GREEK ECONOMIC REVIEW
Vol. 1, No. 1, August, 1979

INTERGENERATIONAL TRANSFERS AND INEQUALITY

By D. L. Bevan and J. E. Stiglitz *

I. Introduction

Why are the rich so rich? And why are the poor so poor? This question has puzzled thinkers for millenia.[1] Yet our understanding of the processes which generate wealth inequality is still extremely limited. Some view inequality as an inherent consequence of the necessity to provide incentives in market economies; attempts to reduce inequality may have, in this view, significantly deleterious effects on the level of national income. Others view inequality as resulting from monopolies (often government-granted), in which case the elimination of monopolies would both increase equality and economic efficiency. Still others view inequality as arising simply out of random processes (luck); there would neither be a gain or a loss in efficiency from reducing it. Thus, the effect of any particular policy measure designed to reduce the degree of inequality will depend on the underlying structural determinants of inequality, and it is important to obtain a better understanding of these.

This task is, however, a difficult one. The data are extremely sparse and empirical work requires considerable perseverance as well as ingenuity. This paucity of evidence does of course leave the theorist a field uncluttered by too many awkward facts; he can be fairly uninhibited in his construction of models. Nonetheless, working models of inheritance are almost as rare as good data. Recently, however, there have been several attempts to remedy this deficiency; this paper attempts to survey and put into perspective a number of these attempts. Most of these models do not, at least as presently formulated, seem completely consistent with the "stylized facts" of wealth inequality. We conclude our discussion with some conjectures on why this may be so and on some suggestions for future research.

* St. Johns and All Souls Colleges of Oxford, respectively. The authors would like to express their appreciation to John Flemming and Tony Atkinson for helpful discussions. Research support from the National Science Foundation is also gratefully acknowledged. Part of the paper was written while Stiglitz held the Oskar Morgenstern Distinguished Fellowship at Mathematica and was visiting professor at the Institute for Advanced Studies, Princeton. An earlier version of the paper was presented at the American Economic Association Meetings, Chicago, August, 1978.

[1] Although, with a sufficiently low discount rate, there is a view that the present discounted value of "welfare" inequality is less than it appears at present. Cf. "The meek shall inherit the earth."

II. Life Cycle Theories of Wealth Distribution

There are two broad categories of models for explaining the wealth distribution: life cycle models and inheritance models[1]. As is well known, differences between the patterns of consumption and income generate patterns of wealth accumulation and decumulation. Then, the distribution of life cycle savings will depend on the age distribution of the population, as well as on the distribution (for any given age) of differences in the patterns of lifetime consumption and earnings. Although the magnitude of this inequality should presumably be directly related to the magnitude of differences in wages[2], it is likely to be affected by a number of features of the economy:

(*a*) The introduction of a social security program should reduce the amount of life cycle savings[3]. Since benefits increase less than proportionately to lifetime incomes, this should have a more marked effect on the poor, thus increasing observed wealth inequality. If, however, we include the (expected) present discounted value of social security payments in the measure of an individual's wealth, the social security program probably decreases wealth inequality.

(*b*) The introduction of privately funded pension schemes will have similar effects.

(*c*) More generally, since in the absence of annuities, individuals have to save against the contingency that they live a long time, improvements in annuity markets are likely to decrease life cycle savings. Since annuities are mainly purchased by the wealthy, this may decrease wealth inequality.[4]

[1] For an early attempt to see the implications of the life cycle model for the wealth distribution, see Stiglitz (1966).

[2] But inequality in life cycle wealth, as measured by (say) the coefficient of variation, will not, even for a given cohort, be precisely equal to that of observed wages. First, individuals differ in their savings rate; secondly, the distribution of lifetime wages is likely to have a lower coefficient of variation than the distribution of wages observed at any moment of time; and thirdly, savings rates are not necessarily constant, independent of income levels. The first effect would imply that wealth inequality is greater than wage inequality, the second that it is smaller, and the third is ambiguous, depending on whether savings rates increase or decrease with (lifetime) income.

[3] Unless retirement ages are reduced significantly.

[4] If individuals knew precisely how long they would live, then annuities would make no difference; individuals could as well simply spread their wealth over their retirement years. But since death is stochastic, individuals always have to keep a reserve against the contingency that they live a long time. The problem (after retirement) is formally analogous

(*d*) The availability of better rental markets for housing and the ability to reduce equity holdings in housing by obtaining a mortgage should again reduce the need for savings for retirement. The effect of this on the wealth distribution is ambiguous: if the income elasticity of housing is less than unity, there will be a proportionately smaller reduction in wealth holding of the wealthier, and this will tend to increase inequality; but to the extent that there is a larger proportion of home owners among the wealthy, this will tend to decrease inequality.

Clearly, some of wealth inequality is due to life cycle inequality. The difficulty is assessing how much. In principle, this is an easy question: all we have to do is to find out the distribution of incomes (over lifetime) and consumption, and calculate from that the distribution of life cycle savings. To do this, properly, of course, would require following a large panel of individuals over their lifetime. Since this has not been done so far, we must rely on indirect evidence for assessing the relative importance of life cycle savings. This is of three sorts: (a) How much of aggregate savings can be explained by a life cycle model? Although several such calculations have been made, (See, e.g., Tobin (1967), Farrell (1970)), they required such strong assumptions as to cast doubt on the credence of any results obtained; (b) Has the wealth distribution changed qualitatively in the way predicted by changes in the characteristics of the economy, of the kind discussed above? Although some changes have occurred, they do not (yet) seem to be as large as they would be if life cycle savings were the only (or primary) determinant of the wealth distribution; (c) Can the life cycle model explain the tail of the wealth distribution, the large percentage of wealth owned by the upper 1-2% of the population? The answer seems to be no, not under any plausible hypothesis concerning the distribution of lifetime wages and savings patterns. (See Flemming 1976)

III. Inheritance Models: Unplanned Bequests

Thus, to understand wealth inequality, some reliance must be placed on models of inheritance. We distinguish here between two kinds of such models:

to the optimal depletion of a stock of a natural resource, when a substitute with zero cost of production will be discovered, at some unknown date in the future. The question is, what effect does uncertainty about the date of death have on the amount of wealth accumulated at the date or retirement? The presumption is that this increases the demand for wealth, but there may be conditions under which the reverse happens. For a discussion within the context of natural resources, see Dasgupta and Heal (1974) and Stiglitz and Dasgupta (1978).

those in which bequests are planned and those in which they are unplanned.

There are two important reasons that individuals have unplanned bequests. First, because of the imperfect annuity markets referred to earlier, and because death is stochastic, at the time of death they will always have some wealth. Secondly, because of imperfect rental markets for housing, and other durable assets, some individuals at death will own houses (or other assets). Since the magnitude of the inequality of wealth is much larger than can be explained by such assets themselves, this may contribute to inequality, but is clearly not important in explaining the tail of the distribution.

The imperfect annuity market can, however, give rise to significant inequality. For individuals who inherit more (because their father died early) will wish to save more for their own retirement, spreading the increased consumption completely over their life. If they have, say, a constant elasticity utility function, they will have, at each date, proportionately more wealth than their friends who inherited nothing. Thus, if they too die early, their children will inherit proportionately more. This will give rise to a wealth distribution which is Pareto in the tail. (See Stiglitz (1977)) Moreover, the degree of inequality will be increased if the size of families, among whom the wealth is divided, is also stochastic. This model can, we suspect, go a long way in explaining the degree of wealth inequality.

The extent that wealth inequality arises out of life cycle savings or unplanned bequests has some important implications for inheritance tax policy: for, *to the extent that inheritances are unplanned, the imposition of an inheritance tax will have no incentive effects* (*e.g.*, *on work effort or risk taking*) *and such taxes are non-distortionary.*

IV. Incidence of Non-Distortionary Inheritance Taxes: Balanced Growth Path Incidence

The fact that such a tax is non-distortionary does not mean that it has no effect on the economy. If the average savings rate out of inheritances differs from that out of income as a whole, an inheritance tax[1] will change the aggregate savings rate of the economy; this change in the aggregate savings rate will change factor prices and the distribution of income between capital and labour. This, in turn, may have a significant effect on the distribution of income and wealth in long-run equilibrium. (See Stiglitz (1966, 1977, 1978))

[1] Keeping government expenditures fixed, and compensating for the increased revenue from the inheritance tax by a *uniform* decrease in income taxes.

For instance, if the savings rate of the wealthy is higher than the average of the population, the decrease in the equilibrium supply of capital will, if unskilled labour has a higher degree of complementarity with capital than does skilled labour, increase relative wage differentials, and this will increase the inequality in (lifetime) wealth. The fact that an inheritance tax could actually increase inequality because of the change in the capital-labour ratio is not necessarily to be taken as an argument against it; for there may be other instruments available to the government to offset this, in particular, government monetary and debt policy. Stiglitz (1978) has argued, accordingly, that in assessing the effects of such taxes, one should employ balanced growth path incidence, where the effects on capital accumulation are neutralized by appropriate government policy.

V. *Planned Bequests*

Although some bequests are of the unplanned variety, many bequests are planned, and the remainder of this paper is devoted to exploring the implications of planned bequests.[1]

First, we need to ask, why would an individual plan to leave something to his heirs? Even if individuals are altruistic, and have a utility function in which the welfare of their immediate heirs enters directly, say in the form:

$$(1) \qquad U_t = u_t + \left(\frac{1+n}{1+\delta}\right) U_{t+1} = \sum_{r=t}^{\infty} \left(\frac{1+n}{1+\delta}\right)^{r-t} u_r$$

where u_t is the utility associated with the direct consumption of the t^{th} generation, U_t is the (total) utility of the t^{th} generation, $(1+n)$ is the number of children an individual has and δ is the discount rate, then an individual would only leave a bequest if, in the absence of the bequest, he believed that his heirs would be worse off than himself. But with technological progress, the descendants are likely to be better off. Thus, planned bequests can only be generated if

a) One expects one's descendants to have a lower wage, i.e., there is a wage distribution, and one's children's relative wage (relative to the mean) is not only lower than one's own, but sufficiently lower to compensate for the secular increase in wage rates.

[1] For a more extensive treatment of some of the topics discussed in this section see Bevan (1974, 1979) and Stiglitz (1977).

b) One has inherited some wealth, the benefits of which, if one is altruistic, one desires to share with one's descendants. But one inherits wealth only under two conditions: (*a*) unplanned bequests from one's parents as described in Section 3; (*b*) planned bequests.

There are two important implications of this line of reasoning: If all bequests are planned, then the distribution of bequests (inheritances) must simply be related to the wage distribution; but even if there were no wage differentials, then the occurence of some unplanned bequests would, in subsequent generations, give rise to planned bequests. More generally, in planning one's pattern of life cycle savings, with stochastic death, the realization that one's children will benefit from an early death may have an important effect on the levels of savings. In such situations, even with unplanned bequests, an inheritance tax has a distortionary effect and may have a significant effect on aggregate savings, possibly more significant than when inheritances are totally planned[1].

To the extent that bequests are planned, they represent intra-family redistribution; thus, *taxing inheritances may actually increase inequality in consumption;* this can be seen most simply by considering a polar case; assume that the interest rate is zero and the growth rate is zero (or more generally, the interest rate is equal to the growth rate); then if inheritances are abolished, the distribution of lifetime consumption will be equal to the distribution of lifetime incomes, while with inheritances, the inequality of lifetime consumption will be smaller, because of intra-family redistribution.

If the interest rate exceeds the growth rate, then an individual who happens to receive a higher than average wage can sustain a permanently higher level of consumption for his descendants, and thus with inheritances, the inequality in consumption may exceed the inequality of wages[2].

[1] The reason for this is that, with completely planned bequests, motivated by an attempt to redistribute income between oneself and one's less fortunate heirs, there is an income effect and a substitution effect from the tax. To leave one's heirs with the same level of welfare requires, before tax, greater savings. There are, as usual, the income and substitution effects which operate in opposite directions. But with stochastic death, the substitution effect would seem to dominate: consider, for instance, individuals who have inherited nothing. They would plan to leave nothing to their heirs; part of the expected return, however, from savings is the expectation that, if one does die early, one's savings will be of some value, although less than if one survived. An inheritance tax reduces the value of savings, and thus discourages lifetime savings.

[2] For instance, those whose wages exceeded mean wages could always buy a perpetuity which grew at the rate of growth of the population (so *per capita*, all heirs would have their consumption increased by the same amount.)

Thus, to assess the effect of inheritance taxation on the distribution of wealth and consumption, we need explicitly to formulate a model exploring the interactions between wages and bequests.

VI. Wealth Distribution Arising From Planned Bequests

We now attempt to formulate a model in which the distribution of wealth arises completely from the planned decisions of altruistic individuals towards their descendants. There are two parts to such a model:

a) First, we need to relate the wages (or ability) of the child to that of the parent. Two polar models have been investigated extensively in the literature, one where the child's ability is unrelated to that of the parent, and the other where it is perfectly correlated. (Cf. Stiglitz (1969)) A more reasonable specification, however, would take account of the stochastic nature of the relationship between the ability of the parent and the ability of the child.

If we describe the economy as if there were a discrete number of ability levels (wage groups), the ability (wage) distribution at time $t+1$ is related to that at time t by

$$x_{t+1} = Dx_t \tag{2}$$

where x_t is the vector describing the ability distribution, i.e. x_{it} is the proportion of the population at date t in the i^{th} ability group and D is a Markov matrix; d_{ij} specifies the proportion of those whose parents are in ability level i who will be at ability level j. The steady state ability distribution is given by the solution to

$$x^* = Dx^*$$

and it can be shown that, under fairly weak conditions,

$$\lim_{t \to \infty} x_t = x^*$$

The nature of the wage distribution can be related to properties of the matrix D. Much of the earlier literature on stochastic models of wealth and income distribution was essentially concerned with finding restrictions on the matrix D which ensured that the distribution which was generated had particular properties, e.g. had a Pareto tail. (Cf. Mandelbrot (1961), Champernowne (1953)) Usually little attempt to provide an explanation of why these conditions should be satisfied in any economic model was made.

The approach which we describe below represents an attempt to relate the notion of "regression towards the mean" to the long-run ability (wage) distribution.

We postulate that[1]

(3) $$A_{t+1} = \beta_t(A_t) + z_t(A_t), \quad Ez_t = 0$$

the ability, A, of the child is related to that of his parent, but there is "noise" in the relation, described by the random variable z.

If β_t is linear and constant over time, (3) becomes

(3') $$A_{t+1} = \beta A_t + \nu + z_t$$

Taking expectations, we can derive a difference equation for mean ability,

(4) $$\bar{A}_{t+1} = \beta \bar{A}_t + \nu$$

Thus mean ability converges to

(5) $$\bar{A}^* = \frac{\nu}{1-\beta}$$

More interesting, however, is the fact that we can derive a similar difference equation for the Variance of A, if

$$E(z_t - \bar{z}_t) = \sigma_z^2 \text{ for all } t.$$

(6) $$V_A^{t+1} = E(A_{t+1} - \bar{A}_{t+1})^2 = \beta^2 E(A_t - \bar{A}_t)^2 + \sigma_z^2$$

$$= \beta^2 V_A^t + \sigma_z^2$$

[1] This specification is Markovian: it is only the parents' ability which affects the child. This seems reasonable, but more general specifications can also be handled. The technique employed below follows closely that used in Bevan (1974).

We can again solve for the steady state solution

$$V_A^* = \sigma_z^2/(1-\beta^2) \tag{7}$$

Hence we obtain the coefficient of variation:

$$\frac{\sqrt{V_A^*}}{A^*} = \frac{\sigma_z}{v}\sqrt{\frac{1-\beta}{1+\beta}} \tag{8}$$

Notice that (3′) can be rewritten as

$$A_{t+1} = \beta A_t + (1-\beta)\overline{A^*} + z_t \tag{3''}$$

and β determines the speed with which ability regresses towards the mean.

Even if it were thought that the difference equation was basically genetically determined, it could be affected by social policy; for instance, coeducational residential colleges may increase the tendency for assortive mating (more able males marrying more able females). This may lead to an increase in β, i.e. the ability of the son is likely to be more highly correlated with the ability of his father. Note that such a policy change will increase the variance of ability.[1] (Conversely compensatory educational policies may reduce the variance in wages, if wages are just a function of ability.)

If we postulate further that

$$A = \ln w \tag{9}$$

and that z is normally distributed, then it immediately follows that in steady state, A will be normally distributed, and wages will be lognormally distributed.

For many purposes, it is easier to work with a continuous approximation to the difference equation (3); thus, we write:

$$dA = \beta(A)\,dt + z\,dt \tag{10}$$

where dA represents the change in ability level (wage). We continue to assume that the "noise" is normally distributed; more specifically, we postulate that (10) is a diffusion process. This allows us to solve explicity for the distribution of wages at each date, and, in particular, for the steady state distribution. This steady state distribution depends on the two functions $\beta(A)$ and $z(A)$. If we set $A = \ln w$ then, if $\beta(A)$ and $z(A)$ are both constants,

[1] Assuming σ_z is invariant.

ndependent of A, we obtain a lognormal wage distribution. (Stiglitz, 1977) This is the continuous time analogue to the earlier result.

This model is consistent with the observed patterns of regression toward the mean and indeed it was this which motivated the original formulation of the model; for our purposes, it makes no difference whether this regression is related to genetic or environmental factors, or to mating patterns which are imperfectly assortive.

(*b*) The bequest behavior. The bequest of a parent will be a function of his lifetime wage, his inheritance, and the return he has received on his capital[1].

$$B_t = B(w_t,\ B_{t-1},\ r_t) \tag{11}$$

(11) is a dependent stochastic process, i.e., the stochastic process for w, given the function B, generates a stochastic process for B_t. Thus, (11) and (3) together with a relation between ability and wages, generate a steady state distribution $F(w, B)$ of wages and wealth.

In principle, the bequest function can be derived from the expected utility maximization behavior of the individual, knowing the stochastic process generating the wage distribution, and knowing that his heirs will be solving a problem analogous to the one he is presently solving.[2] If there are borrowing constraints (since the low ability children would like to borrow against the future higher income of their descendants) these ought to be taken into

[1] This obviously oversimplifies, by assuming that we can reduce the individual's wage earnings to a single number, w, and similarly for his return on capital. If there were no uncertainty, or if there were perfect risk markets, this could clearly be done; but more generally, bequest behaviour will depend on the wage received at each date of his life and on the return received on each investment made.

[2] Some readers of earlier versions of these models have expressed concern over the fact that, in this formulation, parents do not know the ability of the child before they make their bequests. If they did, then the bequest would be a function of the ability (wage) of the child as well as of the parent. This complicates the analysis, but will not alter its structure: it will still be impossible for an individual to know the ability of his children's children (or grandchildren). However, this modification would imply (when the bequest function is derived from an expected utility maximization problem) that parents ought to leave different amounts to children whose abilities differ. There is some fragmentary evidence that parents do this to a limited extent at most. This can be reconciled with the model presented here in several ways: (a) since wages are an imperfect surrogate of ability, if parents distributed according to true ability (which they knew) there might still be only a weak statistical relation between wages and bequests; (b) parents may be concerned with the incentive effects associated with compensatory bequests; (c) parents may feel a 'moral obligation' to treat their children alike (just as firms often treat workers whose productivity differs the same) but they still may take into account the average level of ability of their children in determining the (uniform) bequest to give their heirs.

account in the solution. This turns out to be a difficult problem to solve in general, and so we have (reluctantly) taken two alternative approaches. The first is to investigate the consequances of certain simple rules, e.g., linear bequest functions; the second is to assume that the individual solves the optimal bequest problem, slightly incorrectly, substituting the mean value of each of the variables for the distribution, (i.e., as if, in equation (3), he postulated that $z_t = 0$).

If we postulate that (11) is a linear function (as before), then we may write [1,2]

$$B_t = s_1 w_t + s_2 B_{t-1} + s_3 r_t B_{t-1} \tag{12}$$

We again take expectations to obtain the difference equation for the mean of B,

$$\bar{B}_t = s_1 \bar{w}_t + s_2 \bar{B}_{t-1} + s_3 E r_t B_{t-1} \tag{13}$$

If we now postulate that the distribution of r_t is uncorrelated with the wealth of the individual and invariant over time, we obtain that the mean of B converges to

$$\bar{B}^* = \frac{s_1 \bar{w}^*}{1 - s_2 - s_3 \bar{r}} \tag{14}$$

Similarly, we can derive a difference equation for the covariance of wealth and wages:

$$E(B_t - \bar{B}_t)(w_t - \bar{w}_t) = E\{[s_1(w_t - \bar{w}_t) + s_2(B_{t-1} - \bar{B}_{t-1}) + s_3(r_t B_{t-1} - \bar{r}\bar{B}_{t-1})](w_t - \bar{w}_t)\} \tag{15}$$

To solve this, we have to recall our difference equation for w_t or A_t. Although the more natural specification is (3'), with z assumed normal, and the relationship between A and w given by (9), to solve (15) we need to assume

$$w_{t+1} = \hat{\beta} w_t + \hat{\nu} + z_t \tag{3'''}$$

[1] This assumes that the effect of r_t depends on the wealth the individual has invested; in this simple model, the only relevant wealth which is invested is his inheritance.

[2] We have written this equation without an explicit stochastic term (other than through r). This is a convenient simplification; the reader may easily introduce such a term, which may for instance represent variations in the number of children. Randomness in savings may also be easily incorporated.

(This can be thought of as a linearization of the difference equation obtained by si bstituting (9) into (3').

Then (15) can be rewritten

(16) $$\operatorname{cov}(B_t, w_t) = s_1 V_w^t + \hat{\beta}(s_2 + s_3 \bar{r}) \operatorname{cov}(B_{t-1}, w_{t-1})$$

wher : we have made use of the fact that

$$w_t - \bar{w}_t = z_{t-1} + \hat{\beta}(w_{t-1} - \bar{w}_{t-1})$$

and where we have assumed

$$\operatorname{cov}(z_{t-1}, B_{t-1}) = 0$$

$$\operatorname{cov}(r_t, z_t) = 0$$

Moreover

(17) $$V_B^t = s_1^2 V_w^t + s_2^2 V_B^t + s_3^2 (\sigma_r^2 E B_t^2 + \bar{r}^2 V_B^t) + 2\hat{\beta} s_1 (s_2 + s_3 \bar{r}) \operatorname{cov}(B_{t-1}, w_{t-1})$$

Hence, the variance and covariance converge to

(18a) $$\operatorname{cov}^*(B, w) = \frac{s_1 V_w^*}{1 - \hat{\beta}(s_2 + s_3 \bar{r})}$$

(18b) $$V_B^* = \frac{V_w^* s_1^* \left(1 + \frac{2\hat{\beta}(s_2 + s_3 \bar{r})}{1 - \hat{\beta}(s_2 + s_3 \bar{r}} + \frac{s_1^2 s_3^2 \bar{w}^2 \sigma_r^2}{(1 - s_2 s_3 \bar{r})^2}\right.}{1 - s_2^2 - s_3^2 E r^2 - 2 s_2 s_3 \bar{r}}$$

(7') $$V_w^* = \frac{\sigma_z^2}{1 - \beta^2}$$

This completely describes the asymptotic distributions. Policies can affect the savings parameters (s_1, s_2, and s_3), and the variance of wages and r (redistributive taxation). In the long run, these changes may give rise to changes in the aggregate capital stock (eq. 14) and, in a closed economy, this will change factor prices and shares and hence inequality. These long-run general equilibrium effects may not only alter quantitatively the results, they may actually reverse the sign, so a tax on inheritances which reduced wealth inequality in the short run may increase it in the long run.

To gain some further insight into the model, let us simplify by assuming that individuals save a given fraction of their total lifetime income, and divide it equally among their children. Then

$$B_t = \frac{s(w + rB_{t-1}) + B_{t-1}}{1+n}$$

so

$$s_1 = s_3 = \frac{s}{1+n}, \quad s_2 = \frac{1}{1+n}$$

Moreover, let us assume $\sigma_r^2 = 0$. Then (18a) and (18b) simplify to

(18a′) $$\mathrm{Cov}^*(B, w) = \frac{sV_w^*/(1+n)}{1 - \frac{\beta}{1+n}(1+sr)}$$

(18b′) $$V_B^* = \frac{V_w^* s^2}{(1+n)^2}\left(1 + \frac{2\hat\beta(1+sr)}{1+n-\hat\beta(1+sr)}\right)\Big/\left(1 - \frac{(1+sr)^2}{(1+n)^2}\right)$$

(18c) $$\varrho^2 = \frac{\mathrm{cov}(B,w)}{V_w V_B} = \frac{1 - \frac{(1+sr)^2}{(1+n)^2}}{\left(1 - \frac{\hat\beta(1+sr)}{1+n}\right)^2} \Big/ 1 + \frac{2\beta(1+sr)}{1+n-\hat\beta(1+r)}$$

If $n, s,$ and r are small, we obtain

(18c′) $$\varrho^2 \approx \frac{2(n-sr)}{1-\beta^2}$$

and

(18b″) $$V_B^* \approx \frac{V_w^* s^2 (1+\hat\beta)}{2(n-sr)(1-\hat\beta+n-sr\hat\beta)}$$

Observe that the inequality of wealth increases with the savings rate and interest rate, and decreases with the reproduction rate, and with the speed of regression towards the mean.

The variance in consumption can also be calculated, under these special assumptions, from

(19) $$C_t = (1-s)(w_t + r_t B_{t-1})$$

(20) $$V_C^t = (1-s)^2 (V_w^t + \bar r_t \hat\beta \ \mathrm{cov}(w_{t-1}, B_{t-1}) + EB_{t-1}^2 \bar r^2 + \bar B_{t-1}^2 \sigma_r^2)$$

Simplifying further, as before, by letting $s_1 = s_3 = \dfrac{s}{1+n}$, $s_2 = \dfrac{1}{1+n}$ and $\sigma_r^2 = 0$, we obtain

$$(21) \qquad V_C^* = (1-s)^2 \left[V_w^* \left(1 + \frac{\hat{\beta}\, rs}{1+n-\beta(1+sr)} \right) + \left(\frac{r^2 s^2 (1+\hat{\beta})}{2(n-sr)(1-\beta+n-sr\hat{\beta})} \right) + r^2 \left(\frac{s\bar{w}}{n-sr} \right)^2 \right]$$

This equation makes clear the critical determinants of the variance of consumption (which we would consider, for many purposes, more relevant than inequality in income or even of wealth itself). The first of the critical parameters is the magnitude of inequality of wages. This we have related to the speed of regression towards the mean (coefficient β) and the random component in ability transmission (σ_2^2). The speed of regression towards the mean enters in a second way (besides through its effect on V_w^*) : when abilities are more highly correlated, there is more "compounding" of savings, and, if the rate of interest is positive, this increases inequality. Thus, the higher the rate of interest and the lower the rate of reproduction, the greater the inequality.

The effect of s is itself ambiguous. On the one hand, if the speed of regression towards the mean is fast and the rate of interest is low, then it serves to reduce inequality: able individuals share their good fortune with their descendents, who are likely to be less able than they. But if the rate of interest is high, relative to the rate of reproduction, and the speed of regression towards the mean is low, then it serves to increase consumption inequality.

Thus if r is small, it is clear that the variance of C is reduced as a result of inheritance compared to what it would have been without (which would have been just V_w^*) and *the abolition of inheritances would increase the inequality in consumption.* Of course, prohibitively high inheritance tax rates generate no revenue; they simply force the individual to consume his income during his lifetime. Small tax rates will generate some revenue, and may reduce inequality of consumption.[1]

As we shall note below, one has to be somewhat careful with this kind of comparative statics exercise. As we mentioned, the behavioral parameters

[1] If one's objective is minimizing the variance of consumption, there exists an optimal inheritance tax rate to do this. See Stiglitz (1977).

(e.g. savings rates) may, in fact, be affected both by governmental policies and by the non-behavioral parameters. Thus, changes in r and β may well give rise to changes in savings rates.

Similar results obtain if we use the continuous time approach. Thus, focusing on the special case with zero variability in the rate of return on capital and the rate of reproduction, and where a constant fraction of life time income is left as a bequest, we obtain[1]

$$\dot{B} = sw + (sr - n)B$$

where

$$d \ln w = -(1 - \beta) \ln w \, dt + \tilde{z} \, dt.$$

These equations can be solved simultaneously for the wage and wealth distribution. As before, to obtain solutions, a linear approximation to these equations must be taken. When that is done, results closely parallel to those for the discrete model described above are obtained. (See Stiglitz (1978))

The intractability of the theoretical model in its general form, and the necessity therefore of working with linear approximations, has led us to attempt a numerical simulation. We considered a large number of alternative specifications of the parameters of the problem (the interest rate, the utility function (which we parameterized by a constant elasticity utility function), the discount factor (δ in equation (1)), the regression factor (β), and the coefficient of variation in wages). We have also considered the effect of imposing a borrowing constraint on individuals.[2]

The results of this exercise are complex, and are described in greater detail in Bevan (1979). Here, we draw attention to five aspects of them.

First, the results are very sensitive to the value assumed for the elasticity of the marginal utility of consumption, η. For example, lowering η from 4

[1] To see this, write

$$B_{t+1} \frac{s(w_t + rB_{t-1}) + B_t}{1 + n}$$

$$\Delta B_t = \frac{s(w_t + rB_{t-1}) - nB_{t-1}}{1 + n}$$

$$\approx s(w_t + rB_{t-1}) - nB_{t-1}$$

[2] In the simulations, attention was focused on sets of parameters which yielded reasonable values of aggregate savings. This was an important restriction, implying that we had one less degree of freedom.

to 1 may raise the sensitivity of the average level of bequest to changes in the after tax rate of interest by a factor of 100.

Second, for plausible values of η, the average level of bequest is very sensitive to other parameter values, particularly to the rate of interest (a typical result with $\eta = 2$ would have a rise of one tenth of a percentage point in the interest rate raising the average bequest by an amount equal to average annual earnings).

Third, inequality in wealth is sensitive to the shape of the upper tail of the earnings distribution. Working with a lognormal distribution, it is difficult to generate a tail in the wealth distribution which is sufficiently "fat", for reasonable values of the parameters. An earnings distribution with a Pareto tail (such as the Champernowne distribution) offers a much better fit [1] (see Table).

TABLE 1

% of population in various wealth ranges

wealth range *	Simulated ** (lognormal)	Simulated ** (Champernowne)	Actual ***
10 - 25	2.80	1.34	6.38
25 - 50	1.26	1.12	1.72
50 - 100	.10	.62	.70
100 - 200	0	.14	.16
200 +	0	0	.10
Average wealth	1.18	1.29	2.12
R^2 wealth/wages	.052	.034	?

* Times median wages

** $\beta = .5$, elasticity of marginal utility of 2, $\delta = 1/3$. Lognormal has median 1, standard deviation in the logs 0.39, r = 4.5%. Champernowne has median 1, other parameter 3.37, r = 4.35%. Rate of growth of per capita earnings and population both = 2%.

*** Derived from Atkinson and Harrison's estimates for the U.K. wealth distribution, 1966.

Fourth, the relation between wealth inequality and consumption inequality is very complex; they may easily move in opposite directions. This is partly due to the rather low correlation between inheritance and earnings even when earnings are themselves heritable to a high degree.

Finally, the borrowing constraint plays a crucial role. When it is set at zero, the large majority of individuals (typically 80 - 95%) are constrained

[1] Recall that the model is required to explain the top 1-2% of the wealth distribution only, the bulk of the remainder being adequately explained by models of the life cycle type.

to leave no bequest. If it is relaxed somewhat, so that negative bequests are possible up to some (reasonable) finite amount, the majority then proceed to make the maximum negative bequest, and a very different wealth distribution is obtained. This is not a surprising result under the present assumptions: it is due partly to limited altruism but mainly to the growth in per capita earnings over time. It does mean that analytic models which ignore the constraint (such as the one outlined in our first approach) must be treated with some caution.

VII. Concluding Remarks

These models, simple as they are, do we think enhance our understanding of the process of wealth determination, and have some important implications for policy. They also raise some questions requiring further exploration.

It seems likely that there are several processes at work determining the wage and wealth distribution. Much of wealth accumulation may be for life cycle purposes, but the tail of the distribution —in which a great deal of wealth is concentrated— probably cannot be related to life cycle savings. Secondly, the tails of the wage and wealth distributions are fatter than can be explained by a conventional model of regression towards the mean.[1] A fat tail in the wealth distribution can, through the transmission mechanism we have described here, be related to a fat tail in the wage distribution; but then, how do we explain these extremely high incomes? They may reflect the fact that in certain types of activities, in particular, in certain types of inventive and entrepreneurial activities, the returns have extremely large (possibly infinite) variances. They may, however, simply be the outcome of "lotteries" with no economic effects other than on income distribution.

The point of making these distinctions is that the effect of taxation may be markedly different on different kinds of savings. For instance, the taxation of extremely large inheritances will not have the serious distortionary effects that the critics of the tax seem to claim if these are generated by windfall gains; while the consequences, both positive and negative, of inheritance taxes for upper middle income individuals, in the (log) normal part of the income distribution, may be significantly different from those for the very wealthy. The design of the inheritance tax structure ought to take this into account.

[1] There are other mechanisms which generate Pareto tails, related to primogeniture and stochastic death (unplanned bequests), but it is unlikely that these account for the observed distribution. (See Stiglitz, 1969 and 1978.)

The design of tax policy needs to take some further aspects into consideration. First, the models we have formulated are characterized not only by a steady state distribution, but by a set of stochastic difference (differential) equations. Two economies might have the same steady state distribution, but differ in these dynamic equations; they may differ, in other words, in the degree of intergenerational mobility. There is a widespread view that social mobility is good. Can this be related to more basic ethical postulates (e.g. to utilitarianism), and what implications do these dynamic aspects have for the design of policy?

Secondly, there is a widespread view that "it is better to give than to receive". Gifts, in a utilitarian context, are doubly blessed: for both the giver and the receiver obtain utility from them. Does this suggest that government policy ought to encourage these gifts, rather than discourage them?

Finally, although from a utilitarian point of view, it is natural to focus on the distribution of consumption, it is not clear that this fully captures the nature of inequality in our society; there are reasons why one may oppose inequality of wealth, even if consumption were equalized, e.g. because of the abuses of power to which it may give rise.

These are all considerations which a fuller analysis of public policy towards inheritance and the inequality of wealth would need to take into account.

REFERENCES

Bevan, D. L. (1974) "Savings, Inheritance and Economic Growth in the Presence of Earnings Inequality..." Oxford, mimeographed.

—— (1979) "Inheritance and the Distribution of Wealth." Oxford. *Economica*, forthcoming.

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

Dasgupta, P. and Heal, G. M. (1974) "The Optimal Depletion of Exhaustible Resources." *Review of Economic Studies*, Symposium on the Economics of Exhaustible Resources, pp. 3-28.

Farrell, M. J. (1970) "The Magnitude of 'Rate of Growth' Effects on Aggregate Savings." *Economic Journal*, vol. 80, pp. 873-894.

Flemming, J. S. (1976) in Selected Evidence submitted for Report No. 1 Royal Commission on the Distribution of Income and Wealth. London, HMSO.

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

STIGLITZ, J. E. (1966) "Distibution of Income and Wealth Among Individuals." Paper presented at Econometric Society Meeting, San Francisco, mimeographed.

—— (1969) "Distibution of Income and Wealth Among Individuals." *Econometrica*, vol. 37, pp. 382-397.

—— (1977) "Equality, Taxation and Inheritance." Paper presented at IEA Conference on Income Distribution.

—— (1978) "Notes on Estate Taxes, Redistribution, and the Concept of Balanced Growth Path Incidence." *Journal of Political Economy*, vol. 86, pp. 137-149.

—— and DASGUPTA, P. (1978) "Market Structure and Resource Extraction Under Uncertainty." Mathematica, mimeographed.

TOBIN, J. (1967) "Life Cycle Saving and Balanced Growth." in *Ten Economic Studies in the Tradition of Irving Fisher*, Fellner, W. J. ed., New York: John Wiley & Sons.