Econometrica, Vol. 49, No. 4 (July, 1981)

INTERGENERATIONAL TRANSFERS AND THE DISTRIBUTION OF EARNINGS

BY GLENN C. LOURY[1]

This paper models the dynamics of the earnings distribution among successive generations of workers as a stochastic process. The process arises from the random assignment of abilities to individuals by nature, together with the utility maximizing bequest decisions of their parents. A salient feature of the model is that parents cannot borrow to make human capital investments in their offspring. Consequently the allocation of training resources among the young people of any generation depends upon the distribution of earnings among their parents. This implies in turn that the often noted conflict between egalitarian redistributive policies and economic efficiency is mitigated. A number of formal results are proven which illustrate this fact.

1. INTRODUCTION

EXPLAINING THE DISTRIBUTION OF INCOME is among the central tasks of economics. Indeed, the classical query "What determines the division of product among the factors land, labor, and capital?" has stimulated some of the most profound contributions to economic thought. In modern neoclassical writing the size distribution of income, particularly the distribution of wage income, has become a topic of interest. This topic has many facets, not all of which can be studied theoretically within the same model. The present paper focuses upon one empirically important aspect of the process by which a distribution of earnings capacities arises in a population—the consequences of social status or family background for individual earnings prospects. A commonly observed fact is that earnings correlate positively across generations.[2] One explanation, for which there is some evidence, is that the acquisition in youth of productivity enhancing characteristics is positively affected by parental income.[3]

Here we pursue this notion in some depth. A choice theoretic framework is used to model the effect of parental income on offspring's productivity. An economy is constructed with an elementary social structure. Individuals, who live for two periods, are arranged into *families*. Each family has one young person and one old person. Think of the older family member as the parent and the younger as the offspring. Families exist contemporaneously but are isolated from each other. They each produce a perishable good in a quantity which depends on the *ability* and the *training* of the parent. Each individual of this stationary population is assumed to learn during the second period of life the value of his randomly assigned ability endowment. Each family divides its income of the

[1]This paper derives from the author's Ph.D. Thesis. The advice of Professor Robert Solow is gratefully acknowledged and comments of Carl Futia, Sanford Grossman, and Joseph Stiglitz have also been helpful. All errors are, of course, my own.

[2]See, for example, Jencks, et al. [**15**].

[3]See Duncan et al. [**10**], Datcher [**8**], and Jencks, et al. [**15**]. A discussion of the consequences of this observation for the analysis of racial income differences can be found in Loury [**18**, Ch. 1].

perishable commodity between consumption and training for the offspring. Offsprings' abilities, independent and identically distributed random variables, are unknown when this division is made.

Parents control family resources. They are motivated to forego consumption and invest in their offspring by an altruistic concern for the well-being of their progeny. Because parents are of varying abilities and backgrounds their incomes differ. Consequently the training investments which they make in their offspring also vary. If the "production function" linking ability and training to output has diminishing returns to training, then two parents making different income constrained investment decisions face divergent expected marginal returns to training in terms of their offsprings' earnings. If the parent investing less, facing the higher expected marginal return, could induce the parent investing more to transfer to his son a small amount of the other's training resource, then the offspring of both families could on average (through another inter-family income transfer in the opposite direction) have greater incomes in the next period. If, however, such trades are impossible because the relevant markets have failed, then the distribution of income among parents will affect the efficiency with which overall training resources are allocated across offspring. (This will be true even when abilities are correlated within families across generations, since parents of the same ability will also have different incomes.) The model presented here captures some features, positive and normative, of the situation which arises when such training loans between families are not possible.

Because family income in a particular generation depends on the level of training which the parent received in the previous period and that training investment depended on the family's income in the preceding generation, the distribution of income among a given generation depends on the distribution which obtained in the previous generation. A stochastic process for the income distribution in this economy is thus implied by the parental decision-making calculus and the random assignment of ability. It is natural to think of an equilibrium or stationary distribution as one which, if it obtains, persists.[4] We demonstrate for our model the existence and global stability of such an equilibrium distribution. Moreover, because parents are altruistic regarding their offspring but uncertain about the earnings (and hence welfare) of their children, institutional arrangements which redistribute income among subsequent generations affect the wellbeing of the current generation. Such redistributive arrangements can have insurance-like effects.[5] We demonstrate below that under certain conditions egalitarian redistributive measures can be designed which make all current members of society better off.

The remainder of the paper is organized as follows: In the next two sections the analytical model is described and the main mathematical results are stated. Section 4 subjects the equilibrium distribution to more detailed examination, and

[4] Becker and Tomes investigate a similar notion of equilibrium in [3].

[5] Varian [29] has discussed insurance aspects of redistributive taxation in a single period model.

a number of economically interesting results are obtained. As mentioned, risk spreading opportunities permitted by the independence of the random ability endowments across families are explored. Moreover, inefficiency in the allocation of training resources among the young as a consequence of incomplete loan markets is noted. Public provision of training is shown under some conditions to increase output and reduce inequality. The section concludes with an examination of the joint distribution of ability and earnings in equilibrium. Here we assess the extent to which reward may be presumed to be distributed according to merit. We find the notion of *meritocracy*, appropriately defined, as difficult to defend when social background constrains individual opportunity. In Section 5 we provide a complete solution of our model in the case of Cobb-Douglas utility and production functions, and uniform ability distribution. The equilibrium distribution is explicitly derived and studied. Proofs of the various results appear in Section 6, while concluding observations are mentioned in the final section.

2. THE BASIC MODEL

Imagine an economy with overlapping generations along the lines of Samuelson **[26]**, composed of a large number of individuals. Time is measured discretely and each individual lives for exactly two periods. An equal number of individuals enter and leave the economy in each period, implying a stationary population. Every person in the first period of life (youth) is "attached" to a person in the second period of life (maturity), and this union is called a *family*.

The family is the basic socioeconomic unit in this idealized world. We take the mature member as family head, making all economic decisions. The income (output) of a family depends on the productivity of the head, which in turn depends upon his training and innate ability.[6] Family income must be divided in each period between consumption and training of the youth. We assume that such investment is the only means of transferring goods through time.[7] The training which a mature person possesses this period is simply that portion of family income invested in him last period.

Production of the one perishable commodity is a family specific undertaking. It requires no social interaction among families, nor the use of factors of production other than the mature member's time, which is supplied inelastically. Consumption is also a family activity, with young and mature individuals deriving their personal consumption from the family aggregate in a manner determined by social custom and not examined here. Natural economic ability differs among individuals, each person beginning life with a random endowment of innate aptitudes. Moreover, the level of this endowment becomes known only in maturity. Respectively denoting by x, α, and e a mature individual's output,

[6] "Ability" here means all factors outside of the individual's control which affect his productive capacity.

[7] We abstract from physical capital and real property so as to focus on intergenerational transmission of earnings capacity.

ability, and training, we write the production possibilities as

$$x = h(\alpha, e). \tag{1}$$

The technology $h(\cdot, \cdot)$ is assumed common to all families in all periods.

The decision to divide income between consumption and training is taken by mature people to maximize their expected utility. We assume that all agents possess the same utility function.[8] The utility of a parent is assumed to depend on family consumption during his tenure, and on the wellbeing which the offspring experiences after assuming family leadership.[9] This wellbeing is taken by parents to be some function of the offspring's future earnings. Earnings of offspring, for any level of training decided upon by the parent, is a random variable determined by (1) and the distribution of α. Call the function which parents use to relate the prospective income of their offspring to the child's happiness an *indirect utility function*. Call an indirect utility function *consistent* if it correctly characterizes the relationship between maximized expected utility and earnings for a parent whenever it is taken to characterize that same relationship for the offspring. That is, an indirect utility function V^* is consistent if and only if V^* solves

$$\left[\forall y \geqslant 0, V^*(y) = \max_{0 \leqslant c \leqslant y} E_\alpha U(c, V^*(h(\alpha, y - c)))\right] \tag{2}$$

where $U(c, V)$ is parental utility when family consumption is c and offspring expected utility is V, and E_α is the expectation operator under the random distribution of innate ability. Consistency is assured by (2) because expected utility maximization leads to an indirect utility function for mature individuals identical to that which they presume for their offspring. Along with a solution for (2) will come an optimal *investment function*, $e^*(y)$, which relates parent's income to the training of their offspring. The properties of these functions are studied below.

Notice for now that the earnings of a mature individual (x) depends on his random ability and parent's earnings (y) in the following way:

$$x = h(\alpha, e^*(y)). \tag{3}$$

Equation (3) is the basic relationship which characterizes social mobility in this economy. It establishes a stochastic link between parent's and offspring's earnings. Each individual has limited mobility in that economic achievement varies with aptitude, but in a manner determined by parental success. Moreover, the utility obtained from having a particular level of income, $V^*(y)$, depends

[8] The assumption of identical tastes for everyone is purely for convenience. It may be replaced with the hypothesis of inheritance of preferences within the family only without affecting any of the formal results.

[9] Altruism is modelled in this way (i.e., with the cardinal utility of children entering the parent's utility function) rather than with child's consumption in the parent's utility function (as in Kohlberg [**16**]) to avoid the possibly inefficient growth paths associated with the latter set-up (see Diamond [**9**]).

(through equation (2)) on the extent of social mobility. This is because intergenerational altruism, as modelled here, implies that parents are indirectly concerned about the incomes of all of their progeny. Thus, when there is a great deal of mobility (because α has a large variance and/or $\partial h/\partial\alpha$ is large) a low income will be less of a disadvantage and a high income less of an advantage than when there is very little social mobility. Thus, the graph of $V^*(\cdot)$ will be "flatter" when the society is more mobile. This suggests that our indirect utility function may be useful for assessing the relative merits of social structures generating different degrees of social mobility across generations. We shall return to this point below.

Summarizing, there are three distinct elements which determine the structure of our model economy: the utility function, the production function, and the distribution of the ability endowment. We require some assumptions concerning these elements which are informally described here. First, we shall assume that some consumption is always desirable, parents are risk averse, and parents care about but discount the wellbeing of their offspring. Concerning production we assume diminishing returns to training, a strictly positive marginal product of ability, that the net marginal return to training eventually becomes negative for all individuals, and the net average return is always negative for the least able. Finally, we assume an atomless distribution of ability with connected and bounded support, we rule out autocorrelation in innate endowments,[10] and make extreme values of α sufficiently likely that parents will always desire to invest something in their offspring. The formal statement of these assumptions follows ($\mathbb{R}_+$ is the nonnegative half-line):

ASSUMPTION 1: (i) $U:\mathbb{R}_+^2 \to \mathbb{R}_+$ *is a strictly increasing, strictly concave, twice continuously differentiable function on the interior of its domain, satisfying* $U(0,0) = 0$. (ii) $\forall V \geqslant 0$,

$$\lim_{c\downarrow 0} \frac{\partial U}{\partial c}(c,V) = +\infty.$$

(iii) *There exists* $\gamma > 0$ *such that* $\forall (c,V) \in \mathbb{R}_+^2, (\partial U/\partial V) \in [\gamma, 1-\gamma]$.

ASSUMPTION 2: (i) $h:\mathbb{R}_+^2 \to \mathbb{R}_+$ *is continuously differentiable, strictly increasing, strictly concave in e, satisfying* $h(0,0) = 0$ *and* $h(0,e) \leqslant e, \forall e \geqslant 0$.

(ii) $$\frac{\partial h}{\partial \alpha}(\alpha,e) \geqslant \beta > 0, \quad \forall (\alpha,e) \in \mathbb{R}_+^2.$$

[10] Positive intergenerational correlation of abilities would be a more natural assumption. (For a discussion of the evidence see Goldberger [13].) We have assumed this away for simplicity's sake. One might think this affects the basic character of our argument, since if "rich" parents (and hence their children) are smarter than "poor" parents it may not be inefficient for them to invest relatively more in their offspring. But as our discussion of "meritocracy" in Section 4 illustrates, when parental resources determine human capital investments it is not necessarily true that those with greater income have (probabilistically) greater ability.

(iii) *There exists* $\rho > 0$, $\hat{e} > 0$ *such that*

$$e \geqslant \hat{e} \Rightarrow \max_{\alpha \in [0,1]} \frac{\partial h}{\partial e}(\alpha, e) \leqslant \rho < 1.$$

ASSUMPTION 3: *Innate economic ability is distributed on the unit interval independently and identically for all agents. The distribution has a continuous, strictly positive density function,* $f:[0,1] \to \mathbb{R}_+$.

3. CONSISTENT INDIRECT UTILITY AND THE DYNAMICS OF THE INCOME DISTRIBUTION

Our assumptions (hereafter A1–A3, etc.) enable us to demonstrate the existence of a unique consistent indirect utility function $V^*(y)$, and to make rather strong statements about the long run behavior of the distribution of earnings implied by equation (3). These essentially mathematical results are required before we can proceed to the economic analysis which is developed in the following section. To see that the notion of consistent indirect utility is not vacuous we adopt to our context a dynamic programming argument originated by Bellman. First, however, notice that A2(iii) implies there is a unique earnings level $\bar{y}$ such that $h(1, \bar{y}) = \bar{y}$ and $y > \bar{y} \Rightarrow h(1, y) < y$. Hence, we may, without loss of generality, restrict our attention to incomes in the interval $[0, \bar{y}]$, since no family's income could persist outside this range. Let $C[0, \bar{y}]$ denote the continuous real valued functions on $[0, \bar{y}]$. For $\phi \in C[0, \bar{y}]$ define the function $T\phi(\cdot)$ as follows:

$$\forall y \in [0, \bar{y}], \quad T\phi(y) \equiv \max_{0 \leqslant c \leqslant y} E_\alpha U(c, \phi(h(\alpha, y - c))). \tag{4}$$

An indirect utility function is consistent if and only if it is a fixed point of the mapping T. Our assumptions permit us to show that $T: C[0, \bar{y}] \to C[0, \bar{y}]$ is a contraction mapping, and hence has a unique fixed point. In Section 6 we prove the following:

THEOREM 1: *Under* A1, A2, *and* A3, *for the above defined positive number* $\bar{y}$, *there exists a unique consistent indirect utility function* V^* *on* $[0, \bar{y}]$. V^* *is strictly increasing, strictly concave, and differentiable on* $(0, \bar{y}]$. *The optimal consumption function* $c^*(y)$ *is continuous and satisfies* $0 < c^*(y) < y$, $\forall y \in (0, \bar{y}]$. (*Denote* $e^*(y) \equiv y - c^*(y)$.)

The property of uniqueness established here is very important. For neither positive nor normative analysis would our story carry much force were we required to select arbitrarily among several consistent representations of family preferences. On the other hand we have assumed a cardinal representation of preferences. This is unavoidable if parents are to be taken as contemplating the welfare of their offspring. While an axiomatic justification of our choice criterion might be sought along the lines of Koopmans **[17]**, this task is complicated by the uncertainty inherent in our set-up and is not pursued here. As one might expect V^* inherits the property of risk aversion from U. This is important for much of the discussion of Section 4.

We turn now to the study of the intergenerational motion of the distribution of income. In any period (say nth) the distribution of family income is conceived as a probability measure on $[0, \bar{y}]$, ν_n. The interpretation is that for $A \subset [0, \bar{y}]$, $\nu_n(A)$ is the fraction of all mature agents whose incomes lie in the set A. In what follows we denote by $\mathcal{P}$ the set of probability measures on $\mathcal{B}$, the Borel sets of $[0, \bar{y}]$. μ will denote the Lebesgue measure on $\mathcal{B}$.

The system of family decision making outlined above enables us to treat income dynamics in this economy as a Markoff process. To do so we require the following definition.

DEFINITION 1: A *transition probability* on $[0, \bar{y}]$ is a function $P:[0, \bar{y}] \times \mathcal{B} \to [0,1]$ satisfying: (i) $\forall y \in [0, \bar{y}]$, $P(y, \cdot) \in \mathcal{P}$; (ii) $\forall A \in \mathcal{B}, P(\cdot, A)$ is a $\mathcal{B}$-measurable function on $\mathbb{R}_+$.

To each transition probability on $[0, \bar{y}]$ there corresponds a Markoff process. That is, once a transition probability has been specified one can generate, for any fixed $\nu_0 \in \mathcal{P}$, a sequence of measures $\{\nu_n\}$ as follows:

$$\nu_n(A) = \int P(y,A)\nu_{n-1}(dy) \qquad (n = 1,2,\ldots). \tag{5}$$

Thus, the intergenerational motion of the distribution of income will be fully characterized once we have found the mechanism which relates the probability distribution of an offspring's income to the earnings of its parent. Towards this end define the function $h^{-1}: \mathcal{B} \times \mathbb{R}_+ \to \mathcal{B}$ as follows:

$$h^{-1}(A,e) \equiv \{\alpha \in [0,1] \mid h(\alpha,e) \in A\}.$$

Then it is clear that the probability an offspring has income in A when parent's earnings is y is simply the probability that the offspring has an innate endowment in the set $h^{-1}(A, e^*(y))$. The relevant transition probability is given by

$$P(y,A) \equiv \int_{h^{-1}(A, e^*(y))} f(\alpha)\mu(d\alpha). \tag{6}$$

With P defined as in (6) it is obvious that $P(y, \cdot) \in \mathcal{P}$; moreover, it follows from a result of Futia[11] that $P(\ ,A)$ is $\mathcal{B}$-measurable when $e^*(\cdot)$ is continuous. The last fact was shown in Theorem 1.

We now introduce the notion of an equilibrium distribution.

DEFINITION 2: An *equilibrium income distribution* is a measure $\nu^* \in \mathcal{P}$ satisfying, $\forall A \in \mathcal{B}$,

$$\nu^*(A) = \int P(y,A)\nu^*(dy). \tag{7}$$

[11] See Futia [**12**, Theorem 5.2, p. 74].

Thus an equilibrium distribution is one which, if it ever characterizes the distribution of earnings among a given generation, continues to do so for each subsequent generation. Assumptions A1–A3 already permit us to assert the existence of at least one equilibrium, and convergence to some equilibrium. However, since our focus here is on dynamics it seems preferable to introduce the following simple assumption from which follows the global stability of the unique equilibrium point.

ASSUMPTION 4: $y_1 > y_2 \Rightarrow e^*(y_1) > e^*(y_2)$.

That is, education is a normal good. We now have the following theorem.

THEOREM 2: *Under* A1–A4 *there exists a unique equilibrium distribution* $\nu^* \in \mathcal{P}$. *If* $\{\nu_n\}$ *is a sequence of income distributions originating from the arbitrary initial* ν_0, *then* $\lim_{n\to\infty} \nu_n = \nu^*$, *pointwise on* $\mathcal{B}$. *Moreover,* ν^* *has support* $[0, \tilde{y}]$, *where* y *satisfies* $h(1, e^*(\tilde{y})) = \tilde{y}$.

Thus, by the proof given in Section 6, we establish existence, uniqueness, and stability for our notion of equilibrium. Together with Theorem 1 this constitutes the framework within which we explore the issues raised in the introduction. This examination is conducted in the next two sections.

4. ECONOMIC ANALYSIS OF THE EQUILIBRIUM DISTRIBUTION

We have now constructed a relatively simple maximizing model of the distribution of lifetime earnings among successive generations of workers. The salient features of this model are: (i) the recursive family preference structure reflecting intergenerational altruism; (ii) the capital-theoretic characterization of production possibilities; (iii) the stochastic nature of innate endowments, assumed independent within each family over time, and across families at any moment of time; and (iv) the assumed absence of markets for human capital loans and income risk sharing between families. These factors determine the average level of output in equilibrium and its dispersion, as well as the criterion which families use to evaluate their wellbeing.

Thus, we may think of our model as defining a function M, which associates with every triple of utility, production, and ability distribution functions satisfying our assumptions, a unique pair, the equilibrium earnings distribution and consistent indirect utility:

$$M : (U, h, f) \to (\nu^*, V^*).$$

A natural way to proceed in the analysis of this model is to ask "comparative statics" questions regarding how the endogenous (ν^*, V^*) vary with changes in the exogenous (U, h, f). This is a very difficult question in general, though we are able to find some preliminary results along these lines, presented below.

Before proceeding to this, however, let us notice how the model distinguishes the phenomenon of *social mobility across generations* from that of *inequality in* *earnings within a generation.* The relationship between these alternative notions of social stratification has been the focus of some attention in the literature [**4**, **27**]. Here the two notions are inextricably intertwined; intergenerational social mobility is a property of the transition probability P, while cross-sectional inequality is (asymptotically) a property of the equilibrium distribution to which P gives rise. The basic problem for a normative analysis of economic inequality raised by this joint determination is that the concept of a "more equal" income distribution is not easily defined. For instance, how is one to compare a situation in which there is only a slight degree of inequality among families in any given generation but no mobility within families across generations, with a circumstance in which there is substantial intragenerational income dispersion but also a large degree of intergenerational mobility? Which situation evidences less inequality? In the first case such inequality as exists between families lasts forever, while in the latter case family members of a particular cohort experience significant earnings differences but no family is permanently assigned to the bottom or top of the earnings hierarchy.

The formulation put forward here provides a way of thinking about this trade-off. Our equilibrium notion depicts the intragenerational dispersion in earnings which would persist in the long run. On the other hand, the consistent indirect utility function incorporates into the valuation which individuals affix to any particular location in the income hierarchy the consequences of subsequent social mobility. As noted earlier, a great deal of intergenerational mobility implies a relatively "flat" indirect utility function, making the cross-sectional distribution of *welfare* in equilibrium less unequal than would be the case with little or no mobility. Thus, by examining the distribution in equilibrium of consistent indirect utility (instead of income) one may simultaneously account for both kinds of inequality within a unified framework. Indeed, the statistic

$$W = \int V^*(y)\nu^*(dy) \tag{8}$$

is a natural extension of the standard utilitarian criterion, and provides a complete ordering of social structures (U, h, f) satisfying A1–A4.

Let us consider now the absence of loan and insurance markets in this model. These market imperfections have important allocative consequences. A number of writers have called attention to the fact that efficient resource allocation requires that children in low-income families should not be restricted by limited parental resources in their access to training. For example, Arthur Okun has remarked [**22**, p. 80–81] concerning the modern U.S. economy that "... The most important consequence (of an imperfect loan market) is the inadequate development of the human resources of the children of poor families—which I would judge, *is one of the most serious inefficiencies of the American Economy today*." (Emphasis added.) While it is certainly the case that in most societies a variety of devices exist for overcoming this problem, they are far from perfect. For example, early childhood investments in nutrition or pre-school education are fundamentally income-constrained. Nor should we expect a competitive loan market to completely eliminate the dispersion in expected rates of return to training across families which arises in this model. Legally, poor parents will not be able to constrain their children to honor debts incurred on their behalf. Nor will the newly matured children of wealthy families be able to attach the (human) assets of their less well-off counterparts, should the latter decide for whatever reasons not to repay their loans. (Default has been a pervasive problem with government guaranteed educational loan programs, which would not exist absent public underwriting.) Moreover, the ability to make use of human capital is unknown even to the borrower. Thus genuine bankruptcies can be expected, as people mature to learn they are not as able as their parents gambled they would be. The absence of inter-family loans in this model reflects an important feature of reality, the allocative implications of which deserve study.

Another critical feature of this story is that parents bear theoretically insurable risk when investing in their offspring. Earnings of children are random here because the economic ability of the child becomes known only in maturity. Yet a significant spreading of risk is technologically possible for a group of parents investing at the same level, since their children's earnings are independent identically distributed random variables. Again we may expect that markets will fail to bring about complete risk spreading; the moral hazard problems of income insurance contracts are obvious.[12]

Any attempt by a central authority to "correct" for these market imperfections must cope with the same incentive problems which limit the scope of the competitive mechanism here. Efforts by a government to alter the allocation of training among young people will affect the consumption-investment decision of parents. Moreover parents may be presumed to have better information than a central authority about the distribution of their children's abilities. (This is especially so if ability is positively correlated across generations of the same family.) Yet, as long as a group of parents with the same information have different incomes and there is no capital market, their investment decisions will vary causing inefficiency. This inefficiency could be reduced if the government were to redistribute incomes more equally, but only at the cost of the excess burden accompanying such a redistributive tax.

However, because we have assumed altruism between generations, a program of income redistribution which is imposed permanently has welfare effects here not encountered in the usual tax analysis. The fact that income will be redistributed by taxing the next generation of workers causes each parent to regard his offspring's random income as less risky. Since parents are risk averse, egalitarian redistributive measures have certain welfare enhancing "insurance" characteristics. It may be shown that, as long as the redistributive activity is not too extensive, this "insurance effect" dominates the "excess burden effect." Under such circumstances a permanent redistributive tax policy can be designed in such a way as to make all current mature agents better off.

[12] See Zeckhauser [30] for a discussion of these problems.

DEFINITION 3: An *education specific tax policy* (ESTP) is a function $\tau : \mathbb{R}_+^2 \to \mathbb{R}_+$ with the interpretation that $\tau(y, e)$ is the after-tax income of someone whose income is y and training is e. An ESTP is *admissible* if $h_\tau(\alpha, e) \equiv \tau(h(\alpha, e), e)$ satisfies Assumption 2. An admissible ESTP is *purely redistributive* (PRESTP) if $\int[\tau(h(\alpha, e), e) - h(\alpha, e)]f(\alpha)\mu(d\alpha) = 0$ for every $e \geqslant 0$. Denote by $\mathfrak{L}$ the class of admissible PRESTP. $\mathfrak{L}$ is a non-empty convex set.

DEFINITION 4: The partial ordering "*is more egalitarian than*," denoted $\succcurlyeq$, is defined on $\mathfrak{L}$ by the requirement $\tau_1 \succcurlyeq \tau_2$ if and only if, for each e, the after tax distribution of income under τ_2 is "more risky" in the sense of second order stochastic dominance (see for example [24]) than the after tax distribution under τ_1.

These two definitions allow us to formalize the discussion above. By a result of Atkinson [1], $\tau_1 \succcurlyeq \tau_2$ implies the Lorenze curve of the after tax income distribution for persons of a given level of training under τ_1 lies on or above the corresponding curve under τ_2. However, since the investment function $e^*(\cdot)$ will depend on τ, the same statement will not generally hold for the overall distributions of income. Thus, we must use the term "more egalitarian" with some caution. Let $\bar{\tau}(y, e) \equiv \int h(\alpha, e)f(\alpha)\mu(d\alpha)$ and $\underline{\tau}(y, e) \equiv y$, for all $(y, e) \in \mathbb{R}_+^2$. Then $\bar{\tau}$, representing perfect income insurance, is maximal in $\mathfrak{L}$ relative to $\succcurlyeq$; $\underline{\tau}$ on the other hand represents *laissez faire*. For $\beta \in (0, 1)$ let $\tau_\beta \equiv \beta\bar{\tau} + (1 - \beta)\underline{\tau}$. Then $\tau_\beta \in \mathfrak{L}$. Finally, for every $\tau \in \mathfrak{L}$ let V_τ^* denote the unique solution for (2) when h is replaced by h_τ.

THEOREM 3: *Under* A1–A3 $\tau_1 \succcurlyeq \tau_2 \Rightarrow V_{\tau_1}^*(y) \geqslant V_{\tau_2}^*(y)$, $\forall y \geqslant 0$, $\tau_1, \tau_2 \in \mathfrak{L}$.

Theorem 3 states that the use of a more egalitarian PRESTP among future generations of workers will lead to a higher level of welfare for each mature individual currently. The careful reader will notice that this result depends upon our assumption that labor supply is exogenous. In general we may expect that an income tax will distort both the educational investment decisions of parents and their choices concerning work effort. By allowing the income tax function to vary with the level of educational investment (redistributing income within education classes only), we avoid the negative welfare consequences of the former distortion. It is possible to preserve the result in Theorem 3 in the face of the latter kind of distortion as well, if we restrict our attention to "small" redistributive efforts. To see this, imagine that $h(\alpha, e)$ determines output per unit of time worked, that each parent is endowed with one unit of time, and that parents' utility depends on work effort as well as consumption and offspring's welfare. Under these slightly more general circumstances one can repeat the development of Theorems 1 and 2 with the natural modifications that indirect utility becomes a function of output per unit time (instead of total income), and income is the product of work effort and $h(\alpha, e)$. Now consider introducing the tax scheme τ_β (described above) for "small" β. For fixed labor supply it may be shown that welfare, at any level of income, is a concave function of β. Moreover, Theorem 3 implies that welfare is an increasing function of β. However, from standard tax theory we know that the excess burden of the tax τ_β is zero when $\beta = 0$, and has derivative with respect to β equal to zero at $\beta = 0$. (That is the excess burden of a "small" tax is a second order effect.) Thus, for β sufficiently close to zero the "insurance effect" described in Theorem 3 will outweigh the "excess burden" effect of the tax scheme, leading to an increase in wellbeing for all parents of the current generation.

This result suggests that the standard tradeoff between equity and efficiency is somewhat less severe in the second-best world where opportunities to acquire training vary with parental resources. It is sometimes possible to design policies which increase both equity and efficiency. As a further indication of this last point, let us examine the impact of "public education" in this model. By public education we mean the centralized provision of training on an equal basis to every young person in the economy. For the purposes of comparison we contrast the *laissez faire* outcome with that which occurs under public education with a per-capita budget equal to the expenditure of the average family in the no-intervention equilibrium. We assume for this result that education is financed with non-distortionary tax.

ASSUMPTION 5: *$h(\alpha, e^*(y))$ is concave in y, and $\partial h(\alpha, e^*(y))/\partial \alpha$ is convex in y.*

Assumption 5 posits that the "marginal product of social background" declines as social background (i.e., parent's income) improves, but declines less rapidly for the more able. We have proven in Section 6 the following result.

THEOREM 4: *Under* A1–A5 *universal public education, with a per-capita budget equal to the investment of the average family in the laissez faire equilibrium, will produce an earnings distribution with lower variance and higher mean than that which obtains without intervention.*

Finally, let us turn to an examination of the relationship between earnings and ability in equilibrium. It is widely held that differences in ability provide ethical grounds for differences in rewards. An often cited justification for the market determined distribution of economic advantage is that it gives greater rewards to those of greater abilities. In our model, such is not literally the case because opportunities for training vary with social origin. Thus rewards correspond to productivity, but not ability. Productivity depends both on ability and on social background. Nonetheless, some sort of systematic positive relationship between earnings and ability might still be sought. Imagine that we observe in equilibrium an economy of the type studied here. Suppose that we can observe an agent's income, but neither his ability nor his parent's income. Now the pair $(\tilde{\alpha}, \tilde{x})$ of individual ability and income is a jointly random vector in the population. A very weak definition of a meritocratic distribution of income is that $\tilde{\alpha}$ and $\tilde{x}$ are positively correlated. Under this definition the distribution is meritocratic if a regression of income on ability would yield a positive coefficient in a large sample. A more substantive restriction would require that if $x_1 \geqslant x_2$ then $\mathrm{pr}[\tilde{\alpha} \geqslant a \mid \tilde{x} = x_1] \geqslant \mathrm{pr}[\tilde{\alpha} \geqslant a \mid \tilde{x} = x_2]$ for all $a \in [0, 1]$. Under this stronger definition of a meritocratic distribution of rewards, we require that greater earnings imply a greater conditional probability that innate aptitude exceeds any prespecified level. The following theorem asserts that the economy studied here is generally weakly meritocratic but not strongly so.

THEOREM 5: *Under* A1–A4 *the economy is weakly meritocratic. However, it is not generally true that the economy is strongly meritocratic.*

The reason that a strong meritocracy may fail to obtain can be easily seen. Suppose there exists some range of parental incomes over which the investment schedule $e^*(y)$ increases rapidly, such that the least able offspring of parents at the top earn nearly as much as the most able offspring of parents at the bottom of this range. Then the conditional ability distribution of those earning incomes slightly greater than that earned by the most able children of the least well-off parents may be dominated by the conditional distribution of ability of those whose incomes are slightly less than the earnings of the least able children of the most well-off parents. An example along these lines is discussed in Section 6.

5. AN EXAMPLE

It is interesting to examine the explicit solution of this problem in the special case of Cobb-Douglas production and utility functions. For the analysis of this section we assume that

$$(9)\qquad U(c, V) = c^{\gamma} V^{1-\gamma}, \qquad 0 < \gamma < 1;$$

$$h(\alpha, e) = (\alpha e)^{1/2},$$

$$f(\alpha) \equiv 1, \qquad \alpha \in [0, 1].$$

With these particular functions it is not hard to see that the consistent indirect utility function and optimal investment function have the following forms:

$$(10)\qquad V^*(y) = k y^{\delta}, \qquad k > 0,\ 0 < \delta < 1; \qquad \text{and,}$$

$$e^*(y) = sy, \qquad 0 < s < 1.$$

The reader may verify by direct substitution that the functions V^* and e^* given above satisfy (2) when $s = (1+\gamma)/2$, $\delta = 2\gamma/(1+\gamma)$, and

$$k = \frac{1}{2}(1+\gamma)^2\left[(1+\gamma)^2 - 2\gamma^2\right]^{-1}(1-\gamma)^{(1-\gamma)/(1+\gamma)}.$$

By Theorem 1 we know these solutions are unique.

Thus, in this simple instance, we obtain a constant proportional investment function. This result continues to hold for arbitrary distributions of ability f, and for production functions $h(\alpha, e) = \alpha^a e^{1-a}$, $a \in (0,1)$. However, only in the case given in (9) are we able to explicitly derive the equilibrium distribution, the task to which we now turn.

A parent earning y invests sy in his offspring whose income is then $x = (\alpha sy)^{1/2}$. Notice that $y > s \Rightarrow x < y$, no matter what α turns out to be. Thus no family's income can remain indefinitely above s, and the equilibrium distribution is concentrated on $[0,s]$. Now $h^{-1}(x, e^*(y)) = x^2/sy$, $y \in [x^2/s, s]$, gives the ability necessary to reach earnings x when parental income is y. Let $g_n(\cdot)$ be the density function of the income distribution in period n. Then $g_n(x)\,dx$ is, for "small" dx, the fraction of families in period n with incomes in the interval $(x, x+dx)$. Heuristically,

$$g_n(\bar{x})\,dx \approx \mathrm{pr}\left[x \in (\bar{x}, \bar{x}+dx)\right]$$
$$\approx \sum_y \mathrm{pr}\left[\alpha \in \left(h^{-1}(\bar{x}, sy), h^{-1}(\bar{x}+dx, sy)\right)\right] g_{n-1}(y)$$
$$\simeq \sum_y f\left(h^{-1}(\bar{x}, sy)\right) \frac{\partial h^{-1}}{\partial x}(\bar{x}, sy) g_{n-1}(y)\,dx.$$

This suggests the following formula, shown rigorously to be valid in Lemma 1 of Section 6:

$$g_n(x) = \frac{2x}{s}\int_{x^2/s}^{s} \frac{g_{n-1}(y)}{y}\,dy.$$

Thus, we have that the density of the equilibrium distribution solves the functional equation

$$\hat{g}(x) = \frac{2x}{s}\int_{x^2/s}^{s} \frac{\hat{g}(y)}{y}\,dy. \tag{11}$$

Consider now the change of variables $z = x/s$. Then z varies in $[0,1]$ and the density of z satisfies $g^*(z) = s\hat{g}(sz)$. One may now use (11) (with a change of variables $\tilde{y} = y/s$ in the integrand) to see that $g^*(\cdot)$ must satisfy

$$g^*(z) = 2z\int_{z^2}^{1} \frac{g^*(\tilde{y})}{\tilde{y}}\,d\tilde{y}. \tag{12}$$

A solution of (12) gives immediately the solutions of (11) for all values of $s \in (0,1)$. Theorem 2 assures the existence and uniqueness of such a solution.

We solve (12) by converting it into a functional differential equation, presuming its solution may be written in series form, and then finding the coefficients for this series. Differentiating (12) we have that

$$\frac{dg^*}{dz}(z) = g^*(z)/z - 4g^*(z^2). \tag{13}$$

Suppose that $g^*(z)=\sum_{n=0}^{\infty}a_n z^n$. Substitute this form into (13). Equate coefficients of like powers on the left and right hand sides and deduce the following recursive system for $\{a_n\}$:

$$a_n = 0, \quad n \neq n_k \qquad (k=1,2,\dots),$$

$$a_{n_k} = -2a_{n_{k-1}}/n_{k-1} \qquad (k=2,3,\dots),$$

$$a_{n_1} = \bar{g}, \text{ arbitrary, and}$$

$$n_k = 2^k - 1 \qquad (k=1,2,\dots).$$

Solving directly for $\{a_{n_k}\}$ yields

$$a_{n_k} = \bar{g}(-2)^{k-1}\left(\prod_{j=1}^{k-1}(2^j-1)\right)^{-1} \qquad (k=2,3,\dots).$$

Thus, we have that the unique solution of (12) is given by:

$$g^*(z) = \bar{g}\left\{z + \sum_{k=2}^{\infty}\left[(-2)^{k-1}z^{2^k-1}\left(\prod_{j=1}^{k-1}(2^j-1)\right)^{-1}\right]\right\}, \quad z \in [0,1].$$

The constant $\bar{g}$ is determined by the requirement that g^* integrates to unity ($\bar{g} \approx 6.92$).

The implied solution for $\hat{g}(x)$ may be derived from this formula for any $s \in (0,1)$. The basic character of the equilibrium distribution is unaffected by the savings rate (i.e., the parameter γ in the utility function) which simply alters its scale. The equilibrium distributions are single peaked, just slightly right skewed, with a mean of $(0.422)s$ and a variance of $(0.037)s^2$. Consumption per family is maximized in equilibrium when $s=1/2$, the output elasticity of education. Public education in this example with a per capita budget equal to the training expenditure of the average family in equilibrium, reduces the variance of the equilibrium distribution by 35 per cent, to $(0.024)s^2$. At the same time, mean output rises by 3.4 per cent under the centralized provision of training resources. In this example the efficiency gain of centralized training is small relative to the associated reduction in inequality.

6. PROOFS

PROOF OF THEOREM 1: Recall the map $T: C[0,\bar{y}] \to C[0,\bar{y}]$ defined by

$$(T\Phi)(y) = \max_{0 \leqslant c \leqslant y} E_\alpha U(c, \Phi(h(\alpha, y-c))).$$

For $\Phi \in C[0,\bar{y}]$ define $\|\Phi\| \equiv \sup_{x\in[0,y]}|\Phi(x)|$. We shall show that under this norm T is a contraction on $C[0,\bar{y}]$. Let $\Phi, \Psi \in C[0,\bar{y}]$:

$$\|T\Phi - T\Psi\| \equiv \sup_{\bar{y}\geqslant y\geqslant 0}\Big|\max_{0\leqslant c\leqslant y} E_\alpha U(c,\Phi(h(\alpha,y-c))) - \max_{0\leqslant c\leqslant y} E_\alpha U(c,\Psi(h(\alpha,y-c)))\Big|.$$

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

$$\begin{aligned}
\|T\Phi - T\Psi\| &= \sup_{\bar{y}\geqslant y\geqslant 0}\big|E_\alpha\big(U(\hat{c},\Phi) - U(\hat{\hat{c}},\Psi)\big)\big| \\
&\leqslant \sup_{\bar{y}\geqslant y\geqslant 0}\max\Big\{\big|E_\alpha\big(U(\hat{c},\Phi) - U(\hat{c},\Psi)\big)\big|;\ \big|E_\alpha\big(U(\hat{\hat{c}},\Phi) - U(\hat{\hat{c}},\Psi)\big)\big|\Big\} \\
&\leqslant \max\Big\{\sup_{\bar{y}\geqslant y\geqslant 0}\big|E_\alpha\big(U_2(\hat{c},\Psi)\cdot[\Phi-\Psi]\big)\big|;\ \sup_{\bar{y}\geqslant y\geqslant 0}\big|E_\alpha\big(U_2(\hat{\hat{c}},\Psi)\cdot[\Phi-\Psi]\big)\big|\Big\} \\
&\leqslant \gamma\max\Big\{\sup_{\substack{\bar{y}\geqslant y\geqslant 0\\ a\in[0,1]}}\big|\Phi(h(\alpha,y-\hat{c}(y))) - \Psi(h(\alpha,y-\hat{c}(y)))\big|;\ \sup_{\substack{\bar{y}\geqslant y\geqslant 0\\ a\in[0,1]}}\big|\Phi(h(\alpha,y-\hat{\hat{c}}(y))) - \Psi(h(\alpha,y-\hat{\hat{c}}(y)))\big|\Big\} \\
&\leqslant \gamma\sup_{\bar{y}\geqslant y\geqslant 0}|\Phi(y)-\Psi(y)| = \gamma\|\Phi-\Psi\|
\end{aligned}$$

for some $\gamma < 1$, using A1(iii). Hence T is a contraction and, by the Banach Fixed Point Theorem [**11**, Theorem 3.8.2, p. 119] there exists a unique fixed point $V^* \in C[0,\bar{y}]$. By the definition of T, V^* is the solution of (2).

Define the sequence of functions $\{V^n\} \in C[0,\bar{y}]$ inductively as follows:

$$V^1(y) \equiv \max_{0\leqslant c\leqslant y} E_\alpha U(c, U(h(\alpha,y-c),0)),$$

$$V^n(y) \equiv (TV^{n-1})(y) \qquad (n = 2,3,\dots).$$

Clearly $\{V^n\}\to V^*$ uniformly. Let $\hat{c}^n(y)$ be the optimal policy function corresponding to V^n. It follows from A1(i), A2(i), the Implicit Function Theorem, and an induction that $\hat{c}^n(y)$ are differentiable functions, $n = 1,2,\dots$. Moreover, A1(ii), A2(i), A3, the Envelope Theorem, and another induction imply that V^n is differentiable, and $\lim_{y\downarrow 0} dV^n/dy(y) = \infty$, $n = 1,2,\dots$. Thus $0 < \hat{c}^n(y) < y$, $\forall n$ and $\forall y \in (0,\bar{y}]$.

Now, it follows from the strict concavity of U, the concavity of h in its second argument, and an induction that the functions V^n are strictly concave. Hence V^* is concave. But for $0 < \delta < 1$, using the above-mentioned concavity, we have

$$\begin{aligned} V^*(\delta y_1 + (1-\delta)y_2) &= \max_c E_\alpha U(c, V^*(h(\alpha, \delta y_1 + (1-\delta)y_2 - c))) \\ &\geqslant E_\alpha U(\delta c_1 + (1-\delta)c_2, \\ &\quad \times V^*(h(\alpha, \delta(y_1 - c_1) + (1-\delta)(y_2 - c_2)))) \\ &\geqslant \delta E_\alpha U(c_1, V^*(h(\alpha, y_1 - c_1))) \\ &\quad + (1-\delta)E_\alpha U(c_2, V^*(h(\alpha, y_2 - c_2))) \\ &= \delta V^*(y_1) + (1-\delta)V^*(y_2) \end{aligned}$$

where $c_1 \equiv c^*(y_1)$ and $c_2 = c^*(y_2)$, with equality iff $y_1 = y_2$. Hence V^* is strictly concave, and therefore differentiable almost everywhere. Now the Envelope Theorem implies

$$\frac{dV^{*+}}{dy}(y) = E_\alpha(c^*, V^*)\frac{dV^*}{dy}(h)^+ h_2(\alpha, y - c^*)$$

while

$$\frac{dV^{*-}}{dy}(y) = E_\alpha\left(U_2(c^*, V^*)\frac{dV^*}{dy}(h)^- h_2(\alpha, y - c^*)\right),$$

where "+" and "−" refer to right or left hand derivatives, respectively. It follows from the monotonicity of h in α, the continuity of f, and the fact that $dV^{*+}/dy \neq dV^{*-}/dy$ at most on a set of μ-measure zero, that the right-hand side of the two equations above are equal. Hence V^* is differentiable on $(0, \bar{y}]$, and (by uniform convergence of V^n) $\lim_{y\to 0} dV^*/dy(y) = \infty$. Thus $c^*(y)$ is unique for each y and $0 < c^*(y) < y$. Notice now that $\{\hat{c}^n\} \to c^*$ pointwise on $[0, \bar{y}]$. If this were not true then, for some y, the sequence $\{\hat{c}^n(y)\}$ would possess a convergent subsequence bounded away from $c^*(y)$, contradicting the uniqueness of the latter. The continuity of c^* follows.

PROOF OF THEOREM 2: We want to show that the Markoff chain

$$\nu_n(A) = \int_{[0,\infty)} P(y, A)\nu_{n-1}(dy), \quad \nu_0 \in \mathcal{P} \text{ given}, \qquad (n = 1, 2, \dots),$$

possesses a unique, globally stable invariant measure $\nu^* \in \mathcal{P}$, where the transition probability is given by

$$P(y, A) \equiv \int_{h^{-1}(A, e^*(y))} f(\alpha)\mu(d\alpha).$$

The mathematical question here is precisely that which arises in the theory of stochastic growth (see, for example, Brock-Mirman **[5]**). To answer this question we use the following results reported in the excellent recent survey of Futia:

PROPOSITION (*Futia* [**12**, *Theorems* 3.6, 3.7, 4.6, 5.2, *and* 5.6]): *Under Assumptions* A1–A3, *the process represented by equations* (5) *and* (6) *has at least one invariant measure* $\nu^* \in \mathcal{P}$, *and always converges to some invariant measure. Moreover, if in addition* A4 *and the following condition hold, then the invariant measure is unique*:

CONDITION 1: There exists a $y_0 \in [0, \bar{y}]$ with the property that for every integer $k \geqslant 1$, any number $y \in [0, \bar{y}]$, and any neighborhood U of y_0, one can find an integer n such that $P^{nk}(y, U) > 0$.

The m-step transition probability $P^m(y,A)$ is defined inductively by $P^1(y,A) \equiv P(y,A)$ and

$$P^m(y,A) = \int_{[0,\infty)} P(y,ds)P^{m-1}(s,A) \qquad (m = 2,3,\ldots).$$

Thus, we will have proven the theorem if we can exhibit an earnings level y_0, every neighborhood of which is with positive probability and finite periodicity entered infinitely often from any initial earnings level. We will need the following notation:

$$x_0^1(y) \equiv h(0,e^*(y)); \quad x_0^n(y) \equiv x_0^1\big(x_0^{n-1}(y)\big) \qquad (n = 2,3,\ldots);$$

$$x_1^1(y) \equiv h(1,e^*(y)); \quad x_1^n(y) \equiv x_1^1\big(x_1^{n-1}(y)\big) \qquad (n = 2,3,\ldots).$$

The following Lemmata will also be useful:

LEMMA 1: $P(y,A) = \int_A f(h^{-1}(x,e^*(y)))[\partial h^{-1}/\partial x(x,e^*(y))]\,\mu(dx)$.

PROOF: Implement the change of variables $\alpha = h^{-1}(x,e^*(y))$ in (6). *Q.E.D*

REMARK: Thus our transition probability on the bounded state space $[0, \bar{y}]$ arises from integrating a bounded, measurable density function. It is this fact which enables us to employ Futia's results indicated in the Proposition above.

LEMMA 2: *Let* $Sm(y)$ *be the support of the probability* $P^m(y,\cdot) \in \mathcal{P}$. *Then* $Sm(y) = (x_0^m(y), x_1^m(y))$.

PROOF: Define

$$p^1(y,x) = f\big(h^{-1}(x,e^*(y))\big)\left[\frac{\partial h^{-1}}{\partial x}(x,e^*(y))\right], \quad x \in \big(x_0^1(y), x_1^1(y)\big),$$

$$= 0 \text{ otherwise,}$$

and inductively

$$p^m(y,x) \equiv \int_{[0,\infty)} p^1(y,z)p^{m-1}(z,x)\mu(dz) \qquad (m=2,3,\dots).$$

Now, by Lemma 1 we have

$$P^m(y,A) = \int_A p^m(y,x)\mu(dx) \qquad (m=1,2,\dots).$$

Thus $Sm(y) = \{x \mid p^m(y,x) > 0\}$, and we need only show that $p^m(y,x) > 0$ iff $x \in (x_0^m(y), x_1^m(y))$. This is done by induction. Obviously $p^1(x,y) > 0$ iff $x \in (x_0^1(y), x_1^1(y))$ from the definitions. Suppose the hypothesis holds for $p^{m-1}(y, x)$, and define $Am(x,y) \equiv \{z \in [0,\infty) \mid p^1(y,z)p^{m-1}(z,x) > 0\}$. Now $p^m(y,x) > 0$ iff $\mu(Am(x,y)) > 0$. But $p^1(y,z) > 0$ iff $z \in (x_0^1(y), x_1^1(y))$, while $p^{m-1}(z,x) > 0$ iff $x \in (x_0^{m-1}(z), x_1^{m-1}(z))$. It follows from Assumptions A2(i) and A4 that the functions $x_0^m(\cdot)$ and $x_1^m(\cdot)$ are strictly increasing. Thus

$$Am(x,y) = \left(x_0^1(y), x_1^1(y)\right) \cap \left(\left(x_1^{m-1}\right)^{-1}(x), \left(x_0^{m-1}\right)^{-1}(x)\right).$$

So $\mu(Am(x,y)) > 0$ iff both $(x_1^{m-1})^{-1}(x) < x_1^1(y)$ and $(x_0^{m-1})^{-1}(x) > x_0^1(y)$; that is, iff $x_1^m(y) > x > x_0^m(y)$. *Q.E.D.*

LEMMA 3: $\mathrm{Lim}_{m\to\infty} x_0^m(y) = 0$, $\forall y \in [0, \bar{y}]$. *Moreover, there exists* $\hat{y} > 0$ *such that* $\liminf_{m\to\infty} x_1^m(y) \geqslant \hat{y}$ $\forall y \in [0, \bar{y}]$.

PROOF: For $y > 0$ by A2(i) and Theorem 1, $x_0^1(y) \leqslant e^*(y) < y$. Thus the sequence of numbers $\{x_0^m(y)\}$ is monotonically decreasing and bounded below by zero. Let $\underline{x} = \lim_{m\to\infty} x_0^m(y)$. Then if $\underline{x} > 0$ we have, by continuity of $x_0^1(\cdot)$,

$$\underline{x} > x_0^1(\underline{x}) = x_0^1\left(\lim_{m\to\infty} x_0^m(y)\right) = \lim_{m\to\infty} x_0^1(x_0^m(y)) = \underline{x},$$

a contradiction. The second statement in the lemma follows from the fact that $x_1^1(0) > 0$ (by A2(i)), so by continuity there exists $\hat{y} > 0$ such that $x_1^1(y) > y$, $\forall y \leqslant \hat{y}$. Thus $\{x_1^m(y)\}$ could never have a convergent subsequence approaching a value less than $\hat{y}$. *Q.E.D.*

We can now complete the argument for Condition 1. By Lemma 2, $P^m(y,A) > 0$ if $\mu(A \cap (x_0^m(y), x_1^m(y))) > 0$. By Lemma 3 there is an interval $(0, \hat{y})$ such that $[A \subset (0,\hat{y})] \Rightarrow [A \subset (x_0^m(y), x_1^m(y))]$ with y given and m sufficiently large. Thus for any point $y_0 \in (0, \hat{y})$ we have shown that every neighborhood of y_0 has positive probability of being entered after a finite number of steps from any original point $y \in [0,\infty)$. Moreover we have shown that $P^m(y,A) > 0 \Rightarrow p^{mk}(y, A) > 0$ for $k \geqslant 1$ and $A \subset (0, \hat{y})$. Condition 1 then follows.

It remains to argue that the support of the invariant measure is connected and compact. But this follows from A2(iii) and A4 which imply the existence of a smallest point $\tilde{\tilde{y}} \in [0,\infty)$ with the property that $x_1^1(y) < y$, $\forall y > \tilde{\tilde{y}}$. Thus, when

$y > \tilde{\tilde{y}}$, the supremum of the support of $P^m(y, \cdot)$ monotonically decreases until it falls below $\tilde{\tilde{y}}$. Hence the invariant measure has support contained in the interval $[0, \tilde{\tilde{y}}]$. It is clear that the supremum of the support of the invariant measure $\tilde{y}$ satisfies $\tilde{y} = x_1^1(\tilde{y})$. Moreover $\lim_{m\to\infty} x_0^m(y) = 0\ \forall y$ implies that if y_1 is in the support of the invariant measure and $y_2 < y_1$, then y_2 is in the support. Hence the support is connected. *Q.E.D.*

PROOF OF THEOREM 3: For any $t \in \mathfrak{L}$, the map $T_t : C[0, \bar{y}] \to C[0, \bar{y}]$ is defined by

$$T_t\Phi(y) \equiv \max_{0 \leqslant c \leqslant y} E_\alpha U(c, \Phi(t(h(\alpha, y - c), y - c))), y \in [0, \bar{y}].$$

Inductively define $T_t^n\phi \equiv T_t(T_t^{n-1}\Phi)$, $n = 2, 3, \ldots$. By Theorem 1 we know there is a unique $V_t^* \in C[0, \bar{y}]$ such that $\lim_{n\to\infty} |(T_t^n\Phi)(y) - V_t^*(y)| = 0$, $\forall y \in [0, \bar{y}]$. Now let $\Phi(y) = U(y, 0)$ and suppose $\hat{t} \succcurlyeq t$. Then we have

$$\begin{aligned}(T_t\Phi)(y) &= \max_{0 \leqslant c \leqslant y} E_\alpha U(c, U(t(h(\alpha, y - c), y - c), 0))\\ &\leqslant \max_{0 \leqslant c \leqslant y} E_\alpha U(c, U(\hat{t}(h(\alpha, y - c), y - c), 0))\\ &= (T_{\hat{t}}\Phi)(y)\end{aligned}$$

by the facts that the composition of increasing concave functions is concave, and that $\hat{t} \succcurlyeq t$ implies the expectation of any concave function of after tax income is no less under $\hat{t}$ than under t. Suppose, inductively, that $(T_t^{n-1}\Phi)(y) \leqslant (T_{\hat{t}}^{n-1}\Phi)(y)$, $\forall y \in [0, \bar{y}]$. Then

$$\begin{aligned}(T_t^n\Phi)(y) &= T_t(T_t^{n-1}\Phi)(y)\\ &= \max_{0 \leqslant c \leqslant y} E_\alpha U(c, (T_t^{n-1}\Phi)(t(h(\alpha, y - c), y - c)))\\ &\leqslant \max_{0 \leqslant c \leqslant y} E_\alpha U(c, (T_{\hat{t}}^{n-1}\Phi)(t(h(\alpha, y - c), y - c)))\\ &\leqslant \max_{0 \leqslant c \leqslant y} E_\alpha U(c, (T_{\hat{t}}^{n-1}\Phi)(\hat{t}(h(\alpha, y - c), y - c)))\\ &= T_{\hat{t}}(T_{\hat{t}}^{n-1}\Phi)(y) = (T_{\hat{t}}^n\Phi)(y), \forall y \in [0, \bar{y}],\end{aligned}$$

where the first inequality follows from the induction hypothesis and monotonicity of U, while the second inequality is established by the argument employed just above. Hence we have shown that $(T_t^n\Phi)(y) \leqslant (T_{\hat{t}}^n\Phi)(y)$ $\forall y \geqslant 0$, $\forall n = 1, 2, \ldots$. It then follows from pointwise convergence that $V_t^*(y) \leqslant V_{\hat{t}}^*(y)$, $\forall y \in [0, \bar{y}]$. *Q.E.D.*

PROOF OF THEOREM 4: Let $\bar{x}$ be the average income in the *laissez faire* equilibrium. Denote by x the earnings of an arbitrary mature individual, and by

y the earnings of his parent in the preceding period. In what follows we will think of (x, y) as a jointly random vector. Now it is a well known fact of bivariate distribution theory that

$$\text{(14)} \qquad \operatorname{var}(x) = E(\operatorname{var}(x \mid y)) + \operatorname{var}(E(x \mid y)).$$

Moreover, the variance of income under public education is given by

$$\operatorname{var}_{P.E.} = \operatorname{var}(x \mid y = \bar{x}).$$

The proof proceeds by noting that A5 implies $\operatorname{var}(x \mid y)$ is a convex function of y. One may then apply Jensen's inequality which, along with (14) implies the result.

Consider

$$\begin{aligned}
\frac{\partial}{\partial y}\operatorname{var}(x \mid y) &= \frac{\partial}{\partial y}\int_{[0,1]}\left[h(\alpha, e^*(y)) - \int_{[0,1]} h(\alpha, e^*(y)) f(\alpha)\mu(d\alpha)\right]^2 \\
&\qquad \times f(\alpha)\,\mu(d\alpha) \\
&= 2\int_{[0,1]} h(\alpha, e^*(y))\left[\frac{\partial}{\partial y} h(\alpha, e^*(y))\right] f(a)\mu(d\alpha) \\
&\qquad - 2\left[\frac{\partial}{\partial y}\int_{[0,1]} f(\alpha) h(\alpha, e^*(y))\mu(d\alpha)\right] \\
&\qquad \times \int_{[0,1]} h(\alpha, e^*(y) f(\alpha)\mu(d\alpha) \\
&= 2\operatorname{cov}_\alpha\left(h(\alpha, e^*(y)), \frac{\partial h}{\partial y}(\alpha, e^*(y))\right) \gtrless 0 \qquad \text{as} \qquad \frac{\partial^2 h}{\partial \alpha\, \partial e} \gtrless 0.
\end{aligned}$$

Thus, the conditional variance of offspring's earnings increases (decreases) with parental income iff ability and education are complements (substitutes). Furthermore, it follows from (14) that $\operatorname{var}(x) > \operatorname{var}(x \mid y = \bar{x})$ when $\partial^2 h/\partial\alpha\,\partial e = 0$. Now one may calculate

$$\begin{aligned}
\frac{\partial^2}{\partial y^2}\operatorname{var}(x \mid y) &= 2\left[\operatorname{var}_\alpha\left(\frac{\partial}{\partial y} h(\alpha, e^*(y))\right)\right. \\
&\qquad \left. + \operatorname{cov}_\alpha\left(h(\alpha, e^*(y)), \frac{\partial^2 h}{\partial y^2}(\alpha, e^*(y))\right)\right].
\end{aligned}$$

This term is nonnegative under A5. It follows that $\operatorname{var}(x) \geqslant \operatorname{var}_{P.E.}$. Moreover, the concavity of $h(\alpha, e^*(y))$ in y also implies $\bar{x} \leqslant \int_{[0,1]} f(\alpha) h(\alpha, e^*(\bar{x}))\mu(d\alpha)$, so mean income is no lower under public education. *Q.E.D.*

PROOF OF THEOREM 5: Fubini's Theorem implies that in equilibrium

$$\begin{aligned}\operatorname{cov}(\tilde{\alpha},\tilde{x}) &\equiv E(\tilde{\alpha}\tilde{x}) - (E\tilde{\alpha})(E\tilde{x}) \\ &= \int_{[0,\infty)}\left(\int_{[0,1]}\alpha f(\alpha)h(\alpha,e^*(y))\mu(d\alpha)\right)\nu^*(dy) \\ &\quad - \left(\int_{[0,1]}\alpha f(\alpha)\mu(d\alpha)\right)\left(\int_{[0,\infty)}y\nu^*(dy)\right) \\ &= \int_{[0,1]}\alpha f(\alpha)\left[\int_{[0,\infty)}(h(\alpha,e^*(y)) - y)\nu^*(dy)\right]\mu(d\alpha).\end{aligned}$$

Now ν^* invariant implies

$$\int_{[0,1]}f(\alpha)\left[\int_{[0,\infty)}(h(\alpha,e^*(y)) - y)\nu^*(dy)\right]\mu(d\alpha) = 0,$$

which simply states that per capita earnings is constant across a generation. By A2(i) the bracketed term in the integrand above is strictly increasing in α. Thus, $\exists\bar{\alpha} \in (0,1)$ such that

$$f(\alpha)\int_{[0,\infty)}(h(\alpha,e^*(y)) - y)\nu^*(dy) \gtrless 0 \qquad \text{as} \quad \alpha \gtrless \bar{\alpha}.$$

Hence

$$\operatorname{cov}(\tilde{\alpha},\tilde{x}) > \bar{\alpha}\int_{[0,1]}f(\alpha)\left[\int_{[0,\infty)}(h(\alpha,e^*(y)) - y)\nu^*(dy)\right]\mu(d\alpha) = 0.$$

We now outline an argument showing that the economy needn't be strongly meritocratic. Consider the following:

$$\operatorname{pr}[\tilde{\alpha} \geqslant a \mid \tilde{x} = x] = \operatorname{pr}[\tilde{y} \leqslant y^{-1}(a,x) \mid \tilde{x} = x]$$

where $(\tilde{x},\tilde{y})$ is the random vector of offspring-parent earnings and $y^{-1}(a,x)$ is defined by $h(a,e^*(y^{-1}(a,x))) \equiv x$. That is, an individual observed to have earnings x will have an innate endowment exceeding a if and only if that individual's parent had an income less than $y^{-1}(a,x)$. Let $p^1(y,x)$ be as defined in the proof of Lemma 2 under the proof of Theorem 2 given above. Then, we have that

$$\operatorname{pr}[\tilde{y} \leqslant y^{-1}(a,x) \mid \tilde{x} = x] = \left[\int_{[0,\infty)}p^1(y,x)\nu^*(dy)\right]^{-1} \times \int_{[0,\,y^{-1}(a,x)]}p^1(y,x)\nu^*(dy).$$

Thus the equilibrium distribution is meritocratic in the strong sense if and only if the right-hand side above is an increasing function of x. Notice that this restriction is in principle testable when the distribution of parental income conditional on offspring's income can be observed.

An increase in x has two effects: First, it leads to an increase in $y^{-1}(a, x)$, the critical level below which parental income must lie if $\tilde{\alpha}$ is to exceed a. However, an increase in x also shifts to the right the conditional distribution of parent's earnings $(\tilde{y} \mid x)$, tending to reduce the right-hand side above. That this latter effect can outweigh the former may be seen by the following (admittedly extreme) example. Suppose there are only two possible levels of training investment which parents can make, $\underline{e}$ and $\bar{e}$, with $\underline{e} < \bar{e}$. Imagine further that $\underline{x} \equiv h(\underline{e}, 1) < h(\bar{e}, 0) \equiv \bar{x}$. Now imagine observing two individuals, one of whom earns $\bar{x} + \epsilon$, the other of whom earns $\underline{x} - \epsilon$ where ϵ is a "small" positive number. Then it must be the case (by continuity of $h(\cdot, \cdot)$) that the person with earnings $\bar{x} + \epsilon$ has ability close to zero, while the person who earns $\underline{x} - \epsilon$ has ability close to one. *Q.E.D.*

7. CONCLUSIONS

The analysis presented here brings together two separate strands of the literature on economic inequality. On the one hand we have modelled the income dynamics as a stochastic process, and studied the properties of its ergodic distribution. This approach has a long history (for example, [**6**, **7**, **19**, **25**, **28**]). On the other hand, the underlying structure of this stochastic process results from maximizing decisions of the economic agents. As such, it has much in common with the traditional human capital approach to earnings [**2**, **3**, **20**, **21**]. Yet we have departed from the human capital school in one crucial respect—the assumption that a perfect market for educational loans exists. We have assumed instead a "balkanized" market, where each family must generate internally funds for the training of its young. We believe this assumption parallels more closely actual circumstances. As a consequence of this assumption family background exercises an independent, constraining effect on social mobility. Our analysis is in this regard consistent with the approach to the study of mobility which economists and sociologists have found so fruitful in recent years (for example, [**4**, **8**, **10**, **15**]).

The present approach commends itself in one important additional respect. By grounding training decisions in rational choice, we are able to rigorously analyze some normative issues regarding public policies which would alter the distribution of income. Moreover, this analysis [Section 4] has not presupposed the existence of some social welfare function. Rather, it has proceeded on a purely individualistic basis. We have seen that redistributive policies can serve to improve both the allocation of risks and of training resources when capital markets are incomplete. Indeed, normative prescriptions similar in spirit to those of Harsanyi [**14**] or Rawls [**23**], to the extent that they are derived from an analysis of the risks which an individual faces in some ex ante "original position," are here given a positive, behavioral grounding. We have replaced the "veil of ignorance" with the "generation gap," across which parents cannot look to perceive the position which their offspring will occupy in society.

Finally, we should note some shortcomings of the foregoing analysis. By far the strongest assumption we have made is that innate abilities are independent across generations. While the exact nature of intergenerational correlations has not yet been satisfactorily ascertained[13] the existence of some such effects is impossible to deny. Theorems 1 and 2 survive, with slight modification, generalization to a Markovian structure on abilities. However, the tax analysis of Section 4 is lost once such allowance is made. This seems a stimulating topic for further investigation.

Another weak point is the fact that the welfare implications of income taxation studied in Section 4 are limited to education-specific tax systems. These may not be politically or administratively feasible, and do not involve redistribution across social classes. In any case, a similar analysis of tax systems dependent only upon income would be desirable. This appears to be a very difficult matter in general. Moreover, a strong result like that given in Theorem 3 will no longer hold for such schemes.

Finally, the framework developed here suggests an interesting research problem in the normative economics of social mobility. The transition probability for a society $P(y,A)$ contains all relevant information regarding both cross-section inequality in the long run (if the process is ergodic), as well as intergenerational mobility. One might consider axiom systems to impose on orderings of such transition functions which reflect intuitive notions of what it means to say that one social structure is "more equal" than another. The question then becomes "when can orderings satisfying these intuitive axioms be represented in the simple way suggested in equation (8)?" (That is, when may social structures "P" be ranked by the expected value of some indirect utility function under their ergodic distribution?)

University of Michigan

Manuscript received July, 1978; *revision received June*, 1980.

[13] See Goldberger [13] for an informative discussion of the problems here.

REFERENCES

[1] ATKINSON, A.: "On the Measurement of Inequality," *Journal of Economic Theory*, 2 (1970), 244–263.

[2] BECKER, G.: *Human Capital*, 2nd Edition. New York: National Bureau of Economic Research, 1975.

[3] BECKER, G., AND N. TOMES: "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87 (1979), 1153–1189.

[4] BOUDON, R.: *Mathematical Structures of Social Mobility*, Amsterdam: Elsevier Scientific Publishing Co., 1973.

[5] BROCK, W., AND L. MIRMAN: "Optimal Economic Growth and Uncertainty: The Discounted Case," *Journal of Economic Theory*, 4 (1972), 479–513.

[6] CHAMPERNOWNE, D. G.: "A Model of Income Distribution," *Economic Journal*, 63 (1953), 318–351.

[7] CONLISK, J.: "Can Equalization of Opportunity Reduce Social Mobility?" *American Economic Review*, 64 (1974), 80–90.

[8] DATCHER, L.: "The Effect of Family and Community Background on the Education and Earnings of Black and White Men," Unpublished Manuscript, Institute for Social Research, University of Michigan, 1980.

[9] DIAMOND, P.: "National Debt in a Neoclassical Growth Model," *American Economic Review*, 55 (1965), 1126–1150.

[10] DUNCAN, O., D. FEATHERMAN, AND B. DUNCAN: *Socioeconomic Background and Achievement*. New York: Seminar Press, 1972.

[11] FRIEDMAN, A.: *Foundations of Modern Analysis*. New York: Holt, Rinehart and Winston Inc., 1970.

[12] FUTIA, C., "A Stochastic Approach to Economic Dynamics," Bell Telephone Laboratories Discussion Paper, 1976.

[13] GOLDBERGER, A. S.: "Models and Methods in the I.Q. Debate: Part I," SSRI Workshop Series No. 7710, University of Wisconsin, 1977.

[14] HARSANYI, J. C.: "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, 63 (1955), 309–321.

[15] JENCKS, C. ET AL.: *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Harper and Row, 1972.

[16] KOHLBERG, E.: "A Model of Economic Growth with Altruism between Generations," *Journal of Economic Theory*, 13 (1976), 1–13.

[17] KOOPMANS, T.: "Representation of Preference Orderings over Time," in *Decisions and Organization*, ed. by C. B. McGuire and R. Radner. Amsterdam: North-Holland Publishing Co., 1972.

[18] LOURY, G. C.: "Essays in the Theory of the Distribution of Income," Ph.D. Thesis, Massachusetts Institute of Technology, 1976.

[19] MANDELBROT, B.: "Stable Paretian Random Functions and the Multiplicative Variation of Income," *Econometrica*, 29 (1961), 517–543.

[20] MINCER, J.: "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy*, 66 (1958), 281–302.

[21] ———: "The Distribution of Labor Incomes: A Survey," *Journal of Economic Literature*, 8 (1970), 1–26.

[22] OKUN, A.: *Equity and Efficiency: The Big Tradeoff*. Washington, D. C.: The Brookings Institution, 1975.

[23] RAWLS, J.: *A Theory of Justice*. Cambridge, Massachusetts: Harvard University Press, 1971.

[24] ROTHSCHILD, M., AND J. STIGLITZ: "Increasing Risk I: A Definition," *Journal of Economic Theory*, 2 (1970), 225–243.

[25] ROY, A. D.: "The Distribution of Earnings and Individual Output," *Economic Journal*, 60 (1950), 489–505.

[26] SAMUELSON, P.: "An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money," *Journal of Political Economy*, 66 (1958), 467–482.

[27] SHORROCKS, A.: "The Measurement of Mobility," *Econometrica*, 46 (1978), 1013–1024.

[28] SOLOW, R.: "On the Dynamics of the Income Distribution," Ph.D. Thesis, Harvard University, 1951.

[29] VARIAN, H.: "Redistributive Taxation as Social Insurance," *Journal of Public Economics*, 14 (1980), 49–68.

[30] ZECKHAUSER, R.: "Risk Spreading and Distribution," Chapter 8 in *Redistribution Through Public Choice*, ed. by Hochman and Peterson. New York: Columbia University Press, 1974.