NBER WORKING PAPER SERIES

THE OPTIMAL LEVEL OF SOCIAL SECURITY BENEFITS

Martin Feldstein

Working Paper No. 970

The Optimal Level of Social Security Benefits

ABSTRACT

The optimal level of Social Security benefits depends on balancing
the protection that these benefits offer to those who have not provided
adequately for their own old age against the welfare costs of distorting
economic behavior. The primary such cost is the distortion in private
saving. The present paper derives the level of Social Security benefits
that is optimal in three basic cases.

In the first section of the paper, the optimal level of benefits is
derived for an economy in which all individuals do not anticipate retire-
ment at all and therefore do not save. The second and third sections then
derive the optimal benefits for economies with two different definitions
of attitudes toward retirement and saving.

Martin Feldstein
National Bureau of Economic Research
1050 Massachusetts Avenue
Cambridge, Massachusetts 02138

(617) 868-3905

The Optimal Level of Social Security Benefits

Martin Feldstein[*]

The provision of social security retirement benefits is a major government activity in every industrial nation. In the United States, these public pensions now account for more than 20 percent of the federal budget. The principle rationale for such mandatory programs is that some individuals lack the foresight to save for their retirement years. Since the provision of social security benefits imposes real costs on a nation, the optimal level of benefits requires balancing the protection of the myopic against the costs of distorted real resource allocation.

The primary cost of providing social security benefits is the welfare loss that results from reductions in private saving.[1] In addition, the payment of benefits distorts retirement behavior and the imposition of the tax used to finance the program distorts the labor supply during the preretirement period.[2] If tastes differ and borrowing is restricted, some individuals may be forced to consume more in their retirement years than they would with perfect foresight have chosen to do. Although all three types of welfare cost should be considered

---

[1]The tax used to finance the pension affects the reward for preretirement work effort -- unless the benefits of each individual provide an actuarily fair return based on the net-of-tax rate of return available to that individual.

[2]For a discussion of the nature of this welfare cost, see Feldstein (1982b). There has been substantial research on the effect of social security on private saving. Although there is no agreement on the magnitude of this effort, most studies find that social security reduces private saving. See e.g., Blinder, Gordon and Wise (1981), Diamond and Haussman (1982), Feldstein (1976, 1982a) and Kotlikoff (1979).

in selecting the optimal benefit level, the present analysis focuses exclusively
on the savings distortion.

In principle, the adverse effect of social security benefits on pri-
vate saving can be offset if the government accumulates an adequately large
social security trust fund (Feldstein, 1977). More generally, as Samuelson
(1975) has noted, a social security trust fund could acquire enough capital to
bring the economy to golden rule efficiency. In general, this would require
that the social security obligations are more than fully funded and may require
the trust fund to own the nation's entire capital stock. As a practical matter,
however, the social security programs in the United States and in many other
countries operates on a pay-as-you-go basis without a capital fund. I shall in
this paper derive the optimal benefit level in an unfunded program.

To focus on the welfare cost imposed by the saving distortion, I will
assume that labor is supplied inelastically and that the retirement date is exo-
genous. This also excludes the problem of uncertain health status that Diamond
and Mirrles (1978, 1981) have examined. In the economy that I study, all indivi-
duals also have equal earnings and identical tastes although they differ in
their ability to plan during their working years for their future retirement
period. There is no uncertainty about future rates of return, rates of popula-
tion growth, productivity and demographics.

The paper begins by deriving the optimal level of benefits in an
economy in which individuals do not anticipate their retirement at all and there-
fore do no saving. This establishes the general framework of analysis and pro-
vides a standard of reference for evaluating the effect of more realistic saving
behavior on the optimal level of benefits. The second and third sections then

derive these optimal benefit levels for two different definitions of incomplete myopia. There is a brief concluding section that discusses possible extensions of this analysis.

1. Optimal Social Security with Complete Myopia

The framework of analysis here follows Samuelson (1958) and uses a life cycle model with overlapping generations. Individuals work in the first period of their lives and retire in the second. The population grows at rate n per period. But unlike Samuelson's analysis, I assume that physical capital is a productive resource. To avoid the problem of an endogenous and varying rate of return, I assume that the marginal product of capital remains constant at ρ.

If L_t is the labor force at time t and A_t is the number of aged retirees at that time, population growth at rate n implies:

$$(1.1) \qquad L_t = (1+n) L_t - 1$$

and

$$(1.2) \qquad L_t = (1+n)A_t.$$

Each worker in period t earns a wage of w_t. The government imposes a tax at rate θ_t and therefore collects taxes of

$$(1.3) \qquad T_t = \theta_t w_t L_t.$$

Each aged retiree receives benefits of b_t, implying that total benefits are

$$(1.4) \qquad B_t = b_t A_t.$$

The pay-as-you-go character of the program implies that benefits and taxes are equal ($B_t = T_t$) and therefore that

(1.5) $$b_t A_t = \theta_t w_t L_t .$$

From equation 1.2 it therefore follows that

(1.6) $$b_t = \theta_t w_t (1 + n).$$

Equation 1.6 shows the relation between the social security tax rate and the level of benefits relative to concurrent earnings.

I will write the utility function of the representative individual in separable form as $u[\cdot] + v[\cdot]$ where the argument of the u function is first period consumption and the argument of the v function is consumption during retirement. Since the individual is myopic and does no saving, first period consumption is the net-of-tax wage, $(1-\theta)w_t$, and retirement consumption is the social security benefit, b_t. An additive social welfare function implies that social welfare at time t is

(1.7) $$W_t = L_t \, u[(1-\theta_t)w_t] + A_t v[b_t]$$
$$= \{(1 + n) \, u[(1-\theta_t)w_t] + v[b_t] \} L_{t-1} .$$

Using 1.6 to replace b_t implies:

(1.8) $$W_t = \{(1 + n) \, u[(1-\theta_t)w_t] + v[\theta_t w_t (1 + n)] \} L_{t-1} .$$

The first order condition for a maximum of W_t is $dW_t/d\theta_t = 0$ or

(1.9) $$-(1 + n) \, w_t \, u'_t + (1 + n) \, w_t v'_t = 0$$

or simply

(1.10) $$u'_t = v'_t .$$

It is thus optimal[1] to divide the total income of the working generations between the young and the old until their marginal utilities of consumption are equalized. This egalitarian optimum reflects the assumption that taxes do not distort any type of behavior and is reminiscent of the Edgeworth (1897) and Lerner (1944) conclusions about optimal income taxation and income redistribution.

Obtaining an explicit value for the optimal level of benefits requires making an assumption about the nature of the utility functions of workers and retirees. If we assume that u and v have the same functional form, the optimum condition of equation 1.10 implies that the arguments of u and v are equal and therefore that the optimal tax rate, θ^*_t, satisfies

$$(1.11) \qquad (1-\theta^*_t)\, w_t = \theta^*_t w_t (1+n) \qquad .$$

or

$$(1.12) \qquad \theta^*_t = \frac{1}{2+n}$$

Substituting θ^*_t into equation 6 shows that the ratio of optimal benefits, b^*_t, to concurrent wages is

$$(1.13) \qquad \beta^* = \frac{b^*_t}{w_t} = \frac{1+n}{2+n} \qquad .$$

Note first that the optimal tax rate and the optimal benefit ratio are constants that are independent of time. If there is no population growth (n=0), $\theta^* = \beta^* = 1/2$. In this case, workers give up half of their wages in tax and retirees receive benefits that are equal to half of the annual wage level. When the population is growing (n > 0), there are more workers than retirees. If

_____

[1]It is clear from 1.9 and 1.8 that the second order condition is always satisfied.

each worker gives up in taxes less than half of his wages, the retirees can still receive benefits that equal more than half of the annual wage level. The specific fraction is chosen to make the level of benefits equal to the after-tax wage. The faster the rate of population growth, the lower is the optimal tax rate and the higher is the corresponding ratio of benefits to wages.

To provide a numerical illustration of these optimal values, it is important to recognize that n refers to the growth rate per period and not per year. Since $L_t/L_{t-1} = 1 + n$, the value of n is the growth rate per generation. If the annual population growth rate is 1.4 percent[1] and a generation is 30 years, $1 + n = (1.014)^{30} = 1.52$. Equation 1.13 then implies that $\beta^* = 0.60$ and $\theta^* = 0.40$. A decline in the population growth rate to 0.7 percent a year implies that $1 + n = (1.007)^{30} = 1.23$ and therefore that optimal benefits are lower ($\beta^* = 0.55$) while the optimal tax rate is higher ($\theta^* = 0.45$).

2. Optimal Social Security Benefits with Partial Myopia

In reality, of course, not everyone is completely myopic. The interesting problem is therefore to characterize the appropriate partial myopia and to derive the corresponding optimal level of social security benefits. That optimization involves balancing the advantage of income support for the myopic against the loss caused by reduced saving.

Economic myopia has two aspects. The most important is that some or all individual's have, in Pigou's (1920) words, a "faulty telescopic faculty" that causes them to give too little weight to the utility of future consumption. Such faulty vision may also cause them to ignore or underestimate the size of their future social security benefits, thereby reducing the extent to which

---

[1]This was the average annual growth rate for the three decades beginning in 1950.

those promised benefits change personal saving.

One way to model incomplete myopia is to divide the population into two types, the first of which optimizes according to life cycle principles with perfect foresight while the second is completely myopic and does no saving. The optimal level of social security benefits depends on the relative numbers of life-cyclers and myopes in the population. I will develop this approach in the next section.

The second way to analyze the problem is to assume that all individuals are alike and are neither perfectly foresighted life-cyclers nor completely myopic.[1] Although the individual's "true" utility function is given by $u(c_1) + v(c_2)$ where c_1 and c_2 are consumption during the first and second periods of his life, he makes his saving decision during the first period of his life by maximizing $u(c_1) + \lambda v(c_2)$ where $\lambda < 1$ represents the degree of discounting of future consumption. If $\lambda = 1$, the individual is a proper life-cycler with no myopia. If $\lambda = 0$, the individual is completely myopic and has no reason to save.

Each individual earns a wage of w_t during his working years, pays taxes of θw_t and chooses to save s_t. His first period consumption is therefore $c_t = (1-\theta)w_t - s_t$. His saving earns a rate of return ρ and therefore provides him with $s_t(1+\rho)$ in his retirement period. To simplify the analysis, I shall ignore capital income taxes in this section. The individual's consumption during retirement is thus $c_{2, t+1} = s_t (1+\rho) + b_{t+1}$ where b_{t+1} is the individual's social security benefit. Myopic individuals may not only give too little weight to $v(c_2)$ but may also underestimate their future social security

---

[1] It would of course be possible to analyze a more general case in which there is a distribution of degrees of myopia but this would probably provide no additional analytic insights.

benefits by anticipating future consumption of only $s_t(1+\rho) + \alpha\, b_{t+1}$ with $\alpha < 1$.

To summarize, each individual chooses s_t to maximize $u[(1-\theta)w_t - s_t]$ $+ \lambda \cdot v\, [s_t(1+\rho) + \alpha b_{t+1}]$. Society chooses θ to maximize the sum over all periods of the true utilities $u[(1-\theta)\, w_t - s_t] + v\, [s_t(1+\rho) + b_{t+1}]$ where the s_t must be chosen by the individuals themselves.

The individual's first order condition for maximizing lifetime utility with θ and b_{t+1} taken as parameters is

$$(2.1) \qquad \frac{u'}{v'} = \lambda(1+\rho).$$

The lower the value of λ, the lower is the chosen u' relative to v' and therefore the higher is the chosen c_1 relative to c_2.

To permit an explicit solution for the chosen value of s, I will assume a logarithmic utility function: $u(c_1) = \ln c_1$ and $v(c_2) = \ln c_2$. Equation 2.1 then implies that the chosen value s_t^* satisfies

$$(2.2) \qquad \frac{s_t^*(1+\rho) + \alpha\, b_{t+1}}{(1-\theta)w_t - s_t^*} = \lambda(1+\rho)$$

or

$$(2.3) \qquad s_t^* = \frac{\lambda}{1+\lambda}\, [(1-\theta)w_t] - \frac{\alpha\, b_{t+1}}{(1+\lambda)(1+\rho)}$$

Note that this is an extension of the usual result with a log-linear utility function; with $\lambda = 1$ and $b = 0$, we get the familiar result that $s_t^* = 0.5\,(1-\theta)w_t$ and therefore $c_{it} = 0.5(1-\theta)w_t$. The rate of return matters only in converting the future income (b_{t+1}) to its present value. Low values of λ and high values of b both reduce saving while the underestimation of future social security benefits (a low value of α) raises saving.

As I noted in section 1, the pay-as-you-go nature of social security implies that

(2.4)
$$b_t L_{t-1} = \theta w_t L_t$$

or

(2.5)
$$b_t = \theta w_t (1+n).$$

Substituting 2.5 into 2.3 and noting that $w_{t+1} = (1+g)w_t$ and that $(1+g)(1+n) = 1+\gamma$ yields

(2.6)
$$s_t^* = \frac{\lambda}{1 + \lambda} (1-\theta)w_t - \frac{\alpha\theta(1+\gamma)w_t}{(1+\lambda)(1+\rho)}$$

Equation 2.6 describes the saving behavior of all relevant generations except the generation of retirees who receive unrequited benefits when the social security programs began. Although this group receives per capital benefits of $b_0 = \theta w_0(1+n)$, these were not anticipated at time $t = -1$ when they made their savings decisions. Thus

(2.7)
$$s_{-1}^* = \frac{\lambda}{1 + \lambda} w_{-1}$$

$$= \frac{\lambda w_0}{(1+\lambda)(1+g)}$$

When this generation retires at $t = 0$, their utility during retirement is

$$v[(1+\rho)s_{-1}^* + b_0] = v\left[\frac{\lambda w_0(1+\rho)}{(1+\lambda)(1+g)} + \theta w_0(1+n) \right]$$

Total utility at $t = 0$ is the sum of the utilities of the current workers and the current retirees: $w_0 = (1+n) u(c_{1,0}) + v(c_{2,0})$ or

$$(2.8) \qquad W_0 = (1+n) \cdot \ln[(1-\theta)w_0 - s_0^*] + \ln \left[\frac{\lambda w_0 (1+\rho)}{(1+\lambda)(1+g)} + \theta w_0 (1+n) \right]$$

where I have normalized by dividing all terms by the number of persons in the first generation of retirees.[1]

For all subsequent years, $(t > 0)$, total utility takes the form:

$$(2.9) \qquad W_t = (1+n)^{t+1} \ln[(1-\theta)w_t - s_t^*] + (1+n)^t \ln[(1+\rho)s_{t-1}^* + b_t]$$

$$= (1+n)^{t+1} \ln[1-\theta)w_t - \frac{\lambda}{1+\lambda}(1-\theta)w_t + \frac{\alpha\theta(1+\gamma)}{(1+\lambda)(1+\rho)} \ w_t \ ]$$

$$+ (1+n)^t \ln[\frac{\lambda(1+\rho)(1-\theta)}{1+\lambda} \ w_{t-1} - \frac{\alpha\theta(1-\gamma)}{(1+\lambda)} \ w_{t-1}]$$

This expression can be simplified by noting that $W_t = W_0(1+g)^t$ and that $1 - \lambda/(1+\lambda) = 1/(1+\lambda)$:

$$(2.10) \qquad W_t = (1+n)^t \ \{ \ (1+n) \ \ln[1-\theta+\alpha\theta(1+\gamma)(1+\rho)^{-1}]$$

$$+ \ln \ [\ \lambda(1+\rho)(1-\theta) + \theta(1+\gamma)(1+\lambda-\alpha)] \ \} + C_t$$

where C_t is a function of time that is independent of the policy parameter θ.[2]

It is clear from the form of 2.10 that the value of θ that maximizes W_t for any $t > 0$ also maximizes w_t for all other values of $t > 0$. This value of θ does not correspond to the full social optimum since it ignores the transfer of income from the first generaltion of workers (at $t = 0$) to the concurrent generation of retirees. Nevertheless, unless future utility is discounted at a very high rate, this initial period effect will be unimportant relative to the effect in all future periods.

---

[1] This is of course the "true" utility. The parameters λ and α enter because they influence individual behavior but not social evaluation.

[2] $C_t = (1+n)^{t+1} \ln w_0 (1+g)^t - (1+n)^{t+1} \ln(1+\lambda) + (1+n)^t \ln w_0 (1+g)^{t-1} - (1+n)^t \ln(1+\lambda)$

To see this, note that the true optimum θ^* maximizes the discounted sum of utilities:

$$(2.11) \qquad S = \sum_{t=0}^{\infty} \frac{W_t}{(1-\eta)^t}$$

where η is a pure time preference discount rate.[1] More specifically,

$$(2.12) \qquad S = W_O + f(\theta, \lambda, \alpha, \rho, \gamma, \eta,) \sum_{t=1}^{\infty} \left(\frac{1+n}{1+\eta} \right)^t + \sum_{t=1}^{\infty} \frac{C_t}{(1+\eta)^t}$$

where the f function is specified in equation 2.10 and is independent of time. The second term converges to a finite value if and only if $\eta > n$. This is also a sufficient condition for the third term to converge.[1] Since this term is independent of θ, it can be ignored in the maximization.

Using equation 2.6 and 2.8 to write W_O explicitly and evaluating the infinite sum in equation 2.12 implies that the optimal value θ^* maximizes:[2]

$$(2.13) \qquad S' = (1+n)\ln[1-\theta + \alpha\theta(1+\gamma)(1+\rho)^{-1}] + \ln\left[\frac{\lambda}{1+\lambda} \frac{1+\rho}{1+\gamma} + \theta \right]$$

$$+ \frac{(1-n)^2}{\eta-n} \ln[1 - \theta + \alpha\theta(1+\gamma)(1+\rho)^{-1}]$$

$$+ \frac{1+n}{\eta-n} \ln[\lambda(1+\rho)(1-\theta) + \theta(1+\gamma)(1+\lambda-\alpha)]$$

Note that the first two terms correspond to the utility of workers and retirees

---

[1] It is clear from footnote 2, page 10, that C_t increases as $t(1+n)^t$ and therefore increases more slowly than $(1+\eta)^t$ for $\eta > n$ as t tends to infinity.

[2] In evaluating W_O I have dropped the constant term $(1+n)\ln[w_O/(1+\lambda)]$ $+\ln w_O + \ln(1+n)$.

in the first period (t = 0) while the third and fourth terms correspond to the utility of workers and retirees in all future years. As the utility discount rate approaches the rate of population growth, the relative size of the terms corresponding to the first period tends to zero.

I shall begin therefore by ignoring the first period effect and finding the value of θ that maximizes the steady-state level of utility. I must emphasize that this is not equivalent to ignoring the obviously positive transfer to the initial retirees but ignores the taxes paid by the initial workers as well. After examining the steady-state solution, I shall return to the more general problem and derive the optimal θ^* corresponding to different values of the utility discount rate.

The value of θ that maximizes the steady-state level of utility is the solution of the first-order condition[1]:

$$(2.14) \qquad \frac{(1+n) \; [\alpha(1+\gamma)(1+\rho)^{-1}-1]}{1-\theta+\alpha\theta(1+\gamma)(1+\rho)^{-1}} \;+\; \frac{(1+\lambda-\alpha)(1+\gamma)-\lambda(1+\rho)}{\lambda(1+\rho)(1-\theta)+(1+\lambda-\alpha)\theta(1+\gamma)} \;=\; 0$$

The special case of $\alpha = 0$ is interesting because it corresponds to the situation in which individuals do not reduce their saving because of anticipated social security benefits. Saving is affected only by the reduction in disposable income caused by the tax at rate θ. With $\alpha = 0$, equation 2.14 implies

$$(2.15) \qquad \theta^* = \frac{(1+\lambda)(1+\gamma) - \lambda(1+\rho)(2+n)}{(1+\lambda)(1+\gamma)(2+n) - \lambda(1+\rho)(2+n)}$$

---

[1]This condition is equivalent to $dW_t/d\theta = 0$ where W_t is defined by equation 2.10 or equivalent $dS'/d\theta=0$ with the first two terms of S' ignored. The second order condition is easily checked and is always satisfied.

Note first that with complete myopia ($\lambda = 0$) this reduces to $\theta^* = (2+n)^{-1}$, the value derived in section 1 under the simpler specification of the complete myopia problem. A reduction in myopia, i.e., an increase in λ, necessarily reduces θ^*. It can be shown that

$$(2.16) \qquad \frac{d\theta^*}{d\lambda} = \frac{-(1+n)(1+\gamma)(1+\rho)}{[(1+\lambda)(1+\gamma) - \lambda(1+\rho)]^2(2+n)} \quad < \quad 0$$

This confirms that complete myopia sets the upper bound on the optimal tax and benefit levels.

With anything less than complete myopia, individuals save and social security distorts saving. The welfare loss of this distortion is an increasing function of the marginal product of capital. More specifically, it is clear that an increase in ρ reduces the numerator and denominator by an equal amount [$\lambda(2+n)$ per unit of ρ] and therefore $d\theta^*/d\rho < 0$ if $\lambda > 0$.

The extent of the welfare loss that is caused by reduced saving depends on the implicit return that individuals receive from the unfunded social security program. Since that rate of return is γ, an increase in γ reduces the loss caused by decreased saving and therefore raises the optimal value of θ. Differentiating θ^* with respect to γ shows

$$(2.17) \qquad \frac{d\theta^*}{d\gamma} = \frac{\lambda(1+\lambda)(1+\rho)(2+n)(1+n)}{[1+\gamma+\lambda(\gamma-\rho)]^2(2+n)^2}$$

which is clearly positive for $\lambda > 0$.

Equation 2.15 shows that there is a wide range of parameter values for which it is optimal to have no social security program at all. More specifically, equation 2.15 shows that $\theta^* = 0$ when $(1+\lambda)(1+\gamma) = \lambda(1+\rho)(2+n)$, i.e., when $\lambda = (1+\gamma)[(1+\rho)(2+n) - (1+\gamma)]^{-1}$. Since $\theta^* > 0$ at $\lambda = 0$ and equation 2.16 shows

that $d\theta^*/d\lambda < 0$, $\theta^* < 0$ for all values of $\lambda > (1+\gamma) \ [(1+\rho)(2+n)-(1+\gamma)]^{-1}$.
If social security taxes and benefits are constrained to be non-negative, $\theta^*=0$
for all $\lambda \geq (1-\gamma) \ [(1+\rho)(2+n)-(1+\gamma)]^{-1}$.

To evaluate this critical threshold for λ, it is important to
interpret the values of n, γ and ρ as growth rates per period and not per year.
The model specifies that $1+n$ is the ratio of the labor force in one generation
divided by the labor force in the previous generation. I shall again take a
generation to be 30 years. Since the U.S. labor force grew at an annual rate of
1.4 percent for the three decades beginning in 1950, $1+n = (1.014)^{30} = 2.97$.
Finally, I will use the estimate that the average annual marginal product of
capital in U.S. nonfinancial corporations during these years was 11.4 percent
(Feldstein, Poterba and Dicks-Mireaux, 1981) and write $1+\rho = (1.114)^{30} = 25.5$.
Substituting these values implies that the critical value of λ is 0.049. A
positive social security program is justified under these conditions only if
individuals give a weight of less than five percent to future utility.

Although these numbers are only illustrative, they do indicate that a
social security pension may be inappropriate even if economic myopia is univer-
sal and very substantial. This conclusion is not very sensitive to variations
in the three parameters (ρ, γ, and n) that determine the critical value of λ.
For example, reducing the annual marginal product of capital from 11.4 percent
to 8.0 percent or increasing the growth of income from 3.7 percent to 7.0 per-
cent still implies that $\theta^* > 0$ for any $\lambda > 0.14$.

In these calculations, I have assumed that $\alpha = 0$, i.e., that indivi-
duals ignore future social security benefits when making their savings decision.
Social security therefore affected savings only by reducing the disposable

income of workers. When the analysis is extended to recognize that workers
reduce their saving in anticipation of future benefits, the optimal level of
social security benefits is even lower.

Consider, for example, an economy in which the annual marginal product
of capital is 8.0 percent, population grows at 1.4 percent a year, and real
income rises at 2.3 percent a year. If $\lambda = 0.05$ and $\alpha = 0$, the optimal social
security tax rate is $\theta^* = 0.28$. If individuals take 10 percent of their future
social security benefits into account ($\alpha = 0.10$), the optimal tax rate falls
slightly to $\theta^* = 0.26$. If individuals take half of their benefits into account
($\alpha = 0.50$), $\theta^* = 0.20$. Further increases in α cause θ^* to decline rapidly. At
$\alpha = 0.67$, $\theta^* = 0$.

As I noted earlier, the value of θ^* that maximizes steady state
welfare ignores the effect of taxing the first generation of workers and trans-
ferring these revenues to the first generation of beneficiaries. The importance
of this omission depends on the discount rate, n, used to discount future utility.
We have implicitly been examining the limiting case of $\eta < n$ in which the
first period is irrelevant and the steady state welfare is the criterion. As
η increases, more weight is given to the initial period relative to the steady
state. It is interesting therefore to consider the other extreme in which
$\eta = \infty$ and only the initial period is important. With $\eta = \infty$ and $\alpha = 0$, the
first-order condition for a maximum of S' in equation 2.13 is

$$(2.18) \qquad - \frac{(1+n)}{(1-\theta)} + \frac{(1+\lambda)(1+\gamma)}{\lambda(1+\rho)+(1+\lambda)(1+\gamma)\theta} = 0$$

Using the values of n, γ and ρ that correspond to the last three decades of U.S.
experience implies that $\theta^* > 0$ if and only if $\lambda < 0.083$. Thus taking the first

year effects into account raises the critical value of λ from 0.049 to 0.083 but leaves unchanged the conclusion that a positive social security program is justified only if the universal myopia is extreme. For high but finite rates of utility discount, the critical value of λ lies between these extremes. For example, an annual rate of utility discount of 5 percent implies that the critical value of λ is 0.066 while a 10 percent discount rate makes the critical value 0.078.

Even when these high utility discount rates imply $\theta^* > 0$, the optimal value of θ^* remains low. For example, with an annual utility discount rate of 10 percent, the optimal value of θ is only 0.14 implying an optimal benefit-wage ratio of 0.21.

3. Optimal Benefits with a Heterogeneous Population

The conclusion of the previous section, that even with universal and extreme myopia it may be optimal to have no social security pension, rests on the assumption that no one is completely myopic. If some individuals are completely myopic, some provision must be made to support them in their old age. This might take the form of a means tested program or of private charity, as it did before the introduction of the universal social security program. In this section, however, I will assume that the only such support is the universal social security program. I will divide the population into a group who are completely myopic and a group who are not myopic at all and will examine how the optimal level of social security benefits varies with the fractions of the population who are "myopes" and "life-cyclers".

As in the analysis of the previous section, the initial period differs from all subsequent periods because the initial retirees receive an unpredic-

table windfall so that those who are not myopic have more private retirement wealth than life-cyclers will in subsequent generations. All subsequent periods are identical and the value of θ that maximizes total utility in any one period will maximize it in all periods. To simplify the analysis of this section, I will focus on this steady state level of utility.

As in the previous section, workers earn w_t and pay a social security tax of θw_t. The myopes consume their entire disposable income, $(1-\theta)w_t$. The life-cyclers set their saving optimally and therefore, using equation 2.3 with $\lambda = \alpha = 1$, $s_t = \frac{1}{2}[(1-\theta)w_t - b_{t+1}/(1+\rho)]$. Lifecyclers therefore consume $\frac{1}{2}[(1-\theta)w_t + b_{t+1}/(1+\rho)]$ during their working period. In retirement, the myopes consumption is equal to their social security benefits (b_{t+1}) while the life-cyclers consume $s_t(1+\rho) + b_{t+1} = \frac{1}{2}[(1-\theta)w_t(1+\rho) + b_{t+1}]$.

The rate of growth of the labor force implies that there are always $1 + n$ workers for every retiree. I shall denote the fraction of myopes by μ and normalize the social welfare equation by dividing all terms by the number of retirees in that generation. For this model, total welfare in period t can therefore be written:

(3.1) $W_t = (1+n)\mu \ln[(1-\theta)w_t] + (1+n)(1-\mu)\ln 0.5[(1-\theta)w_t + b_{t+1}/(1+\rho)]$

$+ \mu \ln b_t + (1-\mu)\ln 0.5[(1-\theta)w_{t-1}(1+\rho) + b_t]$.

Since $w_t = (1+g)w_{t-1}$ and $b_t = (1+n)\theta w_t$, equation 3.1 is equivalent to

(3.2) $W_t = (1+n)\mu \ln[(1-\theta)w_t] + (1+n)(1-\mu)\ln 0.5[(1-\theta)w_t + \theta(1+\gamma)w_t/(1+\rho)]$

$+ \mu \ln (1+n)\theta w_t + (1-\mu)\ln 0.5[(1-\theta)w_t(1+g)^{-1}(1+\rho) + (1+n)\theta w_t]$

Since w_t is a factor of every term, it can be eliminated without altering the value of θ that maximizes w_t. Similarly the 0.5 factors can be eliminated and

the last term can be rewritten $(1-\mu)\ln(1+n) + (1-\mu)\ln[(1-\theta)(1+\rho)(1+\gamma)^{-1} + \theta]$.

With these algebraic simplifications, the problem can be defined as selecting θ to maximize

$$(3.3) \quad W' = (1+n)\mu\ln(1-\theta) + (1+n)(1-\mu)\ln[1-\theta) + \theta(1+\gamma)(1+\rho)^{-1}]$$

$$+ \mu\ln\theta + (1-\mu)\ln[(1-\theta)(1+\rho)(1+\gamma)^{-1} + \theta]$$

The first two terms correspond to the utility of the two groups of workers while the second two terms correspond to the utility of the concurrent retirees. It is clear from the third term (the utility of the myopic retirees) that $\theta = 0$ can never be optimal.

The first-order condition for a maximum of W' with respect to θ is:

$$(3.4) \quad 0 = -\frac{(1+n)\mu}{1-\theta} + \frac{(1+n)(1-\mu)[(1+\gamma)(1+\rho)^{-1} - 1]}{1-\theta + \theta(1+\gamma)(1+\rho)^{-1}} + \frac{\mu}{\theta}$$

$$+ \frac{(1-\mu)[1-(1+\rho)(1+\gamma)^{-1}]}{\theta + (1-\theta)(1+\rho)(1+\gamma)^{-1}}$$

Note first that complete myopia for the entire population means $\mu = 1$ and therefore implies

$$(4.5) \quad \frac{-(1+n)}{1-\theta} + \frac{1}{\theta} = 0$$

or

$$(4.6) \quad \theta^* = \frac{1}{2+n} \quad .$$

This is the same optimum condition derived for complete myopia in sections 1 and 2. It again represents the upper bound on the possible values of θ^*. With the 1.4 percent annual population growth rate of the U.S. since 1950, this upper bound is $\theta^* = 0.40$. At the opposite extreme, if everyone is a lifecyler, $\mu = 0$

and condition 3.4 cannot be satisfied for any $0 < \theta$; the optimum θ^* is negative and the feasible optimum is to have no social security program.

Some tedious but straightforward manipulation permits equation 3.4 to be rewritten as:

(3.7) $\theta^2 (2+n)(1-x) - \theta[(2+n)(1-x)(1-\mu) + \mu(3+n-x)] + \mu = 0$

where $x = (1+\gamma)(1+\rho)^{-1}$. This quadratic equation shows that that optimal social security tax rate varies from $\theta^* = 0$ to $\theta^* = (2+n)^{-1}$ as the frequency of myopia varies from $\mu = 0$ to $\mu = 1$. For the values of n, γ and ρ corresponding to the U.S. experience of the past three decades, the optimal value of θ varies almost linearly with μ with $\theta^* = 0.11$ at $\mu = 0.25$, $\theta^* = 0.21$ at $\mu = 0.50$ and $\theta^* = 0.31$ at $\mu = 0.75$. The corresponding optimal benefit-wage ratios are $\beta^* = 0.17$ at $\mu = 0.25$, $\beta^* = 0.32$ at $\mu = 0.50$ and $\beta = 0.47$ at $\mu = 0.75$. Thus more than half of the population must be completely myopic for the optimal benefit-wage ratio to be as high as one-third and more than three-fourths must be completely myopic for the optimal benefit-wage ratio to be as high as one-half.

A realistic description of the population presumably involves a distribution of degrees of myopia. The analysis of this section and of the previous one suggests that even if (say) one-fourth of the population were completely myopic while three-fourths were so myopic that they gave future utility only one-fourth the weight of the current utility, the optimal ratio of social security benefits to wages would be very low, probably less than 0.20.

4. Conclusion

Although the specific numerical results reflect the simplified model and the logarithmic utility structure, the broad qualitative results are likely

to be valid more generally. The analysis has shown that even if every indivi-dual is substantially myopic (and would therefore save less for his retirement than perfect foresight utility maximization would imply), it may be optimal to have either no social security retirement program or a very low ratio of benefits to earnings. If some fraction of the population is completely myopic and would in the absence of a social security pension do no retirement saving, it cannot be optimal to have no social security program (unless some other retirement income is provided for nonsavers.) Nevertheless, the optimal level of benefits may be quite low unless a large fraction of the population is completely myopic.

It would, of course, be desirable to examine the sensitivity of these results to a richer class of models. Such extensions might include a tax on capital income (implying a gap between the marginal product of capital and the rate of return that influences individual savings decisions)[1] and a wider class of individual utility functions. The "mixed case" of a fraction of the popula-tion that is completely myopic while the rest are partly myopic might also be usefully analyzed. The assumption of predetermined retirement behavior should be modified to recognize that social security benefits influence retirement behavior and that planned retirement alters the individual's rate of saving.

The analyses of both models show how the optimal size of the social security program is related to the steady-state rates of growth of population and productivity and to the marginal product of capital. Current policy discussions about the adjustment of social security to the productivity slowdown and the changing demographic structure of the population suggest that these

---

[1]Feldstein (1982b) shows the importance of a capital income tax in evaluating the welfare cost of the saving distortion induced by social security.

models might provide a useful framework for studying how the level of social security benefits should be changed in response to temporary changes in the rates of growth of population and productivity.

Finally, it would be good to relax the assumption that benefits are to be provided uniformly and without reference to accumulated assets. The appropriate role for means-tested benefits for retirees could be derived by including such transfers as an additional parametric option in an extended version of the present models.

July 1982

Bibliography

Blinder, A., R. Gordon and D. Wise (1981) "Life Cycle Savings and Bequest: Cross-sectional Estimates of the Life Cycle Model," in The Determinants of National Savings and Wealth, F. Modligiani (ed.), London and New York: Macmillan, in press.

Diamond, P. A. and J. Hausman (1982) "Individual Retirement and Savings Behavior Decisions," presented at the Oxford Conference on Microdata and Public Economics, June 27-30, 1982.

Diamond, P. A. and J. Mirrlees (1978) "A Model of Social Insurance with Variable Retirement," Journal of Public Economics.

Diamond, P. A. and J. A. Mirrlees (1981) "Social Insurance with Variable Retirement and Private Savings," Journal of Public Economics.

Edgeworth, F. Y. (1897) "The Pure Theory of Taxation," Economic Journal, 7 46-70, 226-238 and 550-571.

Feldstein, M. (1976) " Perceived Wealth in Bonds and Social Security: A Comment," Journal of Political Economy 84, no. 2, April, 331-36.

Feldstein, M. (1977) "The Social Security Fund and National Capital Accumulation," Funding Pensions: The Issues and Implications for Financial Markets, A Federal Reserve Bank of Boston publication.

Feldstein, M. (1982a) "Social Security and Private Saving: Reply." Journal of Political Economy 90, 630-642.

Feldstein, M. (1982b) "The Welfare Cost of Social Security's Impact on Private Saving." National Bureau of Economic Research Working Paper, forthcoming.

Feldstein, M. J., Poterba, J. and L. Dicks-Mireaux (1981) "The Effective Tax Rate and the Pre-Tax Rate of Return," National Bureau of Economic Research Working Paper No. 740, Journal of Public Economics, forthcoming.

Kotlikoff, L. (1979) "Testing the Theory of Social Security and Life Cycle Accumulation," American Economic Review 69, 396-410.

Lerner, A. P. (1944) The Economics of Control: Principles of Welfare Economics, New York: Macmillan.

Pigou, A. C. (1920) The Economics of Welfare, London: Macmillan.

Samuelson, P. A. (1958) "An Exact Consumption-loan Model of Interest with or without the Social Contrivance of Money" Journal of Political Economy, 66, 467-482.

Samuelson, P. A. (1975) "Optimal Social Security in a Life-cycle Growth Model," International Economic Review 16, 539-544.