individuals, c is current family consumption, y is the income of the mature individual, and E_α is the expectation operator over the random distribution of innate economic abilities. Consistency is assured by (2) because expected utility maximization must lead to a consumption function for mature individuals identical to that which they assume their offspring will employ.

It is apparent that (2) is a very difficult problem.[13] It has resisted the author's concerted efforts to resolve it. Accordingly we shall have to be content with a somewhat less elegant formulation of rational family behavior. There have been two alternative approaches adopted in the literature on the theory of bequests. One approach has been to assume that parents are concerned about the level of consumption of their offspring, and thus maximize a utility function dependent on their own consumption, and the consumption of their descendants.[14] This approach is closest to the procedure suggested above. Alternatively, it has been assumed that parents have a direct concern for the well-being of their children.[15] The bequest is viewed as a means to an end. The ultimate objective is to make the children "happy," so their (cardinal) utility enters directly into the parent's utility function. Despite appearance to the contrary, these approaches are not just two different ways of saying the same things. They can have profoundly different implications for the efficiency of competitive markets or the effects of government fiscal policy in intergenerational models of capital accumulation.[16]

Let us imagine then that the well-being of a family head is given as a function of the level of family consumption during his tenure as decision maker, and by the family head's perception of the utility which his offspring will enjoy after the reins of leadership have been passed along. Mature individuals then allocate their available income between consumption and investment in the training of their offspring so as to maximize this well-being. Since the individual is constrained only by the availability of income, his maximum utility must be a function of earnings alone. Let us assume that all individuals in all generations possess the same utility function. Then it is reasonable to suppose that the well-being which a mature individual attributes to his offspring is also a function solely of the offspring's earnings. Moreover, this perceived well-being should be exactly the same function as that which relates the earnings of the mature individual to his maximum utility attainable.

These considerations suggest the following problem: Find the function $V^*(y)$ such that

$$(3) \qquad V^*(y) = \underset{0 \leqq c \leqq y}{\text{Max}}\, E_\alpha U(c, V^*(h(\alpha, y-c))), \quad \forall y \geqq 0$$

where the notation is as before, with U(c, V) being parental utility when family consumption is c and offspring utility is V. The resemblance of (2) and (3) is obvious, though (3) is considerably easier to handle. Given the utility function U(·, ·), $V^*(\cdot)$ has the interpretation of an indirect utility function. It gives the largest expected utility attainable by a mature individual from some specified income, given the supposition that his offspring will also seek to maximize his expected utility in the same way. In this view of family transfers, parents do not concern themselves about consumption in their later life, but rely on their offspring to provide consumption in exactly the manner that the parents would themselves, were they in the position of family responsibility.[17]

Let us consider a more formal treatment of (3). We need to establish the existence of a solution and some of its basic properties before investigating the consequences of the implied intergenerational transfers for the distribution of earnings. The nature of the solution to (3) will depend on the properties of the utility function, the productivity function, and the distribution of economic abilities. Concerning these functions we adopt the following assumptions:

<u>Assumption 1</u>: The utility function $U(\cdot, \cdot)$ is a twice continuously differentiable, strictly concave, real valued function satisfying:

(i) $U(0, 0) = 0$; $U_1 > 0$, $U_2 > 0$

(ii) $\lim_{c \to 0} U_1(c, V) = +\infty$, $\forall V \geqq 0$

(iii) $\exists \gamma > 0$ such that $U_2(c, V) \leqq \gamma < 1$, $\forall (c, V) \in R_+^2$,

where a subscript indicates differentiation with respect to the indicated argument.

<u>Assumption 2</u>: The productivity function $h(\cdot, \cdot)$ is a twice continuously differentiable, real valued function satisfying:

(i) $h(0, 0) = 0$; $\exists \bar{y} > 0$ such that $h(1, y) < y$, for all $y \geqq \bar{y}$.

(ii) $h_2(\alpha, e) > 0$, $h_{22}(\alpha, e) < 0$, and $\exists \beta > 0$ such that $h_1(\alpha, e) \geqq \beta > 0$, for all pairs $(\alpha, e) \in R_+^2$.

<u>Assumption 3</u>: Innate economic ability α is a real number between zero and one. α is distributed among each generation of agents in a temporally independent and identical way. Let $f(\alpha)$ be the density of the distribution of innate ability within any generation. Then $f(\cdot)$ maps the unit interval continuously into R_+, and $f(0) > 0$.

Note that Assumption 3 (hereafter A3, etc.) implies there exists $\bar{F} > 0$ such that $\sup_\alpha |f(\alpha)| \leqq \bar{F}$. We are now in a position to characterize the solution to (3). This is done in the following theorem.

<u>Theorem 1</u>: Under A1, A2, and A3 there exists a unique solution $V^*(\cdot)$ for problem (3). $V^*(\cdot)$ is a strictly concave, differentiable function on $(0, \bar{y}]$. The optimal consumption policy, $C^*(y)$, is a continuous function of y.

<u>Proof</u>: Let $\mathfrak{F}$ denote the set of continuous real valued maps ϕ, such that $\phi:[0, \bar{y}] \to R_+$. Define a norm on $\mathfrak{F}$ by

$$||\phi|| \equiv \max_{0 \leqq y \leqq \bar{y}} |\phi(y)|.$$

Consider the map T on $\mathfrak{F}$ defined by

$$(T\phi)(y) \equiv \max_{c} E_\alpha U(c, \phi(h(\alpha, y-c))).$$

A1(i) and A2(i) imply $T: \mathcal{F} \to \mathcal{F}$. It may also be shown that T is a contraction on $\mathcal{F}$. Let $\phi, \psi \in \mathcal{F}$;

$$||T\phi - T\psi|| \equiv \max_y \left| \max_c E_\alpha U(c,\phi(h(\alpha,y-c))) - \max_c E_\alpha U(c,\psi(h(\alpha,y-c))) \right| .$$

Let $\hat{c}(y)$ give the maximum for $E_\alpha U(c,\phi)$ and $\hat{\hat{c}}(y)$ give the maximum for $E_\alpha U(c,\psi)$. Then

$$||T\phi - T\psi|| = \max_y |E_\alpha(U(\hat{c},\phi) - U(\hat{\hat{c}},\psi))|$$

$$\leqq \max_y \max\{|E_\alpha(U(\hat{c},\phi) - U(\hat{c},\psi))|; |E_\alpha(U(\hat{\hat{c}},\phi) - U(\hat{\hat{c}},\psi))|\}$$

$$\leqq \max\{\max_y |E_\alpha(U_2(\hat{c},\psi)\cdot[\phi-\psi])|; \max_y |E_\alpha(U_2(\hat{\hat{c}},\psi)\cdot[\phi-\psi])|\}$$

$$\leqq \gamma \max\{\max_{\alpha,y} |\phi(h(\alpha,y-\hat{c}(y))) - \psi(h(\alpha,y-\hat{c}(y)))|;$$

$$\max_{\alpha,y} |\phi(h(\alpha,y-\hat{\hat{c}}(y))) - \psi(h(\alpha,y-\hat{\hat{c}}(y)))|\}$$

$$\leqq \gamma \max_y |\phi(y) - \psi(y)| = \gamma||\phi - \psi|| .$$

Hence T is a contraction. By the Banach fixed point theorem $\exists$ a unique function V^* on $[0,\bar{y}]$ such that

$$V^* = TV^* .$$

By the definition of T, V^* is the solution to (3).

Define the sequence of functions $V^N: [0, \bar{y}] \to R_+$ inductively as follows:

$$V^1(y) \equiv \underset{c}{\text{Max}}\ E_\alpha U(c, U(h(\alpha, y-c), 0))$$

$$V^N(y) \equiv (TV^{N-1})(y), \quad N=2,3,\ldots$$

Clearly $\{V^N\} \to V^*$ uniformly. Let $\hat{c}^N(y)$ be the optimal policy function corresponding to V^N. An easy induction using the strict concavity of $U(\cdot, \cdot)$ and A2(ii) shows that $\{\hat{c}^N\}$ are single valued. Another induction using the continuous differntiability of $U(\cdot, \cdot)$ and $h(\cdot, \cdot)$ shows V^N to be differentiable on $(0, \bar{y}]$. Moreover, it follows from A1(ii), A2(i) and A3 that $\lim_{y\to 0} V'(y) = +\infty$. This property may be extended by induction to V^N, N=2,3,..., using the envelope theorem and the assumption $f(0) > 0$. The above considerations also imply that $0 < \hat{c}^N(y) < y$, $y \in (0, \bar{y}]$, N=1,2,... .

Notice that for $0 < \delta < 1$, $y_1, y_2 \in (0, \bar{y}]$ and $c_1 \equiv \hat{c}^1(y_1)$, $c_2 \equiv \hat{c}^1(y_2)$, we have

$$V^1(\delta y_1 + (1-\delta)y_2) = \underset{c}{\text{Max}}\ E_\alpha U(c_1, U(h(\alpha, \delta y_1 + (1-\delta)y_2 - c), 0))$$

$$\geqq E_\alpha U(\delta c_1 + (1-\delta)c_2, U(h(\alpha, \delta(y_1 - c_1) + (1-\delta)(y_2 - c_2)), 0))$$

$$\geqq \delta E_\alpha U(c_1, U(h(\alpha, y_1 - c_1), 0)) + (1-\delta)E_\alpha U(c_2, U(h(\alpha, y_2 - c_2), 0))$$

$$= \delta V^1(y_1) + (1-\delta)V^1(y_2),$$

with equality if and only if $y_1 = y_2$. Hence V^1 is strictly concave. An induction shows that V^N are strictly concave. Hence V^* is concave. But

$$V^*(\delta y_1 + (1-\delta)y_2) = \underset{c}{\text{Max}}\ E_\alpha U(c, V^*(h(\alpha, \delta y_1 + (1-\delta)y_2 - c))).$$

Now using the strict concavity of U and the argument immediately above, it is seen that V^* is indeed strictly concave. It is therefore differentiable almost everywhere. Now the envelope theorem implies

$$\text{(a)} \qquad \frac{d}{dy} V^*(y)^+ = E_\alpha(U_2(c^*, V^*)\frac{d}{dy} V^*(h)^+ h_2(\alpha, y-c^*))$$

and

$$\text{(b)} \qquad \frac{d}{dy} V^*(y)^- = E_\alpha(U_2(c^*, V^*)\frac{d}{dy} V^*(h)^- h_2(\alpha, y-c^*)),$$

where a "+" or "-" refers to right or left hand derivations, respectively. The monotonicity of h in α, the fact that $\frac{d}{dy}V^{*+} \neq \frac{d}{dy}V^{*-}$ at most on a set of measure zero, and the continuity of $f(\cdot)$ imply that the RHS of (a) and (b) are equal. Hence V^* is differentiable.

The continuity of the policy function follows from the continuity of $E_\alpha U(c, V^*(h(\alpha, y-c)))$ in c and y, the continuity of the interval [0, y] in y (when viewed as the image of a set valued map), and the fact that the maximizing c is unique for each y.

Q.E.D.

The proof of Theorem 1 illustrates the crucial role of the assumption (A1(iii)) $U_2 \leqq \gamma < 1$ in securing the existence and uniqueness of the indirect utility function V^*. This uniqueness is an indispensable property, as the description of family behavior would carry much less force if we were required to select arbitrarily among alternative parental perceptions of offspring's well being. The aforementioned assumption essentially requires a kind of discounting of the well being of the next generation. It says that a given perceived increment to offspring's utility causes parental utility to increase by less. In the case in which $U(c, V) = \bar{u}(c) + \gamma V$, the family may be viewed as collectively maximizing $E_\alpha(\sum_0^\infty \gamma^t \bar{u}(c_t))$ over all subsequent generations. Here the necessity that $U_2 = \gamma < 1$ is obvious.

Comment on the assumption $f(0) > 0$ is also in order. The effect of this assumption is to make it sufficiently likely that an offspring will have ability close to zero, so that parents will be unwilling to invest nothing in their offspring. In this way we assure interior solutions for optimal consumption. Some assumption of this kind is necessary if we wish to assure positive transfers. This assumption does cause problems for our concept of equilibrium however, as will be discussed below.

Along with the optimal consumption function $c^*(y)$, (3) also implies an optimal investment-training schedule, $e^*(y) = y - c^*(y)$. It seems natural to assume that "education" is a normal good, so we may (assuming further that c^* is differentiable) take it that $0 < e^{*\prime}(y) < 1$.[18] Moreover, it is clear that $h(1, e^*(0)) > 0$ and that $h(1, e^*(\bar{y})) < \bar{y}$, by virtue of assumption A2(i). Hence there exists an earnings level $\hat{y}$ for which $h(1, e^*(\hat{y})) = \hat{y}$. A mature individual earning $\hat{y}$ will provide his offspring with training in such a way that the offspring will be able to attain the same earnings only if he is among the most able people in the economy. It is apparent then that no family income could be above $\hat{y}$ if any of the ancestors of that family ever produced earnings less than $\hat{y}$. We shall chose income units so that $\hat{y} = 1$. Then in the following discussion, no generality is lost by considering income distributions on the unit interval only.[19] This procedure of bounding the income distribution will perhaps appear more reasonable when the reader recalls that we are excluding considerations of property income here.

The Effect of Parental Status on Offspring's Earnings:

Having deduced the mechanism by which parents decide on the amount of resources to invest in their offspring, we are now able to characterize the earnings of a mature individual as a function of his innate endowment and parent's income alone. Thus, if x denotes the earnings of any individual with endowment α and parent's income y, we have

$$(4) \qquad x = h(\alpha, e^*(y)) \equiv X(\alpha, y).$$

Each mature individual has limited social mobility. The earnings opportunities for any productive agent vary with that agent's economic ability, but over a range which is determined by the economic success of his parent. We shall want to study this social mobility in more detail, and therefore introduce the following definitions:

Let the range of possible incomes of the offspring of an agent with income y be $[x_0^1(y), x_1^1(y)]$. That is, $X(0, y) \equiv x_0^1(y)$ and $X(1, y) \equiv x_1^1(y)$. Moreover, let

$$y_0^1(x) \equiv \max\{y | X(\alpha, y) = x, \text{ some } \alpha \in [0, 1]\}$$

and

$$y_1^1(x) \equiv \min\{y | X(\alpha, y) = x, \text{ some } \alpha \in [0, 1]\}.$$

The n^{th} iterate of a function will be denoted by the superscript "n". Thus, for example,

$$x_0^n(y) = x_0^1(x_0^{n-1}(y)) = x_0^{n-1}(x_0^1(y))$$

is the lowest possible income of an n-generation descendant of someone whose income was y. Similar notation is used for the iterates of x_1^1, y_0^1, and y_1^1. Finally, define

$$\mathcal{A}(x, y) \equiv \max\{\alpha \in [0, 1] | X(\alpha, y) \leqq x\}, \quad x \geqq X(0, y)$$

$$\equiv 0, \quad \text{otherwise.}$$

Before proceeding we shall need to adopt an additional assumption, which formalizes some of the remarks made at the end of the previous section.

Assumption 4: $X(\cdot, \cdot)$ is assumed differentiable in y. Moreover, there exist $\lambda > 0$, $\zeta > 0$ such that

(i) $X_y(0, y) \leqq \lambda < 1, \forall y \in [0, \zeta)$

(ii) $X_y(1, y) \leqq \lambda < 1, \forall y \in (1-\zeta, 1]$.

We also require

(iii) $x_0^1(y) = y$ if and only if $y = 0$ and $x_1^1(y) = 1$ if and only if $y = 1$.

This assumption stipulates that for the least able of individuals with parents sufficiently poor or the most able whose parents are sufficiently well off, a marginal increment to parental earnings will result in a strictly smaller increment in offspring's earnings. These requirements do not seem inordinately restrictive.[20]

Clearly, the largest income possible for the parent of someone whose income is x is $y_0^1(x)$, while $y_1^1(x)$ is the smallest parental income possible. The following lemma summarizes some rather obvious properties of the functions defined above, and will be stated without proof.

Lemma 1: Under A1 - A4 we have

(i) $x_0^n(\cdot)$ and $x_1^n(\cdot)$ are continuously differentiable, strictly monotonically increasing functions on [0, 1], satisfying

$$x_0^1(y) < y, \quad x_1^1(y) > y, \quad \forall y \in (0, 1).$$

(ii) $y_0^n(\cdot)$ and $y_1^n(\cdot)$ are continuous, non-decreasing functions on [0, 1]. For $x \in [0, X(0, 1)]$, $y_0^1(x)$ satisfies

$$X(0, y_0^1(x)) = x,$$

While for $x \in [X(0, 1)]$, $y_0^1(x) \equiv 1$. For $x \in [X(1, 0), 1]$, $y_1^1(x)$ satisfies

$$X(1, y_1^1(x)) = x,$$

while for $x \in [0, X(1, 0)]$, $y_1^1(x) \equiv 0$. $y_0^1(\cdot)$ and $y_1^1(\cdot)$ are continuously differentiable a.e. on $[0, 1]$.

(iii) $\mathcal{A}(x, y)$ is a continuous function on $[0, 1]^2$, increasing in x and decreasing in y. For $x \in [X(0, y), X(1, y)]$, $\mathcal{A}(x, y)$ satisfies

$$X(\mathcal{A}(x, y), y) = x,$$

while for $x \geqq X(1, y)$, $\mathcal{A}(x, y) \equiv 1$. $\mathcal{A}(x, y)$ is continuously differentiable a.e. on $[0, 1]^2$.

For convenience we will state here several other results which will prove useful later on, and which follow readily from Lemma 1.

Lemma 2: $x_0^n(y_0^n(x)) = \min(x, x_0^n(1))$, and

$x_1^n(y_1^n(x)) = \max(x, x_1^n(0))$.

Proof: The proof is inductive. We prove the first statement only, the other being proved in an analogous manner. By Lemma 1

$$x_0^1(y_0^1(x)) = X(0,\ y_0^1(x)) = x, \quad x \leqq X(0,\ 1)$$

$$= X(0,\ 1) \quad \text{otherwise.}$$

Therefore

$$x_0^1(y_0^1(x)) = \min(x,\ x_0^1(1))$$

Suppose

$$x_0^{n-1}(y_0^{n-1}(x)) = \min(x,\ x_0^{n-1}(1)).$$

Then

$$x_0^n(y_0^n(x)) = x_0^1(x_0^{n-1}(y_0^{n-1}(y_0^1(x)))) = x_0^1$$

$$(\min(y_0^1(x),\ x_0^{n-1}(1)))$$

$$= \min(x_0^1(y_0^1(x)),\ x_0^1(x_0^{n-1}(1)))$$

$$= \min(x,\ x_0^1(1),\ x_0^n(1)) = \min(x,\ x_0^n(1)).$$

Q.E.D.

Lemma 2 shows the pseudo-inverse character of the functions $x_0^n(\cdot)(x_1^n(\cdot))$ and $y_0^n(\cdot)(y_1^n(\cdot))$.

Lemma 3: $x \in [x_0^n(y),\ x_1^n(y)]$ if and only if $y \in [y_1^n(x),\ y_0^n(x)]$, $n=1,2,\ldots$.

Proof: Using Lemma 2 and monotonicity of $x_0^n(\cdot)$ and $x_1^n(\cdot)$ we find

$$y \leqq y_0^n(x) \text{ iff } x_0^n(y) \leqq x_0^n(y_0^n(x)) = \min(x,\ x_0^n(1)),$$

while

$$y \geqq y_1^n(x) \text{ iff } x_1^n(y) \geqq x_1^n(y_1^n(x)) = \max\ (x,\ x_1^n(0)).$$

But

$$x_0^n(y) \leqq x_0^n(1) \forall y, \text{ and } x_1^n(y) \geqq x_1^n(0) \forall y.$$

Hence

$$y \leqq y_0^n(x) \text{ iff } x_0^n(y) \leqq x \text{ and } y \geqq y_1^n(x) \text{ iff } x_1^n(y) \leqq x.$$

Q.E.D.

Note that since $x_0^n(\cdot)$ and $x_1^n(\cdot)$ are strictly increasing functions, the lemma holds for open or half open intervals as well. Lemma 3 proves rigorously the obvious fact that x is in the range of possible incomes of n^{th} generation descendents of someone with income y if and only if y is among the possible antecedents, n generations removed, of x.

Lemma 4: The sequence of functions $\{x_0^n(\cdot)\}_{n=1}^{\infty}$ converges to the constant function zero uniformly. Similarly, $\{x_1^n(\cdot)\}$ converges uniformly to 1.

Proof: By monotonicity,

$$x_0^n(1) \geqq x_0^n(y) \qquad \forall y \in [0, 1], \quad \forall n,$$

and $$x_1^n(0) \leqq x_1^n(y) \qquad \forall y \in [0, 1], \quad \forall n.$$

Hence we need show only that $\{x_1^n(0)\}\uparrow 1$ and $\{x_0^n(1)\}\downarrow 0$. Now

$$x_0^n(1) = x_0^{n-1}(x_0^1(1)) < x_0^{n-1}(1), \text{ and}$$

$$x_1^n(0) = x_1^1(x_1^{n-1}(0)) > x_1^{n-1}(0).$$

Hence $\exists$ $\underline{x}$, $\bar{x}$ such that $\{x_0^n(1)\}\downarrow\underline{x}$, and $\{x_1^n(0)\}\uparrow\bar{x}$.

Now suppose $\underline{x} > 0$. Then $x_0^1(\underline{x}) < \underline{x}$. But $\{x_0^1(x_0^n(1))\} = \{x_0^{n+1}(1)\}\downarrow\underline{x}$. By continuity

$$\underline{x} = \lim_{n\to\infty} x_0^1(x_0^n(1)) = x_0^1(\lim_{n\to\infty} x_0^n(1)) = x_0^1(\underline{x}) < \underline{x}.$$

This contradicts $\underline{x} > 0$. Hence $\underline{x} = 0$. Similarly, one proves $\bar{x} = 1$.

Q.E.D.

Lemma 4 essentially illustrates that, independent of initial income, any family may rise to the top (fall to the bottom) of the income hierarchy given that its descendants have sufficiently good (bad) innate endowments.

The Movement of the Distribution of Earnings Across Generations:

In this section we will show how the evolution of the distribution of earnings in the economy may be characterized, given the intergenerational transfers discussed above. It may already be apparent that under the assumption that economic ability is distributed across each generation in a temporally independent and identical manner, the motion over time of the income of any family may be characterized as a Markoff process. This link between stochastic processes, especially Markoff chains, and income distributions has a long history in economic analysis. However, previous economic models have considered the intra-generational problem only.[21]

We shall assume that the initial state is characterized by a number of families with incomes distributed continuously over the unit interval. The normalized frequency distribution function describing the initial distribution of earnings is denoted by $g^0(\cdot)$, with

$$\int_0^1 g^0(y)dy = 1.$$

We will work with densities rather than cumulative distribution functions in the sequel. This requires that we demonstrate the existence of a stochastic density kernel for the process in question. The following theorem exhibits the transition kernel, and shows how the density of the distribution of income in any generation may be found if the density for the previous generation is known. The proof proceeds by deducing the cumulative distribution function, showing that function to be differentiable, and then calculating its derivative.

Theorem 2: Let $g^t(y)$ be the density of the income distribution in period t. Then the density in period t+1 is given by the following formula.

$$g^{t+1}(x) = \int_{y_1^1(x)}^{y_0^1(x)} f(\mathcal{A}(x, y))[X_\alpha(\mathcal{A}(x, y), y)]^{-1} g^t(y)dy \tag{5}$$

Proof: Let $G^{t+1}(x)$ denote the probability that an individual selected at random in period t+1 will have an income less than or equal to x. Since the event $\{x_{t+1} \leq \bar{x}\}$ is equivalent to the event $\{y_t \leq y_0^1(\bar{x})\} \cap \{\alpha_{t+1} \leq \mathcal{A}(\bar{x}, y_t)\}$, the independence assumption (A3) makes it apparent that

$$G^{t+1}(\bar{x}) = \int_0^{y_0^1(\bar{x})} \int_0^{\mathcal{A}(\bar{x}, y)} f(\alpha) g^t(y) d\alpha dy. \tag{5a}$$

One easily verifies that $G^{t+1}(0) = 0$, $G^{t+1}(1) = 1$, and $G^{t+1}(\cdot)$ is monotone increasing and right-continuous. Thus, it is indeed a distribution function. Throughout this paper integration will be intended in the sense of Lebesque, and $\mu(\cdot)$ will denote the Lebesque measure on the Borel sets of the unit interval, $\mathcal{B}[0, 1]$.

The differentiability of $G^{t+1}(\cdot)$ is less obvious, since neither $y_0^1(\cdot)$ nor $\mathcal{A}(\cdot, \cdot)$ are differentiable. Their left and right side derivatives exist everywhere however, and using these we may show the identity of the left and right derivatives of $G^{t+1}(\cdot)$ everywhere on [0, 1]. In what follows let $\frac{dq^+}{dx}$ and $\frac{dq^-}{dx}$ represent respectively the right and left derivatives of some function $q(\cdot)$. Now from the definitions of $y_0^1(\cdot)$ and $\mathcal{Q}(\cdot, \cdot)$, Lemma 1, and the implicit function theorem, we have that:

$$\frac{d}{dx} y_0^1(x)^+ = [X_y(0, y_0^1(x))]^{-1}, \quad 0 \leqq x < X(0, 1)$$

$$= 0, \; X(0, 1) \leqq x \leqq 1$$

$$\frac{d}{dx} y_0^1(x)^- = [X_y(0, y_0^1(x))]^{-1}, \quad 0 \leqq x \leqq X(0, 1)$$

$$= 0, \; X(0, 1) < x \leqq 1$$

while

$$\frac{\delta}{\delta x} \mathcal{Q}(x, y)^+ = 0, \quad 0 \leqq x < x_0^1(y) \text{ and } x_1^1(y) \leqq x \leqq 1$$

$$= [X_\alpha(\mathcal{Q}(x, y), y)]^{-1}, \; x_0^1(y) \leqq x < x_1^1(y).$$

and

$$\frac{\delta}{\delta x} \mathcal{Q}(x, y)^- = 0, \quad 0 \leqq x \leqq x_0^1(y) \text{ and } x_1^1(y) < x \leqq 1,$$

$$= X_\alpha(\mathcal{Q}(x,\ y),\ y)^{-1},\ x_0^1(y) < x \leqq x_1^1(y).$$

Combining these with the differentiation of equation (5a) yields

(6) $$\frac{d}{dx} G^{t+1}(x)^+ = g^t(y_0^1(x)) \int_0^{\mathcal{Q}(x,\ y_0^1(x))} f(\alpha)d\alpha \frac{d}{dx} y_0^1(x)^+$$

$$+ \int_0^{y_0^1(x)} t^t(y) f(\mathcal{Q}(x,\ y)) \frac{\delta}{\delta x} \mathcal{Q}(x,\ y)^+ dy$$

while

(7) $$\frac{d}{dx} G^{t+1}(x)^- = g^t(t_0^1(x)) \frac{d}{dx} y_0^1(x)^- \int_0^{\mathcal{Q}(x,\ y_0^1(x))} f(\alpha)d\alpha$$

$$+ \int_0^{y_0^1(x)} g^t(y) f(\mathcal{Q}(x,\ y)) \frac{\delta}{\delta x} \mathcal{Q}(x,\ y)^- dy$$

Consider the second terms on the right hand sides of equations (6) and (7). These are integrals whose interands are identical except on the set $\{y_0^1(x),\ y_1^1(x)\}$ which has Lebesque measure zero. Hence these terms are equal. The first terms on the RHS of (6) and (7) can differ only at the point $x = X(0,\ 1)$. At this point however,

$$\mathcal{Q}(x,\ y_0^1(x)) = \mathcal{Q}(X(0,\ 1),\ y_0^1(X(0,\ 1))) = \mathcal{Q}(X(0,\ 1),\ 1) = 0.$$

Hence $G^{t+1}(\cdot)$ is differentiable and $g^{t+1}(\cdot)$ may be identified with the RHS of (6) or (7).

Observe now that for $x \notin [x_0^1(y),\ x_1^1(y))$, $\frac{\partial}{\partial x}\mathcal{Q}(x,\ y)^+ = 0$. By Lemma 3 this condition holds iff $y \notin [y_1^1(x),\ y_0^1(x))$. Further,

$$\frac{d}{dx}\, y_0^1(x)^+ \equiv 0, \quad \forall x \geqq X(0,\ 1)$$

while

$$\mathcal{Q}(x,\ y_0^1(x)) \equiv 0, \quad \forall x \leqq X(0,\ 1).$$

Hence

$$g^{t+1}(x) = \int_{y_1^1(x)}^{y_0^1(x)} g^t(y) f(\mathcal{Q}(x,\ y))[X_\alpha(\mathcal{Q}(x,\ y),\ y)]^{-1} dy.$$

Q.E.D.

This theorem illustrates that previous attempts to deduce a simple relationship between the distribution of abilities and the distribution of earnings could not possibly be successful in a world in which social origin influences the acquisition of skills. In such a world there is a natural, though rather complex link between the observed earnings distribution and that which obtained among the previous generation of workers. The distribution of abilities plays a role in this process, though little can be said about the nature of that role a priori.

We may now define the stochastic density kernel $K^1(x, y)$ as follows:

$$K^1(x, y) \equiv f(\mathcal{Q}(x, y))[X_\alpha(\mathcal{Q}(x, y), y)]^{-1}, \forall(x, y) \ni y \in (y_1^1(x), y_0^1(x))$$

$$\equiv 0, \text{ otherwise.}$$

Then (5) becomes

(8) $$g^{t+1}(x) = \int_{[0, 1]} K^1(x, y) g^t(y)dy.$$

Effectively $K^1(x, y)$ is the probability ex ante that the offspring of someone with income y will attain the income x. Notice that $K^1(\cdot, \cdot)$ is a continuous function on $[0, 1]^2$ if and only if $f(0) = f(1) = 0$. Thus the transition kernel is generally discontinuous, and this causes some problems in establishing the uniqueness of equilibrium.[22] These problems are resolved in the next set of results.

Define inductively the n-step stochastic density kernel $K^n(x, y)$ by the equation

(9) $$K^n(x, y) = \int_{[0, 1]} K^1(x, z)K^{n-1}(z, y)dz, \quad n=2,3,\ldots .$$

The following lemma gives some useful properties of the functions $\{K^n(\cdot, \cdot)\}$.

Lemma 5:

(i) $\int_{[0, 1]} K^n(x, y)dx \equiv 1, \quad \forall y \in [0, 1], \quad n=1,2,\ldots$

(ii) $\exists M > 0 \ni \sup_{(x,y)\in[0,1]^2} |K^n(x, y)| \leq M, \quad n=1,2,\ldots$

(iii) $K^n(x, y) > 0$ iff $(x, y) \in S_n \equiv \{(x, y) | x \in (x_0^n(y), x_1^n(y))\}$.

Proof: (i) $\int_{[0, 1]} K^1(x, y)dx = \int_{[x_0^1(y),x_1^1(y)]} f(\mathcal{A}(x, y))[X_\alpha(\mathcal{A}(x, y), y)]^{-1}dx.$

Consider the change of variables $\alpha = \mathcal{A}(x, y)$; $x = X(\alpha, y)$. Then $d\alpha = \frac{\delta}{\delta x}\mathcal{A}(x, y)dx = [X_\alpha(\mathcal{A}(x, y), y)]^{-1}dx$. Now $x = x_0^1(y) \Rightarrow \alpha = \mathcal{A}(x_0^1(y), y) \equiv 0$, and $x = x_1^1(y) \Rightarrow \alpha = \mathcal{A}(x_1^1(y), y) \equiv 1$. Thus

(10) $$\int_{[0,1]} K^1(x, y)dx = \int_{[0, 1]} f(\alpha)d\alpha = 1.$$

One then uses Fubini's theorem with an induction to show that (10) holds for all positive integers n.

(ii) $$\sup_{x,y} K^1(x, y) = \sup_{y\in[0,1]} \sup_{x\in[x_0^1(y),x_1^1(y)]} [f(\mathcal{a}(x, y))(X_\alpha(\mathcal{a}(x, y), y))^{-1}]$$

$$\leqq \sup_{\alpha\in[0,1]} f(\alpha) \inf_{(\alpha,y)\in[0,1]^2} X_\alpha(\alpha, y)^{-1}$$

$$\leqq \bar{F}/\beta \equiv M < \infty$$

Now $\forall n > 1$,

$$K^n(x, y) = \int_{[0,1]} K^1(x,z)K^{n-1}(z,y)dz$$

$$\leqq (\sup_{(x,y)\in[0,1]^2} K^1(x,y)) \int_{[0,1]} K^{n-1}(z,y)dz \leqq M < \infty.$$

(iii) Inductively

$$K^1(x,y) \equiv f(\mathcal{a}(x,y))[X_\alpha(\mathcal{a}(x,y),y)]^{-1} > 0$$

for $y \in (y_1^1(x), y_0^1(x))$, or equivalently $x \in (x_0^1(y),x_1^1(y))$. Now, suppose $K^{n-1}(x,y) > 0$ iff $x \in (x_0^{n-1}(y),x_1^{n-1}(y))$. Then define

$$A_n(x,y) \equiv \{z \in [0,1] | K^1(x,z)K^{n-1}(z,y) > 0\}.$$

Now

$$K^n(x,y) > 0 \text{ iff } \mu(A_n(x,y)) > 0.$$

But

$$K(x,z) > 0 \quad \text{iff} \quad z \in (y_1^1(x), y_0^1(x)), \quad \text{while}$$

$$K^{n-1}(z,y) > 0 \quad \text{iff} \quad z \in (x_0^{n-1}(y), x_1^{n-1}(y)).$$

Hence

$$A_n(x,y) = (y_1^1(x), y_0^1(x)) \cap (x_0^{n-1}(y), x_1^{n-1}(y))$$

and

$$\mu(A_n(x,y)) > 0 \quad \text{iff} \quad \text{both}$$

(a) $y_0^1(x) > x_0^{n-1}(y)$ and

(b) $y_1^1(x) < x_1^{n-1}(y)$

Now (a) is equivalent to

$$\min (x, x_0^1(1)) = x_0^1(y_0^1(x)) > x_0^n(y), \text{ by Lemma 2.}$$

Since $x_0^1(1) > x_0^n(y)$ $\forall y \in (0,1)$, this implies that (a) is equivalent to

$$x > x_0^n(y).$$

Similarly (b) holds iff

$$x < x_1^n(y).$$

Thus, $\mu(A_n(x, y)) > 0$ iff $(x, y) \in S_n$. Q.E.D.

Before moving on to a treatment of the existence, uniqueness and stability of equilibrium we must confront directly the problem of the discontinuity of $K^1(\cdot, \cdot)$. Below we show that, under our initial assumptions, the points of discontinuity of the transition kernels may be "removed" without significant effect. This will enable us to avoid difficult probabilistic arguments in establishing the uniqueness of equilibrium.

Theorem 3:

For every pair of positive numbers (ε, η) there exists a positive number δ and a sequence of set valued mappings

$\{E_n\}$ with the property

$$E_n: [0,1]^2 \to \mathcal{B}[0,1]$$

such that for any pair $(y_1,y_2) \in [0,1]^2$,

(a) $|y_1-y_2| < \delta \Rightarrow \int_{[0,1]-E_n(y_1,y_2)} |K^n(x,y_1)-K^n(x,y_2)|dx < \varepsilon$

and

(b) $\mu(E_n(y_1,y_2)) < \eta, \quad n=1,2,\ldots$

Note that ε and η may be chosen arbitrarily, and while δ depends on both ε and η, it is independent of y_1 and y_2.

Proof: Without loss of generality take $y_1 < y_2$. Define the sequence of set valued functions $\{E_n\}$ as follows:

$$E_n(y_1,y_2) \equiv \{[x_0^n(y_1),x_0^n(y_2)]\cup[x_1^n(y_1),x_1^n(y_2)]\}.$$

We will first show that with E_n so defined, for every $\varepsilon > 0$, $\bar{N} > 0$, there exists a positive number δ_1 (depending on ε and $\bar{N}$) such that

$$\max_{1 \leq n \leq \bar{N}} \left\{ \int_{[0,1]-E_n(y_1,y_2)} \left| K^n(x,y_1) - K^n(x,y_2) \right| dx \right\} < \frac{\varepsilon}{2},$$

whenever $|y_1 - y_2| < \delta_1$.

From Lemma 5 we know that $K^n(x,y)$ is continuous on $S_n \equiv \{(x,y)\,|\,x \in (x_0^n(y), x_1^n(y))\}$, and is bounded on $[0,1]^2$. Hence $K^n(\cdot,\ \cdot)$ is uniformly continuous on S_n. Figure 1 illustrates the argument. The set S_n is the area between the lines $x = x_0^n(y)$ and $x_1^n(y)$. Now for $x \subset \{[0, x_0^n(y_1)) \cup (x_1^n(y_2), 1]\}$ we have that

$$\left| K^n(x,y_1) - K^n(x,y_2) \right| = 0.$$

Clearly then

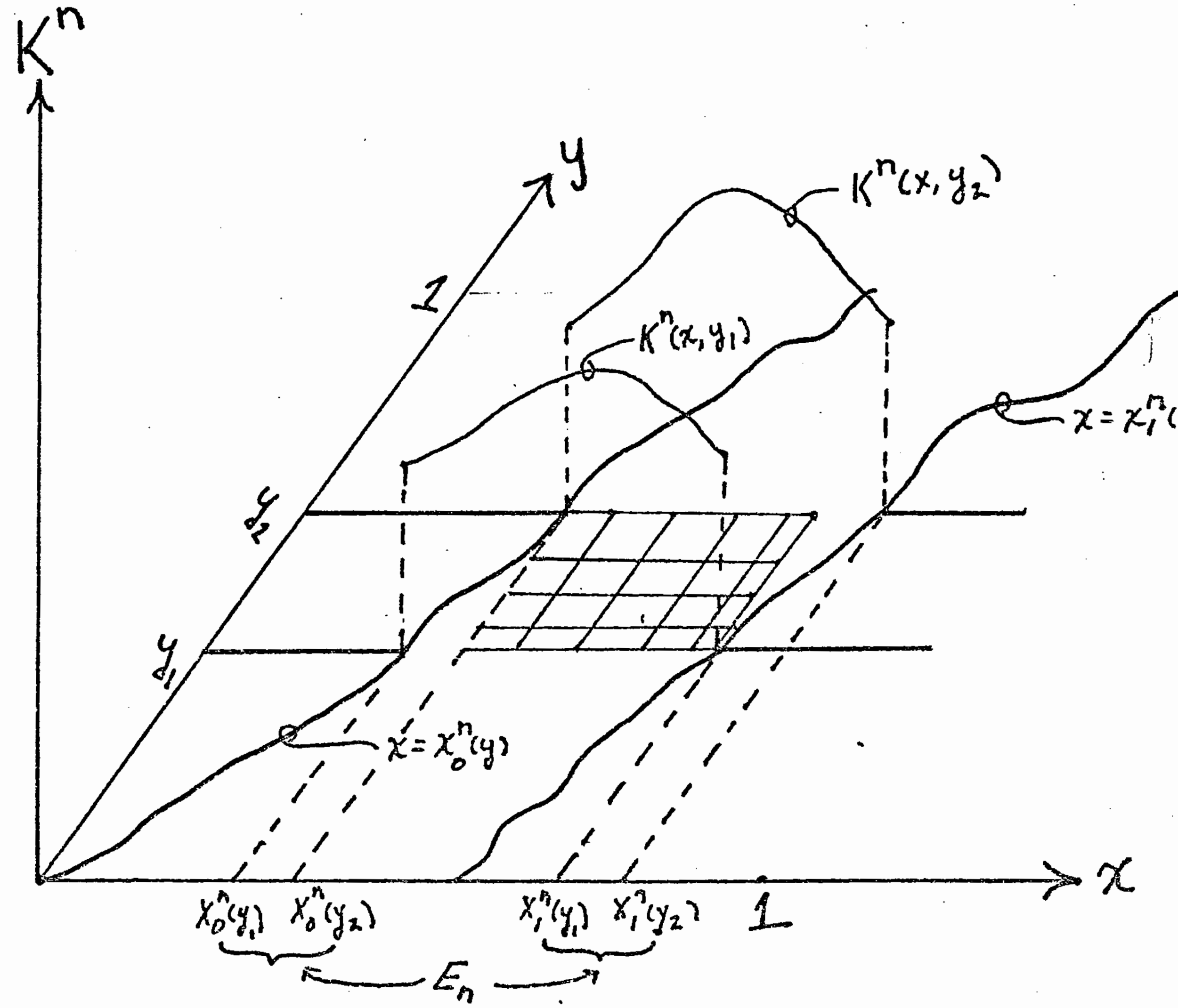
$$(\{[x_0^n(y_1), x_1^n(y_2)] - E_n(y_1,y_2)\} \times (y_1,y_2)) \subset S_n.$$

Hence, by uniform continuity, given $\varepsilon > 0, \exists\, \delta_1^n > 0 \ni$

$$\sup_{x \in \{[x_0^n(y_1), x_1^n(y_2)] - E_n\}} \left| K^n(x,y_1) - K^n(x,y_2) \right| < \frac{\varepsilon}{2},$$

whenever $|y_1 - y_2| < \delta_1^n$.

Hence



K^n
y
$K^n(x, y_2)$
1
$K^n(x, y_1)$
$x = x_1^n(y)$
y_2
y_1
$x = x_0^n(y)$
x
$x_0^n(y_1)$
$x_0^n(y_2)$
$x_1^n(y_1)$
$x_1^n(y_2)$
1
E_n

FIGURE 1

$$\int_{[0,1]-E_n} |K^n(x,y_1) - K^n(x,y_2)|dx$$

$$= [\int_{[0,x_0^n(y_1))\cup(x_1^n(y_2),1]} |K^n(x,y_1) - K^n(x,y_2)|dx +$$

$$\int_{[x_0^n(y_1),x_1(y_2)]-E_n(y_1,y_2)} | K^n(x,y_1) - K^n(x,y_2)|dx$$

$$= 0 + \int_{[x_0^n(y_1),x_1^n(y_2)]-E_n(y_1,y_2)} |K^n(x,y_1) - K^n(x,y_2)|dx < \frac{\varepsilon}{2} ,$$

whenever $|y_1-y_2| < \delta_1^n$. Thus $\delta_1 \equiv \min_{1\leq n\leq \bar{N}} \{\delta_1^n\}$ is the required positive number.

From equation (9) we have that, given (y_1,y_2),

$$\int_{[0,1]-E_n(y_1,y_2)} |K^n(x,y_1) - K^n(x,y_2)|dx =$$

$$\int_{[0,1]-E_n(y_1,y_2)} | \int_{[0,1]} K^1(x,z)[K^{n-1}(x,y_1) -$$

$$K^{n-1}(z,y_2)]dz|dx$$

$$\leq \int_{[0,1]-E_n(y_1,y_2)} \int_{[0,1]} K^1(x,z)\,|K^{n-1}(z,y_1) - K^{n-1}(z,y_2)|\,dz\,dx$$

$$= \int_{[0,1]} \left(\int_{[0,1]-E_n(y_1,y_2)} K^1(x,z)\,dx \right) |K^{n-1}(z,y_1) - K^{n-1}(z,y_2)|\,dz$$

$$\leq \int_{[0,1]-E_{n-1}(y_1,y_2)} |K^{n-1}(z,y_1) - K^{n-1}(z,y_2)|\,dz +$$

$$\int_{E_{n-1}(y_1,y_2)} |K^{n-1}(z,y_1) - K^{n-1}(z,y_2)|\,dz$$

$$\leq 2M\mu(E_{n-1}) + \int_{[0,1]-E_{n-1}(y_1,y_2)} |K^{n-1}(z,y_1) - K^{n-1}(z,y_2)|\,dz$$

where M is the bound on K^n from Lemma 5. Hence we have shown

$$\int_{[0,1]-E_n(y_1,y_2)} |K^n(x,y_1) - K^n(x,y_2)|\,dx \leq 2M\mu(E_{n-1}(y_1,y_2)) +$$

$$\int_{[0,1]-E_{n-1}(y_1,y_2)} |K^{n-1}(x,y_1) - K^{n-1}(x,y_2)|\,dx.$$

Now by A4 we have

$$\mu(E_{n-1}(y_1,y_2)) \leqq |x_0^{n-1}(y_2) - x_0^{n-1}(y_1)| + |x_1^{n-1}(y_2) - x_1^{n-1}(y_1)|$$

$$= |x_0^1(x_0^{n-2}(y_2)) - x_0^1(x_0^{n-2}(y_1))| + |x_1^1(x_1^{n-2}(y_2)) - x_1^1(x_1^{n-2}(y_1))|$$

Now by Lemma 4, given $\xi > 0, \exists \bar{N}$ such that

$n \geqq \bar{N} \Rightarrow x_0^n(y) < \xi$ and $x_1^n(y) > 1-\xi$ for all $y \in [0,1]$.

Hence, it follows from A4(i) and (ii) that $\exists \lambda \in (0,1)$ for which

$$\mu(E_{n-1}(y_1,y_2)) \leqq \lambda(|x_0^{n-2}(y_2) - x_0^{n-2}(y_1)| + |x_1^{n-2}(y_2) - x_1^{n-2}(y_1)|)$$

for all $n \geqq \bar{N} + 2$. Iterating this inequality gives

$$\mu(E_{n-1}(y_1,y_2)) \leqq 2\lambda^{n-\bar{N}-2}\max\{|x_0^{\bar{N}}(y_2) - x_0^{\bar{N}}(y_1)|;\ |x_1^{\bar{N}}(y_2) - x_1^{\bar{N}}(y_1)|\},$$

for $n \geqq \bar{N} + 2$.

From this it follows that

$$\int_{[0,1]-E_n(y_1,y_2)} |K^n(x,y_2)|dx \leq 4M\lambda^{n-\bar{N}-2}\max\{|x_0^{\bar{N}}(y_2) - x_0^{\bar{N}}(y_1)|;$$

$$|x_1^{\bar{N}}(y_2) - x_1^{\bar{N}}(y_1)|\} +$$

$$\int_{[0,1]-E_{n-1}(y_1,y_2)} |K^{n-1}(x,y_1) -$$

$$K^{n-1}(x,y_2)|\,dx,$$

for $n \geqq \bar{N} + 2$. Iterating this inequality we find

$$\int_{[0,1]-E_n(y_1,y_2)} |K^n(x_1y_1) - K^n(x_1y_2)|\,dx \leqq \frac{4M}{\lambda\bar{N}+2}(\lambda^n+\lambda^{n-1}+\ldots+\lambda^{\bar{N}+2})$$

$$\cdot\max\{|x_0^{\bar{N}}(y_2) - x_0^{\bar{N}}(y_1)|;|x_n^{\bar{N}}(y_2) - x_1^{\bar{N}}(y_1)|\} +$$

$$\int_{[0,1]-E_{\bar{N}+2}(y_1,y_2)} |K^{\bar{N}+2}(x,y_1) - K^{\bar{N}+2}(x,y_2)|\,dx$$

$$\leqq \frac{4M}{1-\lambda}\max\{|x_0^{\bar{N}}(y_2) - x_0^{\bar{N}}(y_1)|;|x_n^{\bar{N}}(y_2) - x_1^{\bar{N}}(y_1)|\} +$$

$$\int_{[0,1]-E_{\bar{N}+2}(y_1,y_2)} |K^{\bar{N}+2}(x,y_2) - K^{\bar{N}+2}(x,y_1)|\,dx,$$

for all $n \geqq \bar{N} + 2$.

Given $\varepsilon > 0$, let us now choose $\delta_1 > 0$ so that

$$\max_{0 \leqq n \leqq \bar{N}+2} \int_{[0,1]-E_n(y_1,y_2)} |K^n(x,y_1) - K^n(x,y_2)|dx < \frac{\varepsilon}{2} ,$$

whenever $|y_1-y_2| < \delta_1$.

Furthermore, given the same $\varepsilon > 0$ as above and any $\eta > 0$, choose $\delta_2 > 0$ such that

$$\max_{0 \leqq n \leqq \bar{N}} (\max\{|x_0^n(y_2) - x_0^n(y_1)|;|x_1^n(y_2) - x_1^n(y_1)|\}) <$$

$$\min\{\frac{\varepsilon(1-\lambda)}{8M}, \frac{\eta}{2} \},$$

whenever $|y_1-y_2| < \delta_2$,

which can be done by virtue of the continuity of $x_0^n(\cdot)$ and $x_1^n(\cdot)$. Now recall that

$$\mu(E_n(y_1,y_2)) \leqq 2 \max\{|x_0^n(y_2) - x_0^n(y_1)|;|x_1^n(y_2) - x_1^n(y_2)|\}$$

$$\leqq 2\lambda^{n-\bar{N}}\max\{|x_0^{\bar{N}}(y_2) - x_0^{\bar{N}}(y_2)|;|x_1^{\bar{N}}(y_2) - x_1^{\bar{N}}(y_1)|\}$$

for $n \geqq \bar{N}$. It is apparent then that given $\varepsilon > 0$, $\eta > 0$,

$|y_1 - y_2| < \min(\delta_1, \delta_2) \Rightarrow$

(a) $\int_{[0,1]-E_n(y_1,y_2)} |K^n(x,y_1) - K^n(x,y_2)|dx < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, and

(b) $\mu(E_n(y_1,y_2)) < 2\,\frac{\eta}{2} = \eta$.

Hence $\delta \equiv \min(\delta_1, \delta_2)$ is a number whose existence we asserted and $(\delta, \{E_n(\cdot, \cdot)\})$ satisfy conditions (a) and (b) with ε, η given.

Q.E.D.

The Equilibrium Distribution of Income:

Since the early writings of Pareto, economists have considered the relative stability of the size distribution of income an important empirical fact which theory must explain. Hence, the notion of an "equilibrium" or invariant distribution arose rather early in the attempts to rationalize observed earnings distributions.[23] Champernowne defined "static equilibrium" as "... a state of affairs where individual incomes may change, but where the aggregate distribution remains unchanged..."[24] We shall employ a similar equilibrium notion here.

Because a mature individual's earnings depends on his endowment of innate ability, it will generally not be possible for families to secure a planned stream of incomes over time. A family's earnings (and hence consumption) will tend to fluctuate randomly over generations due to the stochastic nature of the descendants' natural economic aptitudes. Nonetheless, the fact that every family uses the same rule by which to decide upon investments in their young implies that a deterministic relationship exists between the aggregate distributions of earnings at successive periods of time. We shall consider the earnings distribution to be in equilibrium when the parental transfer decisions and the stochastic assignments of ability interact in such a way that successive earnings distributions are all the same.

Put somewhat more formally, $g^*(x)$ is the density of an equilibrium earnings distribution if and only if it satisfies

$$(11) \qquad g^*(x) = \int_{[0,1]} K^1(x,y)g^*(y)dy.$$

That is, an equilibrium is an ergodic distribution of the Markoff process whose stochastic density kernel is $K^1(\cdot, \cdot)$. We are now able to prove that our economy possesses a unique and stable equilibrium earnings distribution. To do this we employ Theorem 3. This allows a non-trivial extension of Feller's classical theorem on Markoff chains to the case where the transition kernel $K^1(\cdot, \cdot)$ is not necessarily continuous on $[0,1]^2$. [25]

<u>Theorem 4</u>: Under A1 - A4, there exists a unique density function $g^*(\cdot)$ satisfying (11). Moreover, $g^*(\cdot)$ has the property that

$$\lim_{t\to\infty} g^t(x) = g^*(x), \quad \forall g^0(\cdot), \quad \forall x \in [0,1],$$

where $\{g^t\}$ is defined inductively by

$$g^t(x) = \int_{[0,1]} K^1(x,y)g^{t-1}(y)dy, \quad t=1,2,\ldots, \quad g^0(\cdot) \text{ given.}$$

This theorem asserts the existence of a unique solution for (11). It also states that no matter what the initial distribution of earnings in the economy, the income distribution will always approach this solution with the passage of time. In this sense $g^*(\cdot)$ is a stable equilibrium. The proof of Theorem 4, which combines the basics of Feller's original method with Theorem 3 above, requires two additional lemmata.

Definition: A family of functions $\{\phi_n\}$, $\phi_n: R \to R$, $n=1,2,\ldots$, is said to be equicontinuous if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|\phi_n(x) - \phi_n(y)| < \varepsilon \quad \text{whenever } |x-y| < \delta, \quad n=1,2,\ldots .$$

Lemma 6: Let $u^0(x)$ be a bounded integrable real valued function on $[0,1]$. Define inductively

(12) $$u^n(y) = \int_{[0,1]} K^1(x,y)u^{n-1}(x)dx, \quad n=1,2,\ldots$$

Then the family of functions $\{u^n\}$, $n=1,2,\ldots$, is equicontinuous.

Proof: Let $\varepsilon > 0$ be given. It is clear that

$$u^n(y) = \int_{[0,1]} K^1(x,y)u^{n-1}(x)dx = \int_{[0,1]} \int_{[0,1]} K^1(x,y)K^1(z,x) u^{n-2}(z)dzdx$$

$$= \int_{[0,1]} (\int_{[0,1]} K^1(z,x)K^1(x,y)dx)u^{n-2}(z)dz$$

$$= \int_{[0,1]} K^2(x,y)u^{n-2}(x)dx = \ldots = \int_{[0,1]} K^n(x,y)u^0(x)dx.$$

Whence it follows that, given $y_1,y_2 \in [0,1]$,

$$|u^n(y_1) - u^n(y_2)| = |\int_{[0,1]} (K^n(x,y_1) - K^n(x,y_2))u^0(x)dx|$$

$$\leqq \int_{[0,1]} |u^0(x)||K^n(x,y_1) - K^n(x,y_2)|dx$$

$$\leqq \hat{N} \int_{[0,1]} |K^n(x,y_1) - K^n(x,y_2)|dx,$$

where $\hat{N} \equiv \sup_{x\in[0,1]} |u_0(x)| < \infty$, by hypothesis.

Now by Theorem 3 there exists $\delta > 0$ and a sequence of sets $E_n(y_1,y_2)$ for which

$$\mu(E_n) < \frac{\varepsilon}{2\hat{N}M}$$

and

$$\int_{[0,1]-E_n} |K^n(x,y_1) - K^n(x,y_2)|dx < \frac{\varepsilon}{2\hat{N}}, \quad n=1,2,\ldots$$

whenever $|y_1-y_2| < \delta$. Thus, for such y_1 and y_2 and all n,

$$|u^n(y_1) - u^n(y_2)| \leqq \hat{N} \int_{[0,1]} |K^n(x,y_1) - K^n(x,y_2)|dx$$

$$= \hat{N}\left(\int_{E_n} |K^n(x,y_1) - K^n(x,y_2)|dx + \int_{[0,1]-E_n} |K^n(x,y_1) - K^n(x,y_2)|dx\right)$$

$$< \hat{N}M\left(\frac{\varepsilon}{2\hat{N}M}\right) + \hat{N}\left(\frac{\varepsilon}{2\hat{N}}\right) = \varepsilon.$$

This proves equicontinuity.

Q.E.D.

<u>Lemma 7</u>: Let $u^0(x)$ be an arbitrary continuous function on $[0,1]$. Define $u^n(y)$ inductively as in (12). Then

(i) $u^n(\bar{y}) \equiv \max_{y\in[0,1]} u^n(y) \leqq \max_{x\in[0,1]} u^0(x)$,

and

(ii) $u^n(\bar{y}) = \max_{x\in[0,1]} u^0(x)$ only if

$$u^0(x) \equiv u^n(\bar{y}), \quad \forall x \in (x_0^n(\bar{y}),\ x_1^n(\bar{y})).$$

Proof:

(i) $$\max_y u^n(y) = \max_y \int_{[0,1]} u^0(x)K^n(x,y)dx$$

$$\leqq [\max_x u^0(x)]\max_y \int_{[0,1]} K^n(x,y)dx = \max_x u^0(x).$$

(ii) Suppose the contrary. Then $\exists\hat{x} \in (x_0^n(\bar{y}),\ x_1^n(\bar{y}))$ and

$\varepsilon > 0$ such that

$u^0(\hat{x}) < u^n(\bar{y}) - \varepsilon.$

Then by continuity there is a δ-neighborhood of $\hat{x}$, $N_\delta(\hat{x})$, for which

(a) $U^0(x) < u^n(\bar{y}) - \varepsilon, \quad \forall x \in N_\delta(\hat{x}).$

Now since $u^n(\bar{y}) = \max_{x\in[0,1]} u^0(x)$, we have

$$u^n(\bar{y}) = \int_{[0,1]} u^0(x)K^n(x,\bar{y})dx = \Big(\int_{[0,1]-N_\delta(x)} + \int_{N_\delta(x)}\Big)u^0(x)K^n(x,\bar{y})dx$$

$$\leqq \int_{[0,1]-N_\delta(\hat{x})} u^n(\bar{y})K^n(x,\bar{y})dx + \int_{N_\delta(\hat{x})} u^0(x)K^n(x,\bar{y})dx\ .$$

That is

$$\text{(b)}\quad u^n(\bar{y}) \leqq \int_{[0,1]} u^n(\bar{y})K^n(x,\bar{y})dx - \varepsilon\int_{N_\delta(\hat{x})} K^n(x,\bar{y})dx$$

$$< u^n(\bar{y}).$$

Inequality (b) follows from inequality (a) and the strict positivity of $K^n(x,\bar{y})$ on $(x_0^n(\bar{y}), x_1^n(\bar{y}))$. But inequality (b) is an obvious contradiction.

Q.E.D.

Proof of Theorem 4: As in Lemma 7, let $u^0(x)$ be an arbitrary continuous function on [0,1], and let $\{u^n(x)\}$ be defined by (12). Lemma 6 proved $\{u^n\}$ to be equicontinuous. Furthermore, $\{u^n\}$ are uniformly bounded as a consequence of Lemma 5(ii). Hence, by the Ascoli-Arzela Lemma,[26] there exists a subsequence $\{u_{n_k}\}$ of $\{u^n\}$ such that $\{u_{n_k}\}$ are uniformly convergent to some function w_0 continuous on [0,1].

Consider now the sequence of functions $\{w_j\}$ defined by

$$w_j(y) = \int_{[0,1]} w_0(x)K^j(x,y)dx, \quad j=1,2,\ldots .$$

It is apparent that for each $j=1,2,\ldots$, since $\{u_{n_k}\}_{k=1}^{\infty} \to w_0$ uniformly, $\{u_{n_{k+j}}\}_{k=1}^{\infty} \to w_j$ uniformly. It is equally clear from uniform convergence that

(13) $$\lim_{k\to\infty} [\max_{x\in[0,1]} u_{n_{k+j}}(x)] = \max_{x\in[0,1]} w_j(x), \quad j=0,1,2,\ldots$$

Now by Lemma 7(i) the sequence of numbers $\{\max_{x\in[0,1]} u^n(x)\}$ is monotonically decreasing. It is obviously bounded below, and hence converges to some number, m. Thus, every subsequence of this sequence converges to m. It follows from this and (13) that

$$\max_{x\in[0,1]} w_j(x) = m \qquad \forall j = 0,1,2,\ldots$$

Then Lemma 7(ii) implies

$$w_0(x) \equiv m, \quad \forall x \in \bigcap_{y\in[0,1]} (x_0^j(y),\ x_1^j(y)), \quad j=1,2,\ldots$$

That is,

$$w_0(x) \equiv m, \quad \forall x \in \bigcup_{j=1}^{\infty} \bigcap_{y\in[0,1]} (x_0^j(y),\ x_1^j(y)) = \bigcup_{j=1}^{\infty} (x_0^j(1),\ x_1^j(0))$$

$$= (0,1)$$

by Lemma 4. Then continuity gives $w_0(x) \equiv M$, $x \in [0,1]$.

Thus we have shown that an arbitrary convergent subsequence (and hence every convergent subsequence) of the functions $\{u^n\}$ converges uniformly to the constant function m on [0,1]. It follows then that $\{u^n\}$ converges uniformly to m. For if this were not so then for some $\varepsilon > 0$ there would be a subsequence $\{u_{n'}\}$, each member of which differed from m by more than ε at some point in [0,1]. This subsequence, being equicontinuous, would have a convergent subsubsequence which could not converge to m, a contradiction. By this reasoning it is established that

$$\lim_{n\to\infty} \{u^n\} = m,$$

where convergence is uniform, and m depends only on u^0.

Let us consider now the sequence of income distributions generated by the transition equation (8):

$$g^n(x) = \int_{[0,1]} K^n(x,y)g^0(y)dy$$

for given g^0. Taking u^0 again as an arbitrary continuous (utility) function on [0,1], its expectation over the distribution of income n-generation removed from g^0 is

$$E^n(u^0(x)) = \int_{[0,1]} g^n(x)u^0(x)dx$$

$$= \int_{[0,1]} \int_{[0,1]} K^n(x,y)\, g^0(y)u^0(x)dydx$$

$$= \int_{[0,1]} \left(\int_{[0,1]} K^n(x,y)u^0(x)\,dx \right) g^0(y)\,dy$$

$$= \int_{[0,1]} g^0(y)u^n(y)\,dy.$$

Now $\{u^n\} \to m$ uniformly. Hence

$$\lim_{n\to\infty} E^n(u^0(y)) = m \int_{[0,1]} g^0(y)\,dy = m.$$

This holds for all initial distributions g^0. Thus the expectation of an arbitrary continuous function converges to a unique number independent of the initial distribution. By a well known theorem of probability theory,[27] the income densities $\{g^n\}$ converge to a density g^* with the property that

$$\int_{[0,1]} g^*(x)u^0(x)\,dx = \lim_{n\to\infty} \int_{[0,1]} g^n(x)u^0(x)\,dx$$

for arbitrary continuous u^0. Furthermore, it is clear that for all u^0 continuous,

$$\int_{[0,1]} g^*(x)u^0(x)\,dx = \lim_{n\to\infty} \left\{ \int_{[0,1]} g^{n+1}(x)u^0(x) \right\}$$

$$= \lim_{n\to\infty} \left\{ \int_{[0,1]} \left(\int_{[0,1]} K^1(x,y)g^n(y)\,dy \right) u^0(x)\,dx \right\}$$

$$= \int_{[0,1]} \left(\int_{[0,1]} K^1(x,y)g^*(y)\,dy \right) u^0(x)\,dx$$

Since u^0 was arbitrary it follows that

$$g^*(x) = \int_{[0,1]} K^1(x,y)g^*(y)dy.$$

Thus, the size distribution always approaches the unique solution of (11).

Q.E.D.

Some Properties of the Equilibrium

Having established that our economy always tends to a unique equilibrium earnings distribution, we now investigate some of the characteristics of this equilibrium. The results of the previous sections imply that the existence of the intergenerational effect of parents' earnings on offsprings' opportunities can strongly influence the nature of observed earnings distributions. We shall examine the impact of certain government policies designed to alter the relationship between parents' and offsprings' earnings. In so doing, the average welfare of individuals in equilibrium may be improved. Moreover, because of the capital market imperfections which cause young people to rely on thier family resources to secure training, such policies frequently have the effect of both reducing inequality and increasing total output. The traditional conflict between equity and efficiency can be much less serious in a world where economic position is partially inherited from previous generations.

The aggregate product of the economy in any period may be either consumed or invested in the training of the young. Investment in the young increases their productive capacity in the subsequent period. Because the ability of an individual to make use of training may only be known ex post, parents must bear the risk of uncertain returns to the investment in their offspring. These are not social risks, however, The Law of Large Numbers guarantees that the output per man among all people receiving a given amount of training is simply the expected value of the random variable representing the output of any one of them ex ante. Thus, the relationship between investment and returns is deterministic in the aggregate.

A corollary observation is that the collective intertemporal production possibilities of this economy are identical to those of a one-sector growth model with training being thought of as circulating capital. Since the marginal productivity of education has been assumed to decline as the investment level increases (A2(ii)), and every young individual in the economy has the same productive potential ex ante, the efficient allocation of a fixed total savings among the young requires that each individual receive the same investment.[28] This would be the outcome of a competitive market in educational loans in which young people bid for training resources to maximize their expected earnings, and to which mature individuals supplied their savings derived from optimal bequest decisions. Thus, the relative productive efficiency of our equilibrium hinges upon two considerations -- the aggregate savings which takes place in equilibrium, and the distribution of that savings as investment among the young.

Concerning the level of aggregate savings, it is well known that even when there exists a perfect loans market, the competitive equilibrium of a one-sector growth model need not be efficient. Over savings of the Phelps-Koopmans variety generally cannot be ruled out. However, the nature of family decision making suggests that this need not be a problem here. Because of the recursive nature of individuals' concern for their offspring, the (possible) consumption of agents many generations into the future affects the welfare of current market participants. Competitive inefficiencies in infinite horizon models seem to require finiteness of the optimization horizons of the acting economic agents.[29] While the uncertainty which individuals face in making their savings decisions might cause inefficiencies to arise, these private risks can be eliminated by the government provision of the appropriate insurance contracts. These contracts would make earnings a function of training alone, guaranteeing to each individual an income equal to the average output of all those with the same level of training. In the presence of such arrangements, each mature individual views the earnings of his offspring as the same function of investment as that which characterizes the aggregate production possibilities of the economy. Thus, his consumption plan is the solution to the optimal accumulation problem which a social planner would face, if all families were to be treated the same and the intertemporal preferences of a mature individual were used in the optimization. Such a plan obviously cannot involve over savings in its asymptotic state(s).[30]

Furthermore, the risk aversion implied by the concavity of the utility function U(c,V) assures us that such an insurance arrangement would be ex ante Pareto superior to the laissez faire scheme. Each parent is made better off because the random distribution of offspring's earnings is replaced by the certain provision of its mean. Perfect risk sharing of this kind would of course require an ex post redistribution of income from those endowed with high ability to those who have low ability. An individual would thus be unable to appropriate the incremental genetic rent associated with having a better than average natural aptitude for production. However, since the indirect utility function $V^{*}(y)$ would remain concave under these circumstances, everyone would be willing to join in the risk sharing arrangement ex ante, before ability is known. Under full income insurance a family's social position would be completely determined by parental investment decisions in the previous generation. Moreover, there need not be complete equalization of all family incomes in the long run (see note 30). Many may therefore find government provision and enforcement of contracts of this kind ethically objectionable. It is nonetheless interesting to note that all families would prefer the implementation of this insurance scheme, even though it might confer permanent advantage on some families relative to others.

Whether or not there is insurance however, it is readily seen that the intergenerational externalities which characterize the investment process in this economy will lead to an inefficient allocation of aggregate investment among the young. Those who belong to high income families receive a larger investment of training than those from low income families. Yet the children of the rich are no better vessels of investment than the children of the poor. A redistribution of investment from high to low income families will therefore lead to a larger social dividend in the subsequent period. Such a move may also be reasonably expected to reduce the inequality of the distribution of earnings among the next generation of workers. Below we shall consider two ways in which this inequality in the distribution of investments can be mitigated. One obvious method is to use taxation to redistribute the earnings of parents in all generations, though this has the disadvantage of affecting investment incentives. An alternative approach is to institutionalize the investment process so that training is done only by the state, with its costs supported by (say) a poll tax. This "public education" scheme avoids the tax distortion by taking all discretionary decisions out of the hands of the individual agents.

If lump sum taxation were possible, any redistribution which takes income from those families where the marginal effect of another dollar of family earnings on the expected earnings of the offspring is low, and gives income to families with a relatively high expected marginal effect, will increase the total output of the next generation of workers. If $X(\alpha,y)$ is a concave function of y, then such redistributions are necessarily from high income families to low income families, thus reducing inequality in the current period as well as increasing efficiency in the subsequent period. In this instance, equity and efficiency considerations do not conflict at all.

This observation is of limited interest however, since lump sum taxation allows any degree of equality to be attained without efficiency costs. The analysis becomes much more complex once we allow taxation to affect incentives. We shall represent a redistributive tax scheme by the net income schedule T(y), which determines the after-tax earnings of an individual whose gross earnings are y. Once a tax schedule is announced, family decision makers will take it into account in assessing the net earnings and associated well-being of their offspring. Their optimal consumption-investment plans will be altered accordingly. Moreover, their well being will be directly affected by the changed distribution of net earnings of their offspring for a fixed level of investment. This structural change will have the effect of causing the economy to tend to a different equilibrium earnings distribution. One can then assess the effect of taxation by comparing this new equilibrium to the old one.[31]

By restricting the net income functions T(y) to the class of continuous concave functions on [0,1], the results of Theorem 1 could be readily established for family behavior under taxation. Limiting consideration to concave net earnings functions enables us to retain the concavity (risk aversion) of the indirect utility function $V^*(y)$. We make no presumption however that an "optimal" net earnings function would be concave. Further restrictions to assure that Assumption 4 still holds after the introduction of taxation would enable one to secure the other results given above, and justify comparative static exercises across equilibria. By considering only small deviations from the no-tax situation, this requirement may be secured. We shall take as admissible only those tax functions satisfying these necessary requirements. Moreover, it is intuitively obvious that one can devise admissible tax schemes which yield positive or negative government revenue in equilibrium. It is plausible, and not difficult to prove, that equilibrium government revenue varies "continuously" with the tax schedule chosen, and that admissible tax schemes form a convex set. Thus, it is always possible to select an admissible tax arrangement which just breaks even in the resulting equilibrium. In view of our requirement that $T(\cdot)$ be concave, i.e., marginal tax rates are increasing, such a tax scheme will be an inequality reducing redistribution of income in equilibrium. By this we mean that before and after tax total incomes are identical, but the Lorentz curve of the after tax income distribution lies everywhere inside of that corresponding to the before tax distribution of earnings.[32]

Thus, we have established that it is possible under certain conditions to impose taxation in this economy in such a way that in the resulting equilibrium, income is being redistributed in an inequality reducing manner. What are the welfare implications of such a move? From the preceding discussion we know when $X(\alpha,y)$ is concave, that such redistribution achieves a more efficient allocation of total investment among the young. Moreover, since parents are risk averse, they may benefit from redistributive taxation which reduces the dispersion of their offspring's random earnings. However, we must distinguish between short run and long run effects. An inequality reducing redistribution of income has some of the risk spreading element of the insurance arrangements discussed above. Once the new equilibrium under admissible balanced budget taxation has been attained, the transfer of income from rich to poor while leaving total earnings unchanged provides an improvement in the average utility of a family head in the steady state. This represents a long run gain for every <u>family</u>, since asymptotically the distribution of earnings of the progeny of any family head is identical to the equilibrium earnings distribution, regardless of the head's initial income. (Recall that this is not necessarily true if the government provides perfect income insurance, making family income independent of head's ability.)

On the other hand, if one considers the welfare effects across a single generation, it is not possible to conclude unambiguously that all <u>parents</u> are made better off from the redistributive effects of taxation. While the variability of offspring's earnings would be reduced for all families, admissible balanced budget taxation would increase the expected earnings of those receiving small amounts of education while reducing the expected income of those whose parents have made large investments in them. The poor benefit on both counts, but the rich must perceive the insurance gains to be worth the cost of lower expected earnings before the redistributive impact of a tax scheme can be said to be Pareto dominating in the short run. However, we have not been able to find a useful characterization of when this can occur.

Yet it must also be the case that taxation affects the optimal investment schedule, $e^*(y)$. Because redistribution simultaneously changes the dispersion of offspring's earnings and the marginal return to investment in a manner which varies for different parental incomes, its effect on optimal investment will be very complex indeed.

An interesting though difficult question is: "Is it possible to impose redistributive taxation which reduces inequality in equilibrium and gives higher total output than the non-tax equilibrium?" Should an affirmative answer be given to this question, it could then be inferred that the intergenerational external effect of parents' earnings on off-springs' opportunities obviates the equity-efficiency trade-off on the margin. A prima facie case for redistribution would thereby be made.

The following result falls considerably short of giving a complete answer to the query raised above. However, it does lend credibility to the belief that an affirmative answer can be given, by exhibiting a plausible set of conditions sufficient for redistribution to increase total product. Let $g^*(\cdot)$ be the density of the equilibrium earnings distribution without intervention, and let $\hat{g}(\cdot)$ represent the corresponding density of the (gross) earnings distribution after the imposition of an admissible tax scheme. Denote by m^* and $\hat{m}$ the respective means of the equilibrium distributions and by V^* and $\hat{V}$ their respective variances. Define the function H(e) as follows:

$$H(e) = \int_{[0,1]} h(\alpha,e)f(\alpha)d\alpha \ .$$

H(e) gives the expected earnings of a mature agent with e units of training. Let $e^*(\cdot)$ and $\hat{e}(\cdot)$ represent the optimal investment functions arising from the solution of (3) without and with taxation, respectively. Finally, define $H^*(y) \equiv H(e^*(y))$, $\hat{H}(y) \equiv H(\hat{e}(y))$. We may now state

Proposition 1: Suppose that T(y) is the net income schedule of a tax plan which is an inequality reducing redistribution of income in equilibrium. Suppose further that

(a) $H^{*''} < 0$, $\hat{H}'' < 0$, $H^{*'} \leqq 1$ and $V^*/\hat{V} > H^{*''}(\hat{m})/H^{*''}(m^*)$;

(b) $\exists\ \tilde{y} \in [0,1]$ such that $e^*(y) \gtrless \hat{e}(y)$ as $y \gtrless \tilde{y}$;

(c) With $\tilde{p} \equiv H'(e^*(\tilde{y}))$,

$$\tilde{p} \int_{[0,1]} [e^*(y)-\hat{e}(y)]g(y)dy < \int_{[0,1]} [\hat{H}(T(y))-\hat{H}(y)]\hat{g}(y)dy .$$

Then total output in the tax equilibrium exceeds that in the no-tax equilibrium (i.e., $\hat{m} > m^*$) .

Proof: Let $\mathbb{P}[0,1]$ be the set of probability densities over [0,1]. Define the operators

$$K: \mathbb{P}[0,1] \to \mathbb{P}[0,1] \text{ such that } (Kg)(x) = \int_{[0,1]} K^1(x,y)g(y)dy$$

and

$$\hat{K}: \mathbb{P}[0,1] \to \mathbb{P}[0,1] \text{ such that } (\hat{K}g)(x) = \int_{[0,1]} \hat{K}^1(x,T(y))g(y)dy$$

where $K^1(\cdot,\cdot)$ is the stochastic density kernel corresponding to investment schedule $e^*(\cdot)$, and $\hat{K}^1(\cdot,\cdot)$ is the kernel resulting from family savings plans $\hat{e}(\cdot)$. Let $E: \mathbb{P}[0,1] \to R_+$ be the expectation operator. By hypothesis $\hat{g}$ and g^* are the unique, stable solutions of (11) whose existence is assured by Theorem 4:

$$g^* = Kg^* \text{ and } \hat{g} = \hat{K}\hat{g} .$$

Now the linearity of E, K, and $\hat{K}$ implies

$$(14) \qquad m^* - \hat{m} = Eg^* - E\hat{g} = E(Kg^*) - E(\hat{K}\hat{g})$$

$$= E(K(g^* - \hat{g})) + E((K - \hat{K})\hat{g}) \ .$$

First we show that hypothesis (a) implies the following "monotonicity" of convergence to g^* under K:

Claim: If $E(K\hat{g}) < E\hat{g}$, then $E\hat{g} > Eg^*$.

To see this, observe that

$$E(K\hat{g}) = \int_{[0,1]} H^*(y)\hat{g}(y)dy \simeq \int_{[0,1]} \{H^*(\hat{m}) + H^{*\prime}(\hat{m})(y-\hat{m}) + \tfrac{1}{2}H^{*\prime\prime}(\hat{m})(y-\hat{m})^2\}\hat{g}(y)dy$$

$$= H^*(\hat{m}) + \tfrac{1}{2}H^{*\prime\prime}(\hat{m})\hat{V} \ ,$$

while

$$m^* = Eg^* = E(Kg^*) \approx H^*(m^*) + \tfrac{1}{2}H^{*\prime\prime}(m^*)V^* \ .$$

Suppose now, contrary to the claim, that $m^* \geqq \hat{m}$ and $E(K\hat{g}) < E\hat{g}$.

Then

$$\hat{m} = E\hat{g} > E(K\hat{g}) \simeq H^*(\hat{m}) + \tfrac{1}{2}H^{*\prime\prime}(\hat{m})\hat{V} \ .$$

It follows that

$$H^*(\hat{m}) - \hat{m} < - \tfrac{1}{2}H^{*\prime\prime}(\hat{m})\hat{V} \ , \text{ and}$$

$$H^*(m^*) - m^* \simeq - \tfrac{1}{2}H^{*\prime\prime}(m^*)V^* \ .$$

Hence, hypothesis (a) implies

$$H^*(m^*) - m^* \leqq H^*(\hat{m}) - \hat{m} < -\tfrac{1}{2}H^{*''}(\hat{m})\hat{V} < -\tfrac{1}{2}H^{*''}(m^*)V^* \approx H^*(m^*) - m^* ,$$

a contradiction. Therefore $\hat{m} > m^*$.

Thus we have shown that $E(K\hat{g}) < E\hat{g}$ is a sufficient condition for $E\hat{g} > Eg^*$. It follows (e.g., from (14)) that $\hat{m} > m^*$ if $E((K - \hat{K})\hat{g} \leqq 0$. Now

$$E((K-\hat{K})\hat{g}) = \int_{[0,1]} [H(e^*(y))-H(\hat{e}(T(y)))]\hat{g}(y)dy$$

$$= \int_{[0,1]} [H(e^*(y))-H(\hat{e}(y))]\hat{g}(y)dy + \int_{[0,1]} [H(\hat{e}(y))-H(\hat{e}(T(y)))]\hat{g}(y)dy$$

$$\leqq \int_{[0,1]} H'(\hat{e}(y))[e^*(y)-\hat{e}(y)]\hat{g}(y)dy + \int_{[0,1]} [\hat{H}(y)-\hat{H}(T(y))]\hat{g}(y)dy$$

$$\leqq H'(\hat{e}(\tilde{y}))\int_{[0,1]} [e^*(y)-\hat{e}(y)]\hat{g}(y)dy + \int_{[0,1]} [\hat{H}(y)-\hat{H}(T(y))]\hat{g}(y)dy$$

$$= \tilde{p}\int_{[0,1]} [e^*(y)-\hat{e}(y)]\hat{g}(y)dy - \int_{[0,1]} \hat{g}(y)[\hat{H}(T(y))-\hat{H}(y)]dy < 0$$

by hypotheses (b) and (c).

Q.E.D.

Thus we may conclude that if a tax policy can be found which satisfies hypotheses (a), (b), and (c) of the proposition, and is also an inequality reducing redistribution of income in equilibrium, then there will be no efficiency sacrifice involved in a marginal redistribution of income. The concavity restrictions of (a) will be satisfied if either the investment functions are concave or the productivity function exhibits sufficiently rapid diminishing returns to education. Essentially what is required is that the marginal effect of family income on the expected earnings of the offspring be a diminishing function of parents' income. Hypothesis (a) also requires that a dollar increase in parental earnings imply less than a dollar expected increase in offspring's income, and that the imposition of taxation lead to a reduction in the variance of the earnings distribution large enough to outweigh any increase in the curvature of H^* at the average income. Since redistributive taxation which reduces inequality leads to a more equal distribution of training among the young, it is reasonable to expect this last requirement to be attainable.

Hypothesis (b) requires that in the face of the taxation of their offsprings' earnings, lower income families invest more and higher income families invest less in their children than in the absence of taxation. It seems that this effect can be secured by having the marginal tax rate near zero at low incomes, but increasing rapidly thereafter. The marginal payoff to investment by the poor would then be only slightly affected by the marginal tax rate, though the redistributive nature of taxation would reduce the riskiness of the earnings of the offspring of low income families. Simultaneously, investment incentives for the wealthy would be reduced. We have not been able to secure general conditions under which the effect of taxation takes this form.

Condition (c) is obviously the critical hypothesis of the proposition. It states directly that taxation to redistribute can improve efficiency if it does not cause aggregate investment to fall by too much. This upper bound on the decline in total investment is given by the requirement that the gains from taxation due to a more efficient allocation of training among the young in the tax equilibrium exceed the value of the tax induced decline in aggregate investment over the tax equilibrium earnings distribution. The value is determined with the shadow price of investment of the offspring of the mature individual whose investment behavior has not changed with the imposition of the tax. Again, this condition suggests that rapidly diminishing returns to education casuing there to be large gains possible through equalizing training among the young, increases the likelihood that redistribution through taxation will actually increase national product.

An alternative policy which can achieve efficiency gains without distortionary taxation is the establishment of universal public education. This policy may be viewed as an even greater intervention into private decision making, because it takes discretionary investment decisions out of the hands of the parents. In such a world equal educational opportunity becomes the government's objective, though individuals are allowed to keep for themselves whatever returns to education which they attain (net of their contribution to the support of the educational process). This mode of procedure enables the economy to attain its full productive potential, though it does involve a significant forfeiture of individual freedoms. Its efficiency effects are so powerful because government can simultaneously determine the optimal aggregate savings, and the optimal distribution of investment among the young. However, if revenue to support education must be raised by distortionary taxes, then the efficiency gains would not be so great.

The question addressed below concerns when the establishment of equal educational opportunity will also reduce the equilibrium dispersion of earnings. It has recently been argued from various quarters[33] that equalizing education will have only a limited effect on income inequality. The following proposition indicates that even in this simple world, a rather strong assumption is needed to assure that universal public education will reduce the dispersion of the distribution of earnings in equilibrium.

Proposition 2: Suppose that $X(\alpha,y)$ is concave in y and that $X_\alpha(\alpha,y)$ is convex in y. Then universal public education with a per capita budget equal to the investment of a family earning the average income in the laissez-faire equilibrium will produce an earnings distribution with lower variance than the no-intervention equilibrium. Aggregate earnings will be increased with the establishment of public education.

<u>Proof</u>: Denote by x the earnings of an individual in the laissez-faire equilibrium, and let y be the earnings of his parent in the preceding period. Let $\bar{x}$ denote the average income in the no-intervention equilibrium. Using previous notation

$$\text{(15)}\qquad \mathrm{Var}(x) = \int_{[0,1]} (x-\bar{x})^2 g^*(x)dx = \int_{[0,1]}\int_{[0,1]} (x-\bar{x})^2 K(x,y)dxg^*(y)dy$$

$$= \int_{[0,1]} g^*(y)\{\int_{[0,1]} K(x,y)(x-E(x|y))^2dx + \int_{[0,1]} K(x,y)[(x-\bar{x})^2 - (x-E(x|y))^2]dx\}dy$$

$$= E(\mathrm{Var}(x|y)) + \mathrm{Var}(E(x|y)) \ ,$$

a standard result of elementary distribution theory. Now it is clear that the variance of earnings under the universal public education scheme is given by

$$\mathrm{Var}_{P.E.}(x) = \mathrm{Var}(x|y = \bar{x}) \ .$$

Moreover

$$\frac{\partial}{\partial y}\mathrm{Var}(x|y) = \frac{\partial}{\partial y}\int_{[0,1]} [X(\alpha,y) - E(x|y)]^2 f(\alpha)d\alpha$$

$$= 2\int_{[0,1]} \{X(\alpha,y)[X_y(\alpha,y) - \frac{\partial}{\partial y}E(x|y)] - E(x|y)[X(\alpha,y)-E(x|y]\}f(\alpha)d\alpha$$

$$= 2\int_{[0,1]} X(\alpha,y)X_y(\alpha,y)f(\alpha)d\alpha - 2E(x|y)\frac{\partial}{\partial y}E(x|y)$$

$$= 2[E_\alpha(X \cdot X_y) - E_\alpha(X) \cdot E_\alpha(X_y)]$$

$$= 2\ \mathrm{Cov}_\alpha(X(\alpha,y),\ X_y(\alpha,y)) \gtreqless 0 \quad \text{as } h_{\alpha e} \gtreqless 0\ .$$

Thus, the conditional variance of offsprings' earnings is increasing (decreasing) with increasing parents' income if and only if ability and education are compliments (substitutes). Furthermore, it follows from (15) that $\mathrm{Var}(x) > \mathrm{Var}(x|y = \bar{x})$ if $h_{\alpha e} = 0$. Thus, uniform education at any level reduces earnings dispersion if ability and education do not interact. Now

$$\frac{\partial^2}{\partial y^2}\mathrm{Var}(x|y) = 2[E_\alpha(X_y^2 + X \cdot X_{yy}) - E_\alpha(X_y)^2 - E_\alpha(X) \cdot E_\alpha(X_{yy})]$$

$$= 2[\mathrm{Var}_\alpha(X_y) + \mathrm{Cov}_\alpha(X, X_{yy})] > 0\ ,$$

when X_α is convex in y. It follows that $\mathrm{Var}(x|y)$ is convex in y, and since $E(y) = \bar{x}$,

$$\mathrm{Var}(x|y = \bar{x}) \leqq E(\mathrm{Var}(x|y)) < E(\mathrm{Var}(x|y)) + \mathrm{Var}(E(x|y))$$

$$= \mathrm{Var}(x)\ .$$

Observe now that the concavity of $X(\alpha,y)$ in y implies

$$\bar{x} = \int_{[0,1]} x g^*(x)dx = \iint_{[0,1]^2} X(\alpha,y)f(\alpha)g^*(y)d\alpha dy$$

$$\leqq \int_{[0,1]} f(\alpha)X(\alpha,\bar{x})d\alpha = \text{output per man under public education} .$$

Hence, the proposition is proved.

Q.E.D.

Solution of a Special Case:

It is interesting to examine the explicit solution of (11) for the equilibrium earnings distribution for a set of special assumptions. We shall suppose that the productivity function is given by

$$h(\alpha,e) = \alpha^{\eta} e^{1-\eta} , \quad 0 < \eta < 1 .$$

Further we assume that each family saves and invests in the young a constant fraction s of the family income in each period. There does not appear to be any utility function which gives a constant savings propensity as a solution for (3) with this technology. Finally, we take it that α is distributed uniformly on the unit interval. Thus, we have

$$X(\alpha,y) = \alpha^{\eta}(sy)^{1-\eta} .$$

The largest sustainable income is $\bar{y} \equiv s^{(1-\eta)/\eta}$. Moreover,

$$y_0^1(x) \equiv \bar{y} , \quad \forall x \in [0,\bar{y}]$$

$$y_1^1(x) = \frac{1}{s} x^{1/1-\eta} , \quad \forall x \in [0,\bar{y}]$$

and

$$\mathcal{A}(x,y) = x^{1/\eta}(sy)^{(\eta-1)/\eta} .$$

Equation (11) then becomes

$$(16) \qquad g^*(x) = \frac{1}{\eta} x^{(1-\eta)/\eta} \int_{s^{-1/\eta} x^{1/1-\eta}}^{1} g^*(s^{(1-\eta)/\eta} y)(s^{(1-\eta)/\eta} y)^{(\eta-1)/\eta} dy .$$

Let us consider the case $\eta = \frac{1}{2}$. This seems to be the only parameter value for which one can actually determine the solution to (16). In this instance (16) simplifies to

$$(17) \qquad g^*(x) = 2x/s \int_{x^2/s}^{s} g^*(y)/y \, dy , \quad x \in [0,s] .$$

Consider now the change of variables $z = x/s$. Then $z \in [0,1]$, and the density of z is given by $\hat{g}(z) = sg^*(sz)$. One may then use (17), with the change of variables $\tilde{y} = y/s$ in the integrand, to see that $\hat{g}$ must satisfy

$$(18) \qquad \hat{g}(z) = 2z\int_{z^2}^{1} \hat{g}(\tilde{y})/\tilde{y} \, d\tilde{y} .$$

A solution for (18) gives immediately the solutions of (17) for all possible savings fractions, s. Differentiation of (18) yields

$$(19) \qquad \frac{d}{dz} \hat{g}(z) = \frac{1}{z} \hat{g}(z) - 4\hat{g}(z^2) , \quad z \in [0,1] ,$$

with the initial condition set so that $\hat{g}$ integrates to one. Equation (19) is a functional differential equation which may be solved by the techniques employed below.

Suppose the solution for (19) (known to exist by virtue of Theorem 4) can be written in the form

$$\hat{g}(z) = \sum_{n=0}^{\infty} a_n z^n .$$

Then (19) becomes

$$(20) \qquad \sum_{n=0}^{\infty} (n+1) a_n z^n = \sum_{n=0}^{\infty} a_n z^{n-1} - 4 \sum_{n=0}^{\infty} a_n z^{2n} .$$

By equation the coefficients of like powers of z on each side of (20) we find that the coefficients a_n must satisfy the following recursive relationship:

(21) $\quad a_n = 0 , \quad n \neq n_k , \quad k = 1,2,\ldots$

$$a_{n_k} = \frac{-2}{n_{k-1}} a_{n_{k-1}} , \quad k=2,3,\ldots , \text{ where}$$

$$n_k \equiv 2^k - 1 , \quad k = 1,2,\ldots , \text{ and}$$

$$a_{n_1} = a_1 \text{ is arbitrary.}$$

Here the constant a_1 must be set so that the initial condition of (19) is satisfied. Solving (21) for a_{n_k} we get

$$(22) \qquad a_{n_k} = a_1(-1)^{k-1}2^{k-1} \Big/ \prod_{j=1}^{k-1} (2^j - 1) , \qquad k=2,3,\ldots$$

Whence it follows that

$$(23) \qquad \hat{g}(z) = a_1\{z + \sum_{k=2}^{\infty} (-2)^{k-1} z^{2^k - 1} \Big/ \prod_{j=1}^{k-1} (2^j - 1)\}$$

and

$$a_1 = 2\{1 + \sum_{k=2}^{\infty} (-1)^{k-1} (\prod_{j=1}^{k-1} (2^j - 1))^{-1}\}^{-1} \approx 6.92 .$$

The implied solutions $g^*(x)$ are depicted in Figure 2 for two representative values of s. The basic character of the equilibrium distribution is unaffected by the savings parameter, which simply alters its scale. The distributions are just slightly right skewed with means equal to (0.422)s and variances of $(0.037)s^2$. Consumption per family is maximized in the equilibrium when the savings fraction s = ½ , the elasticity of education in the productivity function. This result seems to hold quite generally when family investment behavior is characterized by a constant savings propensity.

One may also examine the impact of universal public education in this example. The results indicate that Proposition 2 can probably be strengthened. In spite of the fact that the hypothesis of the proposition requiring X_α to be convex in y is not met here, public education with a per capita budget equal to

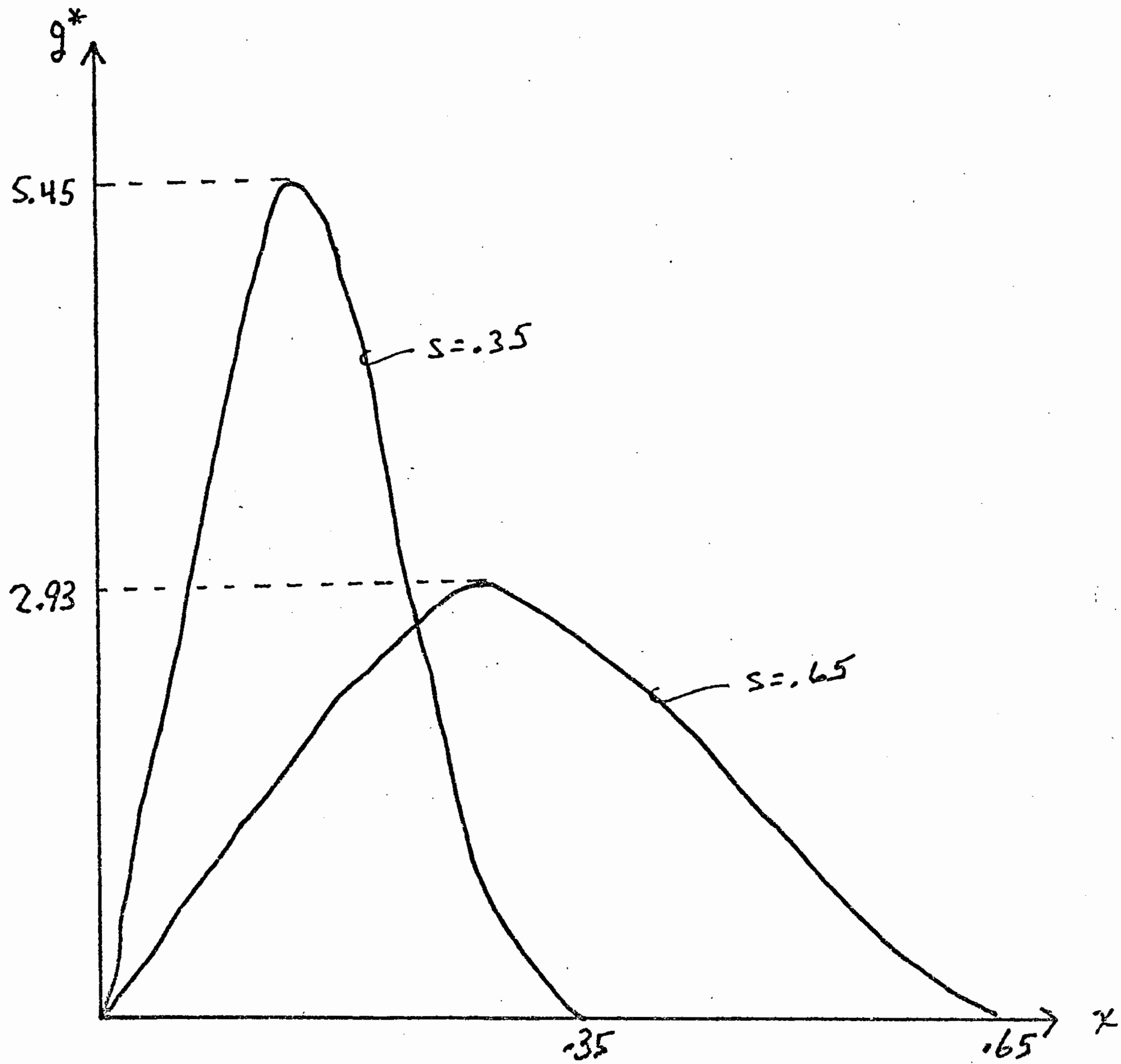


FIGURE 2

the expenditure of the average family reduces the variance of the equilibrium earnings distribution by 35%, from $(0.037)s^2$ to $(0.024)s^2$. At the same time, equal educational opportunity implies an efficiency gain of approximately 3.4% in per capita output. The potential gain from such a policy is somewhat greater than this however, because the budget level may be adjusted toward the optimal investment fraction of ½ .

FOOTNOTES

1. This is a central question of modern capital theory. See, for example, Harcourt [17], or Bliss [5].

2. Study of the size distribution goes back at least to Pareto [25]. An excellent summary of the history of the subject is contained in Blinder [4], Ch. 1.

3. Stochastic elements in the determination of individual earnings have played a dominant role throughout the history of this subject. Examples include the early work of Gibrat [15] and Kalecki [21], Champernowne [10], and Mandelbrot [22]. The authors' aim was to show that certain empirically relevant distributions (usually log-normal or Pareto distributions) could result from some process of random shocks to individual earnings. However, the economic rationale of these random movements remained obscure.

4. Since Pigou's observation [26] that the presumably symmetric distribution of natural abilities should not imply the observed right-skewed earnings distributions, models have been constructed to relate the distribution of earnings to the distribution of innate characteristics of individuals. By introducing several types of abilities which combine to generate earnings in some nonlinear

fashion, Boissevain [6] and later Roy [27] have shown that skewed (log-normal) earnings distributions can result from this process. In a subsequent paper, Roy [28] allows workers to choose among alternative occupations, and still gets the same result.

5. Though the specification of these models is typically ad hoc, some authors have been successful in approximating empirically observed income distributions with some degree of accuracy. See, for example, Rutherford [29].

6. Cf. Mincer [24], and especially Becker's Woytinski Lecture [3], Addendum to Ch. 3.

7. Becker (ibid) has observed that inherited wealth can lower the cost of acquiring human capital, and thus affect the distribution of earnings. Pigou [26], part IV, Ch. II, attributes Pareto's "long tail" of the income distribution in part to the effects of accumulated advantage.

8. The clearest expression of this fact is in the work of sociologists on social mobility. See especially Duncan, et al. [12].

9. This problem is explored by the author in the context of racial income inequality in the first essay of this thesis.

10. An equilibrium of this sort is best viewed as a summary

statistic of the underlying structure which characterizes the intergenerational effects. Like the notion of the steady state, its descriptive power is weakened by the large amount of time required for equilibrium to be achieved. For a discussion of the use of this equilibrium concept in this way, see Boudon [7].

11. What we have is Samuelson's consumption loans model [30] with a twist: an implicit social compact of family commitment acts as the required "social contrivance."

12. This is done by using lifecycle savings motives to rationalize investment in children.

13. I have not seen this problem posed elsewhere, though it is the natural implication of "rational expectations" when applied to parental anticipation of their offspring's behavior.

14. This is the approach most frequently used in life-cycle savings models. One simply posits a direct utility of bequests. See Blinder [4] for an application in a model of income distribution, or Merton [23], where the same assumption is employed while deriving optimal lifetime consumption and portfolio rules.

15. Barro [2] uses this specification in studying the impact of public debt.

16. Contrast the inefficiency results of Samuelson [30] and Diamond

[11] with the competitive efficiency theorem of Hall [16]. Both Samuelson and Diamond find that competitive equilibrium may be inefficient in models where agents have finite planning horizons. Hall, assuming agents' concerns extend into the indefinite future, exhibits a turnpike theorem for competitive paths. Similarly, Diamond [11] finds an impact of public debt when agents have finite horizons. While Barro [2], using an implicit infinite horizon structure, gets a neutrality result.

17. We wish to stress the fact that the utility maximization employed here requires a cardinal representation of individual preferences. An axiomatic justification of this choice criterion is beyond the scope of this endeavor.

18. Normality implies that $e^*(y)$ is strictly increasing and therefore differentiable almost everywhere.

19. With a slight strengthening of the assumption (A4) made below, it may be shown that $(\bar{y},\infty)$ is a transient state for the income distribution. Thus, an economy which has been operating for a "long time" would have only a negligible fraction of its families with incomes in that range.

20. Assumption 4 plays the same role in this theory of income distribution as does the requirement of what Brock and Mirman [8] term a "stable fixed point configuration" in the theory of stochastic optimal growth.

21. The mathematical structure of our problem is very similar to

that of the stochastic growth models mentioned in note 20. The method employed here is closest to that used by Iwai [18].

22. This point seems to have been missed by Iwai [18]. He proceeds as if his density kernel were continuous, without imposing any restrictions on $f(\cdot)$. Our previous assumption (A3) that $f(0) > 0$ forces us to face the problem here.

23. Many writers have exploited the ergodic properties of Markoff processes in explaining the stability of observed earnings distributions. See, for example, Champernowne [9], Solow [31], or Mandelbrot [22].

24. Champernowne [9], p. 82.

25. Feller [13], Thm. 1 , p. 272.

26. Friedman [14], p. 112.

27. Feller [13], Thm. 2 , p. 251.

28. This is only true if there is no other information available about individuals' ability. If there is an observable variable z such that $\mathrm{Cov}(\alpha,z) \neq 0$, then an efficient investment allocation is a function $\hat{e}(z)$ satisfying

$$\int_{[0,1]} h_e(\alpha,\hat{e}(z))f(\alpha|z)d\alpha = \text{constant}.$$

Given the stationary population assumption, consumption per head will be greatest in equilibrium when the constant is unity.

29. Cf. note 16 above.

30. With perfect insurance each family would be solving a deterministic optimal accumulation problem with intertemporally non-separable preferences. Iwai [19] has established turnpike properites for the optimal trajectories in such a problem, though in general the asymptotic state is not unique. This suggests that the income distribution would tend to concentrate at a small number of discrete income levels.

31. Throughout this discussion we shall be comparing long run equilibria while neglecting the problem of what happens in the transition from one equilibrium to another. In this respect, we follow a long-standing practice in the theory of economic growth. See note 10 above.

32. This result is due to Atkinson [1].

33. Cf. Jencks, et al. [20].

REFERENCES

1. Atkinson, A., "On the Measurement of Inequality", Journal of Economic Theory 2 (1970), 244-263.

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

3. Becker, G., Human Capital, 2nd Edition, NBER, 1975.

4. Blinder, A., Toward an Economic Theory of Income Distribution, MIT Press, 1974.

5. Bliss, C., Capital Theory and the Distribution of Income, North Holland, 1975.

6. Boissevain, C. H., "Distribution of Abilities Depending upon Two or More Independent Factors", Metron 13 (1939), 49-58.

7. Boudon, R., Mathematical Structures of Social Mobility, Elsevier Scientific Publishing Co., Amsterdam, 1973.

8. Brock, W. and L. Mirman, "Optimal Economic Growth and Uncertainty: The Discounted Case", Journal of Economic Theory 4 (1972), 479-513.

9. Champernowne, D. G., The Distribution of Income between Persons, Cambridge University Press, 1973.

10. Champernowne, D. G., "A Model of Income Distribution", Economic Journal 63 (1953), 318-351.

11. Diamond, P., "National Debt in a Neoclassical Growth Model", American Economic Review 55 (1965), 1126-1150.

12. Duncan, Featherman, and Duncan, Socioeconomic Background and Achievement, Seminar Press, 1972.

13. Feller, W., An Introduction to Probability Theory and Its Applications, Vol. II, John Wiley and Sons, Inc., 1971.

14. Friedman, A., Foundations of Modern Analysis, Holt, Rinehart and Winston, Inc., 1970.

15. Gibrat, R., "On Economic Inequalities", International Economic Papers 7 (1957), 53-70.

16. Hall, R., "The Allocation of Wealth Among the Generations of a Family Which Lasts Forever -- A Theory of Inheritance", Ch. 1 of unpublished Ph.D. thesis, M.I.T., 1967.

17. Harcourt, G. C., Some Cambridge Controversies in the Theory of Capital, Cambridge University Press, 1972.

18. Iwai, K., "A Theory of Optimal Capital Accumulation Under Uncertainty", Essay II of unpublished Ph.D. thesis, M.I.T., 1972.

19. Iwai, K., "Optimal Economic Growth and Stationary Ordinal Utility -- A Fisherian Approach", Journal of Economic Theory 5 (1972), 121-151.

20. Jencks, C. et al., Inequality: A Reassessment of the Effect of Family and Schooling in America, Harper and Row, 1972.

21. Kalecki, M., "On the Gibrat Distribution", Econometrica 13 (1945), 161-170.

22. Mandelbrot, B., "Stable Paretian Random Functions and the Multiplicative Variation of Income", Econometrica 29 (1961), 517-543.

23. Merton, R. C., "Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case", Review of Economics and Statistics 51 (1969), 247-257.

24. Mincer, J., "Investment in Human Capital and Personal Income Distribution", Journal of Political Economy 66 (1958), 281-302.