Can and should a pay-as-you-go pension system mimic a funded system?*

by

John Hassler

and

Assar Lindbeck

Institute for International Economic Studies‡

Submitted to Economic Journal

June 4, 1998

Keywords: Pension systems, pay-as-you-go, actuarial, funding

Abstract

This paper considers the possibility of letting a pay-go pension system mimic a fully funded pension system. Generically, it turns out to be impossible to make a less than fully funded pension system actuarially fair *on average*. But a non-funded pay-go pension system can provide an actuarially fair implicit return on the margin, which increases economic efficiency. The benefits of this fall entirely on current pensioners as a windfall gain unless compensating transfers are implemented. Such a system can be thought of as a pay-go system that mimics a fully funded pension system in combination with lump transfers to current pensioners from current and future workers.

---

* We are grateful to Avinash Dixit and Hans-Werner Sinn for very useful comments and to Judy Petersen, who copyedited a draft of this paper.

‡ Assar Lindbeck is also a fellow at IUI, Stockholm. Address of authors: Institute for International Economic Studies, Stockholm University, S-106 91 Stockholm, Sweden. Telephone: +46-8-162070, Fax: +46-8-161443, Email: John.Hassler@iies.su.se and Assar@iies.su.se; homepages: http:/www.iies.su.se/

1. Introduction

Obviously, a compulsory pension system violates the (short-term) intertemporal preferences of individuals if they are constrained by liquidity. Indeed, this is often a basic purpose of such systems to prevent free riding and to mitigate asserted myopic behavior of some individuals. It is also well known that such a system distorts labor supply decisions if the system is financed by taxes on labor income, for example, in the form of a proportional payroll tax. Of course this is also the case if the individual is not liquidity constrained, provided the system is not actuarially fair. Here, an *actuarially fair* system means that the expected present value of pension benefits and of fees (contributions) are equal. Usually, compulsory pay-as-you-go pension systems in the real world are not actuarially fair, even disregarding intra-generational transfers (see, for example, Feldstein 1996).

This is the background for various suggestions to make existing pay-go systems more actuarial – perhaps even fully actuarially fair. This paper discusses the possibility and desirability of doing just that. A study of this type is worthwhile because several countries plan to move in this direction, and many observers have argued that it is possible and desirable to mimic an actuarially fair funded system by a properly reformed pay-go system, without making the system funded.

Making a pay-go pension system more actuarial is likely to be *generically* inconsistent with balancing the budget in each period. So the question of the stability of the pension system is naturally raised. More specifically, we ask two questions with respect to this:

1. Under what conditions will an actuarially fair pay-go systems be stable, in the sense non-explosive?
2. Is it possible for a pay-go pension system to be actuarially fair without having a fund of the same size as in a fully funded system?

It is well known that the implicit return in a balanced-budget, pay-go pension system is determined by the growth rate of the tax base. If this growth rate is not much lower than the interest rate, a balanced-budget, non-funded pay-go pension system then provides a return that is close to that of an actuarially fair system. In this case, one might think that a small fund would be sufficient to generate the extra revenues necessary to finance an actuarial pay-go system. This apparently intuitive conjecture turns out to be false—the system must be *fully* funded to provide an actuarially fair return.

So generically, a pay-go pension system is inconsistent with actuarial fairness even if it is partially funded. But from a normative viewpoint, it is also important to ask if it is possible and desirable to mimic a fully funded pension system only *on the margin*, that is, by providing an actuarially fair return on marginal contributions. It is plausible to argue that this would be beneficial because most labor-market distortions depend on the degree of *marginal* actuarial fairness, that is, on the relation between marginal contributions and marginal benefits. So we also investigate the case for introducing marginal actuarial fairness in a pay-go system.

A pay-go pension system, in contrast to a fully funded system, creates intergenerational transfers. In an analysis of the possibility to mimic a fully funded system, it is therefore necessary to calculate how the potential benefits of removing the labor market distortion in the former are distributed between generations. This turns out to be a crucial issue. For example, consider the experiment of increasing the marginal implicit return above the average return while keeping the proportional tax rate that finances the pensions constant. We show that this would reduce the labor market distortions. But the gains from improved efficiency fall entirely as a windfall gain on the generation that is already retired at the time of the policy change. All other generations will actually be worse off. So such a policy change cannot be described as a move in the direction of mimicking a fully funded system.

To highlight our main messages, we make several simplifying assumptions. A general assumption is that factor prices and population growth are exogenous stochastic variables. We make this assumption partly for convenience and analytical tractability, but we believe that our qualitative results do not critically depend on this. It is well known that the introduction of a pay-as-you go pension system may have considerable negative effects on net savings of a country by diverting part of the savings of the working generation into wind-fall gains for the elderly. This may have substantial effects on capital formation and factor prices. In principle, these effects should be considered when determining the optimal size of a pay-go pension system.[1] But this issue is outside the scope of this paper, where we take the size of the pension

[1] Siandra, (1994), analyzes the optimal size of a pay-go pension system when the negative effect on capital formation is considered. Smith (1982) and and Endes & Lapan (1982) analyze optimal intergenerational risk sharing with endogeneous factor prices.

system as given. For this reason, we believe that that an analysis, where effects on factor prices are disregarded, is worthwhile as a step toward an understanding of the difference between a fully funded and a pay-go pension system. In particular, this is true for small economies with highly open, capital markets.

This paper is organized as follow: Section 2 presents the basic model. In section 3, we study the relation between stability and actuarialness of pension systems in both the deterministic and the stochastic case. In section 4, we analyze the issue of marginal actuarial fairness. Section 5 summarizes the main results.

2. The model

We consider a two-period, overlapping generations' model, where individuals work in the first but not in the second period of their lives. All individuals in a generation are identical. The size of the generation born in t is N_t, which is taken to be a large number (in a sense to be made more precise later). To denote a single individual, we use the index i. But because all individuals of a given generation are identical, we can suppress this index most of the time.

We denote the ratio between generations born in t+1 and t, that is, N_{t+1}/N_t, by $1+n_{t+1}$, so n_{t+1} is the rate of population growth.[2] Let consumption in the two periods of life of an individual born in time period t be $c_{1,t}$ and $c_{2,t+1}$, labor supply l_t and the subjective rate of time preference θ. An individual born in period t is assumed to maximize a time-additive, utility function of the following form

$$U_t \equiv u^1\left(c_{1,t}, -l_t\right) + E_t\left(1+\theta\right)^{-1} u^2\left(c_{2,t+1}\right) \tag{1}$$

subject to the budget constraint

$$\frac{c_{2,t+1}}{1+r_{t+1}} + c_{1,t} = w_t l_t (1-\tau) + \frac{B_{t+1}}{1+r_{t+1}}, \tag{2}$$

where τ is a pension fee rate (tax rate) that finances pension benefits denoted B_t. Individuals have access to a capital market where they may invest their savings and receive a return r_{t+1}.

At each time period t, three exogenous stochastic variables are realized: w_t the wage of the young generation in t, r_t the rate of return on the investments in the preceding period of the

---

[2] Aggregate longevity risk could be incorporated in the analysis by interpreting $1+n_t$ as the ratio between the number of working individuals in period t and the number of living retirees.

currently retired, and n_t the rate of growth of the number of working (young) individuals. We denote the growth rate of the aggregate wage income by g_t so that $1+g_{t+1} \equiv N_{t+1}w_{t+1}l_{t+1}/N_t w_t l_t$. We also define the (average) implicit return in the pension system as

$$r^p \equiv \frac{\sum_i B_{t+1}^i}{\sum_i w_t l_t^i \tau} - 1 \tag{3}$$

and the *marginal* implicit return for an individual as

$$r_i^{pm} = \frac{\partial B_{t+1}^i}{\partial \tau w_t l_t^i} - 1, \tag{4}$$

that is, the return on the fees paid on a marginal unit of working time. Note that the marginal implicit return is the marginal return for a single individual, holding the behavior of other individuals fixed.

3. Stability and actuarial fairness

This section explores necessary conditions for the stability of pay-go pension systems with fixed average implicit returns. Here, the notion of stability is that the accumulated stock of debt (claims) in the pension system must be non-explosive. To make this a bit more precise, we postulate two necessary conditions for stability. Letting D_t denote the accumulated debt in the pension system at time *t*, expressed as a share of aggregate wage income, the two conditions are:

Condition 1. A pension system is not stable unless

$$\begin{gathered} \lim_{s\to\infty} E_t D_{t+s} < \infty, \text{ and} \\ \lim_{s\to\infty} E_t (D_{t+1}^2) < \infty. \end{gathered} \tag{5}$$

The first condition states that if the debt share of the system approaches infinity, the system is not stable. The second condition requires that the variance is bounded. Without that, we could have a situation where the debt approaches plus or minus infinity with equal probability, satisfying the first condition. Nevertheless, one could hardly call such a system stable.

Now let us consider the implications of the two stability conditions. We start with the deterministic case.

3.1 The deterministic case

Using (3), we can express the pension benefit in period t for each pensioner in a pay-go pension system as $\left(1+r_t^p\right)\tau w_{t-1}$. So the per-period deficit in the pension system can be written

$$N_{t-1}(1+r_t^p)\tau w_{t-1} - N_t \tau w_t \,. \tag{6}$$

To analyze the behavior of D_t, we express the deficit as a share of the wage bill of the currently young by dividing by $N_t w_t$.

$$\frac{N_{t-1}(1+r_t^p)\tau w_{t-1}}{N_t w_t} - \tau = \tau\left(\frac{1+r_t^p}{1+g_t} - 1\right) = \tau\left(\frac{r_t^p - g_t}{1+g_t}\right). \tag{7}$$

Of course the RHS of (7) that denotes the deficit share is non-zero if the rate of return in the pension system differs from the growth rate of the economy. There is a deficit if r_t^p is larger than g_t and a surplus (negative deficit) otherwise.

Let us now consider what happens if the return in the system differs from the growth rate. The time path of the debt share D_t, that is, the accumulated deficit share, is obviously

$$D_t = D_{t-1}\frac{(1+r_t)}{(1+g_t)} + \frac{\tau\left(r_t^p - g_t\right)}{(1+g_t)}, \tag{8}$$

where r_t is the market interest rate. In the non-stochastic case with constant income growth and returns:

$$D_t = D_{t-1}\frac{1+r}{1+g} + \tau\frac{r^p - g}{1+g}. \tag{9}$$

Equation (9) has one steady state D that satisfies

$$D = -\tau\frac{r^p - g}{r - g}. \tag{10}$$

If the debt share D_t satisfies (10), it remains constant at that level. Thus, if the pay-go system is actuarially fair in the sense that $r^p = r$, a debt share of $-\tau$ is a steady state. A debt share of $-\tau$ implies that the pension system has accumulated surpluses equal to $\tau N_t w_t$. These are of the same size as the fund in a fully funded system, which by construction (in our two-period model) is each period's pension fees, that is, $\tau N_t w_t$. This leads to the following proposition.

Proposition 1. *In an economy with constant income growth and a constant capital market return, an actuarially fair pay-go system is consistent with a constant debt share only if it is fully funded.*

Proof: Follows from the analysis above.

To provide intuition for this result, consider the case when *g*, and thus the implicit return in the balanced budget non-funded system is zero. Then obviously the fund required to provide an actuarial return is equal to the fund in a fully funded pension system. Now consider an increase in *g*. This reduces the difference between the actuarial return *r* and the implicit return in the balanced budget non-funded system. But at the same time, the higher growth rate means that more money must be invested in the fund each period simply to keep its size, expressed as a share of GDP constant. These two effects exactly cancel, and thus the required fund is equal to the fund in the fully funded system, regardless of the difference between *r* and *g*.

An actuarial pay-go system with assets equal in size to the fund in the fully funded pension system is in steady state identical to a fully funded pension system even if the pension payments are not formally tied to the return on the fund in the system. So we can say that any pension system that is actuarially fair and has a constant fund (or debt), expressed as a share of GDP in steady state, is a fully funded system.

Let us analyze the behavior of *D* outside the steady state. The evolution of the debt share given by (9) is stable if the absolute value of $(1+r)/(1+g)$ is smaller than unity. This is the case if the economy is dynamically inefficient, that is, if $r<g$. In that case and that case only, will an actuarially fair pay-go pension system with $r^p = r$ be sustainable. By contrast, in a dynamically efficient economy, where $r>g$, the debt becomes positive in the period immediately after the system has started, and the debt share will explode.[3] The conclusion is:

Proposition 2. *In a dynamically inefficient economy ($r<g$) with constant growth and constant capital market return, an actuarially fair pay-go system converges automatically to a fully funded system.*

Corollary. *If the economy is dynamically efficient ($r>g$), the only stable actuarially fair pay-go pension system is a system that is started with a fund equal to the fund in a fully funded system. The pension system is then a fully funded system.*

Proof: Follows from the analysis above.

---

[3] The fees paid by the young when a pay-go system is started are paid to those who are currently retired. There is then no deficit in the first period. In the actuarially fair pension system and under the assumption that the economy is efficient, we have $r^p = r>$g. It the follows from (9) that the debt is positive and exploding.

The proposition and the corollary could alternatively be formulated: In a dynamically efficient economy, no actuarially fair pay-go pension system can be introduced and in the dynamically inefficient case, an actuarially fair pay-go pension system will converge to a fully funded one. Along the transition phase, there will be differences between the two systems. In particular, the generation that is in retirement when a pay-go system is introduced will receive a windfall gain that they would not receive if a fully funded system had been introduced from the very beginning. This is the *free lunch* first analyzed by Samuelson (1958) and Aaron (1966).

3.2 The stochastic case

When growth and interest rates are stochastic it is more difficult to derive general conditions for stability. So we concentrate on a case where growth and interest rates are i.i.d. over time. Let r_t^f denote the safe interest rate at which the pension system lends and borrows.

Assumption 1. *The ratio* $(1+r_t^f)/(1+g_t)$ *is exogenous, independent of its previous realizations and identically distributed over time.* [4]

The expected deficit share in each period is given by a direct analogue of (7), namely

$$\tau E\left(\frac{1+r_t^p}{1+g_t}-1\right)=\tau E\left(\frac{r_t^p-g_t}{1+g_t}\right). \tag{11}$$

Now consider the expected value of D_{t+1} as a function of D_t. This function is simply the stochastic analogue of (8) and it has only one fixed point D^* at which $E\left[D_t \middle| D_{t-1}\right]=D_{t-1}$. This point is given by

$$D^*=-\tau\frac{E\left[(r^p-g)/(1+g)\right]}{E\left[(r^f-g)/(1+g)\right]}. \tag{12}$$

If the pay-go system provides safe benefits, actuarial fairness requires that $r_t^p = r_t^f$. Then the RHS of (12) is just $-\tau$. Thus, as in the non-stochastic case, the steady-state debt share is equal to the fund in the fully funded system, that is, $\tau N_t w_t$.

Now let us consider the stability of the pay-go system in this setting. We assume that $E[(1+r_t^f)/(1+g_t)]$ and $E[(r_t^p - g_t)/(1+g_t)]$ are constant over time and denoted $\bar{\mu}$ and $\bar{d}$. By iterating on (8) and using the i.i.d. assumption, it is straightforward to show that

---

[4] The assumption of independence over time can be relaxed quite easily.

$$E\left[D_{t+s}\middle|D_t\right] = D_t\bar{\mu}^s + \tau\bar{d}\sum_{i=0}^{s-1}\bar{\mu}^i \ . \tag{13}$$

Equation (13) defines a converging sequence if $\bar{\mu}$ is smaller than unity. In the case when r^p is set equal to r^f, implying that $\bar{d} = \bar{\mu} - 1$, the limit of (13) when s goes to infinity simplifies to $-\tau$. This leads to

Proposition 3. *In an economy where the ratio* $(1+r_t^f)/(1+g_t)$ *follows an i.i.d. stochastic process, an actuarially fair pay-go system with safe benefit is not stable unless the expected value of* $(1+ r_t^f)/(1+g_t)$ *is below unity.*

Corollary. *If the actuarial pay-go system is stable, its expected accumulated fund converges to that of the fully funded system.*

Proof: Follows from the analysis above.

Of course proposition 3 is a stochastic analogue to proposition 2. In the stochastic case, we must also apply condition 2, that is, that the debt share has a non-exploding variance. It turns out that there is a simple sufficient condition for this type of stationarity if we add the assumption that $(1+r_t)/(1+g_t)$ can take on only a finite number of values.

Proposition 4. *Under the assumption that* $(1+r_t^f)/(1+g_t)$ *is i.i.d. and can take only a finite number of values, the debt share of an actuarially fair pay-go system that provides an implicit return equal to the market return has a non-exploding variance if*[5]

$$E\left[\left(\frac{1+r}{1+g}\right)^2\right] < 1. \tag{14}$$

Proof: See Appendix.

Note that this condition for stability is not identical to the condition that the pay-go pension system with a return equal to the market return runs an average surplus, i.e., that $E[((1+r_t^f)/(1+g_t))] <1$. Dynamic inefficiency in the sense of having an expected surplus in the pay-go system does not necessarily imply stability. But stability implies inefficiency because $E[((1+r_t^f)/(1+g_t))^2] = E[((1+r_t^f)/(1+g_t))]^2 + \mathrm{Var}[(1+ r_t^f)/(1+g_t)]$ so $E[((1+r_t^f)/(1+g_t))^2]<1 \Rightarrow E[((1+r_t^f)/(1+g_t))]<1$. The intuitive explanation behind the more stringent conditions for stability in the stochastic case than in the non-stochastic is the following: The stochastic

[5] Results in Warne (1996) suggest that the condition (14) is also necessary for stationarity. It is also straightforward to replace the assumption of i.i.d. by the assumption that the distribution of $(1+r_t)/(1+g_t)$ depends on a finite number of previous states of the world.

analogue to the stability condition in the non-stochastic case is that the debt does not explode *on average*, i.e., that the expected debt converges. But clearly, more is required in the stochastic case, namely that the debt does not explode in any states of the world that have positive probability.

4. Marginal actuarial fairness

In the previous sections, we showed that a pay-go pension system that operates in a dynamically efficient economy cannot mimic a fully funded pension system by setting the average implicit return equal to the market return. But there is still the possibility to set the *marginal* implicit return, as defined in (4), equal to the market return, while keeping a budget balance by having the *average* return equal to the growth rate of the economy. The purpose of this section is to analyze the benefit of such arrangements and, in particular, to analyze to whom the potential benefits accrue.

We consider only balanced budget pay-go pension systems with a linear (affine) relation between fees and benefits. More specifically, we consider systems where B_{t+1} in the individual budget constraint (2) is given by

$$B_{t+1}^i = \tau w_t l_t^i \alpha(1+r^p) + T_{t+1} \tag{15}$$

where r^p is the average implicit return, as defined in (3), and T_{t+1} is a lump sum positive or negative transfer to pensioners in t+1, constant over all i. The parameter α determines the marginal implicit return.

Budget balance in the pay-go pension system determines the average implicit return in the pension system:

$$\begin{aligned} &\sum\nolimits_i B_{t+1}^i = \sum\nolimits_i w_{t+1} l_{t+1}^i \tau = \sum\nolimits_i w_t l_t^i \tau (1+g) \\ &\Rightarrow r^p \equiv \frac{\sum\nolimits_i B_{t+1}^i}{\sum\nolimits_i w_t l_t^i \tau} - 1 = g \end{aligned} \tag{16}$$

Furthermore, budget balance implies that

$$\begin{aligned} &\sum_{i=1}^{N_t} \left(\tau w_t l_t^i \alpha(1+r^p) + T_{t+1}\right) = \sum_{i=1}^{N_{t+1}} N_{t+1} \tau w_{t+1} l_{t+1}^i \\ &\Rightarrow T_{t+1} = \frac{1}{N_t} \sum_{i=1}^{N_t} (1+r^p)(1-\alpha) \tau w_t l_t^i \end{aligned} \tag{17}$$

Now, from the definition of *marginal* implicit return, equation (4), and (15) we derive

$$r^{pm} = \alpha(1+r^{p}) + \frac{1}{N_t}(1+r^{p})(1-\alpha) - 1\,, \tag{18}$$

where the second term is $\partial T_{t+1} / \partial l_t^i$.

Assumption 2. *N_t is sufficiently large for $\partial T_{t+1} / \partial l_t^i$ to be neglected by the individual.*

Of course the assumption that N is large is highly realistic. The effect of an individual on the aggregate budget constraint of the pension system is negligible.

The first-order conditions for the individual are then:

$$\begin{aligned} u_{c_{1,t}} &= (1+\theta)^{-1} E(1+r_{t+1}) u_{c_2,t+1} \\ u_{-l_t} &= \left(w_t(1-\tau) u_{c_{1,t}} + w_t \tau\alpha E_t u_{c_2,t+1} \frac{1+g_{t+1}}{1+\theta} \right) \end{aligned} \tag{19}$$

where subscripts on u denote partials with respect to the relevant period utility function (u^1 or u^2).

Now, what is the optimum α? We want to evaluate a pay-go pension system of a given size from the viewpoint of the current working generation and all future generations. More specifically, we fix the size of the benefits to current pensioners. In other words, we derive an optimal value of α under the restriction that $\tau w_1 l_1$ is fixed. Since variations in α affect labor supply, compensating changes in τ are required.

Because all individuals in each generation are identical, we consider the welfare of a representative individual from each generation. The objective function that we maximize over α then becomes

$$W(\alpha,\tau) = E \sum_{t=1}^{\infty} \delta^t \left(u^1\left(c_{1,t}, -l_t\right) + \left(1+\theta\right)^{-1} u^2\left(c_{2,t+1}\right)\right), \tag{20}$$

where δ^t denote the welfare weight given to the representative individual in a generation born in period t. These weights may reflect both different sizes of generations and social time preferences. Because the benefits to pensioners in period 1 are fixed by construction, we do not have to include their welfare in the objective function.

Now let us consider the welfare effect of varying α. Because τ must be adjusted to variations in α, there is both a direct effect $(\partial W(\alpha,\tau)/\partial\alpha)$ and an indirect effect $(\partial W(\alpha,\tau)/\partial\tau)(\partial\tau/\partial\alpha)$ on welfare. Let us start with the direct effect, that is, when we keep τ fixed.

$$\frac{\partial W(\alpha,\tau)}{\partial\alpha}=E\sum_{t=1}^{\infty}\delta^{t}\left(u_{c_{1,t}}\frac{\partial c_{1,t}}{\partial\alpha}-u_{l_t}\frac{\partial l_t}{\partial\alpha}+\frac{u_{c_{2,t+1}}}{1+\theta}\frac{\partial c_{2,t+1}}{\partial\alpha}\right) \tag{21}$$

Budget balance in the pension system implies that $B_{t+1} = \tau(1+n_{t+1})w_{t+1}l_{t+1} = \tau(1+g_{t+1})w_t l_t$. This means that the budget restriction for a representative individual born in period t is

$$\begin{aligned}c_{2,t+1}&=\left(w_t(1-\tau)l_t-c_{1,t}\right)(1+r_{t+1})+B_{t+1}\\&=\left(w_t(1-\tau)l_t-c_{1,t}\right)(1+r_{t+1})+(1+g_{t+1})w_t\tau l_t\end{aligned} \tag{22}$$

As we see, α, does not enter (22) since a has no direct effect on the amount of resources available for consumption. On the other hand, α does enter the budget constraint under which the individual is optimizing, that is, equations (2) and (15). When α is different from unity, variations in labor supply affect the lump sum transfer. This effect is considered in the welfare analysis but not in the optimization of the (atomistic) individual.

Now we make some additional assumption to simplify the analysis.

Assumption 3. *The economy is on a balanced growth path with constant labor supply and the effect of variations in α on labor supply is the same across generations.*

By the previous assumption, we have

$$dc_{2,t+1}=\left(w_t(1-\tau)dl-dc_{1,t}\right)(1+r_{t+1})+(1+g_{t+1})w_t\tau dl\,. \tag{23}$$

Using this in (21) yields

$$\begin{aligned}\frac{\partial W(\alpha,\tau)}{\partial\alpha}=E\sum_{t=1}^{\infty}\delta^{t}\Bigg(&u_{c_{1,t}}\frac{\partial c_{1,t}}{\partial\alpha}-u_{l_t}\frac{\partial l}{\partial\alpha}\\&+\left(w_t(1-\tau)\frac{\partial l}{\partial\alpha}-\frac{\partial c_{1,t}}{\partial\alpha}\right)\frac{u_{c_{2,t+1}}(1+r_{t+1})}{1+\theta}+w_t\tau\frac{\partial l}{\partial\alpha}\frac{u_{c_{2,t+1}}(1+g_{t+1})}{1+\theta}\Bigg)\end{aligned} \tag{24}$$

Now we can use the first-order conditions from (19). Doing this and using the law of iterated expectations gives

$$
\begin{aligned}
&= E\sum_{t=1}^{\infty}\delta^{t}\left(u_{c_{1,t}}\frac{\partial c_{1,t}}{\partial\alpha}-\left(w_{t}(1-\tau)u_{c_{1,t}}+w_{t}\tau\alpha\frac{u_{c_{2,t+1}}(1+g_{t+1})}{1+\theta}\right)\frac{\partial l}{\partial\alpha}\right.\\
&\left.+\left(w_{t}(1-\tau)\frac{\partial l}{\partial\alpha}-\frac{\partial c_{1,t}}{\partial\alpha}\right)\frac{u_{c_{2,t+1}}(1+r_{t+1})}{1+\theta}+w_{t}\tau\frac{\partial l}{\partial\alpha}\frac{u_{c_{2,t+1}}(1+g_{t+1})}{1+\theta}\right)\\
&= E\sum_{t=1}^{\infty}\delta^{t}\left(u_{c_{1,t}}\frac{\partial c_{1,t}}{\partial\alpha}-\left(w_{t}(1-\tau)u_{c_{1,t}}+w_{t}\tau\alpha\frac{u_{c_{2,t+1}}(1+g_{t+1})}{1+\theta}\right)\frac{\partial l}{\partial\alpha}\right.\\
&\left.+\left(w_{t}(1-\tau)\frac{\partial l}{\partial\alpha}-\frac{\partial c_{1,t}}{\partial\alpha}\right)u_{c_{1,t}}+w_{t}\tau\frac{\partial l}{\partial\alpha}\frac{u_{c_{2,t+1}}(1+g_{t+1})}{1+\theta}\right)\\
&= E\sum_{t=1}^{\infty}\delta^{t}\left((1-\alpha)w_{t}\tau\frac{\partial l}{\partial\alpha}\frac{u_{c_{2,t+1}}(1+g_{t+1})}{1+\theta}\right)
\end{aligned}
\tag{25}
$$

Setting (25) to zero yields the first-order condition for an optimal α, *when* τ*, rather than* τwl*, is held constant.* It is now straightforward to show that labor supply increases in α, i.e., that the derivative $\partial l/\partial\alpha>0$. To do this, assume the opposite, that labor supply decreases in α. Less work means less consumption opportunities, so marginal utility of consumption must increase. But then the RHS of the second equation in (19) must increase in α and thus also in the marginal disutility of work, which contradicts the initial assumption of decreasing labor supply. This means that also the second-order condition is satisfied.

Proposition 5. *If the tax rate* τ *is held constant, setting* $\alpha = 1$ *maximizes the welfare of the current working and all future generations*[6].

If τ is held constant and each generation of workers work an extra hour, every individual in the current young generation get extra income from $w(1-\tau)$ while young and $w\tau(1+g)$ while retired, regardless of α. The same applies to all future generations. It is the value of this extra income that should be set equal to the marginal utility of leisure in a welfare optimum, where the windfall gain in welfare of the initially retired generation is disregarded. Setting α different from unity distorts the labor-leisure choice by creating an externality, because variations in labor supply affect the size of the lump-sum component of pensions. This would reduce welfare. In a sense, we can consider $\alpha = 1$ as the *constrained first best* when the welfare of the current retired generation is disregarded.

So far, we have neglected the utility of the initially retired generation. Setting α to unity results in an allocation that is not Pareto efficient when the initially retired generation is

---

[6] We derived this result for the case of borrowing constrained individuals in Hassler & Lindbeck (1997). As we have shown, the result is more general.

included in the analysis. Increasing α from unity has only second-order negative effects on the current and future young generations. Current pensioners, by contrast, enjoy a positive first-order effect from a marginal increase in α, because labor supply and thus pension benefits increase in α. So increasing α from unity tends to increase economic efficiency, which provides an opportunity for a Pareto improvement. However, without a compensating increase in transfers to the current working and all future generations, the benefit due to the increased efficiency falls fully as a windfall gain just to the currently retired generation. The current working and all future generations would actually be worse off.

Now let us consider the case when τ is varied so as to keep the pension benefits and thus the welfare of the current retired generation constant. In this case, we must consider the indirect effect on taxes due to a change in α. Higher α increases labor supply, so the tax rate can be reduced as α is increased, and this has a positive effect on the welfare of both the current young and all future generations. Holding $wl\tau$ fixed implies that $d\tau = -(\tau/l)dl$.

It follows from the standard envelope theorem that the marginal effect on utility of variations in the tax rate for each generation is given by the direct effect on consumption, that is,

$$\begin{aligned}\frac{dU_t}{d\tau} &= -u^1_{c_1,t} w_t l + w_t l E_t \frac{1+g_{t+1}}{1+\theta} u^2_{c_2,t+1} \\ &= w_t l E_t \left(u^2_{c_2,t+1} \frac{g_{t+1} - r_{t+1}}{1+\theta} \right)\end{aligned} \tag{26}$$

Using this, we can write the first-order condition for optimal α as

$$\begin{aligned}0 &= \frac{\partial W(\alpha,\tau)}{\partial\alpha} + \frac{\partial W(\alpha,\tau)}{\partial\tau}\frac{\partial\tau}{\partial l}\frac{\partial l}{\partial\alpha} \\ &= E\sum_{t=1}^{\infty}\delta^t w_t \tau\left((1-\alpha)\frac{\partial l}{\partial\alpha}\frac{u_{c_{2,t+1}}(1+g_{t+1})}{1+\theta}\right) + E\sum_{t=1}^{\infty}\delta^t w_t \tau\left(\frac{\partial l}{\partial\alpha}u^2_{c_2,t+1}\frac{r_{t+1}-g_{t+1}}{1+\theta}\right) \\ &= E\sum_{t=1}^{\infty}\delta^t w_t \tau\frac{\partial l}{\partial\alpha}\left(\frac{u_{c_{2,t+1}}}{1+\theta}((1+r_{t+1}) - \alpha(1+g_{t+1}))\right)\end{aligned} \tag{27}$$

Setting this to zero yields

$$E\sum_{t=1}^{\infty}\delta^t \left(u_{c_{2,t+1}}(1+r_{t+1})\right) = E\sum_{t=1}^{\infty}\delta^t \left(u_{c_{2,t+1}}\alpha(1+g_{t+1})\right) \tag{28}$$

This condition is straightforward to interpret. To do this, let us focus on an individual belonging to a particular generation born in $t+s$. For him to be indifferent between investing on the capital market and paying an extra dollar to the pension system, it is required that the marginal implicit return in the pension system satisfies[7]

$$E_{t+s}u_{c_{2,t+s+1}}(1+r_{t+s+1}) = E_{t+s}u_{c_{2,t+s+1}}(1+r^{pm}_{t+s+1}) \\ = \left(u_{c_{2,t+s+1}}\alpha(1+g_{t+s+1})\right). \tag{29}$$

If (29) is satisfied, the pension system is marginally actuarially fair for the individual. In other words, (28) implies that the pension system should be marginally actuarially fair *on average*.

Proposition 6. *When $wl\tau$ and thus the welfare of the currently retired is held constant, α should be set so as to yield a marginally actuarially fair implicit return on average.*

Corollary. *In a non-stochastic world with constant growth and interest rates,* α = (1+r)/(1+g) *maximizes the welfare of the currently young and all future generations. The marginal implicit return in the pension system is then equal to* (1+r)*, the return on the capital market. This yields a Pareto efficient allocation corresponding to the case when a lump-sum tax equal to* $wl\tau$(1-(1+g)/(1+r) *is levied on the current and all future young generations.*

The corollary is easily established. First note that when α = (1+r)/(1+g), the individual optimization problem is identical to the one resulting if a lump-sum tax of $w_t l_t \tau$(1-(1+g)/(1+r)) and no pension system are applied. So a government could achieve the same allocation as with the pay-go pension system by imposing a lump-sum tax of $w_t l_t \tau$(1-(1+g)/(1+r)) on each young generation. The remaining part of the pensions, $w_t l_t \tau$(1+g)/(1+r), is borrowed at the capital market.

So we conclude that increasing α from unity to (1+r)/(1+g), so that the pension system is actuarially fair on the margin, maximizes economic efficiency. In combination with a compensating change in the tax rate τ, which keeps the welfare of currently retired generation unchanged, a Pareto improvement is achieved as the current working generation and all future generations are strictly better off. In a sense, this can be seen as a pay-go pension system that mimics a fully funded system *in combination* with lump-sum transfers to the first generation.

[7] This is an application of standard portfolio choice theory under uncertainty.

5. Summary

We analyzed various methods of making a pay-go pension system mimic a fully funded system. But several intuitively plausible methods have not worked out. Simply disregarding budget balance and paying an actuarially fair average return would make the pension debt explode if the economy is dynamically efficient. Any funding less than full provides no remedy against this. By contrast, in a dynamically inefficient economy, an actuarially fair pay-go pension system generates surpluses that *automatically* accumulate into a fund of equal size as in a fully funded system.

We have also seen that inefficiencies created by labor market distortions in a non-actuarial pay-go pension system is removed if the *marginal* implicit return is raised above the average return. But it is then important how the benefits of this are distributed. Unless the tax rate is lowered, or other compensating transfers are used, the benefits of the increased efficiency goes entirely to current pensioners as a windfall gain. The current working generation and all future generations actually lose. Such a change cannot be called a move in the direction of mimicking the fully funded pension system, because it would strengthen the intergenerational transfers created by the introduction of a pay-go pension system.

Lastly, we considered the case when the tax rate in the pension system is adjusted so as to hold the pensions of the currently retired generation constant when the marginal implicit return is varied. In this case, the gains from increased economic efficiency associated with a higher marginal implicit return goes to the current working generation and all future generations. The optimal marginal implicit return should then be such that the pension system is actuarial on the margin. Such a change is Pareto efficient. The reform could be said to mimic a fully funded system combined with a lump-sum transfer to the currently retired generation financed with lump-sum taxes on all future generations.

6. References

Aaron, Henry, (1966), “The Social Insurance Paradox”, *Canadian Journal of Economics,* XXXII (3), 371-377.

Feldstein, Martin, (1996), “The Missing Piece in Policy Analysis: Social Security Reform”, The Richard T. Ely Lecture, *American Economic Review*, 86(2), 1-14.

Hassler, John and Assar Lindbeck, (1997) “Optimal Actuarial Fairness in Pension Systems - a Note", *Economics Letters*, 55 (2), 251-55.

Karlsen, Hans, (1990), *A Class of Non-Linear Time Series Models*, Ph.D. Thesis, Department of Mathematics, University of Bergen, Norway.

Magnus, Jan R. and Heinz Neudecker, (1988), *Matrix Differential Calculus with Applications in Statistics and Econometrics,* John Whiley, Chichester.

Samuelson, Paul, (1958), “An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money”, *Journal of Political Economy* 66 (6), 467-82.

Siandra, Eduardo, (1994), “Optimal Mix of Pension Systems”, Economic Theory Discussion Paper No. 1016, University of Cambridge.

Smith, Alasdair, (1982), “Intergenerational Transfers as Social Insurance”, *Journal of Public Economics* 19, 97-106.

Warne, Anders, (1996), “Autocovariance Functions and Maximum Likelihood in a VAR Model under Markov Switching”, mimeo, IIES, Stockholm University.

Appendix

Proof of Proposition 4

Assume that for all t

$$\frac{1+r_t}{1+g_t} \in [\mu_1, \ldots, \mu_n], \text{ with}$$
$$prob\left(\frac{1+r_t}{1+g_t} = \mu_i\right) = p[i], i = 1, \ldots, n \tag{A.1}$$

where $p[i]$ is an element of a vector of probabilities that sum to unity.

Now we use a result in Karlsen (1990).[8] A sufficient condition for stability of a first-order autoregressive model with state dependent AR coefficients denoted μ_i and with a state transition matrix denoted Π is that the largest eigenvalue of

$$\begin{bmatrix} \mu_1^2 \Pi_{1,1} & \cdots & \mu_1^2 \Pi_{n,1} \\ \vdots & & \vdots \\ \mu_n^2 \Pi_{1,n} & \cdots & \mu_n^2 \Pi_{n,n} \end{bmatrix} \tag{A.2}$$

[8] We are grateful to Anders Warne for showing us Karlsen’s proof.

is smaller than unity. In (A.2) $\Pi_{i,j}$ is the probability of moving from state i to j. In the case of proposition 4, the Π is particularly simple since the probabilities of different states are independent of previous states. This implies that

$$\begin{bmatrix} \mu_1^2 \Pi_{1,1} & \cdots & \mu_1^2 \Pi_{n,1} \\ \vdots & & \vdots \\ \mu_n^2 \Pi_{1,n} & \cdots & \mu_n^2 \Pi_{n,n} \end{bmatrix} = \begin{bmatrix} \mu_1^2 p_1 & \cdots & \mu_1^2 p_1 \\ \vdots & & \vdots \\ \mu_n^2 p_n & \cdots & \mu_n^2 p_n \end{bmatrix}. \tag{A.3}$$

Using a result in Magnus & Neudecker (1988), it can be shown that the only non-zero eigenvalue of the matrix in (A.3) is given by [9]

$$\left[\mu_1^2, \cdots, \mu_n^2\right] \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix}. \tag{A.4}$$

which is the expected value of the square of $(1+r_t)/(1+g_t)$ as stated in the proposition.

[9] See Warne (1996).