MAGYAR TUDOMÁNYOS AKADÉMIA
Közgazdaság- és Regionális Tudományi Kutatóközpont

KRTK
CERS-HAS

Research Centre for Economic and Regional Studies
HUNGARIAN ACADEMY OF SCIENCES

MŰHELYTANULMÁNYOK DISCUSSION PAPERS

MT-DP – 2012/8

Optimal Cap on Pension Contributions

ANDRÁS SIMONOVITS

INSTITUTE OF ECONOMICS, RESEARCH CENTRE FOR ECONOMIC AND REGIONAL STUDIES,
HUNGARIAN ACADEMY OF SCIENCES BUDAPEST, 2012

Discussion papers
MT-DP – 2012/8

Institute of Economics, Research Centre for Economic and Regional Studies,
Hungarian Academy of Sciences

Optimal Cap on Pension Contributions

Author:

András Simonovits
research advisor
Institute of Economics
Research Center for Economic and Regional Studies
Hungarian Academy of Sciences
Mathematical Institute of Budapest University of Technology
Department of Economics of CEU
email: simonov@econ.core.hu

March 2012

ISBN 978-615-5243-01-1
ISSN 1785 377X

Optimal Cap on Pension Contributions

András Simonovits

Abstract

In our model, the government operates a mandatory proportional (contributive) pension system to substitute for the low life-cycle savings of the low-paid myopes. The socially optimal contribution rate is high (equalizing young- and old-age consumption for them), while an appropriate cap on pension contributions makes room for the saving of high-paid far-sighted workers. In our parameterization (with a Pareto earning distribution), the optimal cap can be determined but its aggregate impact is negligible.

Keywords: pensions, contribution rate, contribution cap, maximum for taxable earnings

JEL classification: H53, H24

Acknowledgement:

I express my gratitude to N. Barr, Zs. Cseres-Gergely, J. Köllő and G. Kőrösi for friendly comments. This research has received generous financial support from OTKA K 67853.

A nyugdíjjárulék optimális plafonja

Simonovits András

Összefoglaló

Modellünkben a kormányzat egy kötelező és járulékarányos nyugdíjrendszert működtet a kiskeresetű és rövidlátó dolgozók elégtelen életpálya megtakarításainak pótlására. A társadalmilag optimális társadalombiztosítási járulék kulcsa nagy (egyenlővé teszi az említett dolgozók fiatal és időskori fogyasztását), miközben nyugdíjjárulékra vonatkozó alkalmas plafon helyet biztosít a nagykeresetű és messzelátó dolgozók megtakarításainak. Paraméterezésünkben (Pareto-eloszlású kereseteloszlás) az optimális plafon meghatározható, de aggregált hatása elenyésző.

Tárgyszavak: nyugdíjak, járulékkulcs, járulékplafon, adóztatható keresetek maximuma

JEL KÓD: H53, H24

1. Introduction

Since the publication of World Bank (1994), the debate on privatization and prefunding of the unfunded public pension systems has focused attention to the socially optimal choice of the contribution rates to the arising two pillars. In terms of gross wage, the public contribution rate varies across time and space dramatically: while the US social security contribution rate is as low as 12.4 percent, the Hungarian rate is as high as 34 percent. These differences are due to differences in the breakdown to the employer's and the employee's rates, the dependency rates (the ratios of pensioners to the workers), the replacement rates (the ratios of average benefits to averages wages), the degree of redistribution and the share of the private pillar.

Much less attention has been paid to the socially optimal choice of the *earnings cap*, officially called the maximum for taxable earnings (for a summary statistics, see Table 1 in Valdés-Prieto and Schwarzhaupt 2011). Such a cap (or ceiling) implies an upper limit on the mandatory pension contributions but especially on the future benefits. While the contribution rate affects every worker, a well-designed cap only influences the best-paid; nevertheless, the cap also deserves attention. We give only two examples to illustrate poor design of the cap. (i) In the 1950s, in Great Britain the cap was fixed at the minimum wage (about half the average wage), transforming the usually earning-related contributions into flat ones. It was only realized much later that such a solution reduces excessively the flat benefit and then was replaced by a much higher cap, making the rise of the benefit and the redistribution possible within the public system. (ii) In Hungary, the ratio of the cap to the average gross wage sank from 3.3 to 1.6 between 1992 and 1996 just to grow from 1.6 to 3.1 between 1997 and 2005. (Note also that the cap on the contributions to the tax-favored *voluntary* pension system is much lower and it interacts with the mandatory one; Simonovits, 2011.)

Most experts see the role of the cap as a hidden personal income tax. For example, in some countries, the cap only applies to the employee's contribution and the benefit, while the employer's contribution above the cap is a pure personal income tax. (In Hungary, the former contribution rate is 10 percent of the gross wages, while the latter is 24 percent. The uncapped contributions alone provide 6 percent of the total pension contributions.) Others justify the existence of the cap as follows: the government has no mandate to force high old-age consumption on high-earners.

In contrast to these approaches, we study a third role of the cap in the mandatory public system: to ensure sufficient mandatory savings for the low-earning short-sighted but leave sufficient room for the more efficient voluntary savings for the high-earning far-sighted. (There is a common presumption that one dollar paid to a private fund generates a higher old-age benefit than a public pension system. In fact, the voluntary (private) saving may also be less efficient than the mandatory (public) one, see Barr and Diamond, 2008, Chapter 6. This efficiency assumption is only used here for modeling the incentive problems occurring in the mandatory pension system.)

Correspondingly we start from a very simple model, where workers only differ by their discount factors and wages but have the same age when they start working, retire or die. Using a *paternalistic* social welfare function, where individual discounting is eliminated, a socially optimal pension system successfully combines these two tasks. (Note that in an imaginary world, where the discount factor is a decreasing rather than an increasing function of the wage, the cap should be replaced by a floor.)

For the sake of simplicity, we confine our analysis to *proportional* (defined contribution = DC) pension systems (and write contribution rather than pay-roll tax). To avoid absurd results (suggesting the elimination of any mandatory pension system), we assume realistically that the frequency of lower-paid workers is so high, and their wages and their discount factors are so low that the socially optimal contribution rate is *maximal*, equalizing young and old-age consumption for the myopes. The cap reduces the *effective* contribution rate (i.e. the ratio of contribution to earning) of workers earning above the cap, making room for private savings. In addition to the contribution rate, *the value of the optimal cap can also be determined*, but in the most realistic version of the model, *its aggregate impact is hardly noticeable.* In other words: in a wide interval, when the cap is slightly raised, the marginal gains of the lower-paid and short-sighted workers are canceled by the marginal losses of the higher-paid and far-sighted workers. We are uncertain if this result is only due to the omissions of the model or reflects deeper economic relations.

Formulating our model, we have adopted Feldstein (1987): evaluating the mandatory pension systems, he emphasized the role of different discount factors in the accumulation of voluntary life-cycle savings and used a paternalistic social welfare function for correcting individual myopia. As a short-cut, he neglected wage differences, consequently he confined his attention to flat or means-tested benefits. He found that the optimal means-tested system is typically welfare superior to the flat one, because it limits the benefit provision to the myopes and leaves room for more efficient private savings.

The present paper modifies Feldstein (1987) in several dimensions: (i) as is already mentioned, different discounting types have also different wages, namely higher earners discount the utility of old-age consumption less than lower earners do, (ii) the pensions are proportional to the contributions rather than being flat or means-tested, (iii) contributions and benefits are capped. We have also made two technical modifications: (a) the length of the time spent in retirement is shorter than that in work; (b) the lifetime utility function is of Leontief rather than Cobb–Douglas-type.

Recently, Valdés-Prieto and Schwarzhaupt (2011) have created a model to analyze the issue of the optimal coercion including the choice of cap. They have considered a larger set of errors of judgments on remaining life span and future needs, furthermore, modeled various pension systems. They put the value of the optimal cap near the 80th percent of the earning distribution.

To allude to other approaches, we only mention just one paper (Casamatta, Cremer and Pestieau, 2000) which allowed for mixed (partly flat, partly proportional) systems and replaced social welfare maximization with the median voter's self-optimization of the contribution rate.

Note that models of this type neglect real-life complications like growth, inflation and population aging. Therefore, in our static model, we cannot consider the dynamic problem of carving out a private pillar from a public one (Diamond and Orszag, 2004 vs. Feldstein, 2005). Neither can we evaluate proposals like Diamond and Orszag (2004) who would phase-in a 3 percent tax on incomes above the cap to reduce the long-term imbalance of the US Social Security. The problem of time-inconsistency in private savings (e.g. Laibson, 1997) is also out of scope. (In his dynamic model, Docquier (2002) replaced the static social welfare function by a dynamic one, comprising the life-time utilities of all generations.)

To help the reader and also demonstrate the robustness of our results, we study four versions of the model in a row analytically and illustrate some of them numerically. In the single-type version of the model, there is no role for the cap and the socially optimal contribution rate is either zero (for a sufficiently far-sighted population) or maximal (for an insufficiently far-sighted population): equalizing the young- and old-age consumptions. In the multi-type versions, widespread and strong myopia calls for a high contribution rate. In the two-type version of the model, the role of the cap is to reduce the effective pension contribution rate for the higher-paid far-sighted type, putting the cap at the lower wage. (The end result resembles the British flat contribution–flat benefit system in the 1950s.)

In the three-type version of the model, the first (lowest-earning) type is strongly myope, the second (medium-earning) type is moderately and the third (highest-earning) type is hardly myope. Their frequencies are steeply diminishing. The inevitably high contribution rate seriously undermines even the saving incentives of the medium type. Therefore the socially optimal cap must also fully cover this type. But to ensure some room for saving by the high-earners, the government excludes any further increase in the cap. The socially optimal cap is equal to the medium wage, which may be greater than the average wage. (Note that our medium wage generally differs from the median wage, which is typically lower than the average wage.)

Introducing a continuous-type model, the choice becomes a mathematically intriguing problem. We prove a quite complex optimality condition: marginally raising the optimal cap, the marginal gain of subcritical earners (with zero savings) is equal to the marginal loss of supercritical earners (with positive savings). We have a simple corollary: the optimal cap is a decreasing function of the efficiency of the private savings. For the Pareto-2 distribution (where the minimum wage is just half the average wage and the wage variance is infinite)—the socially optimal contribution rate remains the maximal one—and the socially optimal cap is between 2–3 times the average wage, covering 97-94 percent of the workers. Note, however, that social welfare is practically insensitive to the cap.

Further research is needed to generalize the results for more general setting, for example, to replace Leontief utility functions with general CRRA-ones, and then do the same with our purely utilitarian social welfare function. (The latter is not very significant because relatively quite few workers earn between, say 2 and 3 times the average.) Deterministic and stochastic changes in the relative earnings position may also be important. The dependence of type-specific life expectancy and private saving's efficiency on earning is also important in practice, and may call for progressive benefits.

The structure of the remainder is as follows. Section 2 presents the model, Section 3 displays numerical illustrations. Finally, Section 4 concludes.

2. Model

First we shall outline a framework, and then we shall specify versions of the model with various number of types.

Framework

We consider a very simple pension model, where workers only differ in wages and discount factors but wages do not vary with age. A type can be described by his total wage w and his discount factor δ, the joint distribution function of (w, δ) by F and the corresponding expectations by $\mathbf{E}$. For example, the average wage is normalized as $\mathbf{E}w = 1$. Workers pay contributions τw up to $\tau \bar{w}$, where $0 < \tau < 1$ is the *contribution rate* to the mandatory pension system and $\bar{w}$ is the *wage ceiling or cap*. Hence his *covered total and net wages* are respectively

$$\hat{w} = w \min(w, \bar{w}) \quad \text{and} \quad v = w - \tau \hat{w}.$$

Calculating pension benefits, a related variable is used, the *net covered earning*:

$$\hat{v} = (1 - \tau)\hat{w}.$$

By assumption, the pension benefits are proportional to the net covered wages: $b = \beta \hat{v}$. Everybody works for a unite period, and everybody spends the same time in retirement, namely μ, $0 < \mu \leq 1$. Hence the pension balance is simply $\mu\beta\mathbf{E}\hat{v} = \tau\mathbf{E}\hat{w}$, i.e. $\beta = \mu^{-1}\tau/(1 - \tau)$. Since the two periods' lengths are different, we use *intensities*, i.e. quantities per a unit time period.

In addition to paying pension contributions, workers can also privately save for old-age: $s \geq 0$. Denoting the compound interest factor by $\rho \geq 1$, the intensity of the decumulated saving is $\mu^{-1}\rho s$.

We can now describe the young and old-age consumption (intensities):

$$c = v - s \quad \text{and} \quad d = b + \mu^{-1}\rho s.$$

To determine the individually optimal savings, we introduce lifetime Leontief utility functions:

$$U(c, d) = \min(\delta c, d),$$

where δ is the discount factor, $0 \leq \delta \leq 1$. The Leontief utility function is not only simpler but also more realistic than the Cobb–Douglas-type.

We must separate two cases: either 1) positive saving intention or 2) negative saving intention. At the optimum of type 1, $\delta c = d$, hence equation

$$\delta v - \delta s = b + \mu^{-1}\rho s$$

yields the optimal saving intention:

$$s^i = \frac{\delta v - b}{\delta + \mu^{-1}\rho}$$

which materializes if it is positive: $\delta v > \mu^{-1}\tau\hat{w}$.

Hence the optimal worker's consumption is

$$c_1^* = \frac{\rho v + \mu b}{\mu\delta + \rho}.$$

On the other hand, if $s^i < 0$, then $s_2^* = 0$ hence $c_2^* = v$ and $d_2^* = b = \mu^{-1}\tau\hat{w}$.

Analyzing the indirect utility function, we shall drop the distinction between the two regimes:

$$u(w, \delta) = \min[\delta c^*, d^*].$$

Note that for wages lower than the cap, its existence is indifferent; for wages higher than the cap, i.e. $w > \bar{w}$ this means that the *effective* contribution rate is $\tilde{\tau} = \tau\bar{w}/w < \tau$, reducing the original contribution rate for workers earning above the cap. The function of the cap is as follows: lower-paid myopic workers can be cajoled into a pension system with a high contribution rate, but higher-paid far-sighted workers pay a lower effective rate.

By choosing the contribution rate τ and the wage cap $\bar{w}$, the government maximizes the expected value of the *paternalistic*, undiscounted indirect utility functions, i.e.

$$V(\tau, \bar{w}) = \mathbf{E}u^*(w, \delta) = \mathbf{E}\min[c^*, d^*].$$

Since the utility is linear in consumption, our social welfare function has a direct economic meaning: if a system produces 1% higher social welfare than the other, then the original wages need to be scaled up by 1% in the second system to compensate for suboptimality.

Next we study the single-, the two- and the three-type models and then we turn to the most demanding continuous-type model.

A single-type model

We start the analysis with a single-type model, where the typical individual has a wage $w > 0$ and a discount factor δ, $0 \le \delta \le 1$. Here the cap has no function but the model is a useful ground for preparation.

Using the general formulas above, we can deduce the saving function

$$s^{\mathrm{o}} = \frac{[\delta(1-\tau) - \mu^{-1}\tau]w}{\delta + \mu^{-1}\rho} > 0 \quad \text{if} \quad 0 < \tau < \tau_\delta = \frac{\delta}{\delta + \mu^{-1}},$$

where τ_δ is the maximal contribution rate which allows for positive or zero saving intention.

An optimal contribution rate obviously cannot be higher than the absolute maximum:

$$\bar{\tau} = \frac{1}{1 + \mu^{-1}} = \tau_1,$$

because such a rate would suboptimally raise the old-age consumption above the young-age consumption: $d > c$. Indeed, $(1-\tau)w \le \mu^{-1}\tau w$ implies $\tau \le \bar{\tau}$—to be assumed from now on.

To determine the optimal τ, we need to separate the two cases and to do so, introduce the *critical discount factor*

$$\delta^* = \frac{1}{1 + \mu(1 - \rho^{-1})}.$$

Note that if the private saving is efficient, i.e. $\rho > 1$, then $0 < \delta^* < 1$, which is always assumed from now on.

We shall prove the following theorem.

Theorem 1. *In a single-type model with efficient private savings, no cap is needed and the socially optimal contribution rate is maximal for subcritical discount factors and zero for supercritical discount factors and both for the critical discount factor:*

$$\tau^{\mathrm{o}} = \begin{cases} \bar{\tau} & \text{if } 0 \le \delta < \delta^*; \\ 0,\ \bar{\tau} & \text{if } \delta = \delta^*; \\ 0 & \text{if } \delta^* < \delta \le 1. \end{cases}$$

Remark. In the limit, when the private saving is as efficient as the mandatory pension system, i.e. $\rho = 1$, then $\delta^* = 1$ and the indeterminacy becomes total, extending to the entire interval $[0, \bar{\tau}]$: a perfect substitution arises between mandatory contributions and voluntary savings.

Proof. To find the optimal τ, we must separate the two cases of positive and negative saving intentions.

ad 1) Substituting $v = (1-\tau)w$ and $b = \mu^{-1}\tau w$ into the formula $c_1 = d_1/\delta$, we obtain

$$d_1(\tau) = \frac{\delta[\rho - (\rho - 1)\tau]w}{\mu\delta + \rho}.$$

It is easy to see that $d_1(\tau)$ is a decreasing function in the interval $[0, \bar{\tau}]$, thus its maximum is reached at $\tau = 0$, where

$$d_1(0) = \frac{\delta\rho w}{\mu\delta + \rho}.$$

ad 2) Using the benefit formula $d_2(\tau) = \mu^{-1}\tau w$, $d_2(\tau)$ is obviously an increasing function in the interval $[0, \bar{\tau}]$, and its maximum is reached at $\bar{\tau}$:

$$d_2(\bar{\tau}) = \frac{w}{1+\mu} = c_2(\bar{\tau}).$$

To decide which maximum is higher, one must compare $d_1(0)$ and $d_2(\bar{\tau})$. A simple calculation yields the value of the critical discount factor δ^* which equates the two quantities:

$$\frac{\delta^*\rho}{\mu\delta^* + \rho} = \frac{1}{1+\mu}.$$

For $\delta < \delta^*$, $d_1(0) < d_2(\bar{\tau})$ and $\tau^{\mathrm{o}} = \bar{\tau}$ hold. For $\delta > \delta^*$, $d_1(0) < d_2(\bar{\tau})$ and $\tau^{\mathrm{o}} = 0$ hold. For $\delta = \delta^*$, $d_1(0) = d_2(\bar{\tau})$ holds and τ^{o} is indeterminate: either 0 or $\bar{\tau}$. ∎

We can now formulate a trivial

Corollary. *For multi-type models, a) if all discount factors are lower than the critical one, then the socially optimal contribution rate and the corresponding cap are equal to the maximum, b) if all discount factors are higher than the critical one, then the socially optimal contribution rate is zero, while the corresponding cap is indifferent.*

Note that both cases cover quasi-single type populations.

A two-type model

Our single (or quasi-single)-type model is not only unrealistic but by construction excludes an optimal cap. Therefore we shall continue with a two-type model: low and high types (denoted by subscript L and H) with wages $w_L < w_H$ and discount factors $0 \le \delta_L < \delta_H \le 1$, respectively. To avoid trivialities covered in the Corollary, we assume that the critical discount factor separates the two discount factors:

$$0 \le \delta_L < \delta^* < \delta_H \le 1,$$

suggesting a maximal contribution rate for type L and a minimal (in fact, zero) contribution rate for type H. For practical reasons, such a solution is politically impossible to legislate, therefore as a second-best solution, the government might introduce a cap on the contribution, just at the low wage. To aggregate the type-effects, we define a social welfare function. Introducing the *frequencies* of both types, to be denoted by $f_L > 0$ and $f_H > 0$ with $f_L + f_H = 1$. We assume a *paternalistic* utilitarian social welfare function, without discounting young-age consumption:

$$V(\tau, \bar{w}) = f_L d_L + f_H d_H.$$

First we formulate the optimality condition of that arrangement in a very simple setup.

Theorem 2. *Suppose two types with $0 < w_L < w_H$ and $\delta_L = 0$ and $\delta_H = 1$. If the parameters satisfy the condition*

$$\frac{f_L}{f_H} > \frac{\mu(\rho - 1)}{\mu + \rho},$$

then the socially optimal contribution rate is maximal: $\tau^{\text{o}} = \bar{\tau}$ and the corresponding cap is equal to the lower wage: $\bar{w}^{\text{o}} = w_L$.

Proof. In this extreme setup, $d_L(\tau)$ is increasing and $d_H(\tau)$ is decreasing in the entire interval $[0, \bar{\tau}]$. Then $d_L(\tau) = \mu^{-1} w_L \tau$ and

$$d_H(\tau) = \frac{[\rho - (\rho - 1)\tilde{\tau}]w_H}{\mu + \rho}, \quad \text{where} \quad \tilde{\tau} = \frac{\bar{w}}{w_H}\tau.$$

The social welfare function is

$$V(\tau, \bar{w}) = \frac{f_L w_L \tau}{\mu} + f_H \frac{\rho w_H - (\rho - 1)\bar{w}\tau}{\mu + \rho}.$$

Evidently the lower the cap, the higher is the social welfare, therefore $\bar{w}^{\text{o}} = w_L$. Then the coefficient of $w_L \tau$ in $V(\tau, \bar{w})$ is equal to

$$\frac{f_L}{\mu} - \frac{f_H(\rho - 1)}{\mu + \rho}$$

and this should be positive to have an increasing function of the tax rate τ. ∎

Using extreme assumptions on discounting (á la Feldstein), the displayed condition basically presumes that the efficiency of the private saving cannot be too high. It is time to relax the extreme conditions of full myopia and full far-sightedness. To sharpen Theorem 2, we introduce another critical discount factor, also depending on the ratio of the low to the high wage. Notation:

$$\delta^*_{LH} = \frac{1}{1 + \mu(1 - \rho^{-1})(1 - w_L/w_H)}.$$

It is easy to see that $\delta^* < \delta^*_{LH} < 1$ and in the limit case of $w_L = 0$, $\delta^*_{LH} = \delta^*$.

To avoid absurd results, we assume that the frequency of lower-paid is so high, their wage and discount factor are so low that the socially optimal contribution rate is maximal: $\tau = \bar{\tau}$, yielding $c_L = d_L = w_L/(1 + \mu)$.

We present the generalization of Theorem 2.

Theorem 3. *In a two-type model, we assume a subcritical low discount factors: $0 < \delta_L < \delta^*_{LH}$. Then the socially optimal cap is maximal for subcritical high discount factors and minimal for supercritical high discount factors:*

$$\bar{w}^{\mathrm{o}} = \begin{cases} w_H & \text{if } \delta_H < \delta^*_{LH}; \\ w_L & \text{if } \delta_H > \delta^*_{LH}. \end{cases}$$

Proof. Type H's paternalistic utility function needs some discussion. The cap is only useful if the effective contribution rate $\tilde{\tau}$ yields higher old-age consumption d_H than does $\bar{\tau}$. This is only possible if $s^{\mathrm{o}}_H > 0$. Looking at the formulas:

$$d_{1H}(\tilde{\tau}) = \frac{\delta_H[\rho w_H - (\rho - 1)\bar{\tau} w_L]}{\mu\delta_H + \rho} > \frac{w_H}{1 + \mu} = d_{2H}(\bar{\tau}).$$

Rearranging the inequality,

$$(1 + \mu)\rho w_H \delta_H - \mu(\rho - 1) w_L \delta_H - \mu w_H \delta_H > \rho w_H.$$

Using the definition of δ^*_{LH}, this last inequality is equivalent to the second branch of the optimality condition. In the other case, the opposite holds. ∎

The two-type model is quite illuminating but it gives extreme values; either a too low cap: w_L or a too high cap: w_H. We need to introduce a third, medium type, to obtain a nontrivial socially optimal cap.

A three-type model

We shall work with the following three-type model: low, medium and high types with wages $w_L < w_E < w_H$ and discount factors $0 \approx \delta_L < \delta_E < \delta_H \approx 1$. Now

$$V(\tau, \bar{w}) = f_L d_L + f_E d_E + f_H d_H.$$

The optimal cap does not rise above the medium wage w_E, because then its only function would be punish type H's saving. On the other hand, we know that the optimal cap does not stay below the minimal wage w_L, because it would not relieve any type. Hence the only relevant interval for the optimal cap is $w_L < \bar{w} \leq w_E$.

Now we have the following 'scheme': $v_E = w_E - \tau\bar{w}$ and $\hat{v}_E = (1 - \tau)\bar{w}$, while $v_H = w_H - \tau\bar{w}$ and $\hat{v}_H = (1 - \tau)\bar{w}$. We must ensure that in the relevant interval, $s^{\mathrm{o}}_E < 0$.

At this stage we have only a conjecture rather than a theorem.

Conjecture. *In our three-type model, we have wages $w_L < w_E < w_H$ and discount factors $0 \approx \delta_L < \delta_E < \delta_H \approx 1$. For certain moderate assumptions (most individuals are quite myopic and low-paid), the socially optimal contribution rate is again maximal: $\tau^{\mathrm{o}} = \bar{\tau}$ and the related cap is equal to the medium wage: $\bar{w}^{\mathrm{o}} = w_E$.*

Continuous distributions

Finally we turn to the more realistic *continuous* wage and discount factor distributions. Since typically the higher the earning, the higher is the discount factor, thus we may assume that the joint distribution of (δ, w) is one-dimensional. Let the discount factor $\delta(w)$ be a continuous and monotone increasing function of the wage in the interval $[w_m, w_M]$ with $0 < w_m < w_M \le \infty$, which in turn has a positive density function f and a corresponding distribution function

$$F(w) = \int_{w_m}^{w_M} f(\omega)\, d\omega$$

with $F(w_m) = 0$ and $F(w_M) = 1$.

Furthermore, $\delta_m = \delta(w_m)$ and $\delta_M = \delta(w_M)$, $0 \le \delta_m < \delta_M \le 1$. To avoid trivialities, we make the following two assumptions.

A1. For the lowest wage w_m, the corresponding discount factor δ_m is so low and the value of the density function $f(w_m)$ is so high that the lowest-paid needs to pay full contribution: $\tau^{\mathrm{o}} = \bar{\tau}$.

A2. For the highest wage w_M, the highest discount factor δ_M is so high and the wage ratio w_m/w_M is so low that $\delta_M > \delta^*_{mM}$ (Theorem 3), i.e. even for the maximal contribution rate, there are some room for private saving: $\bar{w}^{\mathrm{o}} < w_M$.

We are looking for conditions which guarantee the existence of at least one cap $\bar{w}$ such that leaves room for private savings. First we define the concept of the critical wage $w^* \in (\bar{w}, w_M)$, above which earners save in addition to their contribution, while below which earners save nothing. Using the saving formula, the critical wage w^* is defined by

$$\delta(w^*)(w^* - \bar{\tau}\bar{w}) - \mu^{-1}\bar{\tau}\bar{w} = 0.$$

Lemma. *If the maximal discount factor δ_M is sufficiently high:*

$$\delta_M(w_M - \bar{\tau} w_m) > \mu^{-1}\bar{\tau} w_m,$$

then there exists a maximal cap $\bar{w}_M$ such that there exists a critical wage for any cap $\bar{w} \in (w_m, \bar{w}_M)$ and it is an increasing function of the cap.

Remarks. 1. Recall that $\bar{\tau} = 1/(1 + \mu^{-1}) = \mu^{-1}\bar{\tau}$. If $\delta_M = 1$, then our condition is simply $w_M > w_m$. Otherwise, w_M/w_m and δ_M together should be large enough to satisfy the condition.

2. For the limiting case of constant discount factor δ, the critical wage is a simple multiple of the cap:

$$w^* = (1 + \mu^{-1}\delta^{-1})\bar{\tau}\bar{w}.$$

Proof. Denote the left hand side of the equation by $g(\cdot)$ which is clearly increasing. Furthermore,

$$g(\bar{w}) = \delta(\bar{w})(1-\bar{\tau})\bar{w} - \mu^{-1}\bar{\tau}\bar{w} < 0$$

and by our condition,

$$g(w_M) > 0.$$

By Bolzano's theorem, there exists a critical w^*.

Replacing the one-variable function $g(w^*)$ by a two-variable function $G(\bar{w}, w^*)$, the implicit function theorem yields

$$w^{*'}(\bar{w}) = -\frac{G_1}{G_2} = \frac{[\delta(w^*) + \mu^{-1}]\bar{\tau}}{\delta'(w^*)(w^* - \bar{\tau}\bar{w}) + \delta(w^*)} > 0,$$

where G_1 and G_2 are the respective partial derivatives. ∎

The main problem of the optimal choice of the cap is that for the maximal contribution rate, raising the cap in the interval $(w_m, \bar{w}_M)$, lower-paid's old-age consumption increases, but higher-paid's old-age consumption decreases. What is the socially optimal balance? To answer this question we need the following notation (to avoid confusion with the differential operator d, we replace here the lower-case letter by its upper-case counterpart):

$$D(w, \bar{w}) = \max[d_1(w, \bar{w}), d_2(w, \bar{w})],$$

where (copying from the single-type model)

$$d_1(w, \bar{w}) = \frac{\delta(w)[\rho w - (\rho - 1)\bar{\tau}\bar{w}]}{\mu\delta(w) + \rho} \quad \text{and} \quad d_2(w, \bar{w}) = \frac{\bar{\tau}\bar{w}}{\mu}.$$

The social welfare function is defined as

$$V(\bar{w}) = \int_{w_m}^{w_M} D(w, \bar{w}) f(w)\, dw.$$

We shall prove

Theorem 4. *Under the assumptions of Lemma, in a continuous-type model with the maximal contribution rate $\bar{\tau}$, the locally optimal cap $\bar{w}^{\mathrm{o}}$ satisfies the following condition:*

$$F(w^{*\mathrm{o}}) - F(\bar{w}^{\mathrm{o}}) = (\rho - 1) \int_{w^{*\mathrm{o}}}^{w_M} \frac{\mu\delta(w)}{\mu\delta(w) + \rho} f(w)\, dw,$$

where

$$\delta(w^{*\mathrm{o}})(w^{*\mathrm{o}} - \bar{\tau}\bar{w}) = \mu^{-1}\bar{\tau}\bar{w}.$$

Remark. Intuitively, our optimality condition states that at a marginal rise of the optimal cap, the marginal gain of subcritical earners ($\bar{w}^{\mathrm{o}} \le w < w^{*\mathrm{o}}$) is equal to the marginal loss of supercritical earners ($w^{*\mathrm{o}} < w < w_M$), those earning below the cap are not affected. It is another question whether there exists such a cap and if there exists, then it is a maximum or not.

Proof. We shall cut the interval $[w_m, w_M]$ into three, therefore

$$V(\bar{w}) = \int_{w_m}^{\bar{w}} d_2(w,\bar{w}) f(w)\, dw + \int_{\bar{w}}^{w^*} d_2(w,\bar{w}) f(w)\, dw + \int_{w^*}^{w_M} d_1(w,\bar{w}) f(w)\, dw.$$

If the social welfare function has a local maximum, then its derivative is equal to zero. Inserting the formulas for d_1 and d_2 into V and omitting the unessential interval $[w_m, \bar{w}]$ and canceling the positive derivative of the upper limit of the first parametric integrals and the negative derivative of the lower limit of the second, provides the necessary condition:

$$V'(\bar{w}) = \int_{\bar{w}}^{w^*} \frac{\bar{\tau} f(w)}{\mu}\, dw - \int_{w^*}^{w_M} \frac{\delta(w)(\rho - 1)\bar{\tau}}{\mu\delta(w) + \rho} f(w)\, dw = 0.$$

Simple transformation yields the optimality condition. ∎

Before presenting a corollary, we make an *additional assumption*: The share $F(w^*(\bar{w})) - F(\bar{w})$ of subcritical earners (who earn above the cap but do not save) decreases with the cap. (We shall see that this assumption holds for any constant discount factor and for our parametric distribution named after Wilfredo Pareto.)

Corollary. *Under the assumptions of Theorem 4 and the additional assumption, the optimal cap is a decreasing function of the private saving efficiency.*

Proof. Surprisingly, the critical earning is independent of the efficiency. Denote the right hand side of the condition of Theorem 4 as

$$R(\rho, \bar{w}) = \int_{w^*}^{w_M} \frac{(\rho - 1)\mu\delta(w)}{\mu\delta(w) + \rho} f(w)\, dw.$$

For a fixed cap, raising ρ, increases the integrand, hence increases R. By our additional assumption, the left hand side, $F(w^*(\bar{w})) - F(\bar{w})$ can only increase if the cap $\bar{w}$ decreases. ∎

Pareto-distribution

For analytical reasons, we use the *Pareto-distribution* (cf. Diamond and Saez, 2011) with density function

$$f(w) = \sigma w_m^{\sigma} w^{-1-\sigma} \quad \text{for} \quad w > w_m,$$

where $\sigma > 1$ is the exponent of the distribution and w_m is the minimum wage. It is easy to give an explicit formula for the distribution function:

$$F(w) = \int_{w_m}^{w} f(\omega)\, d\omega = 1 - w_m^{\sigma} w^{-\sigma} \quad \text{for} \quad w \geq w_m.$$

Hence $F(w_m) = 0$ and $F(\infty) = 1$, and its expectation can explicitly be calculated:

$$\mathbf{E}w = \int_{w_m}^{\infty} w f(w)\, dw = \frac{\sigma w_m}{\sigma - 1}.$$

Since we normalized the expectation as unity, the minimum wage is given as

$$w_m = \frac{\sigma - 1}{\sigma}.$$

In practice, approximately $\sigma = 2$, then $w_m = 1/2$. We also display the variance of the Pareto-distribution:

$$\mathbf{E}w^2 = \frac{\sigma w_m^2}{\sigma - 2} = \frac{(\sigma - 1)^2}{\sigma(\sigma - 2)} \quad \text{for} \quad \sigma > 2.$$

As a detour, we mention that for our distribution function and with a fixed discount factor δ, the additional assumption used in Corollary holds:

$$F(w^*(\bar{w})) - F(\bar{w}) = \{1 - [(1 + \mu^{-1}\delta^{-1})\bar{\tau}]\}\bar{w}^{-\sigma}$$

is a decreasing function of cap $\bar{w}$.

In case of our unlimited distribution, let w_M be the maximal value at which the wage distribution is cut and we represent the wages above w_M by a cleverly chosen w_K. Since the censored expected wage is

$$\mathbf{E}\min(w, w_M) = 1 - \frac{w_m^\sigma w_M^{-\sigma+1}}{\sigma - 1},$$

therefore

$$1 = \int_{w_m}^{w_M} w f(w)\, dw + [1 - F(w_M)]w_K.$$

Hence

$$w_K = \frac{\sigma}{\sigma - 1} w_M.$$

For example, for $\sigma = 2$, the representative highest wage is double of the "maximum": $w_K = 2w_M$.

3. Numerical illustrations

Since our problem is quite involved, therefore we turn to numerical illustrations. When we apply 'very' discrete rather than continuous distributions, then our calculations have only limited numerical use.

Before proceeding, let us recall that unlike others (Feldstein, 1987) we have distinguished the lengths of the working and of the retirement periods. Denoting their ratio by μ rather than 1, we receive more realistic numbers. For example, our socially maximal contribution rate $\bar{\tau} = 1/(1+\mu^{-1})$ drops from 1/2 to 1/3 as we replace 1 by $\mu = 0.5$. On the other hand, if we took into account that the socially optimal discount is less than one, then we could further reduce the contribution rate further, even to 1/4.

We shall choose an annual efficiency of saving $\rho_1 = 1.0233$ for private savings yielding a compound 30-year efficiency factor $\rho = 2$, and search for the socially optimal contribution rate and the cap.

Discrete distribution

We start with discrete distributions. Note that for the single-type distribution, for $\rho = 2$, the critical discount factor $\delta^* = 0.8$ (annually $\delta_1^* = 0.9926$). In the two-type case, the lower limit on $f_H/(1 - f_H)$ of Theorem 2 is quite mild: 0.2, i.e. $f_H > 0.16$

For an illustration, we shall work with three types, with the following frequencies: $f_L = 0.7$, $f_E = 0.25$ and $f_H = 0.05$, when $\delta^*_{EH} = 0.944$. The following 30-year compound discount factors were used: $\delta_L = 0.215$, $\delta_E = 0.545$, and $\delta_H = 0.97$, reducing to annual discount factors $\delta_{1L} = 0.977$, $\delta_{1E} = 0.988$, and $\delta_{1H} = 0.999$. Having found the approximate socially optimal values $\tau^o = 0.333$ (rather than the exact 1/3) and $\bar{w}^o = w_E$, Table 1 displays the gross and net wages, the covered net wages, savings, and the two consumption pairs.

Table 1. *The optimal outcome for 3 types*

Type i	Gross w	Net wage v	Covered $\hat{v}$	Saving s^o	Worker's consumption c^o	Pensioner's consumption d^o
L	0.7	0.467	0.467	0.000	0.467	0.466
E	1.3	0.867	0.867	0.000	0.867	0.866
H	1.7	1.267	0.867	0.073	1.194	1.159

Remark. $\tau^o = 0.333$, $\bar{w}^o = 1.3$.

Table 2 confirms our Conjecture: only type H saves, and apart from numerical rounding-off errors, the low and the medium pairs of intensities are equal: $d_i \approx c_i$, $i = L, E$, while $d_H = \delta_H c_H$ would be assured.

Table 2. *The impact of cap on old-age consumption*

Cap $\bar{w}$	Medium consumption d^o_E	High consumption d^o_H	Social welfare V
1.0	0.666	1.198	0.626
1.1	0.733	1.185	0.641
1.2	0.799	1.172	0.655
1.3	*0.866*	*1.159*	*0.670*
1.4	0.866	1.146	0.668
1.5	0.866	1.133	0.666
1.6	0.866	1.120	0.664
1.7	0.866	1.132	0.666

Remark. See Table 1, $d^o_L = 0.466$.

Note that if δ_{1H} were diminished by 0.003 to 0.996, then the optimal cap would jump from $w_E = 1.3$ to $w_H = 1.7$ and $c_H = 1.214$ and $d_H = 1.077$ would yield a lower utility than the previous optimum.

Pareto distribution

To give a flavor for the behavior of the Pareto distribution, we display the values of the distribution function and the covered expected earnings for selected values of caps, appearing below. Note how fast the probability converges to 1 as the relative value of the cap goes to 4, and how slowly the share of the covered earnings does so. For example, 1.6% of the highest earners still have 12.5% of the total earnings.

Table 3. *Pareto-probabilities and covered earnings for varying caps*

Earning cap $\bar{w}$	Probability $F(\bar{w})$	Share of covered earnings $\mathbf{E}(\bar{w})$
0.707	0.500	0.250
1.0	0.750	0.500
1.5	0.889	0.667
2.0	0.938	0.750
2.5	0.960	0.800
3.0	0.972	0.833
4.0	0.984	0.875

Remark. $\sigma = 2$.

Recall that in our model, the discount factor is an increasing function of the wage: $\delta = \delta(w)$. To map an infinite interval into a finite one, we assume

$$\delta(w) = (\delta_m - \delta_M)e^{\xi(w_m - w)} + \delta_M,$$

where $\xi > 0$ stands for the strength of the dependence. Note that for finite w_M, $\delta(w_M) < \delta_M$, but for high w_M/w_m, the error is small.

We shall divide the interval $[w_m, w_M]$ into $n = 100$ subintervals such a way that the division points w_i form a geometrical sequence: and at integration, the representative points are the geometrical means of the subsequent points: $w_{i+1} = qw_i$ and $z_i = \sqrt{w_i w_{i+1}}$. The mass of the remaining infinite part is $1 - F(w_M) = 0.0001$ (with $w_M = 50$) and the earning $w_K = 100$ represents the average highest wage.

As before, the optimal contribution rate is again maximal: $\tau^{\mathrm{o}} = \bar{\tau}$. According to Table 4, the social welfare function is hardly sensitive to the ratio of cap to average wage in the relevant interval $[1.5, 4]$: the optimum is reached around $\bar{w}^{\mathrm{o}} = 3$ (italicized). Even after multiplying the social welfare function by 10, the variation only appears in the last digits. On the other hand, for a contribution rate slightly below the maximum: $\tau = 0.3$, the dependence is stronger and $\bar{w}^{\mathrm{o}} = 2$.

Table 4. *The welfare impact of the contribution rate and the cap*

Contribution rate τ	Cap $\bar{w}$	10× Social welfare $10U$	Expected saving $\mathbf{E}s$
0.300	1.5	6.194	0.030
	2.0	*6.201*	*0.024*
	2.5	6.190	0.020
	3.0	6.175	0.017
	3.5	6.162	0.015
	4.0	6.152	0.013
0.333	1.5	6.651	0.028
	2.0	6.700	0.022
	2.5	6.716	0.018
	3.0	*6.721*	*0.015*
	3.5	6.719	0.013
	4.0	6.716	0.012

Remark. $\rho = 2$.

With such a flat social welfare function, there is not much practical use in Theorem 4!

Changing the exponent of the wage distribution does not change the results qualitatively, only the cap rises. The influence of the efficiency is visible, its rise diminishes the optimal value of the cap (Corollary to Theorem 4). Table 5 displays the results. The rise of annual private efficiency from 1.02 to 1.04 and then to 1.06 diminishes the *approximate* socially optimal relative value of cap from 3 to 2.5 to 2 (italicized). Using the data from Table 3, our optimal cap covers much higher share of the workers than that of Valdés-Prieto and Schwarzhaupt: 94–97 vs. 80 percent. What is surprising is that at the same time, the expected value of saving diminishes from 0.016 to 0.012 to 0.009. The income effect dominates the substitution effect.

Table 5. *The impact of efficiency on the optimal cap*

Annual efficiency ρ_1	Cap $\bar{w}$	10× Social welfare $10U$	Expected saving $\mathbf{E}s$
1.020	2.5	6.703	0.020
	3.0	*6.709*	*0.017*
	3.5	6.710	0.015
1.040	2.0	6.765	0.014
	2.5	*6.770*	*0.012*
	3.0	6.766	0.010
1.060	1.5	6.797	0.011
	2.0	*6.817*	*0.009*
	2.5	6.812	0.007

Remark. $\tau^{\mathrm{o}} = 0.333$.

4. Conclusions

We have constructed a very simple model of the proportional pension system to analyze the socially optimal contribution rate and especially the contribution cap. We have concentrated on the contradiction between the needs of low-earning myopic and high-earning far-sighted types: the former need a high contribution rate to make up for their low savings; the latter need a low contribution rate to make room for their high and efficient savings. A politically convenient compromise is the introduction of an appropriate cap on the contribution: the well-chosen cap does not diminish the contribution as well as the welfare of the myopes but relieves the far-sighted from a part of the contribution burden.

This model is just the beginning. It neglects very important issues: the heterogeneity of the life spans and of the efficiency of private savings. This may suggest the introduction of progressive pension systems, for example, the proportional part is complemented by a uniform basic benefit. Then the analysis of the progressive personal income tax also comes to the fore. The flexibility of the labor supply and the underreporting of the true labor income are other important issues, have been studied in other papers.

References

Barr, N. and Diamond, P. (2008): *Reforming Pensions: Principles and Policy Choices,* Oxford, Oxford University Press.

Casamatta, G., Cremer, H., Pestieau, P. (2000): "The Political Economy of Social Security", *Scandinavian Journal of Economics 102*, 503-522.

Diamond, P. A. and Orszag, P. (2004): *Saving Social Security: A Balanced Approach,* Washington, D.C., Brookings Institution.

Diamond, P. A. and Saez, E. (2011): "The Case for a Progressive Tax: from Basic Research to Policy Prescriptions", *Journal of Economic Perspectives* 23(4), 165–190.

Docquier, F. (2002): "On the Optimality of Mandatory Pensions in an Economy with Life-cyclers and Myopes", *Journal of Economic Behavior and Organization 47*, 121–140.

Feldstein, M. S. (1987): "Should Social Security Means be Tested?", *Journal of Political Economy 95*, 468–484.

Feldstein, M. (2005): "Structural Reform of Social Security", *Journal of Economic Perspectives 19:1*, 33–55.

Laibson, D. (1997): "Golden Eggs and Hyperbolic Discounting", *Quarterly Journal of Economics 112*, 443-477.

Simonovits, A. (2011): "When are Voluntary Pensions Indifferent?", *Economics Letters 111,* 153–155.

Valdés-Prieto, S. and Schwarzhupt, U. (2011): "Optimal Consumption when Behavioural Biases Vary and the State Errs", CESifo Working Paper 3316.

World Bank (1994): *Averting Old-age Crisis*, Washington, DC, World Bank.

DISCUSSION PAPERS PUBLISHED IN 2012

Judit Karsai: Development of the Hungarian Venture Capital and Private Equity Industry over the Past Two Decades. MT-DP 2012/1

Zsolt Darvas: A Tale of Three Countries: Recovery after Banking Crises. MT-DP 2012/2

Zsombor Z. Méder - András Simonovits - János Vincze: Tax Morale and Tax Evasion: Social Preferences and Bounded Rationality. MT-DP 2012/3

Fertő Imre: Szerződések kikényszeríthetősége a magyar élelmiszerláncban: a kis- és közepes vállalkozások esete. MT-DP 2012/4

Helga Habis and P. Jean-Jacques Herings: Stochastic Bankruptcy Games. MT-DP 2012/5

Štefan Bojnec - Imre Fertő: EU Enlargement and Agro-Food Export Performance on EU Market Segments. MT-DP 2012/6

Judit Markus - Miklos Pinter - Anna Radvanyi: The Shapley value for airport and irrigation games. MT-DP 2012/7

Discussion Papers are available at the website of Institute of Economics Hungarian Academy of Sciences: http://econ.core.hu