Journal of Public Economics 10 (1978) 295–336. © North-Holland Publishing Company

A MODEL OF SOCIAL INSURANCE WITH VARIABLE RETIREMENT

P.A. DIAMOND*

M.I.T., Cambridge, MA 02139, U.S.A.

J.A. MIRRLEES

Nuffield College, Oxford OX1 1NF, England

Received June 1977, revised version received April 1978

Models are studied, in which ability to supply labour is affected by a random variable (health) not observable by government. When ill-health strikes, the consumer must retire, but he may choose to retire in any case. Optimal social insurance policies are found for one-period, two-period, and continuous-time models. It is found that, under plausible conditions, at the optimums consumers are indifferent whether to work or not, but do work when able. Insurance contributions decrease with age, and insurance benefits increase with age of retirement. It is desirable to prevent private saving. Some comments on the U.S. Social Security system are added.

1. Introduction

No-one knows what work he will be capable of in the future. Uncertainty about earning ability in the last years of life is particularly great. The burden of this risk to the individual is eased both by private insurance and by the tax and social insurance system. Complete relief from risk is not available, because neither private insurance nor public arrangements distinguish fully between low income by choice and low income by necessity. Full insurance might defeat itself through moral hazard. For this reason, the design of optimal social insurance is an interesting and difficult problem. In this paper, we consider a simple form of earning-ability risk, and study optimal insurance for a population of identical individuals.[1]

*Financial support by the National Science Foundation during the preparation of this paper is gratefully acknowledged. Computing assistance was provided by the Computing and Research Support Unit of the Oxford University Social Studies Faculty; and research assistance by Jay Helms.

[1]In other papers (not yet published), we consider identical and diverse populations whose savings opportunities are limited only by linear taxation.

We suppose that everyone lives for the same length of time, and that at any date an individual has either full earning capability or none: there is no partial loss of earning ability. Loss of ability strikes randomly. The government is assumed unable to distinguish those who cannot work from those who merely choose not to. Individuals are assumed to maximize their expected utility, and to be willing, without concern for the truth, to claim inability to work when it suits them.

Three models will be used, with one period, with two periods, and with continuous time. The one-period model allows us to develop simply the basic condition determining whether or not there is a moral hazard problem. For the case where there is a problem, we relate the size of the optimal insurance plan to characteristics of the utility function. With the two-period model, we demonstrate an additional aspect of the moral hazard problem, that taxation of alternative commodities can increase expected utility. For our model, we find that, under plausible assumptions, an untaxed individual would try to save too much. Thus the optimal social insurance plan needs to be supplemented by an interest income tax. In the continuous-time model, we examine the optimal consumption path while working, and the optimal relationship between pension and date of retirement. Even in the absence of utility-discounting, and of a positive return to saving, we find that it is optimal for consumption to increase with age, and for retirement benefit to increase with retirement age. This leads us to suggest a way of incorporating these conclusions into the benefit structure of the U.S. Social Security system.[2]

2. The one-period model

Either the individual is capable of work or he is not. He knows the probability θ that he can work. He is an expected-utility maximizer, his utility being specified by three utility functions:

$u_1(c)$ = utility of consumption c when working,
$u_2(c)$ = utility when able to work, but not working,
$u_3(c)$ = utility when unable to work.

When working, he produces one unit of output. Assuming nonlinear taxation of income and one type of individual, the lack of continuous adjustment of labour supply is not critical; but the discrete nature of health outcomes is.

On the assumption that work is unpleasant in the aggregate, we have

$$u_2(c) > u_1(c) \qquad \text{for all } c. \tag{1}$$

[2]The theoretical effects of social insurance on retirement have been studied by Feldstein (1974) and Sheshinski (1978).

We also assume that work plus consumption is preferable to no work:

$$u_1(1) > u_2(0). \tag{2}$$

There are no private insurance markets. The government pools risks well enough to be able to plan to distribute consumption, conditional on work, in such a way that the expected value of consumption equals the expected value of output for any individual.

The government being unable to tell why an individual does not work, its policy is entirely described by c_1, consumption for a person who is working, and c_2, consumption for one who is not. This policy provides the individual with expected utility

$$\theta u_1(c_1) + (1-\theta)u_3(c_2) \qquad \text{if he works when he can,}$$

$$\theta u_2(c_2) + (1-\theta)u_3(c_2) \qquad \text{if he does not work.}$$

Thus he is willing to work if and only if

$$u_1(c_1) \geqq u_2(c_2). \tag{3}$$

It is as well to suppose that, even in the case of indifference, the individual works if he is willing and able to do so. We shall comment on this assumption below.

When individuals who can, work, the aggregate resource constraint is

$$\theta c_1 + (1-\theta)c_2 \leqq \theta \tag{4}$$

since the output of a worker is 1. Otherwise, $c_2 = 0$.

The situation can be portrayed in a diagram, with c_1 and c_2 on the axes. This is done in figs. 1 and 2,* where the plane is divided into two regions by the curve $u_1(c_1) = u_2(c_2)$. By assumption (2), the point $c_1 = 1$, $c_2 = 0$ lies below this curve. Everywhere on and below the curve, the individual works if he can; and above it, he does not work. The feasible region is the shaded area plus the origin. Three indifference curves, I_1I_1, I_2I_2, I_3I_3, are drawn, along which expected utility is constant given the private decision on whether to work. It will be noticed that the curve $u_1 = u_2$ lies entirely below the forty-five degree line: this expresses assumption (1).

Evidently there are two cases to consider, which are shown in the two figures. In the first, the indifference curve is tangent to the line with equation $\theta c_1 + (1-\theta)c_2 = \theta$. Since the slope of the indifference curve is

$$-\frac{\theta}{1-\theta}\,\frac{u_1'(c_1)}{u_3'(c_2)}, \tag{5}$$

*Figures appear at the end of the paper.

and the slope of the budget line is $-\theta/(1-\theta)$, the optimum in this case is given by

$$u_1'(c_1^*)=u_3'(c_2^*) \quad \text{and} \quad \theta c_1^*+(1-\theta)c_2^*=\theta. \tag{6}$$

In this case we have the familiar description of a full optimum. The moral hazard constraint is ineffective, as can be seen from the diagram: in the full optimum, people are willing to work if they can.

If at the allocation described by (6) we have $u_1<u_2$, the situation is different, and must be as shown in fig. 2. Here the optimum is given by the intersection of the budget constraint and the curve $u_1=u_2$:

$$u_1(c_1^*)=u_2(c_2^*) \quad \text{and} \quad \theta c_1^*+(1-\theta)c_2^*=\theta. \tag{7}$$

For this case to occur, the indifference curve at this intersection point must slope down less steeply than the budget line. This will be the case if $u_1'\leqq u_3'$ at all points of the curve $u_1=u_2$. If, on the contrary, $u_1'>u_3'$ at all points of the curve, the first case applies. For convenience, we state all this as a formal theorem.[3]

Theorem 1. If for all x and y

$$u_1(x)=u_2(y) \quad \text{implies} \quad u_1'(x)\leqq u_3'(y), \tag{8}$$

the optimum is given by (7). *If for all x and y*

$$u_1(x)=u_2(y) \quad \text{implies} \quad u_1'(x)\geqq u_3'(y), \tag{9}$$

the optimum is given by (6).

The conditions in the theorem take on a more interesting shape if we make the further assumption that the loss in the ability to work has an additive effect on utility:

$$u_3(c)=u_2(c)-b. \tag{10}$$

Since in this case marginal utility of consumption is the same in the two states, condition (8) can be stated as

$$u_1(x)=u_2(y) \quad \text{implies} \quad u_1'(x)\leqq u_2'(y), \tag{11}$$

for all x and y.

[3]The model under consideration is equivalent to the standard maximization of an additive social welfare function. The particular distribution of characteristics assumed here has not, to our knowledge, been investigated previously.

Thus the moral hazard problem is present if compensating a worker so that he is indifferent to working results in a lower marginal utility of consumption than if he did not work and was not compensated. It seems to us that this condition will much more often be relevant than its contrary. For example, it follows when utility is additively separable in consumption and labour, and labour is disliked.

3. Comparative statics under moral hazard

By means of the diagram, we can examine the way in which the optimum changes when probabilities change, when the disutility of labour is changed, and when the utility of consumption is changed.

When θ increases, the budget line becomes steeper, rotating clockwise about the fixed point $(1,0)$. Thus, if moral hazard is present, c_1^* and c_2^* both increase. Similarly an increase in the disutility of work, i.e. u_2-u_1, moves the curve $u_1=u_2$ downwards, so that c_1^* is increased and c_2^* decreased.

A change in the utility of consumption can also be analysed by considering how it affects the curve $u_1=u_2$. For example, we can show that an increase in risk-aversion can lead to a reduction in the extent of insurance, which may be contrary to some people's intuition. Suppose utility and marginal utility at c_2^* remain fixed, both with and without work, and that risk-aversion increases. u_1 is thereby decreased for $c>c_2^*$, and in particular for c_1^* (see fig. 3). Therefore the curve $u_1=u_2$ is shifted to the right in the neighbourhood of c_2^*. Consequently the optimal c_1^* is increased and the optimal c_2^* is decreased: the extent of insurance is reduced, as claimed. A different kind of increase in risk-aversion can have the opposite effect, for example, by preserving utility and marginal utility at c_1^* while increasing risk-aversion. By analogy to fig. 3 the curve $u_1=u_2$ is shifted up.

The one-period model showed that the government's inability to perceive whether earnings abilities are present may or may not render attainment of the first-best optimum impossible. If this possibility is affected, which is to say that the moral hazard problem is present, then, with individuals identical ex ante and only two levels of earnings ability, at the optimum individuals are indifferent to work. We shall see that this property persists as we extend the time frame of the analysis.

4. The two-period model

With two periods, we assume that everyone can work in the first period, and that an individual is able to work in period two with probability θ. These facts are known to the government. We shall want to use this model to discuss the significance of private saving behaviour, but, for the present, the government is assumed to offer individuals pairs of first and second period consumptions: (c_0, c_1) if work is done in period two, (c_0, c_2) if not. No private saving is allowed. Utility is now in every case a function of consumption in the two periods. The particular utility function will be denoted by superscript instead of subscript, subscripts being reserved for derivatives.

The budget constraint is taken to be

$$c_0 + R[\theta c_1 + (1-\theta)c_2] = 1 + \theta R, \tag{12}$$

it being assumed that aggregate output when everyone works is unity, and that there is a constant interest factor[4] R. In order to portray the situation on a two-dimensional diagram, we use the budget constraint (12) to eliminate one of the variables when everyone works who is able to do so. It is most convenient to use the variables c_0 and

$$z = c_1 - c_2, \tag{13}$$

where z is the return to working and $1-z$ the tax on labour income. Then we have from (12),

$$c_1 = \frac{1-c_0}{R} + \theta + (1-\theta)z, \tag{14}$$

$$c_2 = \frac{1-c_0}{R} + \theta(1-z). \tag{15}$$

Since consumption levels have to be nonnegative, it follows that, so long as everyone works when he can, c_0 and z are constrained by three linear inequalities, $c_0 \geqq 0$, $c_1 \geqq 0$, and $c_2 \geqq 0$. These define a triangle in (c_0, z)-space, as shown in fig. 4.

However, people are willing to work when they can only if

$$u^1(c_0, c_1) \geqq u^2(c_0, c_2). \tag{16}$$

If $z \leqq 0$, $c_2 \geqq c_1$, so that, maintaining our assumption that work has disutility, $u^2(c_0, c_2) > u^1(c_0, c_1)$. Thus the inequality (16) implies that $z > 0$. Notice also that, since an increase in z increases c_1 and decreases c_2, (c_0, z') satisfies (16) whenever (c_0, z) satisfies (16) and $z' > z$. The region in which people work if they can is the one bounded by ABC in fig. 4: AB is the locus of points where people are indifferent whether to work or not. If (c_0, z) is not in this

[4] R equals the inverse of one plus the interest rate.

region, either there is no feasible allocation with these values of c_0 and z, or no-one works in the second period.

In the latter case, $c_0+Rc_2=1$, and expected utility is $\theta u^2(c_0, c_2) + (1-\theta)u^3(c_0, c_2)$. Let

$$v_m = \max[\theta u^2 + (1-\theta)u^3 : c_0 + Rc_2 = 1] \tag{17}$$

be the maximum utility obtainable without work in the second period.

The indifference curves drawn in the diagram are the curves of constant expected utility $\theta u^1 + (1-\theta)u^2$ on the assumption that work is done in the second period. The case shown is one for which the full optimum – the point F – is not feasible; and the maximum attainable utility is greater than v_m, so that moral hazard is present in the optimum, and work is done in the second period. The same argument as in section 2 shows that the full optimum is unattainable if

$$u^1(x_0, x_1) = u^2(x_0, x_2) \qquad \text{implies} \qquad u_1^1(x_0, x_1) \leqq u_2^3(x_0, x_2), \tag{18}$$

where u_i denotes the partial derivative with respect to x_i.

We concentrate on the case where moral hazard is present and there is work in the second period at the optimum. Then at the optimum the indifference curve is tangent to the curve AB. Also a small increase in z would reduce utility. An increase in c_0 may either increase or decrease utility, depending on the sign of the slope of the curves at the optimum. We shall discuss conditions for this in a moment.

The interesting point is that a change in c_0 that would increase expected utility while z is held constant represents a savings opportunity that the individual would wish to avail himself of if he were allowed to; for, as can be seen from (14) and (15), a change in c_0 brings about the changes in c_1 and c_2 that would happen if it were an act of private saving or dissaving. Thus in general, the optimum cannot be obtained without preventing access to the capital market, or, alternatively, imposing an appropriate tax or subsidy on savings.

It is most likely that the curve AB representing the moral hazard constraint will have a negative slope at the point of tangency. If u^2 tends to minus infinity as c_2 tends to zero, then the curve $u^1 = u^2$ can hit the line $c_2 = 0$ only where c_1 is also zero, i.e. on the vertical axis, with $z=0$. Since $z>0$ at the point B, the slope of the curve AB is predominantly negative. We can also find a rather plausible local condition for the optimum to appear as in fig. 4. For this further analysis we assume that u^2 and u^3 differ only by a constant, i.e. have the same derivatives.

Theorem 2. Assume that $u^3 = u^2 - b$,

$$u^1(x_0, x_1) = u^2(x_0, x_2) \quad \text{implies} \quad u_1^1(x_0, x_1) \leqq u_2^2(x_0, x_2), \tag{i}$$

$$u^1(x_0, x_1) = u^2(x_0, x_2) \quad \text{implies} \quad \left(\frac{\partial x_1}{\partial x_0}\right)_{u^1} \leqq \left(\frac{\partial x_2}{\partial x_0}\right)_{u^2}. \tag{ii}$$

Then at the optimum, individuals want to save more.

Proof. The claim is that at the optimum a reduction in c_0 would increase expected utility, z being held constant; i.e. that the curve $u^1 = u^2$ has a negative slope at the optimum. Since $u^1 - u^2$ is an increasing function of z, we have to show that it is an increasing function of c_0,

$$\frac{\partial}{\partial c_0}(u^1 - u^2) = u_0^1 - u_0^2 - \frac{1}{R}(u_1^1 - u_2^2).$$

Here numerical subscripts denote differentiation with respect to the indicated consumption level.

We shall argue that, at the optimum,

$$\frac{1}{R} = \theta \frac{u_0^1}{u_1^1} + (1-\theta)\frac{u_0^2}{u_2^2}. \tag{19}$$

Given this equation, $\partial/\partial c_0\,(u^1 - u^2)$ is positive if u_0^1/u_1^1 is larger than u_0^2/u_2^2 (and conversely). The latter condition is precisely (ii).

It remains to derive (19) given the assumption that at the optimum we have a moral hazard problem (which follows from (i)). From fig. 4 we know that at D an indifference curve is tangent to the moral hazard constraint. Equating these two slopes we have

$$\frac{\theta\left(u_0^1 - \frac{1}{R}u_1^1\right) + (1-\theta)\left(u_0^2 - \frac{1}{R}u_2^2\right)}{\theta(1-\theta)(u_1^1 - u_2^2)} = \frac{\left(u_0^1 - \frac{1}{R}u_1^1\right) - \left(u_0^2 - \frac{1}{R}u_2^2\right)}{(1-\theta)u_1^1 + \theta u_2^2}.$$

Crossmultiplying, we obtain (19), which completes the proof.

The first condition in the theorem is the one that we have earlier claimed to be plausible, and which is pretty much required for moral hazard to be present in the optimum. The second condition says that, at equal utilities whether working or not, the individual has a greater incentive to save when not working. Our conclusion is that it is normal for a tax on saving to be associated with the optimal social insurance policy.

Suppose now that saving is allowed, and is not taxed. This means that optimization is further constrained by a saving equilibrium condition, that the derivative of expected utility with respect to c_0 is zero. Thus the condition is represented by the locus of points where the indifference curves have vertical slope, the curve *FE* in fig. 5. As one moves right along this curve, utility, and the size of the social insurance programme, diminish. If individuals made their savings plans under the assumption that they would work in the future, analysis would be completed by combining the savings constraint with the moral hazard constraint $u^1(c_0,c_1)=u^2(c_0,c_2)$ and the optimum would be at *E*, where the two constraints intersect. However if the individual plans savings and future work at the same time, we have a different moral hazard condition:

$$\max_s \, [\theta u^1(c_0-s,c_1+s)+(1-\theta)u^3(c_0-s,c_2+s)]$$
$$\geqq \max_s \, [\theta u^2(c_0-s,c_1+s)+(1-\theta)u^3(c_0-s,c_2+s)].$$

This is a more stringent condition when a zero level of savings is optimal if the individual works; for the ability to adjust savings increases the utility available to those planning not to work. This is shown in fig. 5 as *GJ* with the optimum occurring at *H*. (This model is explored in more detail in another paper.) Notice that, when individuals are free to save without constraint or taxation, it is still optimal to be on the boundary of the feasible set.[5]

One aspect of these solutions should be further noted. Since they are on the boundary of the feasible set, individuals are in fact indifferent whether to work in period 2 or not. Similarly in the one-period model they are indifferent between working and not working. If this indifference were reflected in random choice rather than an implausible adherence to government wishes, there would, strictly speaking, be no optimum, though the government could adopt policies arbitrarily close to the ones we have discussed under which individuals would certainly wish to work if they could. That is why it is not inappropriate to assume, as we have done, that the government can choose when the consumer is indifferent. Of course a fully satisfactory treatment of these matters would have to model the costs of implementing choices explicitly, and that would take us too far afield.

The chief conclusion of this section is that a government concerned about social insurance would normally want to discourage private saving, for

[5]To achieve the optimum at *D* in fig. 4, the government must either close the capital market, or introduce nonlinear wealth taxation to ensure that individuals do not desire to save.

example by taxing it; and that, if it could not do so, the extent of the social insurance programme would be diminished. This exemplifies a general feature of moral hazard situations, that the provider of insurance would usually want to control trade in related commodities. For example, fire insurance companies should want their clients to buy fire extinguishers; and they like to see high taxes on the consumption of tobacco.

5. The continuous-time model

We turn to a model in which the retirement date is continuously variable, while maintaining the assumption that labour choice at each moment is discrete. This model allows us to ask two further questions: how consumption should be made to vary with age when working, and how the retirement benefit should depend upon the age of retirement. Since the moral hazard constraints can no longer be pictured in a diagram, the mathematics of the model is more complicated than for the previous models, but we shall use the earlier analysis to guide us to the solution.

We make the following assumptions. If everyone were working, output per unit period would be unity. The real interest rate is zero. Utility is the undiscounted integral of instantaneous utility. When the individual is working, his instantaneous utility is $u_1(c)$; when not working, $u_2(c)$ or $u_3(c)$ according to whether he is able to work or not. All three functions are strictly concave. To simplify the analysis, we assume from the outset that u_2 and u_3 differ by a constant:

$$u_2(c)=u_3(c)+b, \qquad b\geqq 0.$$

Saving is controlled by the government. Consequently, consumption when working, a function of age, $c_1(t)$; and consumption when retired can be a function both of age t and of age at retirement, r, $c_2(t,r)$. The date s at which an individual becomes unable to work is a random variable, with density function f and distribution function F. The length of life is denoted by T, and $f(s)>0$ for $0<s<T$. There are no atoms in the distribution.

If the age of retirement is r and the age of disability is $s, s\geqq r$, utility is[6]

$$\begin{aligned}&\int_0^r u_1(c_1(t))\,\mathrm{d}t+\int_r^s u_2(c_2(t,r))\,\mathrm{d}t+\int_s^T u_3(t,r))\,\mathrm{d}t\\&=\int_0^r u_1(t)\,\mathrm{d}t+\int_r^T u_2(t,r)\,\mathrm{d}t-b(T-s),\end{aligned}\tag{20}$$

[6] We do not consider the possibility of a return to work after retirement. An approach to dealing with this complication would be to relate wages and pensions to experience rather than age.

where we have used (19), and introduced the notation

$$u_1(t)=u_1(c_1(t)), \qquad u_2(t,r)=u_2(c_2(t,r)).$$

Neither individual nor government can affect the last part of (20). Therefore we omit it from further considerations. What we are interested in is the utility of work and consumption. For someone who does retire at r, this can be written as

$$v(r)=\int_0^r u_1(t)\,\mathrm{d}t+\int_r^T u_2(t,r)\,\mathrm{d}t. \tag{21}$$

With retirement planned at r, expected utility can be taken to be

$$V(c_1,c_2,r)=\int_0^r v(s)f(s)\,\mathrm{d}s+v(r)[1-F(r)], \tag{22}$$

since the individual has to retire at the date of disability s if s precedes the planned retirement date. Thus r is the age at which the individual decides to retire if he is still able to work at that date.

Eq. (22) expresses the maximand for our problem as an average of v, the expected value of $v(\min(s,r))$. This average occurs frequently in our analysis, and it is convenient to have a special notation for it. For any integrable function h, we define

$$J_r(h)=\int_0^r h(s)f(s)\,\mathrm{d}s+h(r)[1-F(r)]. \tag{23}$$

For later reference we note the following.

Properties of J_r:

(i) J_r is a positive linear functional for each r,
(ii) $J_r(h)$ is monotonically increasing (or decreasing) in r if h is monotonically increasing (or decreasing),
(iii) $J_r(h)$ is a constant function of r if and only if h is a constant,
(iv) If h is differentiable,

$$\frac{\mathrm{d}}{\mathrm{d}r}J_r(h)=h'(r)[1-F(r)]. \tag{24}$$

Proof. (i) and (iv) follow immediately from (23). To prove (ii), we let $r' > r$, and calculate from (23) that

$$J_{r'}(h) - J_r(h) = \int_r^{r'} [h(s) - h(r)] f(s)\,\mathrm{d}s + [h(r') - h(r)] \int_{r'}^{T} f(s)\,\mathrm{d}s. \quad (25)$$

Thus $J_{r'}(h) > J_r(h)$ if h is monotonically increasing. The case where h is monotonically decreasing is proved similarly.

From (25) it also follows that $J_r(h)$ is constant if h is constant. To complete the proof of (iii), suppose that $J_r(h) = c$, a constant, for all r. From (23) it follows that $h(r)$ is differentiable, and (24) then implies $h' = 0$, i.e. h is a constant. The proof is complete.

The net resource cost to the government of an individual who retires at r is

$$z(r) = \int_0^r c_1(t)\,\mathrm{d}t + \int_r^T c_2(t, r)\,\mathrm{d}t - r, \quad (26)$$

recalling that a worker produces unit output. If the government requires A units of output for its own purposes, the aggregate resource constraint is

$$J_r(z) + A \leqq 0, \quad (27)$$

assuming that everyone selects the same retirement date and the government is subject to an expected value resource constraint.

We can now state formally the problem to be addressed. We are to: maximize $J_r(v)$ subject to (27) and

$$J_t(v) \leqq J_r(v), \qquad \text{all } t \leqq r. \quad (28)$$

This last constraint says that individuals do not prefer to retire before r. Since the government can set consumption levels equal to zero for anyone who retires after a date it chooses, we need not be concerned about individual plans to work too long. Thus the effective constraint on government is that it choose r, and the functions c_1 and c_2 so that (27) and (28) hold.

In the full, first-best optimum, the government is constrained only by (27), and marginal utilities of consumption should be equated in all circumstances:

$$c_1(t) = c_1^0, \qquad c_2(t, r) = c_2^0; \qquad \text{with } u_1'(c_1^0) = u_2'(c_2^0). \quad (29)$$

If under these circumstances we had $u_1(c_1^0) \geqq u_2(c_2^0)$, there would be no moral hazard problem, for, as one can see from (21), $v(r)$ would be a nondecreasing function of r, and $J_t(v)$ consequently a nondecreasing function of t.

In cases where no moral hazard problem arises, the optimal retirement date will often be T. From the fact that there is no moral hazard problem, it follows, using (24), that expected utility is a nondecreasing function of the retirement date:

$$\frac{\partial}{\partial r} J_r(v) = [u_1(c_1^0) - u_2(c_2^0)][1 - F(r)] \geqq 0.$$

If it is also the case that expected resource costs decrease with the retirement date, then the optimum occurs at T. Using (24) again, we have

$$\frac{\partial}{\partial r} J_r(z) = [c_1^0 - c_2^0 - 1][1 - F(r)].$$

Thus it is sufficient for the optimum to occur at T that

$$c_1^0 - c_2^0 - 1 < 0.$$

It might be that $u_1'(x) = u_2'(y)$ implies $x \leqq y + 1$. If so the result follows. Even when this is not the case for all x and y, it is true at the optimum if the government has positive resource needs elsewhere ($A > 0$), and the optimum involves some work. For we have $A = -J_r(z)$ and

$$\begin{aligned} z(s) &= sc_1^0 + (T - s)c_2^0 - s \\ &= (c_1^0 - c_2^0 - 1)s + Tc_2^0. \end{aligned}$$

Thus if A is positive, and $J_r(z)$ consequently negative, $c_1^0 - c_2^0 - 1 < 0$, as claimed. By our previous argument, it follows that the optimum has retirement planned to be at T.

As we argued above, the interesting case is where $u_1' = u_2'$ is inconsistent with $u_1 \geqq u_2$, so that there is a moral hazard problem. Accordingly we assume from now on that

$$u_1(c_1) = u_2(c_2) \quad \text{implies} \quad u_1'(c_1) \leqq u_2'(c_2). \tag{30}$$

Before proceeding with the analysis, we note a convenient feature of the optimum. Once a man has retired, there is no advantage from inefficient intertemporal allocation of his consumption. Incentives to work depend on expected utility as a function of the retirement date, so the cost of providing expected utility conditional on retirement should be minimized. With no discounting of utility and a zero interest rate, this implies

$$c_2(t,r)=c_2(r). \tag{31}$$

Then v and z can be written

$$v(r)=\int_0^r u_1(t)\,\mathrm{d}t+u_2(r)(T-r), \tag{32}$$

$$z(r)=\int_0^r c_1(t)\,\mathrm{d}t+c_2(r)(T-r)-r. \tag{33}$$

A bad feature of the maximization problem we have to deal with is a lack of the concavity conditions that would render necessary conditions for the optimum also sufficient. The moral hazard constraints (28) cannot be expected to be concave in any of the control variables c_1, c_2, or r. But if we regard u_1 and u_2 as the control variables, the maximand, and the constraints (28), are linear, and the resource constraint (27) is convex in them. Thus we have a well-behaved programming problem if r is taken as given. We shall first derive necessary and sufficient conditions for the optimum conditional upon a specified value of r, and afterwards derive a necessary condition for optimality with respect to r.

The earlier models yielded solutions in which the individual was indifferent whether to work or not. Correspondingly, one can expect that in the present model, the individual will be indifferent about the age of retirement. We first consider the best path with this property and then show that it is indeed optimal. We derive two differential equations for the growth of $u_1(t)$ and $u_2(t)$. After discussing their properties, we state and prove the optimality result as theorem 3. If $J_t(v)$ is a constant independent of t, then $v(r)$ is a constant independent of r, by property (iii). We write

$$\int_0^r u_1(t)\,\mathrm{d}t+u_2(r)(T-r)=\bar{v}. \tag{34}$$

Given that v is constant, we ask what is the optimal combination of u_1 and u_2. It is the one that for the given value of $J_r(z)$ maximizes $\bar{v}$, or equivalently for given $\bar{v}$ minimizes $J_r(z)$.

Let G_1 and G_2 be the inverse functions of $u_1(c)$ and $u_2(c)$ respectively, so that $c_1(t)=G_1(u_1(t))$ and $c_2(t)=G_2(u_2(t))$. Then we can use the constant utility condition to write resource usage in terms of $u_1(t)$ as

$$z(r)=\int_0^r G_1(u_1)\,\mathrm{d}t+G_2\left(\frac{1}{T-r}\left[\bar{v}-\int_0^r u_1\,\mathrm{d}t\right]\right)(T-r)-r. \tag{35}$$

We can substitute this in the expression for expected cost, $J_r(z)$, and find the first-order conditions for expected cost minimization with respect to the function $u_1(t)$. From (35), we find that

$$\frac{\partial z(r)}{\partial u_1(t)}=\begin{cases}G_1'(u_1(t))-G_2'(u_2(r)), & t\leqq r,\\ 0, & t>r.\end{cases} \tag{36}$$

Consequently, for $t\leqq r$,

$$\begin{aligned}\frac{\partial}{\partial u_1(t)}J_r(z)&=J_r\left(\frac{\partial z}{\partial u_1(t)}\right)\\ &=\int_t^r[g_1(t)-g_2(s)]f(s)\,\mathrm{d}s+[g_1(t)-g_2(r)][1-F(r)],\end{aligned} \tag{37}$$

where we have introduced the notations

$$g_1(t)=G_1'(u_1(t))=1/u_1'(c_1(t)), \tag{38a}$$

$$g_2(t)=G_2'(u_2(t))=1/u_2'(c_2(t)). \tag{38b}$$

For cost-minimization, the derivative (37) should vanish for all t:

$$\int_t^r[g_1(t)-g_2(s)]f(s)\,\mathrm{d}s+[g_1(t)-g_2(r)][1-F(r)]=0. \tag{39}$$

Differentiating with respect to t, and using a dot to denote a time-derivative, we obtain

$$[1-F(t)]\dot{g}_1(t)=[g_1(t)-g_2(t)]f(t), \tag{40}$$

which can be written equivalently as

$$[1-f(t)]\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{u_1'(c_1(t))}=\left[\frac{1}{u_1'(c_1(t))}-\frac{1}{u_2'(c_2(t))}\right]f(t).$$

Also, setting $t=r$ in (39), we have

$$g_1(r)=g_2(r), \tag{41}$$

or, equivalently,

$$u_1'(c_1(r))=u_2'(c_2(r)).$$

We have shown that if there is indifference about retirement age at the optimum (so that v is constant), then (40) and (41) hold. Furthermore, since $v(r)$ is constant, we find on differentiating (34) with respect to r that

$$(T-t)\dot{u}_2(t)=u_2(t)-u_1(t). \tag{42}$$

Thus we are describing the optimal path by two differential equations, (40) and (42), and two further conditions, (41) and the resource constraint (27). Notice that (40) and (42) are time-dependent differential equations.

In fig. 6, we have a phase diagram for this pair of differential equations, in (c_1, c_2)-space. Since the equations are time-dependent, there is an infinity of paths through most points of the diagram, and solution paths can cross one another. Nevertheless a phase diagram is helpful: for, by (40), c_1 is stationary when $g_1=g_2$; and, by (42), c_2 is stationary when $u_1=u_2$. These two curves are invariant with time. Since u_1 and $g_1=1/u_1'(c_1)$ are increasing functions of c_1, and u_2 and g_2 are increasing functions of c_2, both the stationary curves have positive slope.

By assumption (29), the curve $u_1=u_2$ lies below the curve $g_1=g_2$ (i.e. $u_1'=u_2'$). Between the two curves, both c_1 and c_2 are increasing with t. (41) says that the solution we propose has to hit the $g_1=g_2$ curve at $t=r$. Therefore both c_1 and c_2 are increasing functions of t throughout the solution. We now prove that a path satisfying (40)–(42) and the resource constraint is the best path among those inducing retirement at a specified age r. We also give conditions for the existence of an optimum for given r, and show that the optimal path is continuous in the planned retirement age. The reader can, without loss of continuity, turn to the discussion of optimal r in section 7. The formal statement of these results is as follows:

Theorem 3. Assume that $u_1=u_2$ *implies* $u_1'\leqq u_2'$. *Let* $r<T$. *If*

$$\dot{u}_2=\frac{u_2-u_1}{T-t},$$

$$\dot{g}_1=(g_1-g_2)\frac{f}{1-F},$$

$$g_1=g_2 \qquad \text{at} \qquad r,$$

and $J_r(z)+A=0$, *the policies so defined are optimal for the given r. Assume further that* $u_1(0)=u_2(0)=-\infty$; *and that for some constant* $B>0$, $-u_1''/u_1'\geqq B$. *Then a solution with the stated properties exists, for any r between* 0 *and*

T such that

$$\int_0^r [1-F(s)]\,\mathrm{d}s > A \tag{43}$$

(*i.e. such that positive consumption is feasible*).

6. Proof of theorem 3

We first found a proof of this theorem by introducing Lagrange multipliers for the constraints, and following the standard method of proving sufficiency of the first-order conditions for constrained maximization problems. This method is rather involved. The proof we give here, though less clearly motivated in that Lagrange multipliers are not introduced explicitly, is briefer. We begin by establishing some preliminary results. These involve optimal shadow prices and alternative plans satisfying the moral hazard constraint. Using the lemmas, we go on to show that no feasible path has higher expected utility than the one specified. The policies satisfying the conditions of the theorem are denoted by asterisks.

Lemma 1. $J_r(g_2^* v) \geqq g_1^*(0) J_r(v)$ *for any* v *satisfying* $J_t(v) \leqq J_r(v)$.

Since c_1^* is an increasing function of t, g_1^* is an increasing function of t, and we have

$$\int_0^r \dot{g}_1^*(t)[J_r(v) - J_t(v)]\,\mathrm{d}t \geqq 0. \tag{44}$$

Since $J_r(v)$ is independent of t,

$$\int_0^r \dot{g}_1^* J_r(v)\,\mathrm{d}t = [g_1^*(r) - g_1^*(0)] J_r(v). \tag{45}$$

Also,

$$\begin{aligned}\int_0^r \dot{g}_1^* J_t(v)\,\mathrm{d}t &= \int_0^r \dot{g}_1^*(t) \int_0^t v(s) f(s)\,\mathrm{d}s\,\mathrm{d}t + \int_0^r \dot{g}_1^*(t) v(t)[1-F(t)]\,\mathrm{d}t \\ &= \int_0^r [g_1^*(r) - g_1^*(t)] v(t) f(t)\,\mathrm{d}t \\ &\quad + \int_0^r [g_1^*(t) - g_2^*(t)] v(t) f(t)\,\mathrm{d}t,\end{aligned}$$

from the definition of $J_t(v)$, reversal of the order of integration in the first integral, and use of (40) in the second integral. This last expression simplifies to

$$\int_0^r \dot{g}_1^* J_t \, dt = g_1^*(r) \int_0^r vf \, dt - \int_0^r g_2^* vf \, dt = g_1^*(r) J_r(v) - J_r(g_2^* v), \tag{46}$$

since $g_1^*(r) = g_2^*(r)$.

Combining (44), (45), and (46) we obtain lemma 1.

Lemma 2. $J_r(g_2^*) = g_1^*(0)$.

The second and third conditions of the theorem are equivalent to the vanishing of (37) for all t. Putting $t=0$, we obtain the stated result.

We now show that any feasible path has expected utility no greater than the path satisfying the conditions in the statement of the theorem. From the definition of z we have

$$\begin{aligned} z(s) - z^*(s) = \int_0^s [G_1(u_1) - G_1(u_1^*)] \, dt + [G_2(u_2(s)) \\ - G_2(u_2^*(s))](T-s). \end{aligned}$$

Since G_1 and G_2 are convex functions of their arguments, we can use the inequality for convex functions to obtain

$$z(s) - z^*(s) \geqq \int_0^s g_1^*(t)(u_1(t) - u_1^*(t)) \, dt + g_2^*(s)(u_2(s) - u_2^*(s))(T-s).$$

Using the definition of v, (32), and the constancy of $v^*(s)$, we can write this as

$$\begin{aligned} z(s) - z^*(s) &\geqq \int_0^s (g_1^*(t) - g_2^*(s))(u_1(t) - u_1^*(t)) \, dt + g_2^*(s)(v(s) - v^*) \\ &= \int_0^T \frac{\partial z^*(s)}{\partial u_1^*(t)} (u_1(t) - u_1^*(t)) \, dt + g_2^*(s)(v(s) - v^*). \end{aligned} \tag{47}$$

The second step follows from the expression for the derivative, (36). We now wish to apply the linear operator J_r to (47). The first term on the right-hand side vanishes since (37) is zero for all t:

$$J_r\left(\int_0^T \frac{\partial z^*(r)}{\partial u_1^*(t)} (u_1(t) - u_1^*(t)) \, dt \right) = \int_0^T J_r\left(\frac{\partial z^*(r)}{\partial u_1^*(t)} \right) (u_1(t) - u_1^*(t)) \, dt = 0.$$

Thus we have

$$J_r(z)-J_r(z^*)\geqq J_r(g_2^* v)-J_r(g_2^*)v^*$$
$$\geqq g_1^*(0)[J_r(v)-v^*] \tag{48}$$

by lemmas 1 and 2.

Now $J_r(z)+A\leqq 0$, and $J_r(z^*)+A=0$. Therefore (48) implies that

$$J_r(v)\leqq v^*,$$

i.e. expected utility is no greater on the alternative path than on the one satisfying the conditions of the theorem. This proves the sufficiency part of the theorem. Uniqueness follows from a strict inequality in (48) since the G_i are strictly convex.

We next prove the existence part of the theorem. Consider all paths satisfying the differential equations (40) and (42) and satisfying the terminal condition $u_1'(c_1(r))=u_2'(c_2(r))$. Denote the expected utility of any such path by $V(x,r)$, where r is the planned retirement date and x is the value of $c_1(r)$. In terms of fig. 6 we are considering all solution paths which end on $g_1=g_2$ at time r. Denote the expected resource cost of such a path by $Z(x,r)$. We first note without taking the space for formal proof that V and Z are differentiable functions of r and x. If $Z(x,r)+A=0$, then we have an optimal path. We want to show that there exists an x satisfying the resource constraint when $u_i(0)=-\infty$, and $-u_1''/u_1'$ is bounded away from zero. Refer to the diagram. Since marginal utilities go to infinity as consumption goes to zero, g_1 and g_2 then go to zero, and the $g_1=g_2$ curve passes through the origin. Similarly the curve $u_1=u_2$ passes through the origin. Thus we can choose x to be as large or small as we please and find a path ending on $g_1=g_2$ with $c_1(r)$ equal to x.

By choosing x small enough, aggregate consumption, $z(t)+t$, can be made as small as we please, uniformly for $0\leqq t\leqq r$. Thus $J_r(z(t)+(t))$ can be made as small as we please.

$$J_r(t)=\int_0^r tf(t)\,\mathrm{d}t+r-rF(r)=\int_0^r[1-F(t)]\,\mathrm{d}t$$

on integrating by parts. Therefore if (43) holds, $J_r(z)+A$ can be made negative by choosing x small enough.

We now have to show that $J_r(z)+A$ can be made positive by choosing x large enough. By eq. (40), we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\log g_1=\left(\frac{g_2}{g_1}-1\right)\frac{\mathrm{d}}{\mathrm{d}t}\log[1-F(t)]$$
$$\leqq -\frac{\mathrm{d}}{\mathrm{d}t}\log[1-F(t)],$$

since $g_1/g_2 \geqq 0$ and $1-F$ decreases with t. By differentiation of the equation $\log g_1 = -\log u_1'(c_1)$, we obtain

$$\dot{c}_1 = \frac{u_1'}{u_1''} \frac{\mathrm{d}}{\mathrm{d}t} \log g_1.$$

Integrating from 0 to r, and using the assumption that $-u_1''/u_1' \geqq B$, we find that

$$c_1(0) \geqq c_1(r) + B^{-1} \log(1-F(r)). \tag{49}$$

Therefore by choosing x large enough, $c_1(0)$, and all $c_1(t)$ for $0 \leqq t \leqq r$, can be made as large as we wish, thus increasing $J_r(z)$ without limit.

The proof of existence is complete.

The final argument in the theorem, leading to (49), also helps to establish that for all $r \leqq \bar{r}, \bar{r} < T$, the optimal x is bounded. For (49) implies (using definition (33)) that

$$z(t) \geqq [x + B^{-1} \log(1-F(r)) - 1]t.$$

Thus

$$-A = J_r(z) \geqq (x+C) \int_0^r [1-F(t)]\mathrm{d}t \tag{50}$$

where $C = B^{-1} \log(1-F(\bar{r})) - 1$. This observation that x is bounded when r is bounded away from T will prove useful below.

The rather restrictive conditions about utility functions introduced to prove existence are by no means necessary. One would not normally expect any difficulty in raising aggregate consumption. The condition at $c=0$ was introduced merely to exclude cases with zero consumption, and their attendant corner conditions.

Throughout the argument, r has been taken to be less than T. The form of the differential equations makes it clear that the case $r=T$ would need careful handling; and the above existence proof would not go through at all. As we shall see in the next section, in the most plausible cases the optimal value of r is less than T. This will emerge as a by-product of our discussion of the optimization of the planned retirement date r.

Before turning to this, let us establish that the optimal path is continuous in r. We have defined $V(x,r)$ as the expected utility of a path satisfying the differential equations (40) and (42) and ending on the curve $g_1 = g_2$ at time r with $c_1(r) = x$. If, for given r, x also satisfies $Z(r,x) + A = 0$, it follows from theorem 3 that the solution path so defined yields a unique optimum.

Therefore the equation $Z(r,x)+A=0$ has a unique solution, and we can define a function $x(r)$ by

$$Z(r,x(r))+A=0. \tag{51}$$

We have already seen in (50) that $x(r)$ is a bounded function, and Z is continuous in x and r. Thus we have:

Lemma 3. $x(r)$ *is a continuous function for* $r<T$.

It follows from this lemma that the optimal consumption paths and the maximum expected utility level $V(r)=V(x(r),r)$ are continuous in r.

7. The planned retirement date

We have found sufficient conditions for the existence of an optimum when the planned retirement date is specified. We also concluded that both the maximum expected utility $V(r)$, and the optimal plan, which we denote by $c_i^r(t), i=1,2$, are continuous functions of r. In the next section we shall prove that $V(r)$ is in fact a differentiable function of r. In this section we give a heuristic derivation of $V'(r)$. By setting $V'(r)$ equal to zero, we get a necessary condition for the planned retirement date.

We are able to make three further observations about the optimum. Sufficient conditions for the existence of an optimum are given. It is argued that $c_i^r(r)$ at the optimum is unique (although we do not claim that r is necessarily unique). We find sufficient conditions for r to be less than T. We are thus in a position to discuss the properties of the optimum, and to calculate it for any specific example. This is done for a logarithmic example in section 9.

A heuristic calculation of $V'(r)$ can be given as follows. Reducing r by ε, we can, to first order, neglect the change in $c_i^r(t)$ arising from the change in r. Since $v(t)$ is constant, it also follows that the reduction in r has no direct first-order effect on expected utility. But there is an effect on the resource constraint, for consumers still capable of working at r will, when r is reduced, consume c_2 instead of c_1 between $r-\varepsilon$ and r, produce one unit less, and over the time from r until T consume $(T-r)\dot{c}_2(r)\varepsilon$ less than before. Since there are $1-F(r)$ people in this situation the effect on the resource constraint is

$$(1-F(r))(1-c_1+c_2-(T-r)\dot{c}_2)\varepsilon.$$

This resource gain could be used to raise consumption of the worker at the start of his working life without any effect on the moral hazard constraint.

The resulting increase in utility, $u_1'(c_1(0))$ times the increase in resources, is equal to $V'(r)\varepsilon$, to first order.

Thus

$$V'(r)=u_1'(0)[1-F(r)][1-c_1+c_2-(T-r)\dot{c}_2]. \tag{52}$$

To put this in more convenient form, we use (42) to replace $(T-r)\dot{c}_2$ by $(u_2-u_1)/u_2$, which is in turn equal to $(u_2/u_2')-(u_1/u_1')$, since we have an optimality condition that $u_2'=u_1'$ at r. Thus

$$V'(r)=u_1'(0)[1-F(r)][1+h_1(r)-h_2(r)], \tag{53}$$

where we define

$$h_i(r)=\frac{u_i(c_i^r(r))}{u_i'(c_i^r(r))}-c_i^r(r); \tag{54}$$

h_i is a function of r because $c_i^r(r)$ is uniquely determined by r.

From (53), we obtain the first-order condition for optimality of r,

$$h_1(r)-h_2(r)=-1. \tag{55}$$

For the remainder of this section and the next, we consider the sign of h_1-h_2+1 along the curve $g_1=g_2$, since that curve contains the end-points for r-optimal paths as r varies.

A straightforward calculation yields

$$\frac{\mathrm{d}h_i}{\mathrm{d}g_i}=u_i. \tag{56}$$

From this it follows that, as $g=g_1=g_2$ varies,

$$\frac{\mathrm{d}}{\mathrm{d}g}(h_1-h_2)=u_1-u_2 \\ <0, \tag{57}$$

since $u_1<u_2$ on the curve $g_1=g_2$.

The inequality (57) implies that there is at most one point $(\bar{c}_1,\bar{c}_2)$ on the curve $g_1=g_2$ at which $h_1-h_2+1=0$, and at which therefore $V'(r)=0$ for any r satisfying $c_i^r(r)=\bar{c}_i$, $i=1,2$. If $(c_1^r(r),c_2^r(r))$ lies to the left of that point, $V'(r)>0$; if it lies to the right, $V'(r)<0$. If there exists such a point $(\bar{c}_1,\bar{c}_2)$, and if there exists $r^* < T$ for which

$$c_i^{r*}(r^*) = \bar{c}_i, \qquad i = 1, 2, \tag{58}$$

this r^* defines an optimal policy. We have not argued that there is never more than one value of r^* satisfying (58).

In order to give sufficient conditions for the existence of an optimum, we need conditions that imply the existence of a point satisfying the definition of $(\bar{c}_1, \bar{c}_2)$. In the next section we show that it is sufficient to have three conditions (in addition to the existence of an optimal path for each $r < T$), that marginal utilities range over all positive numbers, that the moral hazard problem does not become vanishingly small, and that f, if it tends to zero as $t \to T$, does not do so too rapidly. Specifically, we have:

Theorem 4. Assume that for $i = 1, 2$, $u_i'(0) = \infty$, $u_i'(\infty) = 0$; that there exists $a > 0$ such that

$$u_2 - u_1 \geqq a \qquad \text{when} \qquad u_1' = u_2';$$

and that

$$\frac{(T - t) f(t)}{1 - F(t)} \tag{59}$$

is bounded as $t \to T$. Then there is an optimal r less than T. The optimal path or paths is identified by the unique $(\bar{c}_1, \bar{c}_2)$ satisfying

$$h_1(\bar{c}_1) - h_2(\bar{c}_2) = -1, \qquad u_1'(\bar{c}_1) = u_2'(\bar{c}_2),$$

or by the smallest possible values of $c_i^r(r)$.

The condition (59) is used to prove that optimal r is less than T: it is automatically satisfied if $f(t)$ tends to a nonzero limit as $t \to T$. In cases where T is the optimal r, it has to be optimal for c_1 and c_2 to tend to infinity as t tends to T. Such cases are therefore rather peculiar.

In the next section, proofs of the above propositions are given. It can be omitted without loss of continuity.

8. Proof of theorem 4

We proceed by a sequence of lemmas.

Lemma 4. *$V(r)$ is a differentiable function, and*

$$V'(r)=\frac{1-F(r)}{J_r(g_2)}[h_1(r)-h_2(r)+1], \tag{60}$$

where h_1, h_2 are defined by (54).

Proof. Consider paths which give utility v with retirement planned at r, constructed as follows: u_1^r is chosen to minimize $J_r(z^r_{V(r)})$ no matter what the value of v; and c_2 is then chosen so as to achieve utility v:

$$c_1^r(t)=G_1(u_1^r(t)), \tag{61a}$$

$$c_t^r(t)=G_2\left(\frac{1}{T-t}\left[v-\int_0^t u_1^r(t')\,\mathrm{d}t'\right]\right). \tag{61b}$$

Except when v equals $V(r)$ there is no assurance that such a path is feasible, but we need not be concerned. By the optimality of the path u_1^r used for this purpose, we have, for all r and s,

$$J_r(z^s_{V(r)})\geqq J_r(z^r_{V(r)}), \tag{62}$$

where, using (61a) and (61b), we define for any r and v,

$$z_v^r(t)=\int_0^t c_1^r(t')\,\mathrm{d}t'+c_2^r(t)(T-t)-t. \tag{63}$$

Since the resource constraint is fixed, we have equal resource use on any optimal path:

$$J_r(z^r_{V(r)})=J_s(z^s_{V(s)}). \tag{64}$$

Combining thus (63) and (64) we have the pair of inequalities:

$$J_r(z^s_{V(r)})-J_s(z^s_{V(s)})\geqq 0, \tag{65a}$$

$$J_r(z^r_{V(r)})-J_s(z^r_{V(s)})\leqq 0. \tag{65b}$$

$z_v^r(t)$ is a continuous function of r. Consequently (65) implies that there exists s' between r and s for which

$$J_r(z^{s'}_{V(r)})-J_s(z^{s'}_{V(s)})=0. \tag{66}$$

The two paths in (66) give the same consumption to workers. Thus applying the mean value theorem to $z^r_v(t)$, we have

$$z^{s'}_{V(r)}(t)-z^{s'}_{V(s)}(t)=(T-t)\left[G_2\left(\frac{1}{T-t}\left[V(r)-\int_0^t u_1^{s'}(t')\,dt'\right]\right)\right.$$

$$\left.-G_2\left(\frac{1}{T-t}\left[V(s)-\int_0^t u_1^{s'}(t')\,dt'\right]\right)\right]$$

$$=G_2'\left(\frac{1}{T-t}\left\{v'-\int_0^t u_1^{s'}(t')\,dt'\right\}\right)\left\{V(r)-V(s)\right\}, \quad (67)$$

where v' is a function of r, s, s', t that lies between $V(r)$ and $V(s)$. As s tends to r, G_2' in (67) tends to

$$g_2^r(t)=G_2'\left(\frac{1}{T-t}\,V(r)-\int_0^t u_1^r(t')\,dt'\right),$$

and consequently

$$\lim_{s\to r} J_s(G_2')=J_r(g_2^r). \quad (68)$$

Applying J_s to (67) and substituting from (66), we have

$$J_r(z^{s'}_{V(r)})-J_s(z^{s'}_{V(r)})+J_s(G_2')\{V(r)-V(s)\}=0. \quad (69)$$

The definition of the J-operator shows that

$$\frac{J_s(z)-J_r(z)}{s-r}=\frac{\int_r^s z(t)f(t)\,dt-z(r)\int_r^s f(t)\,dt}{s-r}+\frac{z(s)-z(r)}{s-r}\{1-F(s)\}. \quad (70)$$

Putting $z=z^{s'}_{V(r)}$, we see that the first term here tends to zero as $s\to r$, while

$$\frac{z(s)-z(r)}{s-r}=\frac{\int_r^s [c_1(t')-c_2(r)]\,dt'+(c_2(s)-c_2(r))(T-s)-s+r}{s-r}$$

$$\to c_1^r(r)-c_2^r(r)+c_2(r)(T-r). \quad (71)$$

Dividing (69) by $r-s$, substituting from (70) and (71), and taking the limit as $s \to r$, we see that V is differentiable and has the value given in (60).

Lemma 5. If for $i=1,2$, $u_i'(0)=\infty$, $u_i'(\infty)=0$; and if there exists $a>0$ such that

$$u_2-u_1 \geqq a \qquad \text{when} \qquad u_1'=u_2', \tag{72}$$

then there exists $\bar{c}_1, \bar{c}_2$ such that

$$h_1(\bar{c}_1)-h_2(\bar{c}_2)=-1, \qquad u_1'(\bar{c}_1)=u_2'(\bar{c}_2). \tag{73}$$

Proof. Consider first how $h_i(c_i)$ behaves as $c_i \to 0$, or, equivalently, $g \to 0$. Let ε be an arbitrarily small positive number. Since u_i is concave,

$$u_i(c_i)+u_i'(c_i)(\varepsilon-c_i) \geqq u_i(\varepsilon).$$

Therefore

$$h_i(c_i)=\frac{u_i(c_i)}{u_i'(c_i)}-c_i \geqq \frac{u_i(\varepsilon)}{u_i'(c_i)}-\varepsilon$$

$$\to -\varepsilon \qquad \text{as} \qquad c_i \to 0.$$

It follows that if u_i is negative for small c_i, $\lim h_i(c_i)=0$; while if $u_i(0) \geqq 0$, it is also true that $\lim h_i(c_i)=0$. Thus

$$\lim_{g \to 0} [h_1(c_1)-h_2(c_2)]=0. \tag{74}$$

Now consider what happens as $c_1, c_2 \to \infty$; or, equivalently, since $\lim u_i'=0$, $g \to \infty$. From (57), (72), and the assumption that $g \to \infty$, h_1-h_2 decreases without limit:

$$\lim_{g \to \infty} [h_1(c_1)-h_2(c_2)]=-\infty. \tag{75}$$

Since h_1-h_2 is a continuous decreasing function of g, and (74) and (75) hold, a solution of (73) must exist, and the lemma is proved.

Lemma 6. Let

$$\phi(t)=\frac{(T-t)f(t)}{1-F(t)} \tag{76}$$

be bounded as $t \to T$. *If* $x(r)$ *is defined by*

$$Z(r, x(r)) + A = 0,$$

and $A < \int_0^T [1 - F(t)]\,\mathrm{d}t$ *(so that positive consumption is feasible), and* $u_1'(0) = -\infty$,

$$x(r) \to \infty \qquad \text{as} \qquad r \to T. \tag{77}$$

Proof. If the conclusion does not hold, there exists a sequence $\{r_v\}$ tending to T, and a positive number M, such that

$$c_1^r(r) \leqq M, \tag{78}$$

for $r = r_1, r_2 \ldots$. We show that this implies, for each $t < T$, that

$$c_1^{r_v}(t) \to 0, \qquad c_2^{r_v}(t) \to 0,$$

as $r_v \to T$. Thus, in the limit, resources are not fully utilized, and the supposition that these are optimal plans must be false.

Let ε be a positive number. Then there exists $\alpha > 0$ such that

$$\max\left[\frac{1}{u_1'(c_1)} - \frac{1}{u_2'(c_2)},\ u_2(c_2) - u_1(c_1)\right] \geqq \alpha, \tag{79}$$

whenever $\varepsilon \leqq c_1 \leqq M$, and $c_2 \geqq 0$. For if this were not the case, there would be a limit point c_1', c_2' such that $u_1'(c_1') \geqq u_2'(c_2')$, $u_1(c_1) \geqq u_2(c_2)$, $0 < c_1'$; and that is impossible by assumption.

Consider now an r-optimal plan satisfying (78). Applying (79) to the differential equations satisfied by the optimal path, we have, so long as $c_1^r(t) \geqq \varepsilon$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(g_1 + u_2) &= (g_1 - g_2)\frac{f}{1-F} + (u_2 - u_1)\frac{1}{T-t} \\ &\geqq \alpha \min\left(\frac{f}{1-F}, \frac{1}{T-t}\right).\end{aligned} \tag{80}$$

Since ϕ defined in (76) is bounded, there exists $N > 1$ such that

$$\frac{(T-t)f(t)}{1-F(t)} \leqq N,$$

and therefore (80) implies

$$\frac{\mathrm{d}}{\mathrm{d}t}(g_1+u_2)\geqq\frac{\alpha}{N}\frac{f}{1-F}$$
$$=-\frac{\alpha}{N}\frac{\mathrm{d}}{\mathrm{d}t}\log[1-F(t)].$$

Integrating from t to r we obtain

$$g_1(t)+u_2(c_2^r(t))\leqq g_1(r)+u_2(r)+\frac{\alpha}{N}\log[1-F(r)]-\frac{\alpha}{N}\log[1-F(t)]. \tag{81}$$

Since any optimal plan satisfies $u_2'\geqq u_1'$, we have

$$c_2^r(t)\leqq\psi(c_1^r(t)), \tag{82}$$

where ψ is a continuous increasing function, whose graph is the curve $g_1=g_2$ in fig. 6. Since $u_1'(0)=u_2'(0)=\infty$, $\psi(0)=0$. Applying (78) and (82) to (81), and omitting the positive term $g_1(t)$, we have

$$u_2(c_2^r(t))<u_2(\psi(M))+\frac{1}{u_1'(M)}+\frac{\alpha}{N}\log[1-F(r)]-\frac{\alpha}{N}\log[1-F(t)],$$

provided that $M\geqq c_1^r(t)\geqq\varepsilon$.

If $c_1^r(t)<\varepsilon$, (82) implies that $c_2^r(t)<\psi(\varepsilon)$. Thus we have, in any case,

$$u_2(c_2^r(t))<\max\{D(t)+\frac{\alpha}{N}\log[1-F(r)],\ u_2(\psi(\varepsilon))\}, \tag{83}$$

where $D(t)$ is independent of r. Letting $r=r_\nu\to T$, and $\varepsilon\to 0$ in (83), we deduce that for any $t<T$,

$$c_2^{r_\nu}(t)\to 0. \tag{84}$$

To prove that $c_1^{r_\nu}(t)$ also tends to zero, we use the differential equation

$$\dot{g}_1=\frac{f}{1-F}(g_1-g_2),$$

which can be solved in the form

$$[1-F(t)]g_1(t)=[1-F(r)]g_1(r)+\int_t^r g_2(t')f(t')\,dt'.$$

Therefore if $t<s<r_v$, we have, using (78) and (82),

$$[1-F(t)]g_1^{r_v}(t)\leqq[1-F(r_v)]\frac{1}{u_1'(M)}+\int_t^s g_2^{r_v}(t')f(t')\,dt'$$
$$+[F(r_v)-F(s)]\frac{1}{u_2'(\psi(M))}.$$

Now let $r_v\to T$, and then $s\to T$. We obtain

$$g_1^{r_v}(t)\to 0,\qquad t<T,$$

which is equivalent to

$$c_1^{r_v}(t)\to 0,\qquad t<T. \tag{85}$$

It follows from (84) and (85) that, as $r_v\to T$,

$$J_{r_v}(z)\to-\int_0^T[1-F(t)]\,dt,$$

which is strictly less than $-A$. Therefore, for all r_v close enough to T, there are resources to spare, and the plan cannot be optimal. The assumption that (78) holds for a sequence of r tending to T thus leads to a contradiction, and the lemma is proved.

When the assumptions of these lemmas are satisfied, optimal r^* is less than T. For when $x(r)>\bar{c}_1$, $V'(r)<0$. Since, by lemma 6, $x(r)\to\infty$, $V'(r)<0$ for all r sufficiently close to T. Since V is evidently continuous at T, T cannot be the optimal value of r.

We have found conditions for the existence and uniqueness of the desired point $(\bar{c}_1,\bar{c}_2)$ and for its being less than $(c_1^r(r),c_2^r(r))$ for r sufficiently close to T. There remains one loose end. If the government is putting net resources into the social insurance system, $A<0$, then the smallest value of $c_i^r(r)$ as r varies is strictly positive. If $\bar{c}_1$ is less than the smallest $c_1^r(r)$, the optimum occurs at this corner solution.

9. Features of solution

Several aspects of the optimum emerge from the general analysis and call for explicit notice. In the first place, since v is constant, the consumer is indifferent about the date of his retirement. In effect, he is assumed to retire when the government wishes him to. A small deviation from the optimum would make the consumer strictly prefer r to any other retirement date. But the government wishes consumers to retire before the end of life (and could set consumption for later retirers to zero to achieve this). This latter feature contrasts with the situation where the full optimum is achievable, in which, normally, everyone who can work does. The willingness to continue working beyond r comes from pension growth which is sufficiently rapid to preserve utility (42). As we will note below (see theorem 5), at the planned retirement date, the net cost to the government of having an individual work longer stops shrinking and starts growing. Thus the individual would not choose to work beyond r if the total compensation for additional work did not exceed the marginal product of labour.

Secondly, since the lifetime utility of a man who loses his ability at s is $v(s) - (T-s)b$, the entire cost of ill-health is borne by the sufferer,[7] but this is the only source of difference among the utilities of different individuals. If b were zero, everyone would have the same lifetime utility, while the marginal utility of consumption in different periods would vary from individual to individual. Insurance would be perfect in a naive sense, not the economist's sense.

Thirdly, it may seem curious that the solution is independent of b, that is, independent of the effect of ill-health. The reason is that marginal utilities are, by assumption, unaffected by the state of health of the consumer. In the more general case where the difference between u_2 and u_3 varies with c_2, the manner of dependence does affect the optimum.

Fourthly, it is worth summarizing the basic features of the optimum that emerge from the diagram:

$$u_2(c_2) > u_1(c_1). \tag{86}$$

Thus an individual would always be immediately better off if he retired. He is prepared to continue working because of improved retirement benefits with additional work. It may be helpful to consider why one would not want $u_2 = u_1$. Starting from such a situation, an increase in c_2 at retirement date, financed by a decrease in c_1 earlier, increases expected utility from planning retirement at that date, since it transfers consumption to the state with

[7] We would not expect this result without the assumption that the marginal utility of consumption of a nonworker is independent of his health.

higher marginal utility. Thus work is not discouraged, and utility increased. The argument works at r, and therefore also at earlier dates.[8]

We also have

$$u_2'(c_2) > u_1'(c_1). \tag{87}$$

Thus the individual would prefer more insurance. Moral hazard keeps the extent of social insurance down.

We recall also the basic intertemporal facts, that c_1 and c_2 are differentiable functions of time, and that

$$\frac{\mathrm{d}c_1}{\mathrm{d}t} > 0, \tag{88}$$

$$\frac{\mathrm{d}c_2}{\mathrm{d}t} > 0. \tag{89}$$

There is no difficulty in principle about calculating optimal policies. First the terminal point of the optimal path has to be found, by solving simultaneously $u_1' = u_2'$ and $h_2 - h_1 = 1$. Paths are calculated backwards from there with various values of r. In each case, $J_r(z)$ is calculated, and we have the solution for that particular value of the resource constraint. We illustrate the procedure with a simple case.

Example:

$$u_1 = \log c, \qquad u_2 = a + \log c, \qquad a > 0$$
$$f(s) = 1, \qquad T = 1.$$

It is readily checked that theorems 3 and 4 apply. In terms of c_1 and c_2, the differential equations in theorem 3 are

$$(1-t)\frac{\dot{c}_2}{c_2} = \log\frac{c_2}{c_1} + a,$$
$$(1-t)\dot{c}_1 = c_1 - c_2.$$

[8]To put it another way, higher retirement benefits are greater incentives to work in every earlier period. The later the date the more periods in which work is encouraged. With rising pension benefits, a rising wage is also desirable. Starting with a constant consumption plan, transferring resources to older ages permits more of the resources to go to consumption in the state with higher marginal utility.

The terminal point is given by $c_1 = c_2$ and

$$c_1(\log c_1 - 1) - c_2(\log c_2 + a - 1) + 1 = 0,$$

so that $c_1 = c_2 = 1/a$.

To solve the differential equations, we define

$$w = \log \frac{c_1}{c_2}, \qquad m = \log(1-t) - \log(1-r).$$

In terms of these variables, we have

$$\frac{dw}{dm} = e^{-w} - w + a - 1, \qquad w = 0 \quad \text{at} \quad m = 0,$$

$$\frac{dc_2}{dm} = c_2(w - a), \qquad c_2 = \frac{1}{a} \quad \text{at} \quad m = 0.$$

Solution of these equations gives the optimal path for any value of r. For each r we then compute $J_r(z)$, which is found, after a little manipulation, to be

$$J_r(z) = \int_0^r (1-t)(c_1 + c_2 - 1)\,dt + \frac{(1-r)^2}{a}$$

$$= (1-r)^2 \left[\int_0^{-\log(1-r)} (c_2 + c_2 e^{-w} - 1) e^{2m}\,dm + \frac{1}{a} \right].$$

In this case, then, it is relatively easy to obtain the solution for different resource constraints, since only one solution of the differential equations is required.

In table 1 we give solutions for[9] $a = 0.5$, and three values of the government net subsidy to the scheme, $-A$, 0.5, 0, and -0.2. Table 2 displays solutions for a case with much greater disutility of work, $a = 1$, and the same values of the net subsidy. It should be noted that in table 1 c_1 and c_2 are equal to 2 at r. This is a very high figure relative to the wage, which is unity. Since expected consumption is $r - (r^2/2) - A < \frac{1}{2} - A$, it is even higher relative to mean consumption. A rather implausible assumption about the disutility of work is required to reduce $c_1(r)$ and $c_2(r)$ below unity. But in all cases people receive consumption during most of working life that is less

[9] If the utility function with labour y were $\log c + \log(2-y)$, implying unit labour supply in the absence of lump-sum income, the disutility of work is $\log(2-0) - \log(2-1) = \log 2 = 0.6931$. This Cobb–Douglas model probably overstates the utility of not working at all.

than the wage plus the net subsidy, and only a very small proportion of the population obtain a retirement benefit that is greater than wage plus net subsidy.

Two other features of the numerical results deserve comment. In all cases, $c_1 - c_2$ increases with t until t is very close to r.

The second feature is that in table 1, r is nearly 1, even when there is a very large government deficit ($A = -0.5$). Thus the direct effect of moral

Table 1

$a = 0.5$.

	$A = -0.5$		$A = 0$		$A = 0.2$	
t	c_1	c_2	c_1	c_2	c_1	c_2
0	1.00	0.77	0.50	0.38	0.30	0.23
0.1	1.03	0.78	0.51	0.39	0.31	0.24
0.2	1.05	0.81	0.53	0.40	0.32	0.24
0.3	1.09	0.83	0.54	0.42	0.33	0.25
0.4	1.13	0.86	0.56	0.43	0.34	0.26
0.5	1.17	0.90	0.59	0.45	0.35	0.27
0.6	1.24	0.95	0.62	0.47	0.37	0.28
0.7	1.32	1.02	0.66	0.51	0.40	0.31
0.8	1.45	1.12	0.73	0.56	0.44	0.33
0.9	1.69	1.34	0.86	0.66	0.51	0.39
0.95	1.92	1.64	1.01	0.77	0.60	0.46
0.99			1.46	1.13	0.88	0.67
	$r = 0.969$		$r = 0.998$		$r = 1.000$ (more accurately: $1 - 1.8 \times 10^{-4}$)	

Table 2

$a = 1$.

	$A = -0.5$		$A = 0$		$A = 0.2$	
t	c_1	c_2	c_1	c_2	c_1	c_2
0	0.87	0.59	0.49	0.28	0.30	0.17
0.1	0.90	0.63	0.52	0.30	0.31	0.18
0.2	0.93	0.69	0.54	0.32	0.33	0.19
0.3	0.96	0.75	0.57	0.34	0.35	0.20
0.4	0.99	0.85	0.61	0.36	0.37	0.21
0.5			0.66	0.39	0.40	0.23
0.6			0.72	0.44	0.44	0.25
0.7			0.80	0.51	0.50	0.29
0.8			0.91	0.65	0.59	0.35
0.9					0.78	0.49
0.95					0.96	0.76
0.99						
	$r = 0.494$		$r = 0.882$		$r = 0.963$	

hazard, arising from the consumer's ability to retire healthy without special penalty, is almost entirely eliminated: hardly anyone retires early. Nevertheless, the optimal policy looks very different from the first-best optimum that would be achievable if individuals could be directly prevented from retiring except through ill-health. In that case, consumption would be constant throughout life, and the same whether the individual is working or not.

When expected utility levels for the first- and second-best optima are compared, it is found that the utility loss from moral hazard may be small, despite the very different character of the optimal policy. For each of the six cases described above, $r=1$ in the first-best optimum, and the consumption level is dictated by this fact. To calculate expected utility in the second-best optimum, we use the constancy of utility with respect to planned retirement date to deduce that expected utility is equal to the utility of a man who retires at time zero, viz. $u_2(c_2(0))$. The results are given in table 3. Utility has been transformed exponentially, so as to give figures comparable with consumption. It can be seen that the utility loss from moral hazard is small when $a=0.5$, but significant (up to 6 percent of consumption) when $a=1$.

In addition to comparing first- and second-best optima, it is natural to examine the implications of less complicated policies. As one example of a simpler policy consider selecting the optimal wage and benefit, assuming they are to be constant over the worker's lifetime. The expected utility levels from following the best example of this policy (setting $\log c_1 = a + \log c_2$) is shown in the third column of table 3. The level of exponentiated expected utility equals the level of c_1 actually consumed. The level of c_2 satisfies $c_1/c_2 = e^a$. With $a=0.5$, this policy is close to the other two in expected utility. The logarithmic utility function allows large differences in consumption patterns to yield similar expected utilities: the first best, c_2/c_1 is 1, while their ratio is 0.61 with the optimal

Table 3

Utility at the optimum.[a]

a	A	First best	Second best	Constant policy
0.5	−0.5	1.28	1.27	1.24
0.5	0	0.64	0.63	0.62
0.5	0.2	0.39	0.38	0.37
1	−0.5	1.65	1.60	1.46
1	0	0.82	0.76	0.73
1	0.2	0.49	0.46	0.44

[a]The figures entered in the table are e^u, where u is expected utility, and are therefore the equivalent consumption level for a man working throughout life. No account is taken of utility losses through disability, which are the same in all cases.

constant policy. When $a=1$, the differences among policies are more substantial. In moving from the best constant policy to the first best, the ratio of c_2/c_1 changes from 0.37 to 1.

10. Private saving

As in the simpler models, we can ask whether individuals would attempt to alter their consumption plans if they thought they could lend or borrow at the zero interest rate which the government faces. Calculating the gain from savings, we shall show that individuals would choose to save.

A consumer working at date t wants to lend for repayment at date t' ($<r$) if

$$u_1'(c_1(t))<[u_1'(c_1(t'))(1-F(t'))+\int_t^{t'} u_2'(c_2(s))f(s)\,\mathrm{d}s]/[1-F(t)].$$

When the opposite inequality holds, the consumer would choose to lend for repayment at date t'.

Dividing by $t'-t$ and letting $t'\to t$, we find that the consumer wants to save if

$$\frac{\mathrm{d}}{\mathrm{d}t}u_1'(c_1)>\frac{f}{1-F}(u_1'-u_2'). \tag{90}$$

Only if there were equality here would the individual be content with the saving being done on his behalf by the government. Reference to theorem 3 shows that the optimal policy does not eliminate incentives to save or dissave; for in the optimum,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{u_1'}\right)=\frac{f}{1-F}\left(\frac{1}{u_1'}-\frac{1}{u_2'}\right). \tag{91}$$

Now a simple calculation shows that

$$-(u_1')^2\left(\frac{1}{u_1'}-\frac{1}{u_2'}\right)=\frac{(u_1'-u_2')^2}{u_2'}+u_1'-u_2'$$

$$>u_1'-u_2'. \tag{92}$$

Applying this inequality to (91), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}u_1'>\frac{f}{1-F}(u_1'-u_2'). \tag{93}$$

This proves that the consumer wants to lend, that is, to save. The government can prevent this happening by suitable taxation of saving.[10]

It is interesting to enquire what the government should do if it has chosen, or is constrained, to allow a perfect untaxed capital market. Then, in equilibrium,

$$\frac{\mathrm{d}}{\mathrm{d}t}u_1' = \frac{f}{1-F}(u_1' - u_2'), \tag{94}$$

and it is also the case that

$$u_1' = u_2' \qquad \text{at} \qquad r. \tag{95}$$

This last equality holds because of the possibility of saving for retirement, or borrowing against retirement benefits from the period when one knows one will be retired.

We conjecture that the third-best optimum, when private saving is unconstrained, is defined by (94), (95) and the condition that the consumer be indifferent about retiring age.

11. The magnitude of transfers

An insurance scheme is said to be actuarially fair if variations in the date of retirement have no effect upon the net discounted transfers to the individual. The optimal social insurance scheme we have derived is not in this sense actuarially fair. The net discounted transfer to an individual retiring at s is, in our notation, $z(s)$ (see eq. (33)).

The lifetime transfer declines with longer work if $z'(s) < 0$. From the government's budget equation, we can interpret $-z'(s)$ as the net tax on work at s. That is, the net tax equals the marginal product of labour less the gain from working. That gain equals the extra consumption from working, $c_1 - c_2$, plus the increase in the value of the pension $(T-s)\dot{c}_2(s)$. In addition the rate of decline decreases with age: $z''(s) > 0$. That is, the net tax on work declines with age, reaching zero at the optimal retirement age. Formally, we have:

Theorem 5. For $s < r$, z is a decreasing function of s when social insurance is optimal. Moreover, $z'(s)$ increases with s, reaching 0 at r.

[10] Since $u_2' > u_1'$, undesired saving can be discouraged only by a tax on saving: a tax on wealth would be insufficient since saving followed almost at once by retirement and dissaving would attract negligible tax.

Proof. The first statement follows from the second. From

$$z(s)=\int_0^s c_1(t)\,dt+c_2(s)(T-s)-s,$$

we calculate

$$z'(s)=c_1(s)-c_2(s)+\dot{c}_2(s)(T-s)-1$$

$$=c_1-c_2+\frac{u_2-u_1}{u_2'}-1,$$

and

$$z''(s)=\left(1-\frac{u_1'}{u_2'}\right)\dot{c}_1+\left(1-\frac{u_2-u_1}{u_2'^2}\,u_2''\right)\dot{c}_2. \tag{96}$$

Since $\dot{c}_1,\dot{c}_2>0, u_1'<u_2'$ and $u_1<u_2$ on the optimal path, (96) shows that $z''>0$. When $s=r, u_1'=u_2'$ and $\dot{c}_2=0$. Thus $z''(r)=0$.

The meaning of this result is that those who are unfortunate enough to suffer disability early in life receive a larger net transfer from the State than those able to work until late in life. The optimal social insurance scheme subsidises those who retire early, though only to the extent compatible with maintaining incentives to work. Stated alternatively, there is a net tax on work so that net government revenue increases with individual work. In addition the rate of tax on work decreases with age, reaching zero at the optimal retirement age. Moreover, since $\dot{c}_2>0$, the gain for additional work always exceeds the extra current consumption from working. Thus we have the following:

Corollary to theorem 5. Under the optimal policy, for all $t\leqq r$,

$$c_1(t)-c_2(t)<1,$$

i.e. the extra current consumption obtained by working is always less than the marginal productivity of labour.

12. Social Security in the United States

Despite the level of mathematical complexity, the models we have considered are very special. It would be silly to base policy on the particular equations we have derived. There are two aspects of the current U.S. public-provided pensions which it may not be premature to criticize on the basis of the analysis we have performed. First we shall put the models briefly in perspective. If individuals are saving rationally, the incentive for work comes from the change in their lifetime budget constraints with additional work. It does not matter how that change is divided between current wages and increased future benefits. If they are consuming their net wage, it is necessary to consider separately the incentives which come from current wages for work and from increased future pensions as a result of additional work. If they are following savings rules which differ from both of these models, the two parts of compensation matter differently, although not necessarily in the manner we have analysed. The U.S. population no doubt contains individuals whose behaviour is describable by a wide variety of models, and not just the fully rational model. Since it does not matter for rational individuals how the return to work is divided between wages and increase of pensions, it makes sense to pay attention to those who are not saving rationally in designing these two parts of total compensation.

In the model analysed we found that a growing pension benefit permitted a higher pension, relative to the wage, than would be possible otherwise given the moral hazard problem. However if a pension which grows with work done is to serve as an incentive for work, individuals need to be aware of the relationship between pension and work done. While it would be expensive, a greater flow of information from the Social Security Administration to individuals nearing retirement age about their own pension prospects might well be worth the cost.

Under the newly enacted amendments to the U.S. Social Security system individuals will receive their pensions independent of whether they continue working once they reach age 70. Before that age there is an earnings limitation on pension eligibility. In terms of our notation the payment of benefits independent of retirement represents a large and discontinuous increase in c_1, the consumption enjoyed while working. In our analysis we found that a growing level of c_1 was optimal but that the growth should be continuous. c_1 can be increased continuously either by reducing taxes repeatedly for those continuing to work or by paying a steadily growing fraction of pension benefits independent of any retirement test. For example, 15% of benefits might be paid at age 65 independent of retirement and 85% subject to the retirement test. The former fraction could grow steadily to 100% as in individual ages to 70. Taking account of diversity in the population will affect the optimal total compensation for work. As we argued above, the elements considered in this analysis will play a disproportionate role in the division of that total return between its two components.

13. Conclusions

The main results we have obtained may be summarised by highlighting what they would recommend if they were fully applicable to the real world. Many of the results suggest practicable policy changes, and may be worthy of study in more realistic models.

(1) The presence of moral hazard implies the desirability of policies that leave consumers indifferent about the date of retirement. This conclusion might be modified in models where individuals begin with different abilities, or different probability distributions for the loss of earnings.

(2) It is an essential adjunct to an optimal insurance scheme that capital transactions by consumers be taxed. In most cases, this would be a positive tax, discouraging saving. This conclusion might be modified if we allowed for nonrational saving behaviour.

(3) Post-retirement benefits show a strong tendency to increase with the age of retirement.

(4) Insurance contributions $(1-c_1)$ should diminish with age, sometimes quite rapidly. They may become negative eventually. This is not as odd as it seems, for the optimal allocation can be accomplished alternatively by paying part of benefits independent of retirement. The corollary to theorem 5 shows that the additional current consumption from working rather than retiring should never exceed the marginal product.

(5) There are some indications that contributions diminish more rapidly, and benefits increase more rapidly, as age increases.

(6) Under an optimal scheme, the incidence of retirement for reasons other than ill-health may be very low, even when the immediate utility gain from healthy retirement is quite large. Despite this, the form of the optimal scheme may be very different from a first-best redistribution between people of different ages that ignores all incentive considerations.

(7) An actuarially fair scheme discourages retirement less than an optimal scheme.

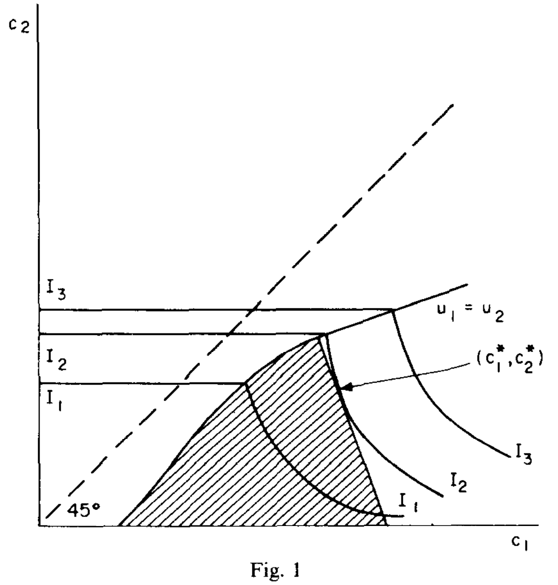

Fig. 1

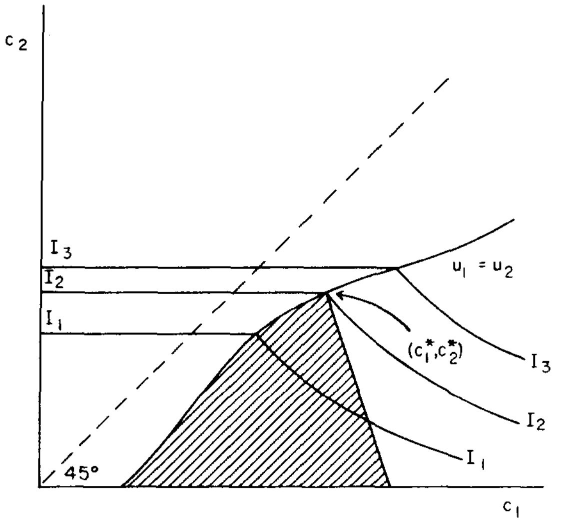

Fig. 2

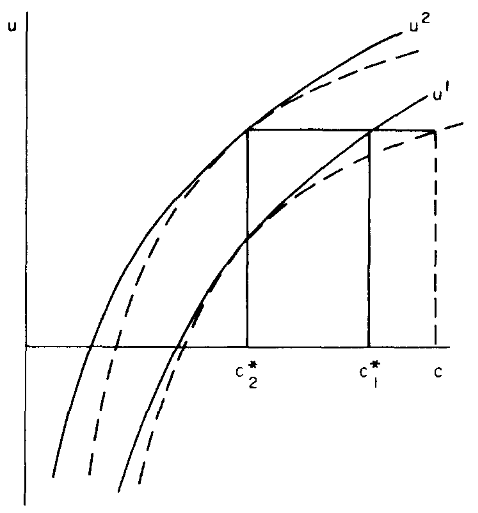

Fig. 3

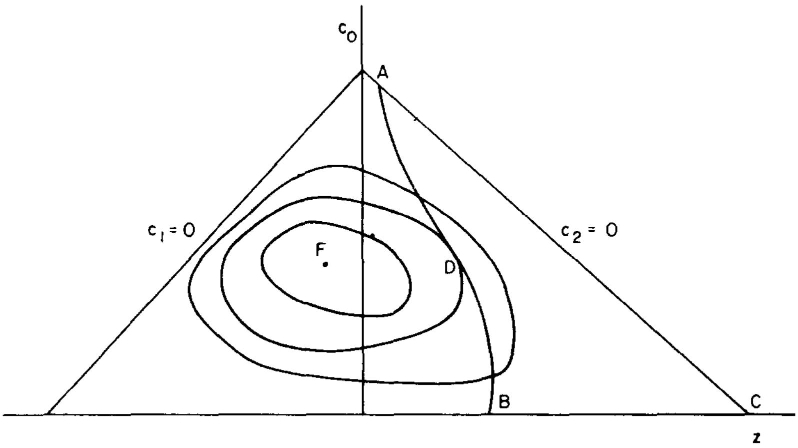

Fig. 4

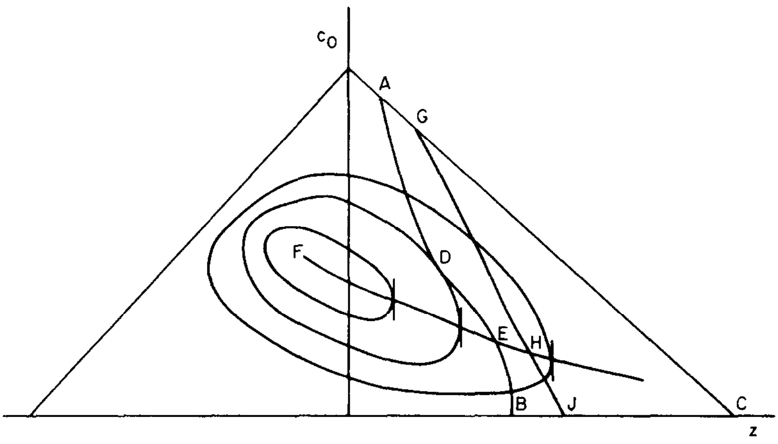

Fig. 5

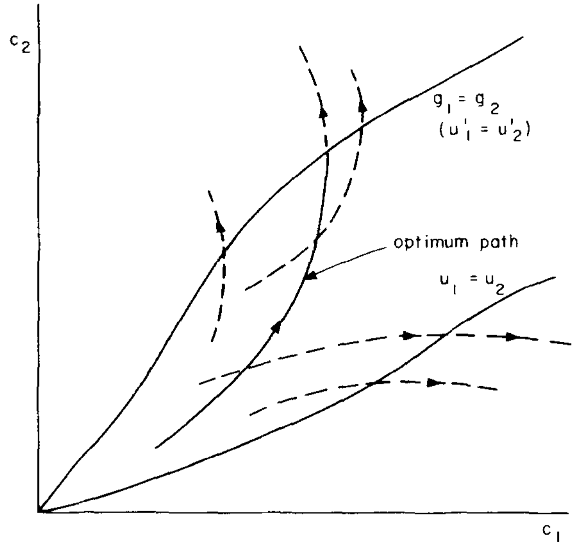

Fig. 6

References

Feldstein, M., 1974, Social security, induced retirement and aggregate capital accumulation, Journal of Political Economy 82, 905–926.

Sheshinski, E., 1978, A model of social security and retirement decisions, Journal of Public Economics, this issue, 337–360.