THE
QUARTERLY JOURNAL
OF ECONOMICS

Vol. LXXXV November 1971 No. 4

THE OPTIMAL DISTRIBUTION OF INCOME

RAY C. FAIR

I. INTRODUCTION

At the end of their debate with R. H. Strotz, F. M. Fisher and J. Rothenberg make a plea for going beyond the Pareto optimality criterion in analyzing issues in welfare economics.[1] The present study is in the spirit of this plea: an attempt is made in this paper to derive bounds on the optimal distribution of income under a particular set of value judgments. Because of a general reluctance of economists to make value judgments and to engage in interpersonal comparisons of utility, there has actually been little work done on the question of the optimal distribution of income. Most of the work that has been done in that area has been concerned with analyzing the *determinants* of income distribution.[2] Karl Marx, of course, felt that income ought to be distributed according to need,[3] but even here it is not clear without knowing how needs are distributed what kind of income distribution this tenet implies. Plato is perhaps one of the few who have been explicit on the question of the optimal distribution of income. He felt that no one in

1. Franklin M. Fisher and Jerome Rothenberg, "How Income Ought to be Distributed: Paradox Enow," *Journal of Political Economy*, LXXX (Feb. 1962), 93. See also Robert H. Strotz, "How Income Ought to be Distributed: A Paradox in Distributive Ethics," *Journal of Political Economy*, LXVI (June 1958), 189–205, and "How Income Ought to be Distributed: Paradox Regained," *ibid.*, LXIX (April 1961), 271–78, and Fisher and Rothenberg, "How Income Ought to be Distributed: Paradox Lost," *ibid.*, LXIX (April 1961), 162–80.

2. For an early review of previous studies of the determinants of income distribution, see Hans Staehle, "Ability, Wages, and Income," *The Review of Economics and Statistics*, XXV (Feb. 1943), 77–87.

3. ". . . only then can the narrow horizon of bourgeois right be fully left behind and society inscribe on its banners: from each according to his ability, to each according to his needs!" (Robert Freedman, ed., *Marx on Economics*; New York: Harcourt, Brace and World, 1961, p. 277.)

a society should be more than four times richer than the poorest member of society, for "in a society which is to be immune from the most fatal of disorders which might more properly be called distraction than faction, there must be no place for penury in any section of the population, nor yet for opulence, as both breed either consequence."[4]

In Section II of this paper a general optimization model is developed, and then in Section III alternative sets of assumptions about the key functions and parameters in the model are formulated. The approach followed in Section III is to make at least two of the sets of assumptions extreme enough in both directions so that the real world situation appears likely to be bounded by them. In Section IV the results of solving the model under the alternative sets of assumptions are presented and analyzed. The sensitivity of the optimal distribution of income to the key parameters of the model is examined, and the actual distribution of income in the United States is compared with the computed optimal distributions. The paper concludes in Section V with a discussion of the limitations of the present analysis and of various extensions that might be made.

II. The General Model

The Individual Utility Functions

This study is concerned with the *lifetime* distribution of income of members of society. Let Y_i denote the after-tax lifetime income of individual i, H_i the number of hours worked by individual i during his lifetime, and T_i the total number of possible working hours in his lifetime. The first assumption of the model is that each individual's lifetime utility, U_i, is a function of his income, Y_i, and leisure, $T_i - H_i$:

(1) $$U_i = f_i(Y_i, T_i - H_i), \qquad i = 1, 2, \ldots, n,$$

where n is the number of individuals under consideration and where the i subscript on f denotes the fact that f may vary from individual to individual. Note that equation (1) ignores the problem of time discounting and implicitly assumes that lifetime utility is independent of the time distribution of income and leisure.

The Individual Earnings Functions

It is next assumed that everyone in society receives at least the equivalent of a high school education — either an academic or

4. Plato, *The Laws*, trans. A. E. Taylor (New York: E. P. Dutton and Co., 1960), p. 127.

vocational education. After high school an individual has the choice of furthering his education or beginning work immediately. E_i will be used to denote the number of hours individual i spends in educating himself beyond high school. The number of hours spent in education beyond high school will be assumed to be hours spent in work.[5]

Each individual is assumed to be born with a certain *innate* ability, denoted as A_i. Innate ability is meant to refer to innate "productivity" and not merely to innate IQ intelligence. Innate ability is likely to be a function of such things as innate physical strength and stamina and innate organizational and administrative ability as well as of innate IQ intelligence.

Let YE_i denote the lifetime, before-tax (earned) income of individual i. Then $YE_i/(H_i-E_i)$ is individual i's average (over his lifetime) hourly productivity. H_i-E_i is the number of hours individual i actually works earning income as opposed to educating himself beyond high school, the latter also being counted as working. The average hourly productivity of individual i is assumed to be a function of his innate ability and of his education beyond high school:

$$(2) \qquad YE_i/(H_i-E_i)=g(A_i,\ E_i), \qquad i=1,\ 2,\ \ldots,\ n.$$

Equation (2) will be referred to as the individual's "earnings function." The functional form of g will be discussed in Section III. Since a given level of education is likely to have more effect on actual productivity for an individual of high innate ability than for one of low innate ability, the specification of g must be made with some care.

The Individual Tax or Distribution Functions

After-tax income, Y_i, is assumed to be related to before-tax income by the following tax equation:

$$(3) \qquad Y_i=YG+YE_i-\int_0^{YE_i} t(YE_i)\,dYE_i, \qquad i=1,\ 2,\ \ldots,\ n.$$

YG in (3) is the minimum guaranteed level of lifetime income for an individual: if an individual earns nothing during his life, he will still receive YG amount of income. $t(YE_i)$ in (3) is the marginal tax rate function, the marginal tax rate being assumed to be a func-

5. Conceivably, E_i should be treated partly as a consumption good as well as an investment good, which would mean that E_i should also enter the utility function (1) in a positive way. It was felt, however, that this addition would unduly complicate the model with little change in the final conclusions and results, and so E_i was kept merely as an investment good.

tion of the level of earned income. The integral in (3) specifies the amount an individual with YE_i income will pay in taxes over his life. Equation (3) can be considered to be a general distribution equation in that it is both a "positive" and "negative" income tax equation. Depending on YE_i, YG, and $t(YE_i)$, $Y_i - YE_i$ can be either negative or positive, and thus individuals may receive "negative" income tax payments under (3).

The Individual Behavioral Assumption

The basic behavioral assumption of the model is as follows. Individuals are assumed to take YG and $t(YE_i)$ as given and to choose E_i and H_i so as to maximize their lifetime utility. Choosing E_i and H_i leads through the earnings function (2) to YE_i, which leads through the distribution equation (3) to Y_i, which then leads through the utility function (1) to U_i. U_i is thus seen to be a function of the endogenous variables E_i and H_i and (from the point of view of the individual) the exogenous variables YG, $t(YE_i)$, and A_i.

The Social Welfare Function and the Government Behavioral Assumption

Social welfare, SW, is taken to be a function of the individual lifetime utilities:

(4) $$SW = h(U_1, U_2, \ldots, U_n).$$

The particular form of h used in this study will be discussed in Section III.

Given the form of h, the government is assumed to choose the minimum guaranteed income level, YG, and the marginal tax rate function, $t(YE_i)$, such that SW is at a maximum, subject to the constraint that each individual maximizes his own utility and subject to the lifetime budget constraint:

(5) $$\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} YE_i.$$

The budget constraint (5) says that the sum of after-tax income is equal to the sum of before-tax or earned income.[6]

Solution of the Model

Aside from specifying the distribution of A_i and the form of the various functions, both of which will be done in Section III, the

6. The budget constraint will be made slightly more complicated in Section III to incorporate the case of public goods.

model is complete. Individuals take YG and $t(YE_i)$ as given and choose the values of E_i and H_i that maximize their utility functions. The government chooses YG and $t(YE_i)$ that maximize SW, subject to the budget constraint and subject to the constraint that each individual maximizes his own utility. The individual maximization equations are

$$\partial U_i/\partial H_i = 0, \qquad i = 1, 2, \ldots, n, \tag{6}$$

$$\partial U_i/\partial E_i = 0, \qquad i = 1, 2, \ldots, n, \tag{7}$$

and the overall constrained maximization problem can be set up as the problem of maximizing the following Lagrangian:

$$\begin{aligned} L = {} & h(U_1, U_2, \ldots, U_n) \\ & - \sum_{i=1}^{n} \psi_i \left(\frac{\partial U_i}{\partial H_i} - 0 \right) - \sum_{i=1}^{n} \phi_i \left(\frac{\partial U_i}{\partial E_i} - 0 \right) \\ & - \lambda \left(\sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} YE_i \right). \end{aligned} \tag{8}$$

The problem in (8) is to maximize h subject to the $2n+1$ constraints. The ψ_i, ϕ_i, and λ variables in (8) are the $2n+1$ Lagrangian multipliers. Theoretically, then, the model can be solved by differentiating L with respect to the $4n+3$ unknowns, H_i, E_i, ψ_i, ϕ_i, $t(YE_i)$, YG, and λ, and solving the resulting set of $4n+3$ equations for the $4n+3$ unknowns. In practice, the model had to be solved by solving a subset of the $4n+3$ equations and scanning for the remaining values. The process by which the model was solved is discussed in the Appendix.

III. Specification of the Model

Both questions of value and questions of fact are involved in specifying the above model. Questions of fact are involved in specifying the form of the individual utility functions, the distribution of innate ability, and the form of the earnings functions; and questions of value are involved in specifying the form of the social welfare function. Value judgments are also involved in the assumption that people have the freedom to maximize their utility subject only to the tax constraint and in the assumption that everyone receives (at public expense) the equivalent of a high school education.

The Individual Utility Functions

The f_i functions in (1) are assumed to be Cobb-Douglas functions in income and leisure, and individuals are assumed to differ only in the weight they attach to income and leisure in their functions. In particular, it is assumed that

$$(9) \qquad U_i = BY_i^{\delta_i}(T_i - H_i)^{1-\delta_i}, \qquad i = 1, 2, \ldots, n,$$

where the parameter δ_i can differ from individual to individual. The constant B in equation (9) is assumed to be the same for all individuals. Individuals with the same δ_i parameter and the same income and leisure, for example, are assumed to have the same utility. Individuals, in other words, are assumed to have (aside from possibly differing δ_i coefficients) the same capacity for absorbing utility; some are not assumed to be more absorptive (better?) than others.

The specification of the δ_i parameters is of considerable importance in the model. δ_i is the measure of how much individual i values income relative to leisure in his utility function. To get an idea of what δ_i might be, assume that there is no tax function (so that $Y_i = YE_i$) and that income is just a simple function of the number of hours worked: $Y_i = w_i H_i$. w_i is the wage rate for individual i. Then Y_i can be eliminated from the utility function to yield $U_i = w_i^{\delta_i} H_i^{\delta_i} (T_i - H_i)^{1-\delta_i}$. On the assumption that the individual maximizes U_i with respect to H_i (w_i and T_i being taken to be exogenous), U_i can be differentiated with respect to H_i, set equal to zero, and solved for H_i. This yields $H_i = \delta_i T_i$, or $\delta_i = H_i/T_i$. δ_i in this simple example is thus the percentage of the total possible working time individual i actually works.

In Table I three values of H_i/T_i are presented for three different

TABLE I
VALUES OF H_i UNDER VARIOUS ASSUMPTIONS

	Assumption about work effect	H_i	H_i/T_i[a]
Below average:	4 hours a day, 20 days a month, 12 months a year, 43 years	41280	0.133
Average:	8 hours a day, 20 days a month, 12 months a year, 48 years	92160	0.297
Above average:	10 hours a day, 25 days a month, 12 months a year, 53 years	159000	0.513

a. T_i = (53 years) (365.25 days) (16 hours) = 309892 hours.

assumptions about work effort.[7] If an individual works the equiv-

7. It is assumed in this study that T_i, the maximum working lifetime of each individual, extends from the age of 18 to the age of 70. It is also assumed that each individual needs 8 hours of sleep a day. This then leaves (53 years) (365.25 days) (16 hours) = 309,892 possible working hours in each individual's lifetime.

alent of 8 hours a day, 20 days a month, 12 months a year, for 48 years (say, ages 18–65), then H_i/T_i is equal to 0.297. If, on the other hand, an individual is less industrious and works only the equivalent of 4 hours a day, 20 days a month, 12 months a year, for 43 years (say, ages 18–60), then H_i/T_i is equal to 0.133. Finally, if an individual is quite industrious and works the equivalent of 10 hours a day, 25 days a month, 12 months a year, for 53 years (say, ages 18–70), then H_i/T_i is equal to 0.513. Under the above simple model, therefore, δ_i would seem to range from about 0.1 for very unindustrious individuals to about 0.5 for highly industrious ones, with the average appearing to be about 0.3. The simple no-tax model is, of course, quite crude, but the above analysis is designed merely to give an indication as to the possible range of values of the δ_i coefficients.

In this study three different assumptions about the δ_i coefficients have been used. The first assumption is that δ_i is equal to 0.3 for all individuals. The second assumption is that for the first one-fifth of the population δ_i is equal to 0.2, for the second fifth to 0.25, for the third fifth to 0.3, for the fourth fifth to 0.375, and for the last fifth to 0.45. The third assumption is similar to the second, with the respective values of δ_i being 0.15, 0.225, 0.3, 0.4, and 0.5. There is thus no variation in individuals' degrees of industriousness under the first assumption, a moderate to large amount of variation under the second assumption, and a rather large amount of variation under the third assumption.

The Distribution of Innate Ability

From the time of Galton on, the assumption has been commonly made that innate characteristics of human beings are normally distributed. Indeed, Hans Staehle has stated that "Since Galton published *Hereditary Genius*, the assumption that individuals are dissimilar as to their 'natural ability,' or 'general aptitudes,' and that their distribution according to these general aptitudes is essentially normal, has never been seriously contested."[8] That assumption will not be contested in this study either, and innate ability, A_i, will be assumed to be normally distributed. Remember, however, that ability is not meant to refer merely to IQ intelligence, but also to such things as physical strength and stamina and organizational ability. The assumption that A_i is normally distributed is thus probably stronger than, say, the assumption that innate intelligence or physical strength is normally distributed.

8. *Op. cit.*, p. 77.

The mean of A_i is arbitrarily taken to be 100. With respect to the standard deviation of A_i, information can perhaps be gleaned from observing the standard deviations that have been measured in various psychological tests. From a perusal of a text on psychological testing by F. S. Freeman,[9] the measured standard deviations (with the mean taken to be 100) seemed to range between about 12 and 20, with a median of about 16. Argument by analogy would thus indicate that the standard deviation of A_i should be about 16, although there is no compelling reason for feeling that innate intelligence (or whatever it is that psychologists test) and innate ability are identically distributed.[1] In line with the methodology of the study, three different assumptions about the standard deviation of A_i have been used: the standard deviation has been assumed to be either 8, 16, or 32.

The Individual Earnings Functions

It is assumed in the general model above that everyone receives the equivalent of a high school education, whether it be an academic or vocational type of education. The quality of education is assumed to be uniform, and education through high school is conceived of as being necessary to keep the distribution of A_i the same at the time of high school graduation as it was at birth. Individuals, in other words, are assumed to be born with a certain *potential* innate ability, where it takes the equivalent of a high school education to make this potential a reality. Education before high school graduation is also taken to include family upbringing, and if for some reason an individual receives an inadequate education or upbringing, then his actual ability at the time of high school graduation will be less than his potential ability was at birth.

Let z_{i0} denote the actual productivity (output per hour) of individual i at the time of high school graduation (E_i equal to zero). Then the above assumption is that z_{i0} and A_i are identically distributed. This is the first step in specifying the earnings function (2).

The next question is how education beyond high school affects productivity. It is quite likely that the degree to which a person's productivity is affected by his education beyond high school is a function of his innate ability. If this possibility is ignored for the

9. Frank S. Freeman, *Psychological Testing* (New York: Henry Holt and Company, 1955).

1. There is also the danger that psychological tests are designed, either explicitly or implicitly, to result in a standard deviation of about 16. If this practice is widespread, then little information about the standard deviation of A_i can be gleaned from the results of psychological tests.

moment, however, some information on the effect of college education on productivity can perhaps be gathered from the studies of G. S. Becker, Z. Griliches, G. Hanoch, H. P. Miller, and T. P. Schultz.[2] In these studies the mean incomes of high school and college graduates were computed for various groups and years. Some of the results of these studies are summarized in Table II. The

TABLE II

RATIO OF THE MEAN INCOME OF COLLEGE GRADUATES TO THE MEAN INCOME OF HIGH SCHOOL GRADUATES COMPUTED FROM VARIOUS STUDIES

A. Becker: Urban, native, white males						
Year/Age	23–24	25–29	30–34	35–44	45–54	55–65
1939	1.04	1.29	1.47	1.56	1.59	1.53
1949	0.84	1.08	1.42	1.86	2.00	1.85
B. Griliches: Males						
All Ages/Year	1939	1949	1958	1959	1963	1966
	1.57	1.63	1.65	1.51	1.45	1.52
C. Hanoch: White, northern males						
Year/Age	27	37	47	57	67	77
1959	1.26	1.44	1.61	1.61	1.53	1.62
D. Miller: White, northern males						
Year/Age	18–24	25–34	35–44	45–54	55–65	25–64
1959	1.21	1.19	1.61	1.83	1.86	1.57
E. Schultz: Males						
Year/Age	25–29	30–34	35–44	45–54	55–64	65–74
1959	1.12	1.34	1.50	1.62	1.62	1.90

Sources: *A*. Becker, *Human Capital*, *op. cit.*, Table 1, p. 71 and Table 3, p. 77. *B*. Griliches, "Notes on the Role of Education," *op. cit.*, Table 6, p. 23. *C*. Hanoch, "An Economic Analysis of Earnings and Schooling," *op. cit.*, Table 2, p. 316. *D*. Miller, *Income Distribution in the United States*, *op. cit.*, Table IV–3, p. 139. *E*. Schultz, *Statistics on the Size Distribution of Personal Income*, *op. cit.*, Table 1, p. 43.

ratio of the mean income of college graduates to the mean income of high school graduates is presented in Table II for various age groups and years from each of the studies. The results presented in Table II are by no means meant to be a complete summary of the work in the field, but are meant to be used only to give a rough indication of how education affects earnings or productivity.

2. Gary S. Becker, *Human Capital* (New York: National Bureau of Economic Research, 1964). Zvi Griliches, "Notes on the Role of Education in Production Functions and Growth Accounting," paper presented to the Conference on Research in Income and Wealth, Nov. 15–16, 1968. Giora Hanoch, "An Economic Analysis of Earnings and Schooling," *The Journal of Human Resources*, II (Summer 1967), 310–29. Herman P. Miller, *Income Distribution in the United States*, a 1960 Census monograph (Washington, D.C.: U.S. Government Printing Office, 1966). T. Paul Schultz, *Statistics on the Size Distribution of Personal Income in the United States*, prepared for the Joint Economic Committee of the U.S. Congress (Washington, D.C.: U.S. Government Printing Office, 1965).

With the figures in Table II as a guide, the assumption was made in this study that a college education (or the equivalent in advanced vocational training) increases the average lifetime hourly productivity of an individual of average ability (A_i equal to 100) by 30 percent. This corresponds to a productivity ratio of college to high school graduates of 1.30 for those of average ability. The 1.30 figure is smaller than most of those in Table II because of the feeling that the figures in Table II partly reflect differences in innate ability.

There is, unfortunately, even less information available on the effect of a graduate school education on productivity. The information that is available from the studies of Hanoch and Miller is summarized in Table III. The ratio of the mean income of those with

TABLE III

RATIO OF THE MEAN INCOME OF THOSE WITH MORE THAN A FOUR-YEAR COLLEGE EDUCATION TO THE MEAN INCOME OF THOSE WITH ONLY A COLLEGE EDUCATION

A. Hanoch: White, northern males						
Year/Age		37	47	57	67	77
1959		1.11	1.20	1.18	1.34	1.59
B. Miller: White, northern males						
Year/Age	18–24	25–34	35–44	45–54	55–64	25–64
1959	1.01	1.01	1.16	1.29	1.22	1.18

Sources: See Table II.

more than a four-year college education to the mean income of those with only a college education is presented in Table III for various age groups from each of the two studies. The figures in Table III are again meant to be used only as a rough guide. With the figures as a guide, the assumption was made in this study that a graduate school education increases the average lifetime hourly productivity (from what it was at the time of college graduation) of an individual of average ability (A_i equal to 100) by 23 percent. This corresponds to a productivity ratio of graduate school to college graduates of 1.23 for those of average ability, or a productivity ratio of graduate school to high school graduates of 1.60 for those of average ability.

The next problem is to decide how the 1.3 and 1.23 ratios are affected by the level of innate ability of the individual. There is again very little evidence one can use for guidance on this question. In his review article, Griliches cites a study by Wolfe, which concludes that those with the highest ability receive the greatest return from their education, but the data from Malmo, Sweden, on income, education, and IQ at the age of ten, which Griliches presents, shows no evidence that returns to education are a function of IQ.[3] Becker presents some evidence from a study of Minnesota males that indicates that those of higher IQ receive a greater return to education,[4] but the evidence presented in the same table on a study of Rochester males gives no indication that this is true. Nevertheless, it will be assumed here that returns to education (or advanced vocational training) do vary by ability, and in line with the methodology of this paper three alternative assumptions will be made.

Let z_{iC} denote the average lifetime hourly productivity of individual i with a college education; let z_{iG} denote the average lifetime hourly productivity of individual i with a graduate school education, and let z_{i0} continue to denote the average lifetime hourly productivity of individual i with a high school education. Since z_{i0} and A_i are assumed to be identically distributed, z_{i0} can be set equal to A_i with no loss of generality. (Remember that A_i is assumed to have mean 100 and standard deviation either 8, 16, or 32.) It will also be convenient to define a new variable, Q_i, which is

$$(10) \qquad Q_i = 100 + \beta(A_i - 100),$$

where β is either 0.5, 1.0, or 2.0, depending on whether the standard deviation of A_i is 32, 16, or 8, respectively. Q_i, in other words, is always assumed to have a standard deviation of 16.

The first assumption that is made concerning the effect of education beyond high school on productivity is the following:

$$(11) \qquad z_{iC}/z_{i0} = \max[1.0,\ 1.3 + 1.3\gamma(Q_i/100 - 1)], \qquad i = 1, 2, \ldots, n,$$

$$(12) \qquad z_{iG}/z_{iC} = \max[1.0,\ 1.23 + 1.23\gamma(Q_i/100 - 1)], \qquad i = 1, 2, \ldots, n,$$

where γ is equal to 1. What (11) and (12) say is that the effect of education on productivity is never less than zero and that the effect of education increases with the level of innate ability. For a person of average ability (Q_i or A_i equal to 100), z_{iC}/z_{i0} and z_{iG}/z_{iC} are 1.30 and 1.23, respectively, whereas for a person of, say, an ability level corresponding to a Q_i of 130, the ratios are 1.69 and 1.60, respectively.[5] (The ratio of z_{iG} to z_{i0} for a person of an ability level corresponding to a Q_i of 130 is thus (1.69) (1.60) = 2.70, as opposed to 1.60 for a person of ability level 100.)

3. Griliches, *op. cit.*, Table 8, p. 35.
4. Becker, *op. cit.*, Table 5.
5. From the definition of Q_i in (10), an ability level corresponding to a Q_i of 130 is either 160, 130, or 115, depending on whether the standard deviation of A_i is 32, 16, or 8, respectively.

The second assumption that is made about the effect of ability on the return to education is that γ is equal to 2 in equations (11) and (12). Under this assumption z_{iC}/z_{i0} and z_{iG}/z_{iC} are still equal to 1.30 and 1.23, respectively, for a person of average ability, but now for a person of an ability level corresponding to a Q_i of 130, the ratios are 2.08 and 1.97, respectively. The third assumption is that γ is equal to 0.5 in equations (11) and (12). In this case, for a person of an ability level corresponding to a Q_i of 130, the ratios z_{iC}/z_{i0} and z_{iG}/z_{iC} are 1.50 and 1.40, respectively, relative to the same 1.30 and 1.23 ratios for a person of average ability. The last two assumptions about γ were felt to be extreme enough that the true situation should fall somewhere in between. The assumption that γ is equal to 2 seems perhaps to be more extreme in the one direction than the assumption that γ is equal to 0.5 is in the other direction.

There are nine different earnings functions implicit in the above specifications. Three assumptions have been made about the distribution of innate ability, and three assumptions have been made about how the level of innate ability (as reflected through the Q_i variable) affects the relationship between education beyond high school and actual productivity. The nine earnings functions are, of course, not analytic, but before discussing how the functions were made analytic, it will be useful to define some units of measurement.

Individuals going to college or graduate school are assumed to work the equivalent of 8 hours a day, 20 days a month, 12 months a year. A college education is thus assumed to require (8 hours) (20 days) (12 months) (4 years) = 7680 hours, and a graduate school education the same amount.[6] An individual with average ability level ($A_i = 100$) and no education beyond high school ($E_i = 0$) is assumed to make 4 dollars an hour. Finally, 50 individuals are assumed to be enough to approximate the normal distribution of abilities, the individuals being placed between equal areas along the normal curve.

The nine earnings functions are presented and numbered in Table IV. To get a better indication of how the functions differ, the average hourly productivity of the least able individual with only a high school education and the average hourly productivity of the most able individual with a graduate school education are presented in Table IV for each of the functions. For earnings function 1 there is very little variation, with the most productive

6. For the computations below, the hours figures have been divided by 1000 to make them more manageable.

individual being only 3.3 times more productive than the least productive. For earnings function 9, on the other hand, the most productive individual is 33.4 times more productive than the least

TABLE IV

THE NINE EARNINGS FUNCTIONS

Number assigned to earnings function	SD of A_i[a]	γ	Average lifetime hourly productivity (in dollars per hour): Least able individual, $E_i=0$[b]	Most able individual, $E_i=15360$[c]
1	8	0.5	3.25	10.68
2	8	1.0	3.25	14.30
3	8	2.0	3.25	23.12
4	16	0.5	2.51	12.36
5	16	1.0	2.51	16.55
6	16	2.0	2.51	26.76
7	32	0.5	1.02	15.72
8	32	1.0	1.02	21.04
9	32	2.0	1.02	34.02

a. SD denotes standard deviation.

b. Least able individual corresponds to an A_i of 25.44, 62.72, or 81.36, depending on whether the standard deviation of A_i is 32, 16, or 8, respectively.

c. Most able individual corresponds to an A_i of 174.56, 137.28, or 118.64, depending on whether the standard deviation of A_i is 32, 16, or 8, respectively. E_i equal to 15360 corresponds to a graduate school education.

productive. For the "median" earnings function 5, the most and least productive individuals differ by a factor of 6.6.

Each of the functions in Table IV consists of a table of 150 values: for each of the 50 values of A_i and for each of the 3 values of E_i ($E_i=0$, 7680, 15360), there corresponds one hourly productivity rate. In order to make the functions usable in the model, each one had to be approximated by a differentiable function. This was done in the following manner. Under the specification of the model, productivity should be zero when E_i is less than zero, and so to approximate this situation, 50 more points (for the 50 values of A_i) were added to each of the nine tables, each point corresponding to an E_i of -3000 and an hourly productivity rate of zero. It should also be the case that further education beyond graduate school (15360 hours) produces no further gain in productivity, and so to approximate this situation, 50 more points were added to each of the nine tables, each point corresponding to an E_i of 18360 and an hourly productivity rate the same as the rate for E_i equal to 15360. Each table then consisted of 250 points, and for these observations for each table, hourly productivity was regressed against various polynomials in A_i and E_i until a good fit was achieved. The simplest polynomial that gave good fits was the following:

$$(13)\quad YE_i/(H_i-E_i)=b_0+b_1E_i+b_2A_i+b_3E_iA_i+b_4E^2{}_iA_i+b_5E_iA^2{}_i+b_6E^2{}_iA^2{}_i+b_7E^3{}_iA_i+b_8E^4{}_iA_i+b_9E^2{}_i+b_{10}E^3{}_i+b_{11}E^4{}_i.$$

In some cases simpler polynomials than the one in (13) gave results almost as good, but (13) seemed to be quite accurate in all nine cases, and it was chosen to be used for all of the cases. There appeared to be no serious outliers or series of outliers for any of the functions that would indicate that the polynomial approximations were diverging from what was required of them.

Equation (13) can be solved for YE_i, and it is now the case that the derivatives of YE_i with respect to H_i and E_i, which are needed in the solution of the model, are well defined. The realism of the model or lack thereof is not likely to be seriously affected by any errors made in the polynomial approximations. What is of much more importance is how accurately the earnings tables have been specified and whether the real world situation has been bracketed by the extreme earnings tables.

This completes the specification of the individual earnings functions, but before continuing with the analysis, it is worthwhile to compare how the work here relates to the work of Thomas Mayer.[7] Mayer presents a number of theoretical arguments for why a normal distribution of ability may lead to a skewed distribution of earnings. He defines ability as the "probability of completing a given task successfully,"[8] and he argues that even if ability is normally distributed, scale of operation effects are likely to lead to a skewed distribution of earnings. The work in this study is not inconsistent with Mayer's arguments. In the final analysis, because of the assumptions made about how education beyond high school affects actual productivity (and thus earnings), the distribution of actual productivity (and earnings) will be skewed.

The Marginal Tax Rate Function

With respect to maximizing social welfare, the best policy the government could follow would be to assign lump sum grants or taxes to individuals and avoid altogether the use of any kind of a distribution equation as in (3). This would have no adverse effect on incentives, so that the sum of earned income would be at a maximum. It is unlikely that a lump sum tax scheme could be carried out

7. Thomas Mayer, "The Distribution of Ability and Earnings," *The Review of Economics and Statistics*, XLII (May 1960), 189–95.

8. *Ibid.*, p. 190.

in practice, however, and that is the reason why a more standard tax structure was used in this model.

The form chosen for the marginal tax rate function, $t(YE_i)$, is the logarithmic form. The equation for $t(YE_i)$ is

$$(14) \qquad t(YE_i) = a_0 \log(YE_i + 1),$$

where a_0 is the tax parameter under the control of the government. Other functional forms for $t(YE_i)$ could have been chosen — a quadratic or cubic equation is an obvious possibility, but (14) has the advantage that it depends on only one parameter. The model becomes much more difficult to solve if there is more than one parameter in (14). It should be kept in mind, however, that were it not for computational constraints, better optima than those achieved below could be achieved by allowing the functional form of $t(YE_i)$ to have more flexibility.[9] Substituting equation (14) into equation (3) and integrating[1] yields

$$(15) \qquad Y_i = YG + YE_i + a_0 YE_i - a_0 (YE_i + 1) \log(YE_i + 1), \quad i = 1, 2, \ldots, n.$$

Equation (15) thus relates after-tax income to before-tax income and the tax parameters YG and a_0. YG is the minimum guaranteed level of income.

The Social Welfare Function

It seems to be part of the national heritage of the United States that all people should be given an equal opportunity in life. If people were given an opportunity to choose that social welfare function they would most like to see maximized from a number of different social welfare functions (with different weights attached to different groups), it is likely that many, if not most, would choose that function that had equal weights for all. One possible choice for the social welfare function is thus the sum of the individual lifetime utilities: $\sum_{i=1}^{n} U_i$. The problem with this function, however, is that it does not guarantee that someone will not receive zero lifetime utility. It is indifferent, for example, between a situation where individuals 1 and 2 each receive 10 utils and a situation where individual 1 receives 0 utils and individual 2 receives 20 utils. This function was thus rejected as being inconsistent with what seem to be most people's ethical views. The function that was chosen as the

9. A linear function for (14) (i.e., $t(YE_i) = a_0 YE_i$) was also tried in some of the initial runs of this study, but the marginal tax rate approached one too quickly to allow a solution to be achieved.

1. Note that $\int \log(x+1)dx = (x+1)\log(x+1) - x + \text{constant}$.

social welfare function is the product of the individual lifetime utilities: $\prod_{i=1}^{n} U_i$. This function avoids the above-mentioned problem and seems to be consistent with commonly held ethical views.[2]

The Budget Constraint

The actual budget constraint was made slightly more complicated than the one specified in (5). The constraint was taken to be

$$\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} YE_i - \frac{1}{5}\sum_{i=1}^{n} YE_i - P_E \sum_{i=1}^{n} E_i. \tag{16}$$

As in (5) the budget constraint (16) basically says that the sum of after-tax income cannot be greater than the sum of before-tax or earned income. It has been made slightly more complicated in the following two respects. First, the society is assumed to have made a decision ahead of time to devote one-fifth of its income to public goods, this income being taken away from the individuals and not given back in the form of Y_i type income. It is instead given back indirectly through the production of public goods. The decision is assumed to have been made ahead of the rest of the analysis, and thus public goods have not been included in the individual utility functions. Once the decision is made, individuals proceed with their utility maximization on the assumption that they can have no further say in how much of society's income is devoted to public goods.

The second complication in (16) is that each hour of education beyond high school (education through high school is counted directly as a public good) costs the society P_E dollars, the total cost to society of education beyond high school being $P_E \sum_{i=1}^{n} E_i$. The cost to the individual of his education beyond high school is thus only the cost of foregone earnings: society pays for the other costs (teachers' salaries, buildings, etc.). This assumption may not be too far removed from the current situation in the United States, since much of college and graduate school education is subsidized. P_E has been taken to be 4 dollars an hour.

The two complications of the budget constraint in (16) are not really critical to the analysis and conclusions of the model; similar conclusions would result if different budget constraints were used. The particular constraint in (16) was chosen as approxi-

2. Solutions of the model were actually obtained using both social welfare functions, and, as discussed below, it turned out that the results were not very sensitive to the particular function used.

mating in some loose sense the present situation in the United States.

The Number of Individuals in the Model

Since 50 individuals are used to approximate the normal distribution of innate ability and since five different values of the δ_i coefficients have been specified under two of the three assumptions above, this means that there are 250 different individuals in the model.[3] The δ_i coefficients were distributed equally by ability: no assumption was made, for example, that those of higher ability tend to have a larger or smaller value of δ_i than do those of lower ability. No particular assumption appeared to be any more reasonable than any other in this regard, and so the simple assumption that δ_i is distributed independently of A_i was made.

Conclusion

This completes the specification of the model. It will be the objective of Section IV to see what kind of income distribution is implied by the model under the various sets of assumptions that have been made. From the utility functions in (9), it can be seen that for a given amount of leisure, the marginal utility of income decreases as income increases, which by itself tends to pull in the direction of an equal distribution of income. Maximizing the product of utilities also tends to pull in a similar direction, that of an equal distribution of utility. Pulling in the other direction, however, is the fact that as the tax parameters YG and a_0 increase (thus making the income distribution more equal), there is less incentive for people to work and educate themselves, which causes earned income (and thus the total amount of income available for distribution) to decrease. The object of the model is thus to find the optimal point between the extremes of a completely equal distribution of income (or utility) and of an income distribution that is the same as the distribution of earned income.

IV. The Results

The Basic Results

Since there are 9 different earnings functions and since 3 different assumptions about the δ_i coefficients have been made, this gives

3. For reasons of programming symmetry, even under the first assumption that δ_i is equal to 0.3 for everyone, 250 individuals were included in the model. In this case, of course, there are really only 50 different individuals.

TABLE V
Optimal Gini Coefficients and Tax Parameters for the 27 Cases

Earnings function	Assumption about the utility functions											
	$\delta_i = 0.3, 0.3, 0.3, 0.3, 0.3$				$\delta_i = 0.2, 0.25, 0.3, 0.375, 0.45$				$\delta_i = 0.15, 0.225, 0.3, 0.4, 0.5$			
	G_{YE}[a]	G_Y[b]	a_0^*	YG^*[c]	G_{YE}[a]	G_Y[b]	a_0^*	YG^*[c]	G_{YE}[a]	G_Y[b]	a_0^*	YG^*[c]
1	0.106	0.099	0.045	0	0.211	0.199	0.044	0	0.271	0.257	0.044	0
2	0.182	0.167	0.048	5	0.257	0.235	0.048	10	0.309	0.284	0.048	10
3	0.284	0.233	0.059	45	0.343	0.283	0.058	50	0.395	0.318	0.060	60
4	0.153	0.143	0.045	0	0.236	0.223	0.044	0	0.295	0.275	0.045	5
5	0.237	0.209	0.053	20	0.303	0.261	0.055	30	0.357	0.304	0.056	35
6	0.340	0.262	0.065	65	0.391	0.304	0.060	70	0.442	0.335	0.065	80
7	0.260	0.215	0.060	35	0.326	0.268	0.059	40	0.382	0.310	0.065	45
8	0.347	0.264	0.068	60	0.398	0.305	0.067	65	0.454	0.336	0.070	75
9	0.448	0.310	0.074	100	0.491	0.340	0.073	110	0.544	0.366	0.076	120

a. G_{YE} is the Gini concentration coefficient of before-tax income.
b. G_Y is the Gini concentration coefficient of after-tax income.
c. YG^* is in thousands of dollars.

27 cases to analyze. The central question for each case is how after-tax income is distributed among the 250 individuals. Also of interest are the optimum values of the tax parameters a_0 and YG and the optimum values of the H_i and E_i variables. The summary results for each of the 27 cases are presented in Table V. Presented in the table are the optimal tax parameters, a^*_0 and YG^*, and the Gini concentration coefficients of before-tax and after-tax income, G_{YE} and G_Y.[4]

The concentration coefficient of before-tax income ranges from 0.106 for earnings function 1 and $\delta_i=0.3$, 0.3, 0.3, 0.3, 0.3 to 0.544 for earnings function 9 and $\delta_i=0.15$, 0.225, 0.3, 0.4, 0.5. The concentration coefficient of after-tax income in turn ranges from 0.099 to 0.366 for the same two cases. For the "median" case of earnings function 5 and $\delta_i=0.2$, 0.25, 0.3, 0.375, 0.45, the before-tax coefficient is 0.303 and the after-tax coefficient 0.261. The tax parameter a_0 in Table V ranges from 0.044 to 0.076 and the minimum guaranteed level YG from 0 to 120 thousand dollars.

Some Detailed Results

Due to space limitations, the detailed results of all of the cases cannot be presented. Of the nine earnings functions, functions 5 and 8 were singled out, and more detailed results for these two functions are presented in Tables VI and VII. Values of H_i/T_i, E_i, YE_i, Y_i, $(YE_i-Y_i)/YE_i$, and U_i are presented in the tables for each value of δ_i for individuals with the lowest, the average, and the highest level of innate ability. H_i/T_i is the percent of possible working hours that the individual spends working, and $(YE_i-Y_i)/YE_i$ is the individual's average tax rate.[5]

The results in the two tables are as expected. Individuals with greater ability work more, other things being equal, than those with lesser ability and also educate themselves more. More industrious individuals (as measured by the δ_i coefficients) work more, other things being equal, than less industrious ones. These two results are quite evident for earnings function 8, where low-ability individuals are considerably less productive than those with high ability. As can be seen in Table VII, for earnings function 8 low-ability individuals with low levels of industriousness work very little: their

4. In a Lorenz diagram, the Gini concentration coefficient is the ratio of the area between the diagonal and the Lorenz curve to the total area below the diagonal. For a perfectly equal distribution the Gini coefficient is zero.

5. For the computations, the constant B in equation (9) was taken to be 1. Also, as mentioned above, the values of H_i, E_i, and T_i have been divided by 1000, which also means that YE_i and Y_i have been divided by 1000.

TABLE VI
MORE DETAILED RESULTS FOR EARNINGS FUNCTION 5

$\delta_i = 0.3, 0.3, 0.3, 0.3, 0.3$: $a_0^* = 0.053$, $YG^* = 20$, $G_Y = 0.209$

	$\delta_i = 0.3$		
Variables \ A_i	62.72	100.4	137.28
H_i/T_i	0.27	0.28	0.32
E_i	2.61	3.42	17.06
YE_i	264.6	444.7	1400.
Y_i	219.8	343.9	955.5
$(YE_i - Y_i)/YE_i$	0.17	0.23	0.32
U_i	223.7	254.5	332.1

$\delta_i = 0.2, 0.25, 0.3, 0.375, 0.45$: $a_0^* = 0.055$, $YG^* = 30$, $G_Y = 0.261$

	$\delta_i = 0.2$			$\delta_i = 0.25$			$\delta_i = 0.3$			$\delta_i = 0.375$			$\delta_i = 0.45$		
Variables \ A_i	62.72	100.4	137.28	62.72	100.4	137.28	62.72	100.4	137.28	62.72	100.4	137.28	62.72	100.4	137.28
H_i/T_i	0.16	0.18	0.22	0.21	0.22	0.27	0.26	0.27	0.32	0.34	0.35	0.38	0.41	0.42	0.45
E_i	2.43	3.12	16.76	2.53	32.90	16.94	2.60	3.41	17.06	2.67	3.54	17.17	2.71	3.62	17.24
YE_i	155.9	276.7	900.7	204.8	357.2	1142.	254.2	438.3	1385.	329.0	561.2	1753.	404.8	685.5	2126.
Y_i	151.1	236.5	644.8	186.1	291.7	794.9	220.8	346.2	943.1	272.4	427.4	1163.	323.8	508.0	1382.
$(YE_i - Y_i)/YE_i$	0.03	0.15	0.28	0.09	0.18	0.30	0.13	0.21	0.32	0.17	0.24	0.34	0.20	0.26	0.35
U_i	232.5	251.2	292.7	227.9	252.0	309.6	226.2	256.0	331.7	228.3	267.6	375.5	235.8	286.2	434.9

$\delta_i = 0.15, 0.225, 0.3, 0.4, 0.5$: $a_0^* = 0.056$, $YG^* = 35$, $G_Y = 0.304$

	$\delta_i = 0.15$			$\delta_i = 0.225$			$\delta_i = 0.3$			$\delta_i = 0.4$			$\delta_i = 0.5$		
Variables \ A_i	62.72	100.4	137.28	62.72	100.4	137.28	62.72	100.4	137.28	62.72	100.4	137.28	62.72	100.4	137.28
H_i/T_i	0.11	0.12	0.18	0.18	0.20	0.25	0.26	0.27	0.31	0.36	0.37	0.41	0.46	0.47	0.50
E_i	2.21	2.81	16.41	2.47	3.20	16.86	2.59	3.40	17.05	2.68	3.56	17.19	2.73	3.66	17.27
YE_i	101.7	190.5	651.8	174.8	310.9	1013.	248.9	432.6	1377.	349.4	497.1	186.9	451.7	764.5	2369.
Y_i	115.9	180.2	488.1	169.1	263.7	714.9	221.1	345.4	935.6	289.8	452.9	1226.	358.1	559.8	1513.
$(YE_i - Y_i)/YE_i$	neg.	0.05	0.25	0.03	0.15	0.29	0.11	0.20	0.32	0.17	0.24	0.34	0.21	0.27	0.36
U_i	242.2	255.0	281.0	231.1	251.9	300.5	227.4	256.6	331.4	231.4	273.5	392.6	244.9	303.3	483.9

TABLE VII

More Detailed Results for Earnings Function 8

$\delta_i = 0.3, 0.3, 0.3, 0.3, 0.3$; $a_0^* = 0.068$, $YG^* = 60$, $G_Y = 0.264$

Variables \ A_i	$\delta_i = 0.3$: 25.44	100.8	174.56
H_i/T_i	0.16	0.25	0.30
E_i	2.76	3.34	17.08
YE_i	68.8	392.3	1632.
Y_i	113.3	318.4	977.1
$(YE_i - Y_i)/YE_i$	neg.	0.19	0.40
U_i	202.1	256.3	340.0

$\delta_i = 0.2, 0.25, 0.3, 0.375, 0.45$: $a_0^* = 0.067$, $YG^* = 65$, $G_Y = 0.305$

	$\delta_i = 0.2$			$\delta_i = 0.25$			$\delta_i = 0.3$			$\delta_i = 0.375$			$\delta_i = 0.45$		
Variables \ A_i	25.44	100.8	174.56	25.44	100.8	174.56	25.44	100.8	174.56	25.44	100.8	174.56	25.44	100.8	174.56
H_i/T_i	0.05	0.15	0.21	0.10	0.20	0.26	0.16	0.20	0.34	0.24	0.32	0.37	0.32	0.39	0.44
E_i	1.88	2.95	16.75	2.51	3.18	16.96	2.74	3.33	17.08	2.89	3.48	17.20	2.97	3.59	17.28
YE_i	19.0	227.9	1045.	42.1	308.2	1339.	65.3	389.5	1636.	101.3	513.6	2087.	138.5	640.0	2547.
Y_i	81.3	225.2	695.1	99.1	275.7	851.6	116.1	325.3	1003.	141.5	398.8	1228.	166.9	471.7	1451.
$(YE_i - Y_i)/YE_i$	neg.	0.01	0.33	neg.	0.11	0.36	neg.	0.16	0.39	neg.	0.22	0.41	neg.	0.26	0.43
U_i	227.4	255.9	300.6	214.7	255.5	318.9	205.0	258.3	342.4	195.1	267.9	380.7	189.7	284.0	451.1

$\delta_i = 0.15, 0.225, 0.3, 0.4, 0.5$: $a_0^* = 0.070$, $YG^* = 75$, $G_Y = 0.336$

	$\delta_i = 0.15$			$\delta_i = 0.225$			$\delta_i = 0.3$			$\delta_i = 0.4$			$\delta_i = 0.5$		
Variables \ A_i	25.44	100.8	174.56	25.44	100.8	174.56	25.44	100.8	174.56	25.44	100.8	174.56	25.44	100.8	174.56
H_i/T_i	0	0.09	0.16	0.06	0.16	0.23	0.13	0.24	0.30	0.24	0.34	0.39	0.36	0.44	0.48
E_i	0	2.45	16.31	1.98	3.03	16.85	2.66	3.31	17.07	2.90	3.51	17.23	2.99	3.63	17.32
YE_i	0	134.0	727.7	21.0	253.0	1165.	55.6	374.4	1606.	104.1	540.6	2207.	155.3	711.8	2826.
Y_i	75.0	172.2	518.6	92.7	247.6	747.0	118.5	320.4	966.0	152.2	415.6	1251.	186.1	510.0	1532.
$(YE_i - Y_i)/YE_i$	neg.	neg.	0.29	neg.	0.02	0.36	neg.	0.14	0.40	neg.	0.23	0.43	neg.	0.28	0.46
U_i	250.9	261.7	287.2	226.0	256.7	308.0	210.0	259.3	340.0	197.3	272.8	402.9	192.5	298.1	495.7

earned income is thus quite small, and their after-tax income is not much larger than the minimum guaranteed level of income.

The level of education beyond high school has a range of 0 to 17270 hours in the two tables. (Remember that 7680 hours is the equivalent of a college education, and 15360 hours the equivalent of a college plus graduate school education.) From the complete set of results, it was observed that the number of hours spent in education tended to change suddenly at a particular level of ability from about 4000 hours to about 16000 hours. The particular level of ability varied from case to case, but a sudden jump always occurred somewhere for each case. This result is no doubt due to the sharp jumps that are inherent in the tables and that are captured quite well by the polynomial approximations. A more detailed specification of the tables would probably have lessened those jumps. It is still true, however, that the results indicate that people choose to receive either about two years of college education or about eight years, with few choosing something in between. The results thus indicate that the specification of the returns for a graduate school education in the earnings tables may be too high relative to the specification of the returns for a college education. Increasing the returns for a college education and decreasing the returns for a graduate school education in the earnings tables would tend to improve matters in this respect and would tend to make the resulting income distribution more equal. Moderate changes in the specification of the relative returns in each of the tables should, however, have only a small effect on the general results for each case.

Sensitivity of the Results to the Different Parameter Values

The sensitivity of the optimal Gini concentration coefficient of after-tax income, G_Y, to the different parameters of the model is examined in Table VIII. The values of G_Y in Table VIII are a rearrangement of the values in Table V. In section A of the table the sensitivity to the standard deviation of A_i is examined; in section B the sensitivity to the values of the coefficient γ in equations (11) and (12) is examined; and in section C the sensitivity to the assumption about the δ_i coefficients in the utility function is examined. The values of G_Y in each section should be examined down columns. From the results in Table VIII, G_Y appears to be the least sensitive to the change in the standard deviation of A_i from 8 to 16 and the most sensitive to the change in δ_i from 0.3 for all individuals to 0.2, 0.25, 0.3, 0.375, and 0.45, respectively, for each fifth of the population. Otherwise, the sensitivity of G_Y appears

TABLE VIII

Sensitivity of G_Y from Table V to the Different Parameter Values of the Model

A. Sensitivity to the standard deviation of A_i:

$\delta_i = 0.3, 0.3, 0.3, 0.3\ 0.3$

Standard deviation of A_i	$\gamma=0.5$ G_Y	$\gamma=0.5$ ΔG_Y	$\gamma=1.0$ G_Y	$\gamma=1.0$ ΔG_Y	$\gamma=2.0$ G_Y	$\gamma=2.0$ ΔG_Y
8	.099	—	.167	—	.233	—
16	.143	.044	.209	.042	.262	.029
32	.215	.072	.264	.055	.310	.048

Average ΔG_Y from 8 to 16 = .038
Average ΔG_Y from 16 to 32 = .058

$\delta_i = 0.2, 0.25, 0.3, 0.375, 0.45$

Standard deviation of A_i	$\gamma=0.5$ G_Y	$\gamma=0.5$ ΔG_Y	$\gamma=1.0$ G_Y	$\gamma=1.0$ ΔG_Y	$\gamma=2.0$ G_Y	$\gamma=2.0$ ΔG_Y
8	.199	—	.235	—	.283	—
16	.223	.024	.261	.026	.304	.021
32	.268	.045	.305	.044	.340	.036

Average ΔG_Y from 8 to 16 = .024
Average ΔG_Y from 16 to 32 = .042

$\delta_i = 0.15, 0.225, 0.3, 0.4, 0.5$

Standard deviation of A_i	$\gamma=0.5$ G_Y	$\gamma=0.5$ ΔG_Y	$\gamma=1.0$ G_Y	$\gamma=1.0$ ΔG_Y	$\gamma=2.0$ G_Y	$\gamma=2.0$ ΔG_Y
8	.257	—	.284	—	.318	—
16	.275	.018	.304	.020	.335	.017
32	.310	.035	.336	.032	.366	.031

Average ΔG_Y from 8 to 16 = .018
Average ΔG_Y from 16 to 32 = .033

Average ΔG_Y from 8 to 16 over all nine cells = 0.027
Average ΔG_Y from 16 to 32 over all nine cells = 0.044

B. Sensitivity to γ:

$\delta_i = 0.3, 0.3, 0.3, 0.3\ 0.3$

Value of γ	$SD=8$ G_Y	$SD=8$ ΔG_Y	$SD=16$ G_Y	$SD=16$ ΔG_Y	$SD=32$ G_Y	$SD=32$ ΔG_Y
0.5	.099	—	.143	—	.215	—
1.0	.167	.068	.209	.063	.264	.049
2.0	.233	.066	.262	.053	.310	.046

Average ΔG_Y from 0.5 to 1.0 = .060
Average ΔG_Y from 1.0 to 2.0 = .055

$\delta_i = 0.2, 0.25, 0.3, 0.375, 0.45$

Value of γ	$SD=8$ G_Y	$SD=8$ ΔG_Y	$SD=16$ G_Y	$SD=16$ ΔG_Y	$SD=32$ G_Y	$SD=32$ ΔG_Y
0.5	.199	—	.223	—	.268	—
1.0	.235	.036	.261	.038	.305	.037
2.0	.283	.048	.304	.043	.340	.035

Average ΔG_Y from 0.5 to 1.0 = .037
Average ΔG_Y from 1.0 to 2.0 = .042

$\delta_i = 0.15, 0.225, 0.3, 0.4, 0.5$

Value of γ	$SD=8$ G_Y	$SD=8$ ΔG_Y	$SD=16$ G_Y	$SD=16$ ΔG_Y	$SD=32$ G_Y	$SD=32$ ΔG_Y
0.5	.257	—	.275	—	.310	—
1.0	.284	.027	.304	.029	.336	.026
2.0	.318	.034	.335	.031	.366	.030

Average ΔG_Y from 0.5 to 1.0 = .027
Average ΔG_Y from 1.0 to 2.0 = .032

Average ΔG_Y from 0.5 to 1.0 over all nine cells = 0.041
Average ΔG_Y from 1.0 to 2.0 over all nine cells = 0.043

C. Sensitivity to δ_i assumptions:

$SD=8$

δ_i assumptions	$\gamma=0.5$ G_Y	$\gamma=0.5$ ΔG_Y	$\gamma=1.0$ G_Y	$\gamma=1.0$ ΔG_Y	$\gamma=2.0$ G_Y	$\gamma=2.0$ ΔG_Y
.3, .3, .3, .3, .3,	.099	—	.167	—	.233	—
.2, .25, .3, .375, .45	.199	.100	.235	.068	.283	.050
.15, .225, .3, .4, .5	.257	.058	.284	.049	.318	.035

Average ΔG_Y from .3 . . . to .2 . . . = .073
Average ΔG_Y from .2 . . . to .15 . . . = .047

$SD=16$

δ_i assumptions	$\gamma=0.5$ G_Y	$\gamma=0.5$ ΔG_Y	$\gamma=1.0$ G_Y	$\gamma=1.0$ ΔG_Y	$\gamma=2.0$ G_Y	$\gamma=2.0$ ΔG_Y
.3, .3, .3, .3, .3,	.143	—	.209	—	.262	—
.2, .25, .3, .375, .45	.223	.080	.261	.052	.304	.042
.15, .225, .3, .4, .5	.275	.052	.304	.043	.335	.031

Average ΔG_Y from .3 . . . to .2 . . . = .058
Average ΔG_Y from .2 . . . to .15 . . . = .042

$SD=32$

δ_i assumptions	$\gamma=0.5$ G_Y	$\gamma=0.5$ ΔG_Y	$\gamma=1.0$ G_Y	$\gamma=1.0$ ΔG_Y	$\gamma=2.0$ G_Y	$\gamma=2.0$ ΔG_Y
.3, .3, .3, .3, .3,	.215	—	.264	—	.310	—
.2, .25, .3, .375, .45	.268	.053	.305	.041	.340	.030
.15, .225, .3, .4, .5	.310	.042	.336	.031	.366	.026

Average ΔG_Y from .3 . . . to .2 . . . = .041
Average ΔG_Y from .2 . . . to .15 . . . . = .033

Average ΔG_Y from 0.3 . . . to 0.2 . . . over all nine cells = 0.057
Average ΔG_Y from 0.2 . . . to 0.15 . . . over all nine cells = 0.041

Note: ΔG_Y denotes the change in G_Y from one cell to the next.

to be fairly similar with respect to the different parameter values: the average change in G_Y is about 0.04 when going from one assumption or value to the next. The results in Table VIII can be used to indicate how the value of G_Y is likely to vary when values of the parameters different from the ones considered in this study are used, although extrapolating outside of the bounds of the parameters considered here should probably not be attempted.

It should also be pointed out that the results did not change much when the social welfare function was taken to be the sum of the individual utilities rather than the product. For earnings function 1 and $\delta_i = 0.2, 0.25, 0.3, 0.375, 0.45$, for example, the optimum point was the same for both welfare functions ($G_Y = 0.199$). For earnings function 3 and $\delta_i = 0.2, 0.25, 0.3, 0.375, 0.45$, G_Y increased by 0.01 (from 0.283 to 0.293) when the sum of utilities was used as the welfare function. For earnings function 7 and $\delta_i = 0.2, 0.25, 0.3, 0.375, 0.45$, G_Y also increased by 0.01 (from 0.268 to 0.278), and for earnings function 9 and $\delta_i = 0.2, 0.25, 0.3, 0.375, 0.45$, G_Y increased by 0.016 (from 0.340 to 0.356).

Comparison of Actual versus Optimal

The data on the actual distribution of income in the United States are not very good, but they can be used to give a general indication of how the actual distribution in the United States compares with the computed optimal distributions. I. B. Kravis has undertaken a very careful study of income distribution in the United States,[6] and his results will be used for the comparisons here. Using U.S. Department of Commerce data on after-tax income for 1950, Kravis reports a Gini coefficient for the United States of 0.38. For before-tax income, the coefficient is 0.41.[7] Both after-tax income and before-tax income include transfer payments; the former differs from the latter only in the exclusion of personal income taxes. There are many problems associated with any estimate of the size distribution of income, as Kravis is well aware, but his general results and the results of others do seem to indicate that the Gini coefficient for the United States is presently around 0.40. Unfortunately, all of these estimates are for a particular year and do not measure the lifetime distribution of income. Kravis reports, however, that lengthening the accounting period to three or four years only decreases the inequality measures between about 10 to 15 percent.[8]

6. Irving B. Kravis, *The Structure of Income* (Philadelphia: University of Pennsylvania, 1962).

7. *Ibid.*, Table 6.1, pp. 184–85.

8. *Ibid.*, p. 275. For similar tests and results, see James Morgan, "The

This means that a 0.40 one-year Gini coefficient would be decreased to about 0.35. This is, of course, still not a measure of the lifetime distribution of income. At the present time there do not appear to be enough data available to measure the actual lifetime distribution of income in the United States.

For lack of more information, it will be assumed for the following comparisons that the lifetime Gini coefficient for the United States is around 0.35. Is this coefficient near the optimum? From the results for G_Y in Table V, the answer appears to be no. Only for the "doubly extreme" case of earnings function 9 and $\delta = 0.15$, 0.225, 0.3, 0.4, 0.5 is the optimal Gini coefficient larger than 0.35, and even for a reasonably extreme case like earnings function 8 and $\delta_i = 0.2$, 0.25, 0.375, 0.45 the optimal Gini coefficient is only 0.305. (Remember that earnings functions 7, 8, and 9 all assume that productivity at the time of high school graduation is distributed with a standard deviation of 32, which is a rather large variation.) For the "median" case of earnings function 5 and $\delta_i = 0.2$, 0.25, 0.3, 0.375, 0.45 the optimal Gini coefficient is, of course, much lower, at 0.261. It is thus apparent that one has to make rather extreme assumptions about the variation of the productivity and industriousness of people before the optimal Gini coefficient approaches 0.35.

Why the Actual May Differ from the Optimal

There are a number of reasons why the actual distribution of income in the United States may differ from the optimal distributions computed above. First, it should be noted that in the model everyone is assumed to face the same earnings function. Individual differences in productivity occur because of differences in ability and educational levels, but not because of different forms of the earnings function. To the extent that in the real world some individuals or groups of individuals face different earnings functions because of discrimination or other social restrictions, the actual distribution of income will be less equal than the optimal distributions computed in this study.[9]

Second, it should be stressed that the model relies heavily on the assumption that the distribution of actual productivity at age eighteen is the same as the distribution of potential produc-

Anatomy of Income Distribution," *The Review of Economics and Statistics*, XLIV (Aug. 1962), 270–83.

9. Note that the figures in Tables II and III, upon which the specification of the earnings function is based, are primarily for white males. White males were used to lessen the chance of the figures being affected by discriminatory practices.

tivity at birth. To the extent that the quality of upbringing and education through age eighteen is not sufficient in society to allow all individuals to achieve their potential, the actual distribution of income will be less equal than the computed optimal distributions. The value judgment that everyone should be given an equal opportunity to achieve his or her potential is thus quite important and may be a major cause of the difference between the actual and computed optimal distributions.

Third, the government behavioral assumption in the model may be a poor approximation of the way the government behaves. To the extent that some people or groups of people of above-average income are given more weight in the social welfare function than others, the actual distribution of income, other things being equal, will be less equal than the optimal distributions computed above.

Finally, it is possible that the real world situation has not been bounded by the extreme sets of assumptions about the parameter values. It may be, for example, that the distribution of potential productivity at birth is skewed or has a larger standard deviation than 32. It may also be that people differ more in their tradeoff between income and leisure than is implied by the assumption that δ_i equals 0.15, 0.225, 0.3, 0.4, and 0.5, respectively, for each fifth of the population. If the real world has not been bounded by these assumptions, then the actual distribution will differ from the computed optimal distributions even if all of the other assumptions and value judgments of the model hold.[1]

V. Conclusions

In this study an attempt has been made to provide bounds on the optimal distribution of income. Much more empirical work is needed on the question of how innate ability is distributed and how education and innate ability affect productivity before more precise answers can be given. More information is also needed on how much people vary in their tradeoff between income and leisure and how their work effort responds to the structure of taxes.

The study has been based on a particular set of value judgments. It has been assumed that everyone should be weighted equally in

1. The concentration of property income in the United States, which the model ignores, may also cause the actual distribution of income to be more unequal than the computed optimal distributions. Kravis, *op. cit.*, p. 197, reports, however, that the concentration of property income in the United States is not very important in determining the overall degree of income inequality.

the social welfare function, that all people should be given an equal opportunity to achieve their potential, and that all people should be given the freedom to maximize their individual utility functions. This set of value judgments seems to be consistent with many people's ethical views, but it cannot be defended in any absolute sense.

Given that the particular set of value judgments is to be accepted, there are a number of possible biases in the model that should be considered for future work. One possible bias relates to the specification of the individual utility functions. An individual's utility has been assumed to be a function of his own income and leisure, but not of other people's income. If his utility is influenced by other people's income in that he dislikes people's having a lot more income than he has and dislikes seeing people with very small amounts of income, then the optimal distribution of income will be more equal than the present model implies it is, since there will be more of a tendency toward the equalization of incomes. On the other hand, if some individuals have increasing marginal utility of income, which they are assumed not to have in the model, this will cause the optimal distribution of income to be less equal than the present model implies it is.

There may also be a bias toward equality in the model because the model does not consider wealth and inheritance taxes. The model essentially assumes that everyone begins life with a zero amount of wealth. The inheritance tax rate, in other words, is assumed to be 100 percent. Work effort is assumed to be unaffected by this rate, and to the extent that work effort does respond to inheritance taxes, the optimal income distribution is less equal than the model implies it is.

The model also avoids any consideration of saving and investment decisions. It may be that the amount of savings and investment in a society is a positive function of the degree of income inequality in the society and thus that the growth rate of income is a positive function of the degree of inequality. If this is true, then leaving the constraint out of the model will result in an income distribution that is more equal than the true optimal distribution. The problem may be less pronounced for a wealthy country like the United States, however, than for less developed countries. It is also possible, of course, for the government to do much of the saving by distributing less of the before-tax income back to the people and then either investing the savings directly or loaning the money to businessmen and corporations to do the investing.

Finally, there may be some bias against equality in the model by using only two parameters to determine the entire tax structure.

Much more work clearly needs to be done in a number of areas before any definitive answers can be given to the question of how equal the distribution of income in a society should be. What this study has tried to do is to provide a general framework for analyzing the question.

Appendix

From the general model in Section II and the detailed specification in Section III, the Lagrangian to be maximized is

$$\text{(A1)}\qquad L=\prod_{i=1}^{n} U_i-\sum_{i=1}^{n}\psi_i\left(\frac{\partial U_i}{\partial H_i}-0\right)-\sum_{i=1}^{n}\phi_i\left(\frac{\partial U_i}{\partial E_i}-0\right)$$
$$-\lambda\left(\sum_{i=1}^{n}Y_i-\sum_{i=1}^{n}YE_i+\frac{1}{5}\sum_{i=1}^{n}YE_i+P_E\sum_{i=1}^{n}E_i\right).$$

There are $4n+3$ unknowns, the H_i, E_i, ψ_i, ϕ_i $(i=1, 2, \ldots, n)$, and YG, a_0, λ. The various variables are defined in the main body of the text. For the work here n was equal to 250. Differentiating L with respect to the $4n+3$ unknowns and setting the results equal to zero gave $4n+3$ nonlinear equations for which no solution could be obtained.[2] The model was solved, however, in the following manner. (For all of the computations, the constant B in the utility functions in (9) was assumed to be 1 and was dropped from the analysis.)

Differentiating L with respect to ψ_i and ϕ_i and setting the results equal to zero yields the $2n$ equations:

$$\text{(A2)}\qquad \partial U_i/\partial H_i=0,\qquad i=1, 2, \ldots, n,$$
$$\text{(A3)}\qquad \partial U_i/\partial E_i=0,\qquad i=1, 2, \ldots, n.$$

From (A2) it can be seen that

2. The technique used to solve (or attempt to solve) the equations was the standard Seidel method. Assume that the following two equations are to be solved for x_1 and x_2:

(i) $f_1(x_1, x_2)=0,$

(ii) $f_2(x_1, x_2)=0.$

The first step is to solve the first equation for x_1 in terms of x_1 and x_2 and the second equation for x_2 in terms of x_1 and x_2:

(iii) $x_1=g_1(x_1, x_2),$

(iv) $x_2=g_2(x_1, x_2).$

Then from initial values of x_1 and x_2, say $x_1^{(0)}$ and $x_2^{(0)}$, equations (iii) and (iv) can be solved to yield new values of x_1 and x_2, say, $x_1^{(1)}$ and $x_2^{(1)}$. These new values can be used to solve the equations again, and the cycle can be repeated until two successive sets of solution values are within a prescribed tolerance level of each other. One modification of this method, which was used in this study, is to take as the change in the value of, say, x_1 in any one step a fraction of the solution change. For example, $x_1^{(1)}$ would be taken to be

(v) $x_1^{(1)}=x_1^{(0)}+\epsilon(g_1(x_1^{(0)}, x_2^{(0)})-x_1^{(0)}),$

where ϵ is positive but less than one. None of the attempts at solving the complete set of the $4n+3$ or 1003 equations by this method resulted in a convergent solution, even for small values of ϵ and (it was thought) cleverly arranged orders of solving the equations.

(A4) $\quad H_i = T_i - Y_i^{\delta_i}(1-\delta_i)/R_{1i}R_{2i}R_{3i}, \qquad i=1, 2, \ldots, n,$

where $R_{1i} = \partial Y_i^{\delta_i}/\partial Y_i$, $R_{2i} = \partial Y_i/\partial YE_i$, $R_{3i} = \partial YE_i/\partial H_i$. Y_i, R_{1i}, and R_{2i} are, of course, functions of H_i. All that has been done in (A4) is to isolate one term in H_i on the left-hand side. From (A3) it can be seen that

(A5) $\quad (T_i - H_i)^{1-\delta_i}R_{1i}R_{2i}\partial YE_i/\partial E_i = 0, \qquad i=1, 2, \ldots, n.$

Since none of the terms left of $\partial YE_i/\partial E_i$ in (A5) can be zero, (A5) implies that $\partial YE_i/\partial E_i = 0$. From equation (13) it can be seen that, given H_i, A_i, and the b_j coefficients, $\partial YE_i/\partial E_i$ is merely a fourth-degree polynomial in E_i, from which one term in E_i can be isolated on the left-hand side. Write this latter equation as

(A6) $\quad E_i = g(A_i, H_i, E_i, b_0, b_1, \ldots, b_{11}), \qquad i=1, 2, \ldots, n.$

Differentiating L with respect to λ and setting the result equal to zero yields an equation that can be solved for α_0 explicitly:

(A7) $$\alpha_0 = \frac{-\sum_{i=1}^{n}\left(YG + P_E E_i + \frac{1}{5}YE_i\right)}{\sum_{i=1}^{n}\left[YE_i - (YE_i+1)\log(YE_i+1)\right]}.$$

Equations (A4), (A6), and (A7) consist of $2n+1$ equations in the $2n+2$ unknowns, H_i, E_i, α_0, and YG. The Lagrangian multipliers, in other words, are not included in any of the equations. Another equation cannot be found by differentiating L with respect to any of the remaining variables in which H_i, E_i, α_0, and YG appear without one or more of the Lagrangian multipliers also appearing. Therefore, an equation for YG in terms of the H_i, E_i, and α_0 alone cannot be found. Since the entire set of $4n+3$ equations could not be solved, the model was solved by solving the $2n+1$ equations for H_i, E_i, and α_0 for given values of YG and choosing that YG and the corresponding values of H_i, E_i, and α_0 that led to the largest value of the objective function, $\prod_{i=1}^{n} U_i$. About 10 values of YG were tried for each of the 27 cases, which required about 270 solutions of the $2n+1$ or 501 equations.

Even the $2n+1$ equations were not easy to solve, since the equations in (A4) and (A6) are highly nonlinear. They were finally solved by first solving for H_i in (A4) using given values of E_i and α_0, then solving for E_i in (A6) using the given value of α_0 and the solution values of H_i, then solving for α_0 in (A7) using the solution values of H_i and E_i, and then going back to (A4) with the new values and repeating the entire process. The value of ϵ (see note 2, page 578) used for H_i had to be as low as 0.1 in many cases; the value of ϵ for E_i was less critical and was generally taken to be 0.3. The computing time on an IBM 360–91 computer averaged about 10 seconds per solution. The 270 solutions thus took about 45 minutes of machine time.

PRINCETON UNIVERSITY