American Economic Association

Can Equalization of Opportunity Reduce Social Mobility?
Author(s): John Conlisk
Source: *The American Economic Review*, Vol. 64, No. 1 (Mar., 1974), pp. 80-90
Published by: American Economic Association
Stable URL: https://www.jstor.org/stable/1814883
Accessed: 18-12-2019 08:49 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms

American Economic Association is collaborating with JSTOR to digitize, preserve and extend access to *The American Economic Review*

This content downloaded from 194.27.101.122 on Wed, 18 Dec 2019 08:49:48 UTC
All use subject to https://about.jstor.org/terms

Can Equalization of Opportunity Reduce Social Mobility?

By John Conlisk*

> "Greater wealth, health, freedom, fairness, and educational opportunity are *not* going to give us the egalitarian society of our philosophical heritage. It will instead give us a society sharply graduated, with ever greater innate separation between the top and the bottom, and ever more uniformity within families as far as inherited abilities are concerned." . . . "What is most troubling about this prospect is that the growth of a virtually hereditary meritocracy will arise out of the successful realization of contemporary political and social goals." . . . " . . . improving the environment raises the heritability. The higher the heritability, the closer will human society approach a virtual caste system, with families sustaining their position on the social ladder from generation to generation as parents and children are more nearly alike . . . "
>
> *Richard Herrnstein*

An important area of overlap between psychology and economics concerns the nature-nurture issue in the intergenerational transmission of intelligence, income, and other characteristics. Psychologists are likely to stress parent-child intelligence effects and discuss income (or socioeconomic status or some such concept) as a related issue; economists are likely to reverse the stress. This paper presents an intergenerational income distribution model, stressing the nature-nurture issue, the relation between IQ and economic success, and the long-run distributional impact of equal opportunity policy.

One of the stimuli for writing this paper may be of interest. The quotes above are from an *Atlantic Monthly* article by Harvard psychologist Richard Herrnstein. The article was presumably widely read, since it was the cover article and received advance publicity through *Newsweek* and *Time* magazines. After a layman's introduction to the usefulness of IQ testing, Herrnstein presents some controversial arguments from which the quotes are taken. To the title of this paper, Herrnstein answers that equalizing opportunity not only can, but will reduce social mobility. At first, it seemed to me that this assertion could be easily "disproved" by signing a certain derivative in the model below. However, discussions with a psychologist suggested that Hernstein was in effect concerned with a different derivative, whose sign provided an easy "proof" of the assertions.

I. The Model

The mathematical formulation is a simple system of three difference equations subjected to additive shock terms. The variables are defined on successive generations of a single family. The distribution of variables over all families in society is assumed to be generated by independent drawings on the shock terms. The equations are

$$(1)\qquad Y_t = \alpha_0 + \alpha_1 I_t + \alpha_2 Y_{t-1} + u_{Yt}$$

$$(2)\qquad I_t = \beta_0 + \beta_1 G_t + \beta_2 Y_{t-1} + u_{It}$$

$$(3)\qquad G_t = \gamma_0 + \gamma G_{t-1} + u_{Gt}$$

Here Y_t, I_t, and G_t are the income, IQ, and genetic IQ potential of the t-th

* Associate professor of economics, University of California, San Diego. The research was supported by National Science Foundation grant GS-3201. For helpful suggestions, special thanks are due to Marjorie Honzik, Richard Schmalensee, the managing editor, and an anonymous referee.

This content downloaded from 194.27.101.122 on Wed, 18 Dec 2019 08:49:48 UTC
All use subject to https://about.jstor.org/terms

generation of the typical family. The Greek letters are parameters. The u's are random shock terms assumed to have zero means, to have the same variance matrix at each generation t, and to be serially uncorrelated. These assumptions are adequate to determine the variances and covariances among (Y_t, I_t, G_t) and $(Y_{t-1}, I_{t-1}, G_{t-1})$ for a given family tree at each generation t. Given the assumption that all family trees obey the same model (with independent shock term drawings), these variances and covariances will, for a large population, be virtually identical to sample variances and covariances computed over all family trees in society. Thus, by analyzing the equilibrium values of the variances and covariances, we can analyze the equilibrium income dispersion in society, the equilibrium income correlation between generations, and similar questions. The approach has been discussed in my 1969 paper.

If humans were unisexual and each person had exactly one offspring, then each family tree would be a straight line (no branching) and the interpretations of variables would be neat. As it is, some convention must be presumed which associates with a given family exactly one previous generation, or parental, family. That is, once (Y_t, I_t, G_t) is measured on a typical family, some convention is needed to decide which parental family to measure $(Y_{t-1}, I_{t-1}, G_{t-1})$ on. A simple convention would be to follow the male, or the female, line. A possibly more sensible convention would start with a set of rules for determining a family's economic "head." Then the straight line family tree needed by the model could be traced through family heads—family head, head parent of family head, head parent of head parent of family head, and so on. Also needed are conventions about which family members to define (Y_t, I_t, G_t) on. Though not the only choice, the following discussion assumes the IQ variables I_t and G_t are measured on the family head, while Y_t includes all family members' income.

Equation (1) states that a family's income is determined, subject to random shock, by the family head's IQ and by the family income of the head's parents. Here "income" may be used loosely. It might be interpreted as any of a number of general success measures—dollar income, occupational status, schooling attained, and so on. Since the time period is a generation, Y_t should be defined as lifetime income. Furthermore, since equilibrium variances and covariances are the objects of interest here, Y_t is assumed to be measured as a trendless index of some sort (absolute income as a ratio to society's average perhaps). Equations (2) and (3) state that the head's IQ is determined by genetic and environmental factors. The genetic component G_t is determined by equation (3) as the sum of a systematic inheritance from the head's parent and a random shock. Equation (2) states that the head's genetic IQ potential (G_t) is translated into actual IQ (I_t) subject to a systematic environmental effect determined by the family income Y_{t-1} of the head's parents, and by a random environmental effect u_{It} which captures omitted variables.

The apparent n-equation generalization of the model, discussed in the Appendix, is straightforward. The simplifications going into the three equation form (1)–(3) are deliberate. Even slightly greater complication would prevent explicit parametric statement of equilibrium magnitudes below (see the Appendix). Nonetheless, the model as it stands does allow a test of the logic of certain important propositions, and does allow plausible numerical illustration.

The following parameter restrictions, some of which are units normalizations, will be assumed. Here σ_Y^2, σ_I^2, σ_G^2, σ_{YI}, σ_{YG}, and σ_{IG} denote the variances and covari-

This content downloaded from 194.27.101.122 on Wed, 18 Dec 2019 08:49:48 UTC
All use subject to https://about.jstor.org/terms

ances of u.

$$(4) \quad \alpha_0 = \beta_0 = \gamma_0 = 0 \quad \sigma_G^2 = 1 \quad \alpha_1 = \beta_1 = 1$$
$$0 \leq \alpha_2, \beta_2, \alpha_2 + \beta_2 < 1 \quad 0 < \gamma < 1$$
$$\sigma_{YG} = \sigma_{IG} = 0$$

Since Y_t, I_t, and G_t are all index numbers, it may be assumed without loss of generality that their origin is chosen to make their equilibrium values zero; the restriction $\alpha_0 = \beta_0 = \gamma_0 = 0$ accomplishes this. Furthermore, it may be assumed without loss of generality that the genetic IQ index G_t is multiplicatively scaled to give the shock term u_{Gt} unit variance; hence the restriction $\sigma_G^2 = 1$. It is assumed that IQ contributes positively to income after controlling on parental income (so $\alpha_1 > 0$); that there is a genetic component to IQ determination (so $\beta_1 > 0$); and that the parent-child genetic intelligence correlation is positive (so $\gamma > 0$). Empirical bases for the assumptions α_1, β_1, $\gamma > 0$ are sketched in the following two paragraphs. Given that α_1, $\beta_1 > 0$, there is no loss of generality in assuming that Y_t and I_t are multiplicatively scaled so that $\alpha_1 = \beta_1 = 1$. So the substantive assumption in $\alpha_1 = \beta_1 = 1$ is that α_1, $\beta_1 > 0$. The assumptions $\alpha_2 \geq 0$ and $\beta_2 \geq 0$ state that it can't hurt in terms of income and intelligence to have more affluent parents, all else equal. The assumptions $\alpha_2 + \beta_2 < 1$ and $\gamma < 1$ are stability conditions (see the Appendix). The determination of the genetic variable G_t for a person takes place at his conception; and the random term u_{Gt} is an accident of conception. It seems safe to assume that u_{Gt} is uncorrelated with u_{Yt} and u_{It}, which mainly involve different forces occurring much later in time. Hence $\sigma_{YG} = \sigma_{IG} = 0$ is assumed.

The assertion $\beta_1 > 0$ (normalized here to $\beta_1 = 1$) and $\gamma > 0$ amount to saying that genetic factors do contribute to the variance of IQ. Psychologists appear to have strong evidence for these assertions in the form of sample correlations between pairs of individuals with differing degrees of blood relationship and differing similarities of environmental background. For example, the IQ correlations of (genetically) identical twins separated from infancy run much higher (about .75) than for unrelated children raised together (about .25), which suggests a genetic contribution to the variance of IQ. Such correlation evidence can be used in genetic models to estimate the "heritability" of IQ—a zero-to-one measure of the relative importance of the genetic contribution to IQ's variance. In terms of equation (2), heritability roughly translates as the ratio of the variance of the G_t term to the total variance of I_t. The IQ heritability estimates run about .80, suggesting that the genetic contribution is very important. Harold Jones and, more recently, Arthur Jensen present lengthy review articles on this type of evidence. (The Jensen article is the one which stimulated a storm of controversy concerning compensatory education and racial IQ differentials, though that controversy is not to the point here.)

The assertion $\alpha_1 > 0$ (normalized here to $\alpha_1 = 1$) says that IQ has a positive partial effect on family income Y_t, controlling on parent's income Y_{t-1}. The most direct type of evidence—that based on direct observations of the triplet (Y_t, I_t, Y_{t-1})—seems very hard to find. My 1971 paper presents such evidence, though based on a small sample; IQ does have the asserted effect, operating through education. Measuring indirectly, a partial correlation of nearly .50 between Y_t and I_t can be pieced together by the formula relating this partial correlation to simple correlations from many separate studies in the literature: about .55 simple correlations between occupational status Y_t and IQ (I_t), (studies briefly reviewed in Jensen, p. 15); about .40 between son's IQ (I_t) and father's status Y_{t-1} (studies reviewed in Jensen,

This content downloaded from 194.27.101.122 on Wed, 18 Dec 2019 08:49:48 UTC
All use subject to https://about.jstor.org/terms

$$\text{(5)} \quad \operatorname{var}(Y_t) = [\sigma^2 + \lambda(2\mu - 1)]/[1 - (\alpha_2 + \beta_2)^2]$$

$$\text{(6)} \quad \operatorname{var}(I_t) = \sigma_I^2 + \lambda + \beta_2^2 \operatorname{var}(Y_t) + 2\beta_2\lambda\mu\gamma$$

$$\text{(7)} \quad \operatorname{var}(G_t) = \lambda$$

$$\text{(8)} \quad r(Y_t, Y_{t-1}) = \alpha_2 + \beta_2 + \lambda\mu\gamma/\operatorname{var}(Y_t)$$

$$\text{(9)} \quad r(I_t, I_{t-1}) = \frac{\lambda\gamma + \beta_2[\sigma_I^2 + \sigma_{YI} + \mu(2\lambda - 1) + \beta_2 \operatorname{var}(Y_t) r(Y_t, Y_{t-1})]}{\operatorname{var}(I_t)}$$

$$\text{(10)} \quad r(G_t, G_{t-1}) = \gamma$$

$$\text{(11)} \quad r(Y_t, I_t) = \frac{\sigma_I^2 + \sigma_{YI} + \lambda\mu + \beta_2 \operatorname{var}(Y_t) r(Y_t, Y_{t-1})}{[\operatorname{var}(Y_t)\operatorname{var}(I_t)]^{1/2}}$$

$$\text{(12)} \quad r(Y_t, G_t) = \mu[\lambda/\operatorname{var}(Y_t)]^{1/2}$$

$$\text{(13)} \quad r(I_t, G_t) = (1 + \mu\gamma\beta_2)[\lambda/\operatorname{var}(I_t)]^{1/2}$$

pp. 74–77); and about .40 between son's status Y_t and father's status Y_{t-1} (see Peter Blau and Otis Duncan, p. 110), which is about the same as father-son years of schooling correlations (see Beverly Duncan). Confidence in the assumption $\alpha_1 > 0$ would also follow from general confidence in (i) the many psychologists' studies of IQ's independent contribution to predictions of educational success, and (ii) the many economists' studies of the contribution of educational to economic success.

The *equilibrium* variances and selected *equilibrium* correlation coefficients implied by the model are shown in equations (5) through (13), (see the Appendix for derivations). The parameter transforms $\lambda = (1-\gamma^2)^{-1}$, $\mu = [1-\gamma(\alpha_2+\beta_2)]^{-1}$, and $\sigma^2 = \sigma_Y^2 + \sigma_I^2 + 2\sigma_{YI}$ have been used to simplify ($\lambda, \mu > 1$ and $\sigma^2 > 0$). The units normalizations $\alpha_1 = \beta_1 = \sigma_G^2 = 1$ and the assumptions $\sigma_{YG} = \sigma_{IG} = 0$ have also been used to simplify.

II. Impacts of Equal Opportunity Policy

Hypotheses about the impact of successful policies to equalize opportunity can be stated in terms of sensitivities of the equilibrium magnitudes (5)–(13) to parameter changes. The size of the parent-child income correlation $r(Y_t, Y_{t-1})$ is a measure of the degree of social immobility in a society described by the model. In these terms, the question posed by the title of this paper is whether equalizing opportunity can increase $r(Y_t, Y_{t-1})$, that is, increase the extent to which a family's income is determined by the income of the head's parents.

If "equalizing opportunity" means reducing the special advantages of being born to affluence and the special disadvantages of being born to poverty, then equalizing opportunity might be interpreted as reducing α_2 or β_2 in the model. The derivative of interest is the following (since α_2 and β_2 appear only as the sum $\alpha_2 + \beta_2$ in the expression for $r(Y_t, Y_{t-1})$).

$$\text{(14)} \quad \frac{\partial r(Y_t, Y_{t-1})}{\partial(\alpha_2+\beta_2)} = \frac{\sigma^2\{\sigma^2 + 2\lambda\mu + \mu^2(\lambda-1)[1-(\alpha_2+\beta_2)^2]\} + \lambda\mu^2}{[\sigma^2 + \lambda(2\mu-1)]^2}$$

In view of parameter constraints (4), this derivative must be positive; that is, opportunity equalization must increase social mobility. This contradicts the opening quotes, and responds "no" to the title question.

This content downloaded from 194.27.101.122 on Wed, 18 Dec 2019 08:49:48 UTC
All use subject to https://about.jstor.org/terms

However, there is a second sensible interpretation of equalizing opportunity. Many sources of unequal opportunity are not related to parental affluence or poverty. Reduction of these would be reflected in the model by a reduction in the variances σ_Y^2 and σ_I^2 of the shock terms u_{Yt} and u_{It}. Since σ_Y^2 and σ_I^2 appear in the expression for $r(Y_t, Y_{t-1})$ only within $\sigma^2=\sigma_Y^2+\sigma_I^2+2\sigma_{YI}$ (the variance of $u_{Yt}+u_{It}$), the relevant derivative is

$$(15)\quad \partial r(Y_t, Y_{t-1})/\partial\sigma^2 = -\lambda\mu\gamma[1-(\alpha_2+\beta_2)^2]^{-1}\operatorname{var}(Y_t)^{-2}$$

which must be negative in view of parameter constraints (4); that is, opportunity equalization in the form of a reduction of σ^2 decreases social mobility. This agrees with the opening quotes, and responds "yes" to the title question.

Summarizing, the two types of opportunity equalization—(I) reduction in $\alpha_2+\beta_2$ and (II) reduction in σ^2—will have opposite effects on social mobility. Intuitively, α_2 and β_2 represent systematic parent-child connections in the model, so reducing $\alpha_2+\beta_2$ will reduce the equilibrium parent-child connection measure $r(Y_t, Y_{t-1})$. On the other hand, σ^2 represents random noise in the model, so reducing σ^2 increases the systematic connection measure $r(Y_t, Y_{t-1})$. An example of predominantly type I equalization (reducing $\alpha_2+\beta_2$) might be a compensatory preschool program for children of the poor. An example of predominantly type II equalization might be the replacement of random military conscription by an all-volunteer system. Many equalization policies which come to mind are not clearly of one or the other type, but a mixture. For example, federal subsidization of genetic counseling might reduce the incidence of birth defects at all income levels, thus reducing σ^2; but, to the extent that more affluent persons already received the counseling, there might be a bigger effect at lower income levels, thus reducing $\alpha_2+\beta_2$. For many opportunity equalization policies, therefore, the net effect on social mobility would seem to be an open issue.

The same conclusions apply for an apparent alternate measure of social mobility. In view of the parameter normalizations $\alpha_0=\beta_0=\gamma_0=0$ and $\alpha_1=\beta_1=1$, substitution of (2) and (3) in (1) yields the reduced form equation for Y_t.

$$(16)\quad Y_t = \gamma G_{t-1} + (\alpha_2+\beta_2)Y_{t-1} + (u_{Yt}+u_{It}+u_{Gt})$$

The equilibrium coefficient of determination, call it R_Y^2, of this equation is a reasonable measure of social immobility. In view of the parameter normalization $\sigma_G^2=1$ and the assumptions $\sigma_{YG}=\sigma_{IG}=0$, the error term variance of (16) is $1+\sigma^2$. Hence R_Y^2 is

$$(17)\quad R_Y^2 = 1-(1+\sigma^2)/\operatorname{var}(Y_t)$$

from which the derivatives of R_Y with respect to $\alpha_2+\beta_2$ and σ^2 follow.

$$(18)\quad \partial R_Y/\partial(\alpha_2+\beta_2) = (1/R_Y)(1-R_Y^2)[\lambda\gamma\mu^2+(\alpha_2+\beta_2)\cdot\operatorname{var}(Y_t)]/[\sigma^2+\lambda(2\mu-1)]$$

$$\partial R_Y/\partial\sigma^2 = -[\lambda(2\mu-1)-1]/\{2R_Y[1-(\alpha_2+\beta_2)^2][\operatorname{var}(Y_t)]^2\}$$

In view of parameter constraints (4) (and their implication that $\lambda, \mu>1$), these two derivatives must be positive and negative, respectively. Therefore, the preceding discussion of types I and II equalization apply to the social immobility measure R_Y as well as to the measure $r(Y_t, Y_{t-1})$.

The paradoxical statement that opportunity equalization might reduce social mobility is in a sense a contrived paradox. To many advocates of opportunity equalization, the major goal is a reduction in

This content downloaded from 194.27.101.122 on Wed, 18 Dec 2019 08:49:48 UTC
All use subject to https://about.jstor.org/terms

the income inequality measure var (Y_t), rather than the social immobility measures $r(Y_t, Y_{t-1})$ and R_Y. The relevant derivatives are

$$(19)\quad \partial\,\mathrm{var}\,(Y_t)/\partial(\alpha_2+\beta_2) = 2[\lambda\gamma\mu^2 + (\alpha_2+\beta_2)\,\mathrm{var}\,(Y_t)]/[1-(\alpha_2+\beta_2)^2]$$

$$\partial\,\mathrm{var}\,(Y_t)/\partial\sigma^2 = 1/[1-(\alpha_2+\beta_2)^2]$$

Since both of these magnitudes must be positive, then either type of opportunity equalization will reduce income inequality. The paradox is thus a result of casual verbal equations between income equality and social mobility, whereas the two concepts are distinct (at least as measured here) and can move in opposite directions.

The heritability concept introduced above—the share of genetic factors in the variance of a measure—would best be represented here by var (G_t)/var (I_t) and var (G_t)/var (Y_t), for IQ and income, respectively. Barring eugenic changes in var (G_t), these heritability measures are just inverse measures of inequality in IQ and income. Thus, the statement that opportunity equalization decreases inequality is equivalent to the statement that opportunity equalization increases heritability, though this equivalence is not apparent in some contexts (in the opening quotes for example).

III. Rough Orders of Magnitude

The model, in the simple form given or in much expanded form, would present no unusual estimation problems if data were available. Unfortunately, intergenerational data of the sort needed are not available (G_t is not even measurable). Nevertheless, rough orders of magnitude for the derivatives discussed above, those of $r(Y_t, Y_{t-1})$, R_Y, and var (Y_t) with respect to $\alpha_2+\beta_2$ and σ^2, can be inferred from evidence in the literature. These derivatives can be expressed as functions of the three correlations $r(Y_t, Y_{t-1})$, $r(G_t, G_{t-1})$, and $r(Y_t, G_t)$;[1] and there is evidence in the literature suggesting the following specifications for the three correlations.

$$(20)\quad r(Y_t, Y_{t-1}) = .4 \qquad r(G_t, G_{t-1}) = .5$$
$$.40 \le r(Y_t, G_t) \le .75$$

Table 1 presents values of the derivatives for these specifications. If the magnitudes of the derivatives are judged by the sizes of the elasticities (in parentheses on the table), it appears that type I opportunity equalization (reduction of $\alpha_2+\beta_2$) has a larger impact (in absolute value) on social mobility than type II equalization (reduction in σ^2). This is encouraging since type I equalization increases mobility, whereas type II diminishes it. The elasticities of var (Y_t) suggest a reverse pattern; type II equalization appears to have a larger impact on inequality than type I equalization.

Specifications (20) need justification. Here "income" Y_t may be interpreted as any general economic success measure. Father-son dollar income data for Y_t are not available; but father-son occupational status correlations and years of schooling correlations[2] cluster remarkably closely about the value $r(Y_t, Y_{t-1}) = .4$ specified in (20).

[1] Equations (5), (8), (10), (12), (14), (15), (18), and (19) form a system of 10 equations determining the 10 magnitudes $\alpha_2+\beta_2$, γ, σ^2, $\partial r(Y_t, Y_{t+1})/\partial(\alpha_2+\beta_2)$, $\partial r(Y_t, Y_{t-1})/\partial\sigma^2$, $\partial R_Y/\partial(\alpha_2+\beta_2)$, $\partial R_Y/\partial\sigma^2$, ∂var $(Y)_t/\partial(\alpha_2+\beta_2)$, ∂var$(Y_t)/\partial\sigma^2$, and var(Y_t) as functions of the three correlations $r(Y_t, Y_{t-1})$, $r(G_t, G_{t-1})$, and $r(Y_t, G_t)$. Explicit solutions are available, but are excessively tedious to present here.

[2] B. Duncan reports father-son years of schooling correlations of about .4; unpublished calculations yield a similar result for the sample analyzed in Morgan et al. Peter Blau and Otis Duncan, p. 110, report about the same again for father-son occupational status correlations. In each case the sample is large, representative of the United States, and apparently of high quality. Furthermore, the correlations for subsamples by son's age group do not vary much around the .4 benchmark (which lends some credence to the use of equilibrium analysis).

This content downloaded from 194.27.101.122 on Wed, 18 Dec 2019 08:49:48 UTC
All use subject to https://about.jstor.org/terms

TABLE 1—DERIVATIVES

Assumed Value of $r(Y_t, G_t)$	Derivatives and (in Parentheses) Elasticities					
	of $r(Y_t, Y_{t-1})$ with respect to		of R_Y with respect to		of var (Y_t) with respect to	
	$\alpha_2+\beta_2$	σ^2	$\alpha_2+\beta_2$	σ^2	$\alpha_2+\beta_2$	σ^2
.75	.92	−.10	.73	−.08	2.34	1.02
	(.32)	(−.28)	(.22)	(−.20)	(.12)	(.43)
.70	.93	−.07	.75	−.06	2.85	1.03
	(.41)	(−.28)	(.29)	(−.21)	(.15)	(.50)
.65	.94	−.05	.77	−.05	3.49	1.05
	(.50)	(−.27)	(.35)	(−.22)	(.19)	(.56)
.60	.95	−.04	.79	−.04	4.29	1.06
	(.57)	(−.25)	(.42)	(−.22)	(.22)	(.62)
.55	.96	−.02	.81	−.03	5.33	1.08
	(.64)	(−.22)	(.49)	(−.22)	(.24)	(.68)
.50	.97	−.02	.84	−.02	6.71	1.09
	(.71)	(−.20)	(.56)	(−.20)	(.27)	(.73)
.45	.97	−.01	.86	−.01	8.56	1.11
	(.77)	(−.17)	(.63)	(−.19)	(.29)	(.78)
.40	.98	−.01	.89	−.01	11.16	1.13
	(.81)	(−.14)	(.69)	(−.16)	(.31)	(.83)

Note: All entries assume $r(Y_t, Y_{t-1}) = .4$ and $r(G_t, G_{t-1}) = .5$

Since the genetic IQ potential G_t is not measurable, there are no reports in the literature on $r(G_t, G_{t-1})$. However, L. Erlenmeyer-Kimling and Lissy Jarvik, in surveying parent-child correlations of measured IQ, suggest a consensus figure of about $r(I_t, I_{t-1}) = .5$. Three reasons may be advanced for speculating that $r(G_t, G_{t-1})$ is about the same, as specified in (20). First, estimates of the heritability var (G_t)/var (I_t) of intelligence run quite high (about .8 as surveyed by Jensen, pp. 46–51), suggesting that G_t and I_t vary closely together. Second, the forces causing a divergence between $r(I_t, I_{t-1})$ and $r(G_t, G_{t-1})$ push in both directions, thus possibly cancelling. In equation (2), the noise term u_{It} tends to make $r(I_t, I_{t-1})$ smaller than $r(G_t, G_{t-1})$, while the systematic term $\beta_2 Y_{t-1}$ tends to make $r(I_t, I_{t-1})$ larger. Third, $r(G_t, G_{t-1}) = .5$ is the theoretical value predicted by very simple genetic models (see Jensen, p. 49).

The specification $.40 \leq r(Y_t, G_t) \leq .75$ in (20) cannot be verified directly from the literature, since G_t is not measurable; but reported values of $r(Y_t, I_t)$ are suggestive proxies. In surveying the literature, Jensen, pp. 15–16, reports $r(Y_t, I_t)$ values from .42 to .57, where Y_t is measured as occupational status; and Jacob Miner, p. 67, reports $r(Y_t, I_t)$ values from .44 to .75, where Y_t is measured as years of schooling.

Table 2 provides further checks on the plausibility of specifications (20) by presenting the implied values of some other magnitudes. The tabled income heritability var (G_t)/var (Y_t) figures are plausible. There is evidence in the literature (reviewed in Jensen, pp. 58–59) that scholastic achievement has on the order of half the heritability of IQ, presumably because many environmental factors intervene between potential and achievement. For the same reason, one might guess that income will have substantially lower heritability than IQ (where, as noted in the paragraph

This content downloaded from 194.27.101.122 on Wed, 18 Dec 2019 08:49:48 UTC
All use subject to https://about.jstor.org/terms

TABLE 2—OTHER MAGNITUDES

Assumed Value of $r(Y_t, G_t)$	Heritability var (G_t)/var (Y_t)	σ^2	$1/(1+\sigma^2)$	$\alpha_2+\beta_2$	R_Y
.75	.49	1.15	.47	.14	.21
.70	.41	1.58	.39	.18	.21
.65	.34	2.12	.32	.21	.21
.60	.28	2.81	.26	.24	.21
.55	.23	3.71	.21	.27	.20
.50	.18	4.90	.17	.29	.19
.45	.14	6.52	.13	.32	.19
.40	.11	8.80	.10	.33	.18

Note: All entries assume $r(Y_t, Y_{t-1})=.4$ and $r(G_t, G_{t-1})=.5$

following (4), IQ has estimated heritability of about .8). The heritability figures on Table 2 are consistent with this guess. To interpret the σ^2 figures on Table 2, it is helpful to recall that $1+\sigma^2$ is the variance of the shock terms $u_{Yt}+u_{It}+u_{Gt}$ of the reduced form equation (16) for Y_t. In view of the normalization $\sigma_G^2=1$, the ratio $1/(1+\sigma^2)$ is a measure of the genetic contribution to the variance of the Y_t contemporaneous shock term—a short-run heritability of sorts. The tabled values of $1/(1+\sigma^2)$ run about the same as var (G_t)/var (Y_t). In this sense, one might say that short- and long-run heritability are about the same, which does not appear to violate any intuitive expectations. The tabled values of $\alpha_2+\beta_2$ add slightly to the credibility of specification (20) by not going negative. The tabled values of R_Y^2 are consistent with evidence from the economic literature. Regressions of income on long lists of explanatory variables yield R^2 in the neighborhood of .35. Since R_Y^2 is based on an equation with only two explanatory variables (father's potential IQ and father's income), R_Y^2 might be expected to run substantially less than .35, as it does.

IV. A Redistributive Tax Application

Let Y_t represent dollar income after all existing government tax and spending mechanisms have been accounted for. Now suppose the government adds a linear redistributive income tax, specified so taxes are negative below the mean income and positive above the mean income. Such a specification assures that the net revenue of the tax is zero; it is purely a redistributive tax. If r is the marginal tax rate and Y_t^a is income after this additional tax, then the equation $Y_t^a=rE(Y_t)+(1-r)Y_t$ describes the tax; taking expected values yields the zero net revenue condition $E(Y_t^a)=E(Y_t)$. The entire model naturally rewrites as

$$
\begin{aligned}
Y_t^a &= rE(Y_t) + (1-r)Y_t \\
Y_t &= I_t + \alpha_2 Y_{t-1}^a + u_{Yt} \\
I_t &= G_t + \beta_2 Y_{t-1}^a + u_{It} \\
G_t &= G_{t-1} + u_{Gt}
\end{aligned}
$$

where the units normalizations $\alpha_0=\beta_0=\gamma_0=0$ and $\alpha_1=\beta_1=1$ have been used. The new expressions for var (Y_t), $r(Y_t, Y_{t-1})$, and R_Y are the same as the old except that $\alpha_2+\beta_2$ must be replaced by $(1-r)(\alpha_2+\beta_2)$ wherever it appears. Since $(1-r)(\alpha_2+\beta_2)$ now plays the role of $\alpha_2+\beta_2$, it follows from (14), (18), and (19) that the imposition of the redistributive tax both reduces inequality and increases social mobility. To get a notion of magnitude, let specifications (20) hold with $r(Y_t, G_t)$ fixed at .5. Then $\alpha_2+\beta_2=.29$, $\gamma=.5$, and $\sigma^2=4.90$. For

This content downloaded from 194.27.101.122 on Wed, 18 Dec 2019 08:49:48 UTC
All use subject to https://about.jstor.org/terms

these parameter magnitudes, the following measures of mobility and inequality follow for $r=0$ and $r=.15$.

	$r(Y_t, Y_{t-1})$	R_Y^2	var (Y_t)
$r=0$	.40	.19	7.32
$r=.15$	.36	.16	7.05

V. Conclusion

Though the model has been kept simple, the linear form can accommodate a large number of additional variables and still be numerically operational. The data problem cannot easily be overcome so long as the model involves parent-child observations on intelligence variables (especially an unmeasurable one like G_t). However, a useful model of this type could be constructed for more available data. For example, a cross-family sample of data on income, occupation, education, IQ, wealth, family size, and other variables is not hard to collect. Matched data for the head's parents would not be available for all variables; but the head could provide enough of them (parental occupation, education, family size, and others) to allow sensible modelling of intergenerational mechanisms.

As it stands, the model has no provision for the special social immobilities which face various subgroups of society (ethnic minorities for example). A straightforward way to handle such immobilities is to specify different parameter values for the various subgroups. The model would then generate a separate income distribution for each subgroup.

Another natural extension of the model, given recent interest in random coefficient regression models, is to add a (zero mean) shock term to each coefficient of the model as well as to each equation of the model. In the case of n equations (and thus n^2 coefficients), this would add $n^2(n^2+1)/2$ additional parameters to the model—the variances and covariances of the n^2 coefficient shocks. Thus, in the general case, the extension would be expensive in number of parameters. However, one special case may be of interest. In most of the expressions of interest above, the crucial parameters α_2 and β_2 appear only as the sum $\alpha_2+\beta_2$. Suppose this sum is subjected to a serially uncorrelated, additive shock of constant variance; while other coefficients remain unshocked. Then, as the Appendix shows, all of the derivatives discussed in Section II retain their signs. That is, relative to the issues raised in this paper, the addition of a shock to $\alpha_2+\beta_2$ leaves the model qualitatively unchanged.

Finally, the model's answer to the title question and to the opening quotes may be summarized. Opportunity equalization may increase or reduce social mobility, depending on the type of equalization. Either way, however, opportunity equalization will reduce inequality. Therefore, the opening quotes seem unreasonably pessimistic. If opportunity equalization reduces inequality, it is hard to see how this equalization will "give us . . . ever greater separation between the top and the bottom." And, even if opportunity equalization does reduce social mobility, a "virtual caste system" seems unlikely. In the extreme case of total opportunity equalization (represented in the model by $\alpha_2=\beta_2=\sigma^2=0$), the immobility measures $r(Y_t\ Y_{t-1})$ and R_Y would reduce to $r(G_t, G_{t-1})$; and psychological evidence suggests a value for $r(G_t, G_{t-1})$ of about .5 (see Section III). A society where only 25 percent of children's income variance was explained by parents' incomes would not seem well described by virtual caste system.

Appendix

Derivation of Equilibrium Expressions

The model of the text is a special case of the following linear difference equation system with additive shock term.

This content downloaded from 194.27.101.122 on Wed, 18 Dec 2019 08:49:48 UTC
All use subject to https://about.jstor.org/terms

$$\text{(A1)} \qquad y_t = Ay_{t-1} + u_t, \qquad E(u_t) = 0,$$

$$E(u_t u_s') = \begin{cases} U & \text{if t} = s \\ 0 & \text{if t} \neq s \end{cases}$$

Here A is an (n, n) constant matrix; U is an (n, n) nonnegative definite constant matrix; and y_t and u_t are $(n, 1)$ random variable vectors at time t. The system (A1) is defined on a typical family tree. The elements of y_t are measurements on n characteristics of the t-th generation of the family (income, IQ, and so on). The distribution of these characteristics over all families is generated by independent drawings on the shock vector u_t.

Corresponding to the stochastic system (A1) are two nonstochastic systems, one n equation system for the mean of y_t, call it μ_t, and one n^2 equation system for the variance matrix of y_t, call it S_t:

$$\text{(A2)} \qquad \mu_t = A\mu_{t-1}$$

$$\text{(A3)} \qquad S_t^* = U^* + (A \otimes A)S_{t-1}^*$$

Here an asterisk indicates the "stacking" operation; for example, S_t^* is the $(n^2, 1)$ vector gotten by stacking the columns of S_t (first column on top, second column next, and so on).

Taking expectations of (A1) leads immediately to (A2). Taking the variance matrix of (A1) leads to $S_t = AS_{t-1}A' + U$, and (A3) is simply the stack of this. The condition that the eigenvalues of A be less than one in modulus is sufficient for the existence of unique equilibria $\mu = 0$ and $S^* = (I - A \otimes A)^{-1}U^*$ for (A2) and (A3), and is necessary and sufficient for their global stability. (The coincidence of asymptotic properties for (A2) and (A3) follows from the theorem that the n^2 eigenvalues of the Kronecker product $A \otimes A$ equal the n^2 pairwise products of the eigenvalues of A.) Letting R and R_{-1} be the equilibrium correlation matrices of y_t with itself and with y_{t-1}, respectively, the relevant formulas for the magnitudes in the text are

$$\text{(A4)} \qquad S^* = (I - A \otimes A)^{-1}U^*,$$
$$R = WSW, \qquad R_{-1} = WASW$$

Here $W = \text{diag}\,(s_{11}^{-1/2}, \ldots, s_{nn}^{-1/2})$, where the s_{ii} are the diagonal elements of S.

The three-equation model (1)–(3) in the text can be reduced to a two-equation model in Y_t and G_t by using (2) to substitute out I_t. The resulting system is of the form (A1) with $y_t = (Y_t, G_t)'$ and

$$\text{(A5)} \qquad A = \begin{pmatrix} \alpha_2 + \beta_2 & \gamma \\ 0 & \gamma \end{pmatrix}$$

$$U = \begin{pmatrix} 1 + \sigma^2 & 1 \\ 1 & 1 \end{pmatrix}$$

Here the units normalizations $\alpha_1 = \beta_1 = \sigma_G^2 = 1$ and the assumptions $\sigma_{YG} = \sigma_{IG} = 0$ have been used to simplify. The eigenvalues $\alpha_2 + \beta_2$ and γ of A lie between zero and one by assumptions (4); hence the model has a unique globally stable equilibrium. Plugging (A5) into (A4) and manipulating yields the equilibrium expressions var (Y_t) and var (G_t) as the diagonal elements of S, $r(Y_t, G_t)$ as the off-diagonal of R, and $r(Y_t, Y_{t-1})$ and $r(G_t, G_{t-1})$ as the diagonals of R_{-1}. The equilibrium expressions involving I_t can then be derived using equation (2).

Even though the condensed system of (A5) is only of size $n = 2$, the matrix to be inverted in (A4) is of size $n^2 = 4$, which would already be too large for explicit inversion in terms of parameters if A were not triangular. This explains the comment in the text that even slightly greater complexity in the model would prevent explicit parametric statement of equilibrium magnitudes.

Now suppose equation (A1) is expanded to allow random coefficients. Specifically, let

$$\text{(A6)} \qquad y_t = (A + Z_t)y_{t-1} + u_t$$

where Z_t is an (n, n) matrix of zero mean coefficient shocks at time t. Let each element of Z_t be serially uncorrelated with itself, with all other elements of Z_t, and with all elements of u_t. And let each element of Z_t have constant variance, and constant covariance with all other elements of Z_t and all elements of u_t. For this expanded model, my 1972 paper showed that the equilibrium variance matrix of y_t, call it S_+, is given in stacked form by

$$\text{(A7)} \qquad S_+^* = (I - A \otimes A - C)^{-1}U^*$$

Here C is (n^2, n^2) constant matrix displaying

This content downloaded from 194.27.101.122 on Wed, 18 Dec 2019 08:49:48 UTC
All use subject to https://about.jstor.org/terms

the variances and covariances of the elements of Z_t in the form $C=E(Z_t \otimes Z_t)$ (which is not usual variance matrix form). The necessary and sufficient global stability condition for the equilibrium (A7) is that every eigenvalue of $A \otimes A + C$ have modulus less than one. The formulas for the contemporaneous and one lag correlation matrices, call them R_+ and $(R_+)_{-1}$, are similar to those in (A4).

$$\text{(A8)} \qquad R_+ = W_+ S_+ W_+ \quad \text{and} \quad (R_+)_{-1} = W_+ A S_+ W_+$$

where $W_+ = \text{diag}\,(s_{+11}^{-1/2}, \ldots, s_{+nn}^{-1/2})$. In the special case when a_{11} is the only element of A with a positive variance shock term, C takes the form $c_{11}E_{11}$, where E_{11} is an (n^2, n^2) matrix of all zeros except for a one in the upper left position. Then (A7) manipulates to the form

$$\text{(A9)} \quad S_+^* = \{I + [c_{11}/(1 - a_{11}^2 - c_{11})]E_{11}\} \cdot (I - A \otimes A)^{-1} U^* = \{I + [c_{11}/(1 - a_{11}^2 - c_{11})]E_{11}\} S^*$$

where, as above, $S^* = (I - A \otimes A)^{-1} U^*$ (the value of S^* when $C=0$).

The formulas of the last paragraph may be applied to the special case discussed in the concluding section of the text. In this special case, $n=2$; A and U are given by (A5); and C takes the form $C=c_{11}E_{11}$, where c_{11} is the variance of the shock term added to $a_{11}=\alpha_2+\beta_2$. To assure stability, the parameter restrictions of the text must be strengthened to include the restriction $(\alpha_2+\beta_2)^2+c_{11} < 1$. Plugging into formulas (A7), (A8), and (A9) yields the equilibrium magnitudes $\text{var}\,(Y_t)_+$, $r(Y_t, Y_{t-1})_+$, and $(R_Y^2)_+$ for the expanded model.

$$\text{(A10)} \quad \text{var}\,(Y_t)_+ = \text{var}\,(Y_t)[1 - (\alpha_2 + \beta_2)^2] / [1 - (\alpha_2 + \beta_2)^2 - c_{11}]$$

$$r(Y_t, Y_{t-1})_+ = \{r(Y_t, Y_{t-1})[1 - (\alpha_2 + \beta_2)^2 - c_{11}] + c_{11}(\alpha_2 + \beta_2)\} / [1 - (\alpha_2 + \beta_2)^2]$$

$$(R_Y^2)_+ = 1 - (1 + \sigma^2)/\text{var}\,(Y_t)_+ = 1 - (1 - R_Y^2)[1 - (\alpha_2 + \beta_2)^2 - c_{11}] / [1 - (\alpha_2 + \beta_2)^2]$$

where $\text{var}\,(Y_t)$, $r(Y_t, Y_{t-1})$, and R_Y^2 are the expressions given by (5), (8), and (17) in the text. The derivatives of each line of (A10) with respect to $\alpha_2+\beta_2$ and σ^2 take the same signs as the corresponding derivatives in the text for the nonexpanded model ((14), (15), (18), (19)).

REFERENCES

P. M. Blau and O. D. Duncan, *The American Occupational Structure*, New York 1967.

J. Conlisk, "An Approach to the Theory of Inequality in the Size Distribution of Income," *Western Econ. J.*, June 1969, *7*, 180–6.

———, "A Bit of Evidence on the Income-Education-Ability Interrelation," *J. Hum. Resources*, summer 1971, *6*, 358–62.

———, "Stochastic Stability of Random Coefficient Models," mimeo. 1972.

B. Duncan, "Education and Social Background," *Amer. J. Sociology*, Jan. 1967, *72*, 363–72.

L. Erlenmeyer-Kimling and L. F. Jarvik, "Genetics and Intelligence: A Review," *Science*, Dec. 13, 1963, *142*, 1477–8.

R. Herrnstein, "I.Q.," *Atlantic*, Sept. 1971, *228*, 43–64.

A. R. Jensen, "How Much Can We Boost IQ and Scholastic Achievement?" *Harvard Educ. Rev.*, winter 1969, *39*, 1–123.

H. E. Jones, "The Environment and Mental Development," in L. Carmichael, ed., *Manual of Child Psychology*, 2d ed., New York 1954.

J. B. Miner, *Intelligence in the United States*, New York 1957.

J. N. Morgan, M. H. David, W. J. Cohen, and H. E. Brazer, *Income and Welfare in the United States*, New York 1962.

Newsweek, Aug. 23, 1971, p. 57.

Time, Aug. 23, 1971, p. 33.

This content downloaded from 194.27.101.122 on Wed, 18 Dec 2019 08:49:48 UTC
All use subject to https://about.jstor.org/terms