DESIGNING BENEFIT RULES FOR FLEXIBLE RETIREMENT: WELFARE VS. REDISTRIBUTION*

P. ESŐ – A. SIMONOVITS – J. TÓTH

With flexible (variable) retirement every individual determines his optimal retirement age, depending on a common benefit–retirement age schedule and his life expectancy. The government maximises the average expected lifetime utility *minus* a scalar multiple of the variance of the lifetime pension balances to achieve harmony between the maximisation of welfare and the minimisation of redistribution. Since the government cannot identify types by life expectancy, it must take the individual incentive compatibility constraints into account. Second-best schedules strongly reduce the variances of benefits and of retirement ages of the so-called actuarially fair system, thus achieving higher social welfare and lower redistribution.

Corresponding author: A. Simonovits, Institute of Economics of the Hungarian Academy of Sciences, Budaörsi út 40, H-1112 Budapest, Hungary. E-mail: simonov@econ.core.hu

1. INTRODUCTION

Every public pension system has a *normal* retirement age – yielding the full bene-fit –, but workers are also often allowed to retire before or after reaching this age; retirement is *flexible (variable)*. Different public systems, however, provide different incentives to retire sooner or later than normal. According to the dominant view (e.g. Gruber – Wise 1999; Lindbeck – Persson 2003; Fenge – Pestieau 2005; Cremer et al. 2008), in most countries these incentives are dysfunctional: they hardly sanction early retirement and they hardly support late retirement, causing a low average retirement age. As the population is aging, however, low average retirement age is becoming less and less tolerable. As a reaction, in some countries (e.g. Sweden and Poland) a system of *notional defined contribution* (NDC) has been introduced which allegedly achieves *proportionality* between lifetime contributions and lifetime benefits, implying *actuarial fairness*. Here the annual benefit is equal to the ratio of the (notional) stock of lifetime contributions to the remaining average life expectancy. (For simplicity, in the present formulation complicating factors like valorisation and discounting are omitted. For a rich discussion of NDC, see Holzmann – Palmer 2006).

This approach, however, overlooks another important dimension of heterogeneity, namely life expectancy. In the present paper, we will concentrate on this neglected dimension. We assume that the individuals know their own life expectancy, while the government does not. Then we show by Theorem 2 that the NDC loses part of its appeal: it heavily redistributes among different types and does not maximise social welfare. Therefore NDC must be replaced by optimally *designed* benefit rules. Barr – Diamond (2006) also criticised the oversimplification behind NDC in a more general way. They underlined that (i) there is no guarantee that minimal redistribution is socially optimal and (ii) there is no guarantee that NDC achieves minimal redistribution.

In a paradoxical way, the importance of this heterogeneity is stressed when evaluating *progressive* public pension systems. (Progressive systems provide proportionally lower monthly benefits to those who contribute more in an average month during a limited period of their careers.) Their critics (e.g. World Bank 1994, p. 131) and their defenders (e.g. Orszag – Stiglitz 2001) equally emphasise that "there [is] little [lifetime] redistribution from the rich to the poor, despite progressive benefit formulas, ... because [they are] earnings-related ... [and]... upper-income people enter the labour force later in life and live longer after retirement". To see the independence of proportionality (between contributions and benefits) and fairness, note that in a group of countries (e.g. Germany) proportionality and weak actuarial adjustment coexist, while in another group of countries (e.g. the US) progressivity and strong actuarial adjustment coexist. From now on

we will neglect earning heterogeneity and confine our attention to proportional systems.

The empirical literature (e.g. Gruber – Wise 1999) assumes that under given pension rules, workers choose their retirement ages to maximise their social security wealth. More sophisticated theoretical analyses (from Sheshinski 1978 to Samwick 1998) describe the individuals' retirement decisions as the constrained maximisation of lifetime utility. During the first part of their life-cycles individuals work, in the second one they get retired. Pensioners receive life annuities, the values of which depend on the lifetime contribution and the retirement age. Workers optimise their retirement ages by maximising their lifetime utility functions. Of course, the government does not know the individual characteristics: the *information is asymmetric*. It is true that this type of asymmetry does not matter if workers only differ in their labour disutility.

The assumption of common life expectations is, however, inconsistent with the following observation: There is a strong positive relationship between the individual life span and the individual length of employment (or almost interchangeably, adult retirement age). Waldron (2001) established this relation empirically for large US categories. Recent Hungarian data firmly establish a strong positive correlation between the life expectancy and the age at retirement, displayed in *Tables A.1–2* of *Appendix 1*. Here we only note that among male old-age pensioners who died in 2004, those retired at 57 lived only 12.2 years in retirement, while those who retired at 65, lived another 24.3 years. For comparison, we display the conditional life expectancy: for aged 57 it was 18.0 years, while for aged 65, it was 13.1. Note also that until 1997 the full benefit retirement age was 60 for males.

The first to model such a relation was Wolfe (1983).[1] There is also direct evidence that everybody can predict his own life expectancy, the expected value of the life span, quite well (Hamermesh 1985 and Smith et al. 2001). For a recent survey see Heidler et al. (2006).[2]

This suggests the consideration of heterogeneity of life expectancy. Assuming only that the higher the life expectancy, the later somebody retires, we have an elementary result: the 'actuarially fair' rule overpays types having above average life expectancy and underpays the rest. Moreover, the balance of the gains and

---

[1]   "...actuarial reduction equates expected present values of early and normal retirement benefit streams only for those with typical life expectancy. ... Adverse selection can ... increase the average expected present value of benefits for the population as a whole" (p. 544).

[2]   "Empirical research suggests that private information on subjective survival probability exists, and that it actually predicts mortality. We furthermore find evidence of adverse selection ... but the relative importance of this effect is by no means unique as the heterogeneous literature shows other determinants of retirement choice significant, as well" (p. 27).

losses is negative. Therefore, under heterogeneous life expectancy and asymmetric information, the so-called 'actuarially fair' benefit schedule is not only unfair but may not be sustainable without uniformly reducing the benefits by adjustment (cf. Theorem 2 and also Simonovits 2003b, Chapter 12). As a consequence, we will call these rules *traditionally fair* or 'fair' rather than actuarially fair and add the adjectives *unadjusted* and *adjusted,* respectively.

Introducing utility maximising workers, one can derive optimal retirement ages for any benefit–retirement-age rule. In this paper, we will deliberately neglect the heterogeneity in labour disutility and concentrate on the heterogeneity in life expectancy. Though both types of heterogeneity are important, we cannot analyse them together, except for linear schedules (cf. Simonovits 2003a).[3] Furthermore, we also neglect the dependence of labour disutility on age, or a little more generally, we assume that the labour disutility has only two values: a low value on the age-interval relevant for working, and a high value on the remaining age-interval, close to death.

Since the seminal paper of Mirrlees (1971) on optimal income taxation has been published, mechanism design provides a tool for the analysis of such problems. In our case, the design concerns optimal flexible (variable) pension rules under the assumption that individuals have private information regarding their life expectancy and other individual characteristics. The government's goal is to design a pension system (a payroll tax or contribution and a function relating benefits to retirement ages), which maximises a social welfare function (a simple function of the individual lifetime utility) and satisfies a social budget constraint. Since individuals with different life expectancy optimise their retirement ages conditional on the benefit schedule, the government must also take into account incentive constraints.

In the pension literature, the first papers to study mechanism design problems were Diamond – Mirrlees (1978, 1986), namely disability retirement (for the latest treatment, see Golosov – Tsyvinsky (2006)). Much later Fabel (1994), Diamond (2003), Simonovits (2003a, b, 2006), Sheshinski (2006), Bommier et al. (2007) and Cremer et al. (2008) considered old-age retirement with heterogeneity in individual life expectancy.[4]

Fabel (1994, Chapter 9) derived the optimal benefit rule in a model, where there are *two types* (expectedly short- and long-lived), who may choose their ages of retirement. Diamond (2003, Chapter 7) considered a multi-type model, where

[3]   Note that allowing for an inverse relation between labor disutility and life expectancy (cf. Diamond 2003, Chapter 7), our qualitative results would still hold in this more general setting though their quantitative impact would be weaker.
[4]   Fabel (1994), Diamond (2003, Chapter 6) and the full version of Simonovits (2006) also considered the classical case, when people only differ in labour disutility.

individuals are only allowed to choose between two *pre-specified* retirement ages. As was mentioned above, Simonovits (2003a, b) studied a model, where individuals are heterogeneous not only in life expectancy but also in labour disutility, they can choose freely their retirement ages, but the analysis was restricted to the special case of *linear* benefit schedules, quite a good description of most formulas applied in practice.

The main contribution of this approach to the existing literature on pension system reforms is to extend the analysis of optimal pension benefit rules in an important new direction, by assuming that individuals have private information regarding their life expectancy. The equations, that determine an optimal (so-called second-best) benefit rule, are analytically derived. These benefit rules appear to be very much different from the 'fair' schedule (which would be acceptable if individuals differed in their disutility of labour, but not in their life expectancy).

Our present model is similar to the models mentioned above except that we allow many types of life expectancy, free choice of retirement ages, non-linear benefit–retirement age rules and an objective function depending not only on individual utility at the same time. Probably we are the first who – following Mirrlees (1971) – explicitly formulate pension mechanism design as an *isoperimetric optimal control problem*. Independently of our work, Bommier et al. (2011) analysed a similar model, where the distribution function of life expectancy is continuous rather than discrete, the disutility of labour is increasing with the age rather than being constant or two-valued, the contribution rate is type-dependent and the social welfare function is generalised utilitarian. Their results are qualitatively similar to ours but they do not present numerical examples.

Our study discovers new relations in optimal pension design and presents a technique which can be used in related fields as well. We emphasise the following results.

(i) It is expected that any reasonable benefit rule preserves some loose proportionality between life expectancy and retirement age. For example, the ratio of the difference of subsequent retirement ages to the difference of the corresponding life expectancy is definitely less than unity. We can demonstrate that such pension rules *redistribute* from individuals with lower life expectancy to individuals with higher ones though under appropriate objective functions this redistribution is probably more moderate than under the traditional rule (Theorem 1 and Conjecture 1).

(ii) Previous studies (including Eső – Simonovits) considered either *utilitarian* or *generalised utilitarian* social welfare functions, defined as the average of individual utility or their concave monotone transformations, respectively (Mas-Colell et al. 1995). It can be shown that if n is the number of types, where

$n \geq 2$, then the dimension of the set of second-best solutions is typically $n - 2$. How to choose among the continuum of optima for more than two types? Or, how to modify the model to eliminate multiplicity? One can introduce further complexities into the model, i.e. making per-period labour disutility smoothly dependent on the age (see above), but it may make the model intractable.

(iii) Numerical experimentation in Eső – Simonovits suggested that by applying stronger concavification of individual lifetime utility appearing in the social welfare function, one obtains more egalitarian solution with less redistribution. *But even the Rawlsian social welfare function does not eliminate redistribution entirely* (for the two-type case, see Simonovits 2006). Therefore, in the present version we introduce a *social objective function*, as the difference between the utilitarian social welfare function and a scalar multiple of the variance of individual lifetime pension balances. Though this choice can be objected on a purely welfare basis, we adopt it as an appropriate tool for a multi-objective optimisation. In other words: our choice makes the conflict between welfare maximisation and minimal redistribution explicit. While this modification reduces the degree of indeterminacy, it does not eliminate it for $n > 3$.

(iv) In the limit, when the multiplier of the variance goes to infinity, we obtain a non-redistributive, *neutral* pension rule (Simonovits 2004, 2006). (Note that several authors use neutrality in a different sense, namely for marginal fairness.) The existence of a Pareto-optimal utilitarian second-best neutral pension rule was established, but it may be Pareto-inferior to the second-best redistributive rule.[5] The reason is heuristically simple: due to the incentive compatibility, in the latter the type-specific retirement ages and benefits are probably much less diverse and on average much higher than in the former.

To aid understanding, we display numerical examples of second-best benefit rules for parameter sets showing some similarity to real life. The second-best schedules benefit as a function of retirement age-increase more slowly than the 'fair'[6] and the resulting (benefit, retirement age) pairs are located much more densely in the former than in the latter case. Among the tables, we call the reader's attention to *Table 2,* displaying a set of efficient second-best solutions for the 3-type case and comparing them to the NDC solution.

Summing up: although our model is rather special (individuals only differ in life expectancy but not in labour disutility), it warns us on the dangers of excessive reliance on the so-called actuarial fairness without incentive compatibility

---

[5]   In a variant of the basic insurance model, Rothschild – Stiglitz (1976, p. 638) have considered neutral rules and found similar results: "The [neutral] separating equilibrium ... may not be Pareto optimal even relative to the information that is available".

[6]   This observation is similar to Diamond – Mirrlees (1986, p. 27): "Optimal benefits rise with the age of retirement but more slowly than would actuarially be fair".

(IC). It confirms earlier findings of the mechanism design literature: since the types of the population may differ not only in their preferences (here: the labour disutility) but in their other characteristics (here: the life expectancy) as well, *the incentives of the NDC rules should be dampened*.

The structure of the paper is as follows: Section 2 presents the model. Sections 3 and 4 derive the first-best and the second-best solutions, respectively. Section 5 outlines several algorithms for numerical solution and displays the numerical results. Section 6 concludes. *Appendix 1* presents empirical material on the strong positive correlation between retirement age and life expectancy. *Appendix 2* discusses the generalised utilitarian approach of Eső – Simonovits.

2. THE MODEL

We analyse the following problem. There is a (stationary) population of individuals who have private information regarding their life expectancy (denoted by a positive integer t, calculated from starting their careers). Every individual enters the labour market at age 0, and produces 1 unit of goods per year while he is active, 0 when he is inactive (retired or dead). As is frequent in models of old-age pension systems, we assume that workers cannot save for retirement.[7]

The pension systems we consider will be realistic in the following aspects. The first ingredient of a pension scheme is an annual social security contribution rate, τ, $0 < \tau < 1$ which is levied on active workers (we exclude other taxes). Because of unit wage, the worker's consumption is equal to $1 - \tau$. When a worker retires (say, at age R), he stops producing goods and paying contribution, and receives a yearly retirement benefit, $b > 0$, until he dies. The government designs the contribution rate, τ, and the benefit schedule as a function of the age of retirement, $b(R)$. We require that the pension system be financially sound (the benefit payments cannot exceed the amount of social security contributions paid in). We do not allow the pension system to cut off or reduce the benefits of individuals over time, or give out the benefit as a lump-sum transfer at retirement. Tricks like these would not only make the solution trivial, but more importantly, contradict the purpose of a social security system.[8]

We consider a discrete-type model. Types (life expectancy) range from S to T (both positive integers): $t = S,\ldots,T$. To avoid triviality, we assume that there are at

---

[7]   We assume that the contribution rate is high enough to exclude private life-cycle savings. For an extension to private savings, we recommend Diamond – Mirrlees (1995).

[8]   For example, in the latter case, the individual would have to get private life annuities for the lump-sum transfer at retirement, which would be equally prone to adverse selection due to the asymmetry of information regarding life expectancy.

least two types, i.e. $S < T$. Let f_t be the relative frequency of individuals with life expectancy t: $f_t > 0$ and $\sum_{t=S}^{T} f_t = 1$. The average life expectancy is denoted by $m = \sum_{t=S}^{T} t f_t$.

Type t's *balance* is the difference between the expected lifetime contributions and expected lifetime benefits:

$$z_t = \tau R_t - b_t \cdot (t - R_t) = (\tau + b_t) R_t - t b_t.$$

The cohort's average lifetime balance or equivalently, the population's cross-sectional budget constraint should be met:

$$Z = \sum_{t=S}^{T} z_t f_t = 0.$$

As was mentioned in the *Introduction,* the optimal design literature generally considers redistributive schemes, while other approaches claim neutrality: $z_t = 0$, for all ts. Note that in any menu, neutrality and benefit b_t uniquely determine the type-specific retirement age: $R_t^N = b_t t / (\tau + b_t)$. We will argue that neutrality is undesirable and to demonstrate the inevitability of some redistribution in any reasonable system, here we will show that z_t is a decreasing function of t for a large class of systems. We make the following weak assumptions: the benefits do not decrease with the life expectancy and the difference ratio of retirement age to life expectancy is small enough, and is definitely less than 1. Note that our result is independent of any individual or social optimisation and it takes the retirement ages and benefits as given.

Theorem 1. *If the benefits do not decrease with life expectancy t and if the difference of retirement ages is less than or equal to the ratio of the t's benefit to the sum of the contribution rate τ and the same benefit:*

$$b_{t+1} \geq b_t \ \text{ and } \ 0 \leq R_{t+1} - R_t \leq \frac{b_t}{\tau + b_t}, \quad t = S, \ldots, T - 1,$$

where either the first or the third inequality is sharp, then the individual balances decrease with life expectancies:

$$z_S > \cdots > z_t > z_{t+1} > \cdots > z_T.$$

Proof. Write $z_{t+1} = (\tau + b_{t+1})R_{t+1} - b_{t+1}(t+1)$. With simple calculation we obtain the following condition for $z_{t+1} < z_t$:

$$(\tau + b_t)(R_{t+1} - R_t) + (b_{t+1} - b_t)(R_{t+1} - R_t) < b_t + (b_{t+1} - b_t)(t+1 - R_t).$$

Because of our assumptions and $R_{t+1} - R_t \leq t+1 - R_t$, the first and the second terms on the left-hand side of the inequality are less than or equal to those on the right-hand side, respectively.

Remark 1. Assumptions $b_t \leq b_{t+1}$ and $R_t \leq R_{t+1}$ are natural and hold in our joint first- and second-best optima (*Lemma 1* below). Assumption $R_{t+1} - R_t \leq b_t / (\tau + b_t)$ has no clear economic content but it holds with equality for the neutral solution R_t^N.. Only its relaxation into $R_{t+1} - R_t < 1$ is a natural requirement. Indeed, why should type $t + 1$ work at least one year longer than type t just because he is going to live one year longer? In the numerical explorations, all these conditions, even the stricter ones, hold.

Remark 2. Theorem 1 and $Z = 0$ imply that the longest-lived type is subsidised: $z_T < 0$ and the shortest-lived type is a net contributor to the system: $z_S > 0$. Indeed, this was shown in Simonovits (2006, Corollary 4) for the two-type case within an optimisation framework.

Since the 'fair' benefits play an important role in the literature, we touch them. In our deterministic framework, the formula for 'unadjusted fair' benefits is given by

$$b^F(R) = \frac{\tau R}{m - R}, \quad R < m. \tag{1}$$

Some readers may ask how we dare exclude $R \geq m$. Note that in reality, a stochastic version is used, where the unconditional life expectancy m is replaced by the conditional life expectancy m_R, which is always greater than R. The numerator also takes into account the inherited accounts of those who passed away before reaching age R.

Here we cite a theorem concerning the fairness of the so-called 'fair' benefits.

Theorem 2. *(Simonovits 2003b, Theorem 12.3). If the retirement ages in an unadjusted 'fair' system* (1) *form an increasing sequence:* $R_t < R_{t+1}$, *then the lifetime balances are positive/negative for types with shorter/higher than average life expectancies:*

$$z_t > 0 \ \ for \ \ t < m \ \ \ and \ \ \ z_t < 0 \ \ for \ \ t > m.$$

Furthermore, the expected balance is negative:

$$Z < 0.$$

To eliminate at least the aggregate loss, we can adjust the 'fair' benefit rule as follows:

$$b^A(R) = \frac{\widetilde{\tau}\, R}{m - R}, \quad R < m, \tag{1A}$$

where $\widetilde{\tau}$ is an adjusted rate, yielding $\widetilde{Z} = 0$.

At this point we introduce individual optimisation. During work, an individual has an instantaneous utility (felicity) function $u(\cdot)$, during retirement another one: $w(\cdot)$, both defined on the interval $[0, 2]$.[9] A worker of type t with retirement age R_t and life annuity $b_t > 0$ enjoys utility or felicity $\bar{u} = u(1 - \tau)$ for R_t years and $w(b_t)$ for $(t - R_t)$ years, and the lifetime utility v_t is the sum of his utility when active and retired, respectively. In formula:

$$v_t = R_t \bar{u} + (t - R_t) w(b_t). \tag{2}$$

We assume that both utility functions are twice differentiable, strictly concave and increasing and the worker's utility is lower than the pensioner's: $u(x) < w(x)$. Similarly, retirement benefit b_t should yield more utility to pensioners than does the net wage $1 - \tau$ to workers: $\bar{u} < w(b_t)$. We only make altogether three restrictions on u and w:

$$w(0) - w'(0)\tau < \bar{u} < w(2) - w'(2)(\tau + 2), \quad \bar{u} < w(b_t). \tag{3}$$

The twin assumptions are useful to have a first-best optimum (in Theorem 3 below) and they are quite mild.

The government's goal is to design an optimal pension system, $\{(b_t, R_t, \tau)\}$, maximizing a social welfare function corrected with the degree of redistribution. We concentrate on $\{(b_t, R_t)\}$ rather than on the choice of τ, because that is the part where asymmetric information on individual life expectancy plays any role. In our model, the social security contribution rate is the same for everyone, thus it is only the benefit–retirement age schedule that enables the mechanism designer to sort individuals according to their life expectancy.[10] The social planner does not

---

[9]   The benefit may also be greater than the total wage.

[10]   As opposed to our model, in reality the personal income tax rate depends on the age and the benefit (cf., Diamond – Mirrlees 1978) but we neglect this issue.

observe the individuals' private information, therefore the pension system has to ensure that every type of workers be interested in choosing a contract designed to him. Like in the optimal income taxation literature, the participation constraints are superfluous in a mandatory system.

3. FIRST-BEST PROBLEM

In this section, we derive the optimal benefit–retirement age schedule under the assumption that every worker's life expectancy is commonly observable and the individuals retire at the age determined by the government. This result will serve as a benchmark for the solution under asymmetric information, which is the subject of the next section.

The government wants to maximise the social welfare function and minimise the redistribution under constraints. We will define the utilitarian social welfare by the frequency-weighted sum of the individual utility and the degree of the aggregate lifetime redistribution by the variance of lifetime balances, respectively, i.e.

$$V = \sum_{t=S}^{T} v_t f_t \quad and \quad D^2 z = \sum_{t=S}^{T} z_t^2 f_t.$$

In *Appendix 2,* we will compare this approach to the usual one, which implicitly harmonises welfare and redistribution via a generalised utilitarian social welfare function.

Choosing an appropriate positive scalar δ, the two-objective problem can be unified into a single social objective function:

$$\max_{\{(b_t, R_t)\}_t} \sum_{t=S}^{T} [v_t - \delta z_t^2] f_t;$$

subject to

$$v_t = [\bar{u} - w(b_t)]R_t + w(b_t)t, \quad z_t = (\tau + b_t)R_t - t b_t, \quad t = S, \ldots, T,$$

$$\sum_{t=S}^{T} z_t f_t = 0.$$

This problem is called the social planner's *first-best* problem. Note also that averaging utility of individuals and lifetime balances with different life expectancy means that we consider either the average lifetime utility and balances of a

cohort or the average felicities and the average net contributions of the whole population, respectively. Denoting the first-best solution by *, we obtain

Theorem 3. *In any first-best schedule* $\{(b_t^*, R_t^*)\}_{t=S}^T$, *the benefit* b_t^* *is independent of the life expectancy t, and satisfies*

$$\bar{u} - w(b^*) + w'(b^*)(\tau + b^*) = 0. \tag{4}$$

a) For $\delta = 0$, *there exists a continuum of vectors of first-best retirement ages* $(R_t^*)_{t=S}^T$, *satisfying*

$$\sum_{t=S}^T R_t^* f_t = \frac{b^* m}{\tau + b^*} = R^*. \tag{5}$$

As a special case, every retirement age may be the same and thus equal to their average:

$$R_t^* = R^*, \quad t = S, \ldots T.$$

b) For $\delta \geq 0$, *the neutral solution* $(z_t = 0)$

$$R_t^{N*} = \frac{b^*}{\tau + b^*} t, \quad t = S, \ldots T$$

is a first-best solution, and for $\delta > 0$, *it is the unique first-best solution.*

Proof. Assign λ to the aggregate budget constraint and write the Lagrangian as

$$L = \sum_{t=S}^T ([\bar{u} - w(b_t)]R_t + w(b_t)t - \delta[(\tau + b_t)R_t - tb_t]^2) f_t +$$

$$\lambda \sum_{t=S}^T \{(\tau + b_t)R_t - tb_t\} f_t.$$

The first-order conditions are

$$L'_{b_t} = \{w'(b_t)(t - R_t) + (-2\delta z_t + \lambda)(R_t - t)\} f_t = 0 \Leftrightarrow w'(b_t) = \lambda - 2\delta z_t,$$

$$L'_{R_t} = \{[\bar{u} - w(b_t)] + (-2\delta z_t + \lambda)(\tau + b_t)\} f_t = 0.$$

$L'_{R_t} = 0$ implies (4). Because the derivative of the left-hand-side of (4) is negative, this yields the unicity of $b_t = b^*$. Formulas for a) and b) can be proved by substitution into Z and $z_t = 0$, respectively.

Remark 1. The intuition behind (4) is as follows: if the government increased b_t above (below) b^*, then worker t must work too long (short) and suffer too much (little) disutility. Note that for a positive benefit solution to exist we rely on (3), implying $\bar{u} < w(b^*)$. The solution is unique because the derivative of the left hand side of (4) is negative.

Remark 2. The degeneracy of optimal retirement ages for $\delta = 0$ may be surprising at first sight but it is a direct consequence of the uniformity of the benefits. Indeed,

$$V^* = [\bar{u} - w(b^*)] \sum_{t=S}^{T} R_t^* f_t + w(b^*)m = [\bar{u} - w(b^*)]R^* + w(b^*)m,$$

$$m = \sum_{t=S}^{T} t f_t.$$

On the other hand, if $\delta > 0$, then the second-term $D^2 z$ singles out the neutral solution.

At this point we turn to the optimality of the contribution rate. For simplicity, we discuss the uniform case, and therefore we confine our attention to $S = T = m$. Then at the optimum, $1 - \tau^* = b^*$. For $\tau < \tau^*$, $R > R^*$, while for $\tau > \tau^*$, $R < R^*$.

4. SECOND-BEST PROBLEM

We return now to the model's original informational assumption and assume that individuals have private information regarding their life expectancy, and only the distribution of these data is commonly known. Therefore the optimal benefit–retirement schedule will have to satisfy all incentive compatibility (IC) constraints, as usual in mechanism design.

To avoid complications, we assume that within any cohort, everybody works for a while and everybody retires before dying: $0 < R_t < S$, $t = S,\ldots,T$. This is a reasonable assumption for old-age pension.[11]

Incentive compatibility of $\{(b_t, R_t)\}_{t=S}^{T}$ means that every type t prefers to choose (b_t, R_t) from the schedule.

---

[11]   Without this assumption, the death of shortest-lived persons would provide new information on the life expectancy of the surviving types to the government, even if there were no uncertainty in individual life expectancy.

The upward and downward IC constraints state that type t is not interested in pretending that he is type $t + k$ and vice versa. In formula, for $t = S, \ldots, T - 1$, $k = 1, 2, \ldots, T - t$,

$$v_t \geq [\bar{u} - w(b_{t+k})]R_{t+k} + w(b_{t+k})t = v_{t+k} - w(b_{t+k})k, \tag{6a}$$

$$v_{t+k} \geq [\bar{u} - w(b_t)]R_t + w(b_t)(t + k) = v_t + w(b_t)k. \tag{6b}$$

Let us consider the following modification of the first-best problem, where (6a) and (6b) are added to the original constraints. We will call this optimisation problem the social planner's *second-best problem* and analyze its solution(s) in the rest of the paper. We will distinguish it (them) by hat, $(\hat{R}_t, \hat{b}_t)$.

We shall see that only the adjacent IC constraints ($k = 1$) are relevant. In formula, for $t = S, \ldots, T - 1$,

$$v_t \geq v_{t+1} - w(b_{t+1}), \tag{6c}$$

$$v_{t+1} \geq v_t + w(b_t), \tag{6d}$$

i.e.

$$v_t + w(b_t) \leq v_{t+1} \leq v_t + w(b_{t+1}), \quad for \quad t = S, \ldots, T - 1. \tag{6}$$

We will show a natural property of the second-best optima.

Lemma 1. *In any second-best solution, the benefits and the retirement ages are weakly increasing with life expectancy:*

$$\hat{b}_t \leq \hat{b}_{t+1} \quad and \quad \hat{R}_t \leq \hat{R}_{t+1} \quad for \quad t = S, \ldots, T - 1.$$

Proof. (6c)–(6d) imply $w(\hat{b}_t) \leq v_{t+1} - v_t \leq w(\hat{b}_{t+1})$, hence the increasing property of $w(\cdot)$ implies $\hat{b}_t \leq \hat{b}_{t+1}$. We will show that in turn, $\hat{R}_t \leq \hat{R}_{t+1}$. If the latter were untrue, then using assumption $\bar{u} < w(b_t)$, (6c) and our last inequality would yield

$$\hat{v}_t \geq [\bar{u} - w(\hat{b}_{t+1})]\hat{R}_{t+1} + w(\hat{b}_{t+1})t > [\bar{u} - w(\hat{b}_t)]\hat{R}_t + w(\hat{b}_t)t = \hat{v}_t.$$

a contradiction.

We now show that non-adjacent IC constraints (i.e. $k > 1$) can be ignored. For example, by mathematical induction, (6d) and the inductive assumption $v_{t+k} \geq v_t + kw(b_t)$ imply (6b) with $k + 1$:

$$v_{t+k+1} = v_{(t+k)+1} \geq v_{t+k} + w(\hat{b}_{t+k}) \geq [v_t + kw(\hat{b}_t)] + w(\hat{b}_t) = v_t + (k+1)w(\hat{b}_t).$$

Similarly, for (6a).

The neutral first-best scheme is not incentive compatible because under a constant $b_t^* = b^*$, type T would like to retire at R_S^N rather R_T^N. We will see that for the no-penalty objective function, the first-best and the second-best solutions are identical.

Given τ, the problem of the social planner now becomes

$$\max_{\{(b_t, R_t)\}_t} \sum_{t=S}^{T} [v_t - \delta z_t^2] f_t$$

subject to

$$v_t = [\bar{u} - w(b_t)] R_t + w(b_t)t, \quad z_t = (\tau + b_t) R_t - tb_t, \quad t = S, \ldots, T,$$

$$\sum_{t=S}^{T} z_t f_t = 0,$$

$$v_t + w(b_t) \leq v_{t+1} \leq v_t + w(b_{t+1}), \quad t = S, \ldots, T-1.$$

Since the qualitative results are markedly different depending on whether the objective function has penalty for redistribution or not, we will deal with these two cases separately.

For the no-penalty case, the second-best solution is clearly an optimal solution to the first-best problem: $(\hat{R}_t, \hat{b}_t) \equiv (R^*, b^*)$. Therefore we have

Theorem 4. *For the no-penalty case* ($\delta = 0$)*, there is a unique incentive compatible retirement plan and it attains the first-best outcome.*

The best incentive compatible flexible retirement plan under no-penalty is completely rigid, but it attains the first-best. We will see, however, in the numerical explorations that in practice this total rigidity can be significantly weakened by slightly relaxing the social budget constraint and allowing for a minor negative balance.

Having discussed the no-penalty case, we turn to the more general case, where redistribution is penalised. We rewrite this problem with Mirrlees' (1986, Section 6) method of change of variables, expressing the retirement age by

$$R(v_t, b_t, t) = \frac{w(b_t)t - v_t}{w(b_t) - \overline{u}} \tag{7}$$

and the balance of type t by

$$z(v_t, b_t, t) = (\tau + b_t)R(v_t b_t, t) - tb_t = (\tau + b_t)\frac{w(b_t)t - v_t}{w(b_t) - \overline{u}} - tb_t. \tag{8}$$

Assumption $\hat{R}_t < S$ obviously implies $\hat{R}_t < t$. We have already seen that the disutility of labour is low enough that every type work: $R_t > 0$.

In the following, we will restrict v_t to the interval $v_t \in (\overline{u}t, w(b_t)t)$, thus we keep R_t in $R_t \in (0, t)$.

Also, we prove

Lemma 2. *In any second-best optimum, (i) the longest lived's benefit is equal to the first-best: $\hat{b}_T = b^*$ and (ii) the downward IC constraints (6d) hold with equalities while the upward IC constraints (6c) can be ignored.*

Proof. We will show that in either case, if the result does not hold, then it is possible to change $(\hat{b}_t, \hat{R}_t)$ so that $\hat{z}_t$ remains unchanged and $\hat{v}_t$ increases. Indeed, using the implicit function theorem for $\hat{z}_t = \tau \hat{R}_t - \hat{b}_t \cdot (t - \hat{R}_t) = \text{const.}$, we obtain

$$\hat{R}'_t(\hat{b}_t) = \frac{t - \hat{R}_t(\hat{b}_t)}{\tau + \hat{b}_t}.$$

Inserting this expression into the derivative of $v_t(\hat{b}_t)$ [cf. (2)] yields

$$\hat{v}'_t(\hat{b}_t) = -w'(\hat{b}_t)\hat{R}_t(\hat{b}_t) + [\overline{u} - w(\hat{b}_t)]\hat{R}'_t(\hat{b}_t) + w'(\hat{b}_t)t =$$

$$= \left\{ \frac{\overline{u} - w(\hat{b}_t)}{\tau + \hat{b}_t} + w'(\hat{b}_t) \right\}(t - \hat{R}_t).$$

In the continuation we discuss the two statements (i)–(ii) separately.

(i) We will prove that $\hat{b}_T \neq b^*$ implies a contradiction. In fact, under this assumption, the expression in brackets $\neq 0$, i.e. $\hat{v}'_T \neq 0$, i.e. changing $\hat{b}_T$ in the appropriate direction increases $\hat{v}_T$ [cf. (4) and Remark 1 to Theorem 3], a contradiction.

(ii) It is sufficient to show that strict inequality cannot hold in (6d) for $t = S, \ldots, T - 1$. Assume the contrary for some t and choose the smallest t for which

$\hat{v}_t + w(\hat{b}_t) < \hat{v}_{t+1}$. Since $\hat{b}_t \leq \hat{b}_T = b^*$, the expression in the bracket is positive, implying $\hat{v}_t' > 0$. If the change in $\hat{v}_t$ is small enough, then the inequality remains valid and $\hat{v}_t$ increases. Since this change does not have any influence on the remaining inequalities, this disproves social optimum. If (6d) holds with equality, then (6c) trivially holds, thus can be ignored.

Replacing the two inequalities by one equality in (6), the transformed problem then becomes

$$\max_{\{(b_t, v_t)\}_t} \sum_{t=S}^{T} [v_t - \delta z^2(v_t, b_t, t)] f_t$$

subject to

$$Z = \sum_{t=S}^{T} z(v_t, b_t, t) f_t = 0$$

$$v_{t+1} - v_t - w(b_t) = 0, \quad t = S, \ldots, T-1 \tag{6e}$$

and

$$z(v_t, b_t, t) = (\tau + b_t) \frac{w(b_t) t - v_t}{w(b_t) - \bar{u}} - t b_t, \quad t = S, \ldots, T.$$

This is an optimal control problem, where the independent variable is not time but type t, the state variable is the lifetime utility v_t and the control variable is the benefit b_t. The state space equation is given by the effective adjacent incentive compatibility constraint (6e). The objective function is equal to the social welfare less, the scalar multiple of the variance of the lifetime balances and the isoperimetric condition is given by the social budget constraint.

Our formulation uses discrete time, and hence one cannot expect a smooth benefit–retirement age schedule. A continuous time formulation (Bommier et al. 2011) yields analogous conditions, and results in a continuous benefit–retirement age schedule. Nevertheless, we have chosen the discrete-time formulation because it is more realistic.

We are looking now for a characterisation of a second-best solution, harmonising social welfare and redistribution. Assigning multipliers λ and $(\mu_t)_t$ to the constraints $Z = 0$ and (6e), respectively, the standard technique yields

Theorem 5. *The necessary first-order conditions for the second-best problem with penalised redistribution are for $t = S, \ldots, T$,*

$$L'_{v_t} = \{1 + [-2\delta z(v_t, b_t, t) + \lambda]z'_{v_t}(v_t, b_t, t)\}f_t - \mu_t + \mu_{t-1} = 0,$$

$$L'_{b_t} = [-2\delta z(v_t, b_t, t) + \lambda]z'_{b_t}(v_t, b_t, t)f_t - \mu_t w'(b_t) = 0,$$

$$L'_{\mu_t} = v_{t+1} - v_t - w(b_t) = 0, \quad \mu_t > 0,$$

$$L'_\lambda = \sum_{t=S}^{T} z(v_t, b_t, t)f_t = 0, \quad \lambda > 0,$$

where the formal definitions $\mu_{S-1} = 0$, $\mu_T = 0$ are applied.

By the definition of $z(v_t, b_t, t)$, the partial derivatives that appear in the first-order conditions are

$$z'_{v_t}(v_t, b_t, t) = -\frac{\tau + b_t}{w(b_t) - \overline{u}},$$

$$z'_{b_t}(v_t, b_t, t) = \frac{v_t - t\overline{u}}{[w(b_t) - \overline{u}]^2}\{(\tau + b_t)w'(b_t) - [w(b_t) - \overline{u}]\}.$$

Finally, we will demonstrate that the joint first- and second-best solution is less flexible than the traditional fair system (cf. (1)).

We will start from the simple observation that for the type of average life expectation (assuming m is an integer), the retirement age and (unadjusted) traditional fair benefit are equal to those of the corresponding utilitarian solution, respectively. Then we show analytically that the traditional fair benefit is more flexible than the first- and second-best rule is.

Theorem 6. *Assume that in the 'fair' solution, the lifetime utility is a concave function of the retirement age. Then the corresponding longer-lived's retirement ages are higher than the average first-best retirement age and they receive higher benefit than the first- and second-best, while the opposite holds for the shorter-lived:*

$$R_t^F > R^* \quad and \quad b_t^F > b^* \quad for \quad m < t \leq T$$

and

$$R_t^F < R^* \quad and \quad b_t^F < b^* \quad for \quad S \leq t < m.$$

Remark 1. We conjecture that a similar relation with $\hat{R}_t$ and $\hat{b}_t$ also holds for the penalty case, where the first- and second best solutions are different. Never-

theless, what is already clear is the following: the longest-lived types' optimal benefit in the unadjusted 'fair' solution is greater than that in the second-best: $b_T^F > b^* = \hat{b}_T$, therefore the former cannot be second-best!

Remark 2. Note that Theorem 6 concerns the unadjusted 'fair' solution, which produces negative average balances *(Theorem 2)*. We have no similar analytical statements on feasible, i.e. adjusted 'fair' solutions, mentioned there but it can be presumed that in realistic cases, the qualitative relations survive.

Remark 3. We do not see yet how one can derive the concavity of the lifetime utility-retirement age analytically, but it appears to be a mild restriction.

Proof. We will characterise the 'fair' optimum of a longer-lived, i.e. of $v_t^F(R)$ for $m < t \le T$. Taking the derivative $v_t^{F\prime}$ with respect to R and equating it to zero, yields the optimal 'fair' retirement age R_t^F:

$$v_t^{F\prime}(R) = \bar{u} - w(b^F(R)) + w'(b^F(R))b^{F\prime}(R)(t - R) = 0, \tag{9}$$

where (1) implies

$$b^{F\prime}(R) = \frac{\tau\, m}{(m - R)^2}. \tag{10}$$

Because v_t^F is concave by assumption, to prove $R_t^F > R^*$, it is enough to show $v_t^{F\prime}(R^*) > 0$, or equivalently, since the first and the second terms are identical in (4) and (9), and the third term in (9) is larger than the corresponding term in (4). Dropping the common multiplier $w'(b^*)$, we end up with $b^{F\prime}(R^*)(t - R^*) > \tau + b^*$. Inserting (5) and (10), our inequality reduces to $t > m$. By symmetry, the opposite inequalities hold for $t < m$.

We have made a statement in the Introduction: redistribution is more moderate in any second-best optimum than it is under the traditional rule. This is now presented as

Conjecture 1. *For any penalty, the variance of the lifetime balances under any second-best rule is less than that of 'fair' rule:*

$$\sum_{t=S}^{T} \hat{z}_t^2 f_t < \sum_{t=S}^{T} z_t^{F^2} f_t.$$

5. NUMERICAL EXPLORATIONS

Having reached the limits of our analytical possibilities, we turn to numerical explorations. We will extend our knowledge beyond these limits and present the quantitative features as illustrations.

Two algorithms for finding the second-best solution numerically

In this subsection we mention two algorithms determining second-best solutions: the first algorithm is to examine the local approximate indeterminacy of the case $\delta = 0$ *(Theorem 4)* and the second algorithm is to find the second-best solution *(Theorem 5)* in case $\delta \geq 0$.

   The idea of the first algorithm is the following. Let $\delta = 0$ and $n > 2$, where n is the number of types. Starting from the (first- and) second-best optimum $\{(b_t^*, v_t^*)\}_t$, we will construct a continuum of (possibly) approximate second-best optima around it. Choose any other increasing sequence $(v_t)_t$ with the same $V = V^*$ close to the original vector $(v_t^*)_t$, and assume that v_T is determined from $(v_S, \ldots, v_{T-1})$. From (6e), $w(b_t) = v_{t+1} - v_t$, $t = S, \ldots, T-1$, i.e. denoting the inverse of $w(\cdot)$ by $w^{-1}(\cdot)$, the benefits are $b_t = w^{-1}(v_{t+1} - v_t)$ except for $b_T = b^*$. Then (7) yields the corresponding retirement ages $\{R_t\}_t$. Note, however, that a second v_t, say v_{T-1} must be chosen so that the resulting R_{T-1} ensure $Z = 0$, presuming that $Z'_{v_{T-1}} \neq 0$. We will see that, unfortunately, $Z'_{v_{T-1}} \approx 0$, yielding only very close approximate solutions with slightly negative Zs. Note that for $n = 2$, there are no undetermined components. It is of interest that our algorithm applies to generalised utilitarian social welfare functions as well (cf. *Appendix 2*).

   The second algorithm uses the techniques of optimal control theory. Similarly to the original paper (Eső – Simonovits 2002), here we can use our own numerical algorithm. Unfortunately, for $n > 3$, there is more than one stationary point and we cannot be sure if we have found the global maximum. Any ambiguity can be eliminated by professional programs like *Mathematica*.

Runs

We turn to the description of our explorations. We assume that life expectancy is uniformly distributed: $f_t \equiv 1/n$. We concentrate on the 3-type case but also discuss the 5-type case. We only refer to the case $n = 11$, analysed in Eső – Simonovits (2002) for a generalised utilitarian social welfare function.[12] Assume that the indi-

---

[12]   As a *tour de force*, Pál et al. (2007) have determined *the* true optimum in this case.

viduals' adult life expectancy (counting from entering the work-force at age 20) is between $S = 5 \times 10$ and $T = 5 \times 12$, quite a *rough resolution*. The finer the resolution, the more difficult it is to solve the optimisation problem, presumably because of local indeterminacy.

We shall connect the two felicity functions u and w by a constant disutility of labour ε: $u(x) = w(x) - \varepsilon$. Let the pensioner's felicity function be of inhomogeneous CRRA (Constant Relative Risk Aversion)-type, $w(x) = \theta + x^{\sigma}/\sigma$, $1 - \sigma > 0$ being the coefficient of relative risk aversion.

We will display several runs.

• Run 1. We set $\theta = 4.1$, $\sigma = -0.5$ and $\varepsilon = 1.3975$. In the first-best case, the optimal contribution rate implies that the worker's consumption is equal to the pensioner's. (This is an unwanted consequence of our assumption: the worker's and the pensioner's felicity functions only differ in an additive constant ε.) Therefore let the contribution rate be $\tau = 0.2$. Then $\bar{u} = 4.1 - 0.8^{-0.5}/0.5 - 1.3975 = 0.466$, and the first-best retirement benefit is $b^* = 0.8$. Note that our choice means that one's 0.8 unit consumption as an active worker is equivalent to a 0.303 unit consumption as a pensioner. The difference is due to the increased leisure (or equivalently, decreased disutility) after retirement. Observe that in a neutral solution, the longest-lived person should retire having worked $R_T^N = Tb^* / (\tau + b^*) = 48$ years (at the age of 68).

As proved in *Theorem 4,* one optimal incentive compatible pension system will send everybody to retirement after working $R^* = 44$ years with the first-best benefits. This cannot be improved even under complete information regarding t and differs from neutrality only in that expectedly long-lived individuals are subsidised by the expectedly short-lived individuals.

• Run 2. This run displays several calculations done by the first algorithm which also helps understand the quasi or genuine indeterminacy of the optimum.

The results in *Table 1* demonstrate that indeed, in addition to the unique (italicised) second-best solution, there exist a lot of approximate second-best solutions in its vicinity. For example, the last row in *Table 1* shows an approximate second-best solution which is quite far away from the unique second-best, and yields a fairer solution by incurring a small deficit. (One can determine D^2z as is done in *Table 2*.) *Appendix 2* will present a similar calculation for a generalised utilitarian social welfare function.

• Run 3. Modifying the first algorithm, here we give up the welfare equivalence but pay attention to penalising redistribution. Delaying the arbitrary choice of the penalty coefficient δ, we try to calculate the efficient or dominant optima, using a variant of algorithm 1.

Table 1

True and approximate second-best solutions for 3 types

Lifetime utility			Benefit			(Adult) retirement age			Average balance
short	average	long	short	average	long	short	average	long	
v_{50}	v_{55}	v_{60}	b_{50}	b_{55}	b_{60}	R_{50}	R_{55}	R_{60}	Z
31.7066	41.0263	50.3459	0.8000	0.8000	0.8	44.0000	44.0000	44.0000	0.0
31.7566	41.0263	50.2959	0.7929	0.7929	0.8	43.9207	43.9207	44.0358	−0.0003
31.8566	40.9763	50.2459	0.7721	0.7929	0.8	43.7127	43.9568	44.0716	−0.0017
31.9066	40.9763	50.1959	0.7654	0.7859	0.8	43.6289	43.8766	44.1073	−0.0032

Table 2

Efficient second-best solutions for 3 types

Benefit			Retirement age			Social welfare	Balance's variance	Utility short
short	average	long	short	average	long			
b_{50}	b_{55}	b_{60}	R_{50}	R_{55}	R_{60}	V	D^2	v_{50}
0.44	0.50	0.80	34.36	39.34	48.00	38.69	0	32.96
0.48	0.56	0.80	36.80	40.85	46.83	39.58	0.82	33.16
0.50	0.58	0.80	37.70	41.17	46.43	39.90	1.49	33.21
0.52	0.62	0.80	38.50	42.02	45.93	40.21	2.46	33.19
0.56	0.68	0.80	39.90	42.99	45.29	40.56	4.31	33.01
0.64	0.72	0.80	41.80	43.28	44.72	40.83	6.82	32.60
0.70	0.76	0.80	42.80	43.68	44.36	40.96	8.56	32.25
0.72	0.78	0.80	43.10	43.90	44.23	40.98	9.18	32.11
0.80	0.80	0.80	44.00	44.00	44.00	41.00	10.67	31.68
0.59[a]	0.79	1.19	42.35	44.95	47.90	40.21	13.42	31.24
0.66[b]	0.72	0.80	42.12	43.18	44.65	40.90	7.22	32.55

Notes: a) Adjusted traditional benefit system. b) Generalised hyperbolic utilitarian social welfare function without penalty.

We choose now b_ts rather than v_ts and express the remaining variables. Consider a non-decreasing series $b_S, \ldots, b_T$, where $w(b_S) > \overline{u}$, $b_T = b^*$, and choose a real number $R_S > 0$. Substitute definitions $v_t = [\overline{u} - w(b_t)]R_t + w(b_t)t$ and $v_{t+1} = [\overline{u} - w(b_{t+1})]R_{t+1} + w(b_{t+1})(t+1)$ into $v_{t+1} = v_t + w(b_t)$:

$$[\overline{u} - w(b_{t+1})]R_{t+1} + w(b_{t+1})(t+1) = [\overline{u} - w(b_t)]R_t + w(b_t)(t+1).$$

Express R_{t+1} with R_t:

$$R_{t+1} = \frac{[w(b_{t+1}) - w(b_t)](t+1) + [w(b_t) - \overline{u}]R_t}{w(b_{t+1}) - \overline{u}}, \quad t = S, \ldots, T-1.$$

According to our earlier choice, we have determined all the $2(T - S + 1)$ components, however, the value of R_S is to be determined from $Z = 0$.

Returning to $n = 3$, we run b_{50} from 0.44 to b_{55}, b_{55} from 0.5 to 0.8 with step size 0.02 and R_{50} from 34.4 to 44 with step size 0.1. (In the first row, we start with 34.36 to obtain a more precise approximation for neutral second-best.) We drop those outcomes (V, D), which either have $Z < 0$ or are dominated by a (V', D'), i.e. if $V' > V$ and $D' < D$ or not very characteristic. (More precisely: if $V' \geq V$ and $D' \leq D$, but only one inequality can stand with equality.)

Table 2 is divided into three parts: upper, middle and lower parts. The upper and middle parts contain efficient solutions. It is noteworthy that moving from the top of the upper part to the bottom of the middle part how quickly the variance D^2 increases and how slowly the welfare V increases. Rows of the upper part, including row 1, representing the neutral second-best, are Pareto-inefficient since v_{50} drops below the maximal value, 33.21, obtained in the first row of the middle part. Thus it is optimal to choose a row in the middle part, for example, the italicised row. The last row of the middle part contains the data of the rigid first- and second-best optimum.

Turning to the lower part of *Table 2,* we see the adjusted traditional 'fair' solution in the first row: its variance is 26 percent higher than the rigid second-best and its welfare is also much lower. The next row displays a second-best solution with a hyperbolic social welfare function, borrowed from Eső – Simonovits, now for 3-type. It is close to the italicised row.

It is true that in the 3-type case, for any positive penalty coefficient, the second-best solution is theoretically unique. While we cannot rely on algorithm 1 now, nevertheless, a numerical experiment will show that the second-best solutions are *practically* indeterminate. We choose a tight neighbourhood around the social welfare value 40.86 and the variance value 6.82 of the selected solution *(Table 3).*

The emerging picture is not too clear, but undoubtedly, there is a wide interval of essentially equally attractive second-best solutions with values of objective functions in the intervals $40.80 \leq V \leq 40.86$ and $6.12 \leq D^2 \leq 6.82$, respectively. For instance, b_{50} and b_{55} vary between 0.60–0.66 and 0.68–0.74, respectively. The ranges of the retirement ages are tighter but except for the almost constant R_{60}, still formidable: 41–42.1 and 42.5–44.3 years for R_{50} and R_{55}, respectively.

• Run 4. We turn now to the case of five types. We will apply the second algorithm to the case $\delta > 0$ and study the dependence of the second-best solutions on the penalty coefficient. Let $S = 5 \times 10$, $T = 5 \times 14$, i.e. life expectancy is uniformly distributed on 50, 55, 60, 65 and 70 years. In *Table 4,* we calculate the second-best solutions for two penalty coefficients: $\delta = 0.02$ and 0.04. (Since V and D^2 are not homogeneous, in the scale transformation, δ should be multiplied by 5!) It is remarkable that for $\delta = 0.042$, condition $R_T < S$ does not hold (omitted).

Both solutions are sensible. In the more neutral case (right side of *Table 4*), the benefits and the retirement ages are more dispersed than in the other: 0.58–0.8 vs. 0.65–0.8 and 44.8–49.4 vs. 46.2–48.8 years. The lifetime utility is higher for any lower life expectancy, but is already lower for the average type and even more so for any higher life expectancy. It can be shown again that our results are not robust.

6. CONCLUSION

In this paper we have made a step to apply mechanism design to determine pension benefit rules, in the case where each individual has private information regarding his own life expectancy. We have determined the second-best benefit schedules by a set of necessary first-order conditions. We have worked out two practical algorithms to compute the optimal incentives and endowed the program with data showing some similarity to real life. Apart from numerical indeterminacy, we were able to characterise the conflict between utilitarian welfare maximisation and minimal redistribution.

Our numerical explorations, however, glossed over many important details, including the heterogeneity and age-dependence of labour disutility and of earnings, and the interaction of the pension system with income taxation. Although we were able to prove some analytical results, we have only provided numerical (and not yet analytical) confirmation for the inferiority of the 'traditional fair' optimum. Further research will have to clarify our exploratory results.

Table 3

A neighbourhood of a second-best solution for 3 types

Benefit			Retirement age			Social welfare	Balance's variance	Utility
short	average	long	short	average	long	welfare	variance	short
b_{50}	b_{55}	b_{60}	R_{50}	R_{55}	R_{60}	V	D^2	v_{50}
0.60	0.74	0.80	41.00	43.75	44.78	40.80	6.12	32.79
0.62	0.72	0.80	41.40	43.35	44.79	40.83	6.36	32.73
0.64	0.70	0.80	41.80	42.96	44.85	40.80	6.55	32.62
0.64	0.72	0.80	41.80	43.28	44.73	40.86	6.82	32.62
0.66	0.68	0.80	42.10	42.49	44.86	40.83	6.71	32.58

Table 4

Second-best solutions for 5 types

Life span	$\delta = 0.02$			$\delta = 0.04$		
	Utility	Benefit	Ret. age	Benefit	Utility	Ret. age
t	v_t	b_t	R_t	b_t	v_t	R_t
50	27.682	0.652	46.210	0.584	28.624	44.762
55	35.794	0.700	46.828	0.636	36.036	45.756
60	44.343	0.738	47.458	0.694	43.995	47.001
65	53.205	0.771	48.107	0.751	52.492	48.249
70	62.317	0.800	48.753	0.800	61.500	49.374

REFERENCES

Arrow, K. J. – Intrilligator, M. D. (eds) (1986): *Handbook of Mathematical Economics.* Vol. III. Amsterdam: Elsevier Science Publisher.

Barr, N. – Diamond, P. (2006): The Economics of Pensions. *Oxford Review of Economic Policy,* 22(1): 15–39.

Bommier, A. – Leroux, M. L. – Lozachmeur, J. M. (2007): *Uncertain Lifetime, Redistribution and Nonlinear Pricing of Annuities.* CESifo Area Conference, April 20–22.

Bommier, A. – Leroux, M. L. – Lozachmeur, J. M. (2011): Differential Mortality and Social Security. *Canadian Journal of Economics,* 44(1): 273–289.

Cremer, H. – Lozachmeur, J. M. – Pestieau, P. (2008): Social Security and Retirement Decision: A Positive and Normative Approach. *Journal of Economic Surveys,* 22: 213–233.

Diamond, P. (2003): *Taxation, Incomplete Markets and Social Security. Munich Lectures*. Cambridge, MA: MIT Press.

Diamond, P. – Mirrlees, J. (1978): A Model of Social Insurance with Variable Retirement. *Journal of Public Economics,* 10: 295–336.

Diamond, P. – Mirrlees, J. (1986): Payroll-tax Financed Social Security with Variable Retirement. *Scandinavian Journal of Economics,* 88: 25–50.

Diamond, P. – Mirrlees, J. (1995): *Social Insurance with Variable Retirement and Private Saving*. Cambridge, MA: MIT Press.

Eső, P. – Simonovits, A. (2002): Designing Optimal Benefit Rules for Flexible Retirement. *Discussion Paper,* CMS-EMS 1353. Evanston, IL: Northwestern University.

Fabel, O. (1994): *The Economics of Pensions and Variable Retirement Schemes*. New York: Wiley&Son.

Fenge, R. – De Menil, G. – Pestieau, P. (eds) (2006): *Strategies of Pension Reforms*. Cambridge, MA: MIT Press.

Fenge, R. – Pestieau, P. (2005): *Social Security and Early Retirement*. Cambridge, MA: MIT Press.

Golosov, M. – Tsyvinsky, A. (2006): Designing Optimal Disability Insurance: A Case for Assets Testing. *Journal of Political Economy,* 114: 257–279.

Gruber, J. – Wise, D. A. (eds) (1999): *Social Security and Retirement around the World*. Chicago: Chicago University Press.

Hamermesh, D. (1985): Expectation, Life Expectancy and Economic Behavior. *Quarterly Journal of Economics,* 100: 389–408.

Heidler, M. – Leifels, A. – Raffelhüschen, B (2006): Heterogeneous Life Expectancy, Adverse Selection and Retirement Behavior. *Discussion Paper* 13. Freiburg University.

Holzmann, R. – Palmer, E. (eds) (2006): *Pension Reform through NDCs: Issues and Prospects for Non-financial Defined Contribution Schemes*. Washington, D.C.: World Bank.

Holzmann, R. – Stiglitz, J. E. (eds) (2001): *New Ideas about Old-age Security: Toward Sustainable Pension Systems in the 21st Century.* Washington, D.C.: World Bank.

Lindbeck, A. – Persson, M. (2003): The Gains from Pension Reform. *Journal of Economic Literature,* 41: 74–112.

Mas-Colell, A. – Whinston, M. O. – Green, J. (1995): *Microeconomic Theory*. Oxford: Oxford University Press.

Mirrlees, J. A. (1971): An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies,* 38: 175–208.

Mirrlees, J. A. (1986): The Theory of Optimal Taxation. In: Arrow – Intrilligator (eds): *Handbook of Mathematical Economics.* Amsterdam: Elsevier Science Publisher, pp. 1197–1249.

Orszag, P. – Stiglitz, J. E. (2001): Rethinking Pension Reform: Ten Myths about Social Security Systems. In: Holzmann – Stiglitz (eds): *New Ideas about Old-age Security: Toward Sustainable Pension Systems in the 21ˢᵗ Century,* Washington, D.C.: World Bank, pp. 17–56.

Pál, L. – Csendes, T. – Sendin, O. H. – Bangha, J. R. (2007): *Improvements on the Global Optimization Algorithm with Numerical Test*. 7th International Conference on Applied Informatics, Eger, Hungary.

Rothschild, M. – Stiglitz, J. E. (1976): Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *Quarterly Journal of Economics,* 90: 629–649.

Samwick, A. A: (1998): New Evidence on Pensions, Social Security and the Timing of Retirement. *Journal of Public Economics,* 70: 207–236.

Sheshinski, E. (1978): A Model of Social Security and Retirement Decisions. *Journal of Public Economics,* 10: 337–360.

Sheshinski, E. (2006): Optimum Delayed Retirement Credit. In: Fenge et al. (eds): *Strategies of Pension Reforms*. Cambridge, MA: MIT Press, pp. 27–36.

Simonovits, A. (2003a): Designing Optimal Linear Rules for Flexible Retirement. *Journal of Pension Economics and Finance,* 2: 273–293.

Simonovits, A. (2003b): *Modelling Pension Systems*. Houndsmill: Palgrave Macmillan.

Simonovits, A. (2004): Designing Benefit Rules for Flexible Retirement with or without Redistribution. *CESifo Discussion Paper,* 1370.

Simonovits, A. (2006): Optimal Design of Pension Rule with Flexible Retirement: The Two-type Case. *Journal of Economics,* 89: 198–222.

Smith, V. K. – Taylor, D. H. – Sloan, F. A. (2001): Longevity Expectations and Death, can People Predict their Own Demise? *American Economic Review,* 91: 1126–1134.

Waldron, H. (2001): Links between Early Retirement and Mortality. *ORES Working Paper* 93, Division of Economic Research, Social Security Administration.

Wolfe, J. R. (1983): Perceived Longevity and Early Retirement. *Review of Economic Statistics,* 65: 544–555.

World Bank (1994): *Averting the Old-age Crisis*. New York: Oxford University Press.

APPENDIX 1

On the correlation between life expectancy and retirement age[13]

Tables A.1 and *A.2* contain relevant data for males and females, respectively.

As a background to the evaluation of the data, we mention that the male normal retirement age has increased from 60 to 62 between 1997 and 2001, while the corresponding female variable has increased from 55 to 62 between 1997 and 2009.

Obviously, the bulk of the current deceased had retired well before 1997, i.e. with the prevailing normal retirement ages 60/55. During this era, the incentives to late retirement or punishment for early retirement hardly existed. Note the ex-

---

[13] Rudolf Borlói and Judit Marosi (Hungarian Pension Headquarters) kindly released the following data on the connection between retirement age and average life span of those deceased in 2004 and László Hablicsek sent us his life table of the same year, Hungary.

traordinary concentration of actual retirement ages around the normal one, especially for males.

Here we only repeat from the Introduction that among male old-age pensioners, who died in 2004, those retired at 57 lived only 12.3 years in retirement, while those who retired at 65, lived another 24.3 years. For comparison, the conditional life expectancy for aged 57 was 18.0 years, while for aged 65, it was 13.1. Similar story is depicted in *Table A.2* about females.

Table A.1

Years spent in old-age retirement, males

Age at retirement	Number of cases	Average years spent in retirement	LEXP
57	2089	12.3	18.0
58	1737	13.5	17.3
59	1254	14.2	16.7
60	*17152*	*17.2*	*16.1*
61	3641	20.9	16.4
62	1149	18.1	14.9
63	609	22.4	14.3
64	450	23.4	13.7
65	417	24.3	13.1

Table A.2

Years spent in old-age retirement, females

Age at retirement	Number of cases	Average years spent in retirement	LEXP
52	687	14.8	27.4
53	631	15.8	26.6
54	705	19.7	25.7
55	*14002*	*20.7*	*24.9*
56	5128	23.6	24.1
57	2376	24.3	23.3
58	1756	24.7	22.5
59	1238	25.4	21.7
60	1266	26.7	20.9
61	757	25.6	20.1
62	566	24.5	19.3
63	451	23.5	18.5
64	327	23.0	17.7
65	288	21.5	16.9

In both cases, the adverse selection is clear. Luckily, most pensioners retired at the normal retirement age, thus the distortions would not have been significant if flexible retirement had been available.

APPENDIX 2

Generalised utilitarian social welfare functions

There might be a temptation to derive less redistributive second-best optima by applying generalised utilitarian social welfare functions rather than introducing *ad hoc* corrections like deducting the variance of lifetime balances from the utilitarian one. In this Appendix we will consider and evaluate such an approach. Following Simonovits (2006), it can be shown that in a 2-type model there is a perfect equivalence between the two approaches. We will show, however, that in a 3-type model the equivalence disappears. There exist not only practically but theoretically welfare equivalent second-best solutions, thus we can choose one with minimal redistribution. Consequently, the purity of the welfare approach cannot and need not be preserved.

Turning to the details, let ψ be an increasing and concave scalar-scalar function. The simplest generalised utilitarian social welfare function is the average value of the ψ-transform of individual utility: $V = \sum_{t=S}^{T} \psi(v_t) f_t$. The more concave ψ is, the more egalitarian the social welfare function is. (Note that a scalar-scalar function $\psi(v)$ is called *more concave* than function $\varphi(v)$ if there exists a third strictly concave scalar-scalar function χ, such that $\psi(v) \equiv \chi(\varphi(v))$.)

Comparing the present paper with Eső – Simonovits (2002) reveals that replacing the present paper's utilitarian approach by a generalised utilitarian one does not affect the qualitative characteristics of the second-best solutions. We adopt here the generalised utilitarian function with $\psi(v) = v^{\sigma}/\sigma$, more specifically $\sigma = -1$, i.e. $\psi(v) = -1/v$ from Eső – Simonovits but otherwise we retain the previous parameter values and set δ to 0.

Table A.3 shows three numerical examples of second-best solutions. We omit the identical social welfare values ($V = -0.1258$) but display three different degrees of redistribution.

To construct *Table A.3*, first we have determined one second-best solution by our own program developed by Eső – Simonovits (middle row) and then we have computed others by algorithm 1. Since we have obtained not only practically but also theoretically welfare equivalent second-best solutions, we can choose one with lower redistribution. In *Table A.1*, this is the last row.

Table A.3

Second-best solutions with the same generalised utilitarian welfare

Benefit			Retirement age			Utility			Balance
short	average	long	short	average	long	short	average	long	variance
b_{50}	b_{55}	b_{60}	R_{50}	R_{55}	R_{60}	v_{50}	v_{55}	v_{60}	D^2
0.693	0.717	0.800	42.624	43.021	44.556	32.397	40.881	49.568	7.844
0.659	0.719	0.800	42.118	43.180	44.648	32.547	40.731	49.440	7.223
0.628	0.723	0.800	41.557	43.348	44.735	32.697	40.581	49.318	6.586