

# Socially optimal contribution rate and cap in a proportional (DC) pension system

András Simonovits

Received: date / Revised version: date

**Abstract** In our model, the government operates a mandatory proportional (DC) pension system to substitute for the low life-cycle savings of the lower-paid myopic workers, while maintaining the incentives of the higher-paid far-sighted ones in contributing to the system. The introduction of an appropriate cap on pension contribution (or its base)—excluding the earnings above the cap from the contribution base—raises the optimal contribution rate, helping more the lower-paid myopic workers and reserving enough room for the saving of higher-paid far-sighted ones. The social welfare is almost independent of the cap in a relatively wide interval but the maximal welfare is higher than the capless welfare by 0.3–4.5%.

Keywords: proportional (DC) pensions, contribution rate, contribution cap, maximum for taxable earnings

JEL Numbers: H53, H24

## 1 Introduction

The revenue of any pension system depends primarily on two factors: on the contribution rate and the cap (on the contribution base) above which workers do not pay further contributions. These two factors vary in the three pillars: public, mandatory private and voluntary private in a given country in a given year (World Bank, 1994). Confining our attention to the public system, the contribution rates

---

An earlier version of the present paper was published as a discussion paper, Simonovits (2012). I am indebted to N. Barr, Zs. Cseres-Gergely, G. Kézdi, J. Köllő, G. Kőrösi, M. C. Lovell, Th. Matheson (my discussant at the IIPF Congress in Dresden held in 2012), J. Pál and G. Varga for friendly comments. H. Fehr deserves a special acknowledgment for his steady contribution to developing this paper. I express my gratitude to anonymous referees, especially of the POJE's; who greatly helped to improve the paper. This research has received generous financial support from OTKA K 81483.

---

Institute of Economics, CERS, Hungarian Academy of Sciences, Budapest, Budaörsi út 45, Hungary 1112 (E-mail: [simonovits.andras@krtk.mta.hu](mailto:simonovits.andras@krtk.mta.hu)) · Budapest University of Technology, Budapest, Hungary

and caps also vary across countries and years.<sup>1</sup> The contribution rate affects every worker, while a cap high enough only influences a minority, i.e. the highest-earners. Nevertheless, due to the strong inequality in earnings, the share of total earnings above the cap is much higher than the share of the corresponding workers.

While the literature has devoted a lot of attention to the contribution rate, it has almost totally neglected the cap. To analyze the problem, in the present paper I set up a simple OLG model, where the workers pay pension contributions (up to the cap) yielding a *proportional* (DC) benefit and may also (privately) save for their old age. To reflect the problems caused by restrained labor supply and underreported earnings, it is assumed that the private saving is more efficient than the public one (cf. Barr and Diamond (2008, Chapter 6) on its limited validity). Here a continuum of workers with heterogeneous wages and discount factors (the latter being an increasing function of the former) determine their savings under a credit constraint to maximize their discounted lifetime utility. Anticipating the workers' insufficient savings, the government maximizes a *paternalistic* social welfare function (without discounting the old-age consumption's utility) by choosing the contribution rate and the cap.

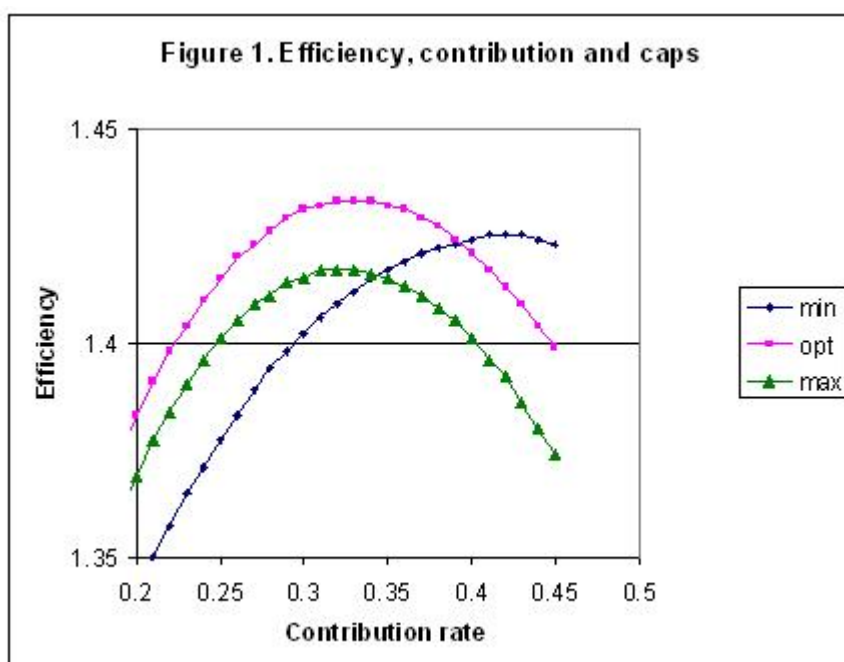
For totally myopic workers, the rise in the contribution rate raises their paternalistic utility, at least until the old-age consumption reaches the young-age consumption at zero saving. We shall call this equalizing rate *myopic*. For totally far-sighted workers, however, any rise in the contribution rate diminishes their welfare. Depending on the distribution of the workers' characteristics and the value of the external parameters like interest factor, the socially optimal contribution rate will be close to or far from the myopic rate. It is also possible that the welfare function first declines, then rises and then again declines. In such a case, the government must *jump the arising welfare gap* when choosing the contribution rate.

Before turning to the second key parameter, we normalize the average wage to unity, giving the cap a standard value (in terms of the average wage). Assuming that the cap is between the minimum and the maximum wages, its existence reduces the *effective* contribution rate—the ratio of the contribution to the earning—of those who earn above the cap. In a proportional system with high enough contribution rate, for the lower-paid short-sighted workers the resulting high effective contribution rate ensures sufficient public pension and the resulting low enough effective contribution rate leaves enough room for the higher-paid far-sighted ones.<sup>2</sup>

The apparently simple model is too complicated to obtain sharp analytic results. I must often rely on numerical illustrations, using Pareto wage distribution (Diamond and Saez, 2011). Since the model is very rudimentary, there is no point to calibrate it. Nevertheless, similarly to Pestieau and Ponthiere (2014), but deviating from others, at least I scaled down the old-age stage rather than identified its length with that of the working stage.

<sup>1</sup> Recently, the US Social Security contribution rate is as low as 12.4%, while in Italy it is as high as 34%. There are a number of causes of this difference, e.g. the higher systemic dependency in Italy and the stronger redistribution in the US, see Disney (2004). Taking the average gross wage as unity, the cap is as low as 1.3 in Sweden and as high as 2.5 in the US, cf. also Table 1 in Valdes-Prieto and Schwarzhaupt (2011). Lovell (2009) gives a useful dataset and analysis of the long and sometimes turbulent history of the US Social Security.

<sup>2</sup> The lowered contribution rate for the self employed can be explained along this line, too, but in fact, it stems from the difficulties of auditing rather than paternalistic differentiation.



I hope that the combination of analytics and numerical illustrations helps the reader understand the message: in addition to the contribution rate, the choice of the cap is also important. Measuring social welfare with respect to the no pension, we shall use its transformation, called *efficiency*, i.e. a scalar by which multiplying every worker's earning, the no pension system gives the same social welfare as the pension system does with the original earnings. Figure 1 illustrates the dependence of efficiency on the contribution rate for three different caps: the minimal cap (diamond, 0.5), the optimal cap (square, here the average wage) and the maximal one which is effectively no cap (triangle). It can be seen that the optimal cap's curve is quite flat around the myopic rate ( $\bar{\tau} = 1/3$ ), which is its optimum argument. The maximal (i.e. no) cap's curve lies definitely below the optimal one, and attains its optimum slightly earlier. It is remarkable that for the minimal cap, the efficiency curve is rising farther and only reaches its maximum around 0.4, but remains below the maximum.

Having outlined the present paper, we turn to the review of the literature. In the empirical literature relatively little explanation has been given on the desirability of a cap. As a rare exception, Barr and Diamond (2008, p. 63) mention two roles for the cap. First, in some countries, the cap only applies to the employee's contribution and to the proportional benefit, therefore from an economic point of view—assuming that the employer's contribution is eventually also paid by the employee—the employer's contribution above the cap is essentially a hidden personal income tax.<sup>3</sup> Similarly, for progressive benefits, the cap limits the otherwise

<sup>3</sup> In Hungary, the former contribution rate is 10% of the gross wage, while the latter is 24%. Until 2013, the uncapped contributions alone provided 6% of the total pension contributions and gave 10% of the *de facto* personal income tax. From 2013, there is no cap at all, decreasing

unbounded redistribution from the higher-paid workers to the lower-paid ones.<sup>4</sup> Second, the government has no mandate to force high old-age consumption on high-earners and in a proportional pension system, removing the cap would further increase the perverse redistribution from the poor to the rich caused by the strong correlation between lifetime earning and life expectancy.

Turning to normative approaches, the seminal papers of Feldstein (1985) and (1987) were the first to discuss pension design with elementary OLG models. We assume that workers with heterogeneous discount factors and a government with a paternalistic social welfare function—originally suggested by Samuelson (1975). Using our vocabulary, we might say that Feldstein (1985) proved the optimality of the myopic rate for totally myopic workers. Calibrating an unrealistically high real interest rate, however, he proved that otherwise at the social optimum, the universal public pension should be eliminated. Next Feldstein (1987) proved that in general the means-tested public system (a loose analogy to the capped system) is socially preferable to a universal one.

Docquire (2002) improved Feldstein’s social welfare function, while Cremer, De Donder, Maldonado and Pestieau (2008) simplified the analysis to a representative generation’s welfare. The latter paper introduced flexible labor supply, wage heterogeneity and progressive pensions but made the distribution of wages and of discount factors independent, moreover, took the efficiencies of public and private savings equal.<sup>5</sup>

At this point it is worth comparing two approximations to the distribution of the wage and discount pairs. It is true that in reality, there are not only low-paid workers who are short-sighted and high-paid ones who are far-sighted (as in the present paper) but also low-paid workers, who are far-sighted and high-paid ones who are short-sighted (as in Cremer et al. 2008). Nevertheless, the correlation between discount factors and wages appears to be strongly positive, therefore the atypical combinations can be neglected in a first approximation. (Becker and Mulligan (1997) provide a rich theory on the endogenous determination of time preferences.) Cremer and Pestieau (2011) is an excellent survey of the field.<sup>6</sup>

Recently, Valdes-Prieto and Schwarzhaupt (2011) analyzed the issue of the optimal coercion including the choice of cap. They considered a larger set of heuristics on remaining life span and future needs than I do; furthermore, modeled various pension systems. They put the value of the optimal cap near the 80th percentile of the earning distribution.

Note that models of this type neglect real-life *dynamic* 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% tax on incomes above the cap

---

the social welfare with respect to the maximum. At the same time, the weak progressivity of the system becomes stronger.

<sup>4</sup> In the US, the progressivity of the Social Security disappears above the cap. In Great Britain of post-WWII the cap was set at the minimum wage, transforming the flat benefit system into a flat contribution one, see also Example 1 below.

<sup>5</sup> One can also argue that their flexible labor supply is replaced by more efficient private saving in the present paper.

<sup>6</sup> Studying lifetime income redistribution, Tenhunen and Tuomala (2010) connected the incomes and discount factors and compared the welfarist and the paternalistic social optima.

to reduce the long-term imbalance of the US Social Security. The problem of time-inconsistency in private savings (e.g. Laibson, 1997 and Findley and Caliendo, 2009) is also out of scope, though the paternalistic government makes up this omission. Most recently, Fehr, Kallweit and Kindermann (2013) analyzed related pension design problems in a dynamic general equilibrium framework.<sup>7</sup>

The structure of the remainder of the paper is as follows. Section 2 presents the OLG model of a proportional pension system with cap. Section 3 displays numerical illustrations, showing the robustness and sensitivity of the qualitative results to certain key parameter values. Finally, Section 4 concludes.

## 2 Model

First we shall outline a framework, and then discuss the two versions of the model, one without the cap, the other with a cap.

### 2.1 Framework

We consider a very simple pension model, where workers only differ in wages and discount factors but their other characteristics do not vary with age. A type can be described by his total wage  $w$  and his discount factor  $\delta$ , the joint distribution function of  $(w, \delta)$  denoted by  $F$  and the corresponding expectations by  $\mathbf{E}$ . Everybody works for a unit period, and everybody spends a shorter (or equal) period in retirement with a common length  $\mu$ ,  $0 < \mu \leq 1$ . Workers pay contributions  $\tau w$  up to  $\tau \bar{w}$ , where  $\tau$  is the *contribution rate* to the mandatory pension system ( $0 < \tau < 1$ ) and  $\bar{w}$  is the *wage ceiling or cap*.<sup>8</sup>

We shall introduce the *effective* contribution rate  $\tilde{\tau}$ , which is the ratio of the contribution to the wage. Distinguishing the wages below and above cap, we have

$$\tilde{\tau} = \begin{cases} \tau & \text{if } w \leq \bar{w} \\ \tau \bar{w}/w < \tau & \text{if } w > \bar{w}. \end{cases} \quad (1)$$

In a proportional system, the pension benefit is equal to the ratio of the contribution  $\tilde{\tau}w$  and the length ratio  $\mu$ :

$$b = \mu^{-1} \tilde{\tau} w. \quad (2)$$

Since the two periods' lengths are different, we use *intensities*, i.e. quantities per a unit time period even if it is not always mentioned.

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

<sup>7</sup> It would be interesting to see how our partial equilibrium result change in their general equilibrium model; for example, if the introduction of the cap raises savings, then the endogenous interest rate diminishes, weakening the advantage of the cap.

<sup>8</sup> The use of various control characters may seem superfluously complicated but in fact this practice helps to understand the notations. For example, writing  $\bar{w}$  rather than an unrelated parameter, e.g.  $\theta$  reminds the reader that it is related to wage  $w$ .

We can now describe the young and the old-age consumption (intensities):<sup>9</sup>

$$c = w - \tilde{\tau}w - s \text{ and } d = \mu^{-1}(\tilde{\tau}w + \rho s). \quad (3)$$

To determine the individually optimal savings, we introduce a discounted lifetime utility function:

$$U(w, \delta, c, d) = u(c) + \mu\delta u(d), \quad (4)$$

where  $u(\cdot)$  is an increasing and concave per-period utility (felicity) function,  $u'(0) = \infty$ ,  $u'(\infty) = 0$ , and  $\delta$  is the discount factor,  $0 \leq \delta \leq 1$ . Inserting (3) into (4), the optimum condition is  $u'(c) = \delta\rho u'(d)$ , the so-called Euler-equation. The transformed equation

$$u'(w - \tilde{\tau}w - s) = \delta\rho u'(\mu^{-1}(\tilde{\tau}w + \rho s)) \quad (5)$$

yields the optimal saving intention  $\hat{s}$ . We must distinguish two cases: either 1) nonnegative saving intention or 2) negative saving intention. The saving intention materializes if it is positive or zero; otherwise it turns to zero:  $s = \max(\hat{s}, 0)$ .

The function of the cap is as follows: lower-paid myopic workers can be locked into a pension system with a high contribution rate, but higher-paid far-sighted ones pay a lower effective rate. Note that there is no reason to set the cap below the minimum wage  $w_m$ , since by multiplying the contribution rate  $\tau$  by a scalar:  $\bar{w}/w_m < 1$ , resulting in  $\tau' = \tau\bar{w}/w_m < \tau$ ,  $(\tau, \bar{w})$  is equivalent to  $(\tau', w_m)$ . Nevertheless, Figure 2 below includes the interval  $[0, w_m]$  as well, to make room for our basic case of no pension at all, when  $\tau > 0$ . Similarly, if there is a finite maximal wage  $w_M$ , then any cap above  $w_M$  can be reduced to  $w_M$  without any impact, i.e. no cap.

The optimal consumption pair are

$$c(\tau, \bar{w}, w, \delta) = w - \tilde{\tau}w - s(\tau, \bar{w}, w, \delta), \quad d(\tau, \bar{w}, w, \delta) = \mu^{-1}[\tilde{\tau}w + \rho s(\tau, \bar{w}, w, \delta)]. \quad (6)$$

To obtain explicit results, Feldstein (1985), (1987) and Simonovits (2012) analyzed two- and three-type models. We shall see that using a discrete wage and discount factor distribution leads to a kinked social welfare function, the analysis of which is cumbersome, even if it yields a closed-form solution. To avoid kinks, we shall generally assume *continuous* wage and discount factor distributions. (Cremer et al. (2008) worked with continuous wage distribution and binary discount factor distribution.)

Let the discount factor  $\delta(w)$  be a monotone increasing continuous function of the wage  $w$  in the interval  $[w_m, w_M]$  with  $0 < w_m < w_M \leq \infty$ . Let the wage distribution have a positive density function  $f$  and a corresponding distribution function  $F(w)$  with  $F(w_m) = 0$  and  $F(w_M) = 1$ . Furthermore,  $\delta_m = \delta(w_m)$  and  $\delta_M = \delta(w_M)$ ,  $0 \leq \delta_m < \delta_M \leq 1$ . The social welfare function is utilitarian:<sup>10</sup>

$$V(\tau, \bar{w}) = \mathbf{E}U^*[\tau, \bar{w}, w] \rightarrow \max.$$

<sup>9</sup> In a median voter model with a progressive pension system, Borck (2007) analytically investigated the important problem when the life expectancy grows with income but at the cost of distorting (3).

<sup>10</sup> Using a Rawlsian social welfare function would defy the purpose of the study, rendering the socially optimal cap indeterminate and the corresponding contribution rate myopic (cf. (10) below)

We shall denote the optimal pair by  $(\tau_C, \bar{w}_C)$ . For a given pair  $(\tau, \bar{w})$ , we shall need the notion of *critical wage*  $w^o = w(\tau, \bar{w})$ , which makes the saving intention 0:  $\hat{s}(\tau, \bar{w}, w^o) = 0$ . If there is no such a wage, i.e. for a high enough contribution rate, the saving intention is negative for any wage, then by definition, the critical wage is identified with the maximal wage plus 1. For another given pair  $(\bar{w}, w)$ , we can also speak of a critical contribution rate  $\tau^o(\bar{w}, w)$ , defined by the implicit equation  $\hat{s}(\tau^o, \bar{w}, w) = 0$ . Note that depending on the value of the contribution rate, we must choose between the two branches of the consumption functions (6):

$$c(\tau, \bar{w}, w) = \begin{cases} (1 - \tilde{\tau})w - s(\tau, \bar{w}, w) & \text{if } \tau < \tau^o; \\ (1 - \tilde{\tau})w & \text{if } \tau \geq \tau^o \end{cases} \quad (7)$$

and

$$d(\tau, \bar{w}, w) = \begin{cases} \mu^{-1}(\tilde{\tau}w + \rho s(\tau, \bar{w}, w)) & \text{if } \tau < \tau^o; \\ \mu^{-1}\tilde{\tau}w & \text{if } \tau \geq \tau^o \end{cases} \quad (8)$$

where  $\tilde{\tau}$  is the effective contribution rate introduced in (1).

To go forward, first we shall consider the simpler, capless system, and then the more complex, capped one.

## 2.2 Capless pension system

If  $\bar{w} = w_M$ , then the government chooses the contribution rate  $\tau$  to maximize the expected value of the paternalistic, undiscounted indirect utility functions, i.e.

$$V[\tau] = \mathbf{E}U^*[\tau, w] \rightarrow \max.$$

We shall denote the (sub)optimal contribution rate by  $\tau_N$  (no cap).

To understand the behavior of  $V[\tau]$ , first we fix an arbitrary wage, and with the fixed wage, we shall study the simpler, second branch  $U_2^*$ . Substitute (7b)–(8b) into the paternalistic utility function:

$$U_2^*[\tau, w] = u((1 - \tau)w) + \mu u(\mu^{-1}\tau w). \quad (9)$$

Take its partial derivative with respect to the contribution rate:

$$U_{2\tau}^*[\tau, w] = -wu'((1 - \tau)w) + \mu u'(\mu^{-1}\tau w).$$

It is easy to see that this becomes zero just when nobody saves and the government equalizes the consumption intensities of the young and of the old:

$$(1 - \tau)w = \mu^{-1}\tau w, \text{ i.e. } \bar{\tau} = \frac{1}{1 + \mu^{-1}}. \quad (10)$$

We shall call this contribution rate *myopic*, since this is the social optimum for totally myopic workers in a capless system. Note, however, that in a system with a low cap, it may be reasonable to choose a contribution rate above the myopic rate (see Figure 1 above)!

We have arrived to

**Lemma 1** *The zero-saving paternalistic utility function  $U_2^*[\tau, w]$  in (9) is increasing in  $[0, \bar{\tau})$  and decreasing in  $(\bar{\tau}, 1)$ .*

Turning to the first branch,

$$U_1^*[\tau, w] = u((1 - \tau)w - s[\tau, w]) + \mu u(\mu^{-1}(\tau w + \rho s[\tau, w])), \quad s[\tau, w] > 0. \quad (11)$$

Take its derivative with respect to  $\tau$ :

$$U_{1\tau}^{*\prime}[\tau, w] = -(w + s'_\tau[\tau, w])u'((1 - \tau)w - s[\tau, w]) + (w + \rho s'_\tau[\tau, w])u'(\mu^{-1}(\tau w + \rho s[\tau, w])).$$

Here  $s'_\tau[\tau, w] < 0$ . For  $c > d$ ,  $u'(c) < u'(d)$  and for  $\rho > 1$ , the absolute value of the multiplier of  $u'(c)$  is lower than that of  $u'(d)$ . Hence  $U_{1\tau}^{*\prime}[\tau, w] < 0$ .

We have arrived to

**Lemma 2** *The paternalistic utility function  $U_1^*[\tau, w]$  in (11) with positive saving is decreasing in  $[0, \tau^\circ[w]]$  and is increasing in  $(\tau^\circ[w], 1)$ .*

Although under normal circumstances,  $\tau^\circ[w] < \bar{\tau}$  holds, complex dependence can arise after taking expectations of  $U^*[\tau, w]$ . For sufficiently myopic workers, the myopic contribution rate would be optimal; for sufficiently far-sighted workers, the zero contribution rate would be optimal.

**Theorem 1** *a) If everybody is totally myopic:  $\delta \equiv 0$ , then the social welfare is an increasing function of the contribution rate in  $[0, \bar{\tau}]$  and decreasing in  $(\bar{\tau}, 1)$ , having the optimum at  $\bar{\tau}$ :  $\tau_N = \bar{\tau}$ . b) If everybody is totally far-sighted:  $\delta \equiv 1$ , then the social welfare is a decreasing function of the contribution rate in  $[0, 1)$ , therefore the social optimum is no-pension:  $\tau_N = 0$ . c) In the remaining cases,  $0 \leq \tau_N \leq \bar{\tau}$ .*

*Remark 1* Feldstein (1985, Section 1) proved similar results for Cobb–Douglas-utility functions.

*Proof* See Lemmas 1 and 2.

In general, there are three domains in the parameter space of the model: 1) where  $\tau_N = \bar{\tau}$ , 2) where  $0 < \tau_N < \bar{\tau}$ , 3) where  $\tau_N = 0$ . In domains 1 and 2, the government must choose a high enough contribution rate to jump the welfare gap arising because the welfare function is decreasing for low enough contribution rates. From now on we assume that this is the case. We cannot determine its condition in general, in terms of the primitive data of the model but we can give a rather trivial condition when the myopic contribution rate provides higher welfare than the zero rate does. Simply substituting into (11) ( $\tau = \bar{\tau}$ ) and (9) ( $\tau = 0$ ), respectively and taking expectations yields

**Theorem 2** *The pension system with myopic contribution rate is welfare superior to the no pension if and only if*

$$(1 + \mu)\mathbf{E}u(w/(1 + \mu)) > \mathbf{E}u(w - s[0, w]) + \mu\mathbf{E}u(\mu^{-1}\rho s[0, w]), \quad (12)$$

where  $s[0, w]$  is the optimal saving function defined in (5) above.



### 2.3 Return to the capped pension system

Having studied the capless system, we return now to the capped system:  $w_m \leq \bar{w} \leq w_M \leq \infty$  and  $w_m < w_M$ . We start with two elementary observations presented as one theorem and one conjecture. The following theorem states that for high enough contribution rates, introducing an appropriate cap, the social welfare can be raised.

**Theorem 3** *Choosing the socially suboptimal contribution rate  $\tau_N$  (the conditional optimum in the capless system), and introducing a cap equaling to the critical wage  $w^\circ[\tau_N]$ , the social welfare is increased:*

$$V(\tau_N, w^\circ[\tau_N]) > V(\tau_N, w_M) = V[\tau_N].$$

*Proof* By definition, the introduction of this cap into the capless system, affects exactly those workers, who earn above the critical wage  $w^\circ[\tau_N]$ , and they would save from their newly found disposable wage more efficiently than they did in the mandatory public system.

*Remark 2* Note that if the contribution rate were lower than the suboptimal value or the cap were lower than the critical wage, then the introduction of the corresponding cap would reduce the paternalistic welfare of those who earn close but below the critical wage. In both cases, complicated calculation would be needed to judge the welfare implications.

Next we formulate a conjecture comparing the optimal contribution rates of the capless and the capped systems and providing an upper bound on the optimal cap.

*Conjecture 1* Assume that the share of well-paid far-sighted workers is sufficiently high that the suboptimal contribution rate in the capless system be lower than the myopic rate:  $0 \leq \tau_N < \bar{\tau}$ . Then the socially optimal contribution rate is higher than the suboptimum:  $\tau_C > \tau_N$  and the optimal cap is lower than the critical wage in the capless system:  $\bar{w}_C < w^\circ[\tau_N]$ .

The easiest way to visualize the conjecture is as follows: raise the cap  $\bar{w}$  continuously from  $w_m$  to  $w_M$  and determine the conditionally optimal contribution rate  $\tau[\bar{w}]$ . One expects that the resulting function  $\tau[\bar{w}]$  is decreasing and before reaching  $\tau_N$  it goes through the unconditionally optimal  $\tau_C$  at the corresponding cap  $\bar{w}_C$ .

Figure 1 above shows, however, that the conditionally optimal contribution rate can be even higher than the myopic one, at least for low enough caps:  $\tau[\bar{w}] > \bar{\tau}$ . In such cases, however, system  $(\tau[\bar{w}], \bar{w})$  can be Pareto-improved with a system of myopic rate and an appropriately raised cap  $w^* = \tau[\bar{w}]\bar{w}/\bar{\tau}$ . In fact, workers earning below  $w^*$  need not pay effective contribution rates above the myopic one, and the others pay the same rate as in the former one.

Making a detour and leaving the realm of the continuous distributions, we outline the simplest case, namely when there are only two types.

*Example 1 (Two types)* Let us assume that there are only two types,  $m$  and  $M$  with wages  $w_m$  and  $w_M$ ,  $w_m < 1 < w_M$  and  $\delta_m = 0$  and  $\delta_M \leq 1$ . Let  $f_m > 0$  and  $f_M > 0$  be the relative frequencies of these types,  $f_m + f_M = 1$ . Then  $s_m = 0$  and

$s_M > 0$  for any  $0 < \tau < \bar{\tau}$ . Denoting the socially (sub)optimal contribution rates in the capless and the capped systems by  $\tau_N$  and  $\tau_C$  with the socially optimal cap  $\bar{w}_C$ , respectively, the optimal contribution rate is higher than the suboptimal one and the optimal cap is equal to the minimal wage:

$$\tau_N < \tau_C < \bar{\tau} \text{ and } \bar{w}_C = w_m.$$

This is a glaring but degenerate example of the claim presented in Conjecture above.

#### 2.4 Cobb–Douglas-utility function

To sharpen our results above and prepare for the numerical calculations, we specialize on a Cobb–Douglas-utility function (cf. Feldstein, 1985):  $u(x) = \log x$ . Then our formulas (e.g. (5)) simplify as follows:

$$\hat{s}[\tau, w] = \frac{\delta(w)\rho(1-\tau) - \mu^{-1}\tau}{[\delta(w) + \mu^{-1}]\rho} w, \quad c(\tau, w, \delta(w)) = \frac{1 - (1 - \rho^{-1})\tau}{1 + \mu\delta(w)} w, \quad w > w[\tau].$$

Hence the critical wage  $w^\circ[\tau]$  is a root to  $\delta(w) = \tau/[\mu\rho(1-\tau)]$ . If there is a cap, then  $\tau$  is replaced by  $\bar{\tau}$ .

For this special case, everything is very simple. Most noteworthy, Theorem 1 can be generalized as follows. For a moment forget about the dependence of the discount factor on wage and introduce the critical discount factor  $\delta^\circ$  for which the paternalistic utility function's two end-values, namely  $V[0]$  and  $V[\bar{\tau}]$  are equal to each other:

$$\log c(0, w, \delta^\circ) + \mu \log d(0, w, \delta^\circ) = \log c(\bar{\tau}, w) + \mu \log d(\bar{\tau}, w).$$

Because the optimal consumptions are proportional to wages, the critical discount factor is independent of the wage:

$$\log c(0, 1, \delta^\circ) + \mu \log d(0, 1, \delta^\circ) = \log c(\bar{\tau}, 1) + \mu \log d(\bar{\tau}, 1).$$

Returning to wage-dependent discount factors, now Theorem 1 on a capless system is transformed into

**Theorem 4** *Assume that the workers have Cobb–Douglas-utility functions with wage-dependent discount factors  $\delta(w)$ s. a) If everybody is sufficiently myopic (subcritical):  $\delta(w_M) < \delta^\circ$ , then the social welfare is an increasing function of the contribution rate in  $[0, \bar{\tau})$  and decreasing in  $(\bar{\tau}, 1)$ , having the optimum at the myopic rate  $\bar{\tau}$ . b) If everybody is sufficiently far-sighted (supercritical):  $\delta(w_m) > \delta^\circ$ , then the social welfare is a decreasing function of the contribution rate in  $[0, 1)$ , therefore the social optimum is no-pension at all. c) If there are both subcritical and supercritical workers:  $\delta(w_m) < \delta^\circ < \delta(w_M)$ , then the optimal contribution rate is between the two extrema:  $0 \leq \tau_N \leq \bar{\tau}$ .*

Similarly, Theorem 2 is transformed into

**Theorem 5** *Under the assumption of Theorem 4, the pension with myopic contribution rate is welfare superior to the no pension if and only if*

$$(1 + \mu) \log(1 + \mu) + \mu \log \rho + \mu \mathbf{E} \log \delta(w) < (1 + \mu) \mathbf{E} \log(1 + \mu\delta(w)). \quad (13)$$

*Proof* Indeed, inserting

$$c(0, w, \delta(w)) = \frac{w}{1 + \mu\delta(w)} \text{ and } d(0, w, \delta(w)) = \delta(w)\rho c(0, w, \delta(w))$$

into (12), simplifies to (13).

We have no analog to Theorem 3.

It is not enough to determine the social optimum, we must evaluate the efficiency gain of having a pension system  $(\tau, \bar{w})$  rather than having no pension at all. We formulate the *relative efficiency*  $\varepsilon$  of system  $(\tau, \bar{w})$  with respect to no pension system as follows:  $V[\varepsilon, 0, 0] = V[1, \tau, \bar{w}]$ .<sup>11</sup> Due to our choice of Cobb–Douglas utility function,  $(1 + \mu) \log \varepsilon$  can be separated in the LHS, i.e.

$$V[1, 0, 0] + (1 + \mu) \log \varepsilon = V[1, \tau, \bar{w}], \text{ i.e. } \varepsilon = \exp \frac{V[1, \tau, \bar{w}] - V[1, 0, 0]}{1 + \mu}.$$

## 2.5 Pareto-distribution

We assume that wages follow a *Pareto-distribution* with a density function

$$f(w) = \sigma w_m^\sigma w^{-1-\sigma} \text{ for } w \geq w_m > 0,$$

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:

$$F(w) = \int_{w_m}^w f(w) dw = 1 - w_m^\sigma w^{-\sigma} \text{ for } w \geq w_m > 0. \quad (14)$$

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

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

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

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

In practice,  $\sigma \approx 2$ , therefore  $w_m \approx 1/2$ . We also display the second moment of the Pareto-distribution:

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

For our unbounded distribution, let  $w_Q$  be the maximal value at which the wage distribution is cut in the numerical illustrations and we represent all the wages above  $w_Q$  by a cleverly chosen  $w_K$ . By definition,

$$1 = \int_{w_m}^{w_Q} wf(w) dw + [1 - F(w_Q)]w_K. \quad (15)$$

<sup>11</sup> Note that for zero contribution rate, the value of the cap is as irrelevant as for zero cap, the value of the contribution rate. Therefore  $V[\varepsilon, \tau, 0] = V[\varepsilon, 0, \bar{w}] = V[\varepsilon, 0, 0]$ .

The expected covered wage (given in the integral in (15)) is equal to

$$\mathbf{E} \min(w, w_Q) = 1 - \frac{w_m^\sigma w_Q^{-\sigma+1}}{\sigma - 1}, \quad (16)$$

hence (14), (15) and (16) yield

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

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

### 3 Numerical illustrations

Even though our framework is very elementary, our problem is quite involved analytically, therefore we turn to numerical illustrations. First we make the necessary preparations, then we display our tables for the capless and the capped systems, respectively.

#### 3.1 Preparations

We have distinguished the lengths of the working and of the retirement periods but we have confined attention to a stationary population and economy! (If we had complicated the model by introducing the growth factors of population  $\nu$  and of wage  $g$ , then we should have calculated with a relative interest factor:  $\rho/(\nu g)$ .) Assuming 40-year working and 20-year retirement periods, the length-ratio is 1/2 rather than 1. This way, we receive more realistic numbers. For example, even in our stationary economy and population, the myopic contribution rate  $\bar{\tau} = 1/(1 + \mu^{-1})$  in (10) drops from 1/2 to 1/3 as we replace  $\mu = 1$  by 0.5.<sup>12</sup> We calculate as if the whole saving and dissaving occurred at the middle points of the working and retirement periods, namely at adult ages 20 and 50 years, respectively.

We start with the tabulation of the critical discount factor as a function of the interest factor. Recall that for workers with subcritical discount factors (where the saving intention is negative), the myopic contribution rate is optimal; while for workers with supercritical discount factors (where the saving intention is non-negative), the zero contribution rate is optimal (cf. Theorem 4). It is qualitatively obvious that the higher the interest factor, the more advantageous is to save, i.e. the lower is the critical discount factor. What Table 1 gives is the quantitative estimation. For example, for the modest annual interest factor of 1.02, the critical annual discount factor is quite high:  $\delta^\circ[1] = 0.952$ , while for the super high annual interest factor of 1.08, the critical annual discount factor is quite low: 0.891.

**Table 1.** *Critical discount factor as a function of interest factor*

|   |       |       |       |       |
|---|-------|-------|-------|-------|
| Annual interest factor $\rho[1]$                  | 1.02  | 1.04  | 1.06  | 1.08  |
| Annual critical discount factor $\delta^\circ[1]$ | 0.952 | 0.928 | 0.909 | 0.891 |

<sup>12</sup> If we took into account that the socially optimal discount factor is less than one (e.g. labor disutility, reduced family size in old-age, etc., as postulated by Cremer et al. (2008)), then we could reduce the contribution rate further, even to 1/4.

Note that for Feldstein's parametrization, where  $\rho[1] = 1.114$ ,  $\delta = 1$ , the socially optimal  $\tau = 0$  yields  $d = \rho c = 25.5c$ , a totally unrealistic ratio.

We shall assume that wages are distributed along a Pareto-distribution. To give a flavor of the behavior of the Pareto-2 distribution, we display selected values of the distribution function and the share of covered earnings. The median wage is about 0.71. Note how fast the probability of being fully covered 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 all the earners still have 12.5% of the total earnings (last row).

**Table 2.** *Pareto-probabilities and covered earnings for varying caps*

| Earning cap<br>$\bar{w}$ | Probability of being<br>fully covered<br>$F(\bar{w})$ | Share of covered<br>earnings<br>$F(\bar{w})\mathbf{E}(w w < \bar{w})$ |
|--------------------------|---|---|
| 0.707 (median)           | 0.500   | 0.250   |
| 1.0 (average)            | 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 the simple relation

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

where  $\xi > 0$  measures the dependence of the discount factor on the wage, shortly and imprecisely, elasticity. Note that for any finite  $w_Q$ ,  $\delta(w_Q) < \delta_M$ , but for high  $w_Q/w_m$ , the error is small. We shall work with  $\delta_m = 0$ ,  $\delta_M = 1$ . For this special choice,  $\delta(w) = 1 - e^{\xi(w_m - w)}$ , i.e.  $\xi = -\delta'(w)/(1 - \delta(w))$  is something of a discount factor–wage elasticity. When fixed, we shall work with  $\xi = 0.2$ .<sup>13</sup>

In the remainder we shall investigate the sensitivity of the social optimum to the parameter values, namely to the wage elasticity of the discounting factor and to the interest factor.

### 3.2 (Sub)optimal contribution rate without cap

We start our numerical investigations with the socially (sub)optimal contribution rate in a capless system. Note that for  $\xi = 0$ , every worker would be totally myopic, therefore by Theorem 1a, nobody would save, thus the myopic contribution rate

<sup>13</sup> We shall divide the interval  $[w_m, w_Q]$  into  $n = 100$  subintervals such a way that the division points  $w_i$  form a geometrical sequence. 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  $1 - F(w_Q) = 0.0001$  with  $w_Q = 50$  and the earning  $w_K = 100$  represents the average highest wage.

would be optimal. Fixing first the annual interest factor as  $\rho[1] = 1.04$ , Table 3 shows that as the discount factor–wage elasticity  $\xi$  rises from 0.1 to 0.7, the socially (sub)optimal contribution rate sinks from 0.330 to 0.235, and the corresponding net replacement rate  $\beta = b/[(1-\tilde{\tau})w]$ —where  $\tilde{\tau}$  stands for the effective contribution rate—drops from 0.985 to 0.614. ( $\tilde{\tau} = 1/3$  and  $\tau_N = 1/3$  for  $\xi = 0$ .) The relative efficiency with respect to no pension drops from 1.74 to 1.05.

**Table 3.** *Socially (sub)optimal contribution rates, changing elasticity*

| Wage elasticity of discount factor<br>$\xi$ | (Sub)optimal contribution rate<br>$\tau_N$ | Net replacement rate<br>$\beta$ | Relative efficiency<br>$\varepsilon$ | Expected saving<br>$\mathbf{E}s_N$ |
|---|--|---------------------------------|--------------------------------------|------------------------------------|
| 0.1   | 0.330                                      | 0.985                           | 1.738                                | 0.007                              |
| 0.2   | 0.324                                      | 0.959                           | 1.417                                | 0.015                              |
| 0.3   | 0.312                                      | 0.907                           | 1.269                                | 0.024                              |
| 0.4   | 0.299                                      | 0.853                           | 1.181                                | 0.033                              |
| 0.5   | 0.282                                      | 0.786                           | 1.122                                | 0.042                              |
| 0.6   | 0.263                                      | 0.714                           | 1.080                                | 0.053                              |
| 0.7   | 0.235                                      | 0.614                           | 1.049                                | 0.066                              |

Remark.  $\rho[1] = 1.04$ . In Table 4, we discuss the sensitivity of the outcomes to the

interest factor for a fixed elasticity  $\xi = 0.2$ . As the annual interest factor  $\rho[1]$  rises from 1.02 to 1.08, the socially (sub)optimal contribution rate  $\tau_N$  drops from 0.332 to 0.198 and the corresponding net replacement rate sinks from 0.994 to 0.498. Note, however, that higher interest factors are unrealistic and produce too high consumption ratio  $d/c = \rho$  for  $w \approx w_M$ . The relative efficiency drops from 1.72 to 1.03.

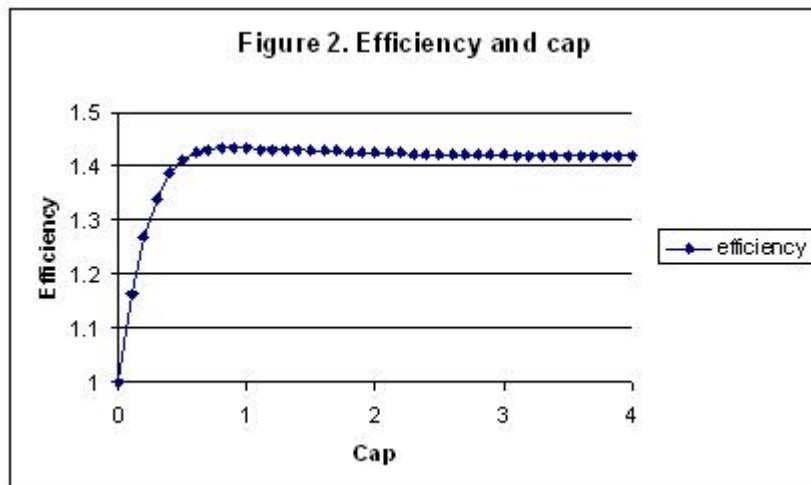
**Table 4.** *The socially (sub)optimal rates, changing interest factor*

| Annual interest factor | (Sub)optimal contribution rate | Net replacement rate | Relative efficiency | Expected saving |
|------------------------|--------------------------------|----------------------|---------------------|-----------------|
| $\rho[1]$              | $\tau_N$                       | $\beta$              | $\varepsilon$       | $\mathbf{E}s_N$ |
| 1.02                   | 0.332                          | 0.994                | 1.716               | 0.007           |
| 1.03                   | 0.329                          | 0.981                | 1.557               | 0.011           |
| 1.04                   | 0.324                          | 0.959                | 1.416               | 0.017           |
| 1.05                   | 0.312                          | 0.907                | 1.291               | 0.023           |
| 1.06                   | 0.291                          | 0.821                | 1.182               | 0.030           |
| 1.07                   | 0.261                          | 0.706                | 1.087               | 0.039           |
| 1.08                   | 0.198                          | 0.494                | 1.030               | 0.051           |

Remark.  $\xi = 0.2$ .

### 3.3 Optimal contribution rate and cap

Extending the calculations on the capless system to the capped system and anticipating Tables 5 and 6, we present Figure 2. This displays the dependency of the relative efficiency  $\varepsilon$  on the cap with the approximate constant optimal contribution rate  $\tau_C = 0.33$ ; for  $\rho[1] = 1.04$  and  $\xi = 0.2$ . Note that in this example, we allow for caps below the minimum wage, when every worker pays the same contribution. That way we can present the no pension even at the myopic contribution rate. The relative efficiency quickly rises from 1 at zero cap (no pension) till 1.41 at the minimum wage  $\bar{w} = 0.5$ , it reaches its maximum 1.433 at the average wage and then slowly declines to 1.417. Note that the numerical difference between the capped and the capless systems does not appear to be significant, but—as has been mentioned in the Introduction—this is typical in such calculations.



We repeat the sensitivity analysis, and the impacts on the optima rather than suboptima of elasticity and of interest factor are displayed in Tables 5 and 6, respectively. Comparing Table 5 to Table 3, our conjecture is “confirmed”, at least in this numerical case. When  $\xi$  rises from 0.1 to 0.7, the socially optimal contribution rate stays close to the myopic one rather than dropping, while the corresponding cap drops from 1.46 to 0.5. This is the quantitative reflection of our qualitative idea: the higher the share of far-sighted workers in the population, the less is the need for mandatory saving; but to defend the remaining myopes, the statutory contribution rate remains high. Here we also display the relative efficiency of the capped system: it sinks from 1.74 to 1.09.

**Table 5.** *The impact of elasticity on the socially optimal cap*

| Wage elasticity of discount factor<br>$\xi$ | Contribution rate<br>$\tau$ | Optimal cap<br>$\bar{w}_C$ | Relative efficiency<br>$\varepsilon$ | Expected saving<br>$\mathbf{E}s_C$ |
|---|-----------------------------|----------------------------|--------------------------------------|------------------------------------|
| 0.1   | 0.333                       | 1.46                       | 1.745                                | 0.028                              |
| 0.2   | 0.333                       | 0.92                       | 1.433                                | 0.048                              |
| 0.3   | 0.333                       | 0.70                       | 1.294                                | 0.065                              |
| 0.4   | 0.333                       | 0.61                       | 1.212                                | 0.079                              |
| 0.5   | 0.335                       | 0.55                       | 1.158                                | 0.090                              |
| 0.6   | 0.330                       | 0.51                       | 1.119                                | 0.100                              |
| 0.7   | 0.321                       | 0.50                       | 1.090                                | 0.109                              |

Remark

$\rho[1] = 1.04$ .

Finally, we check the impact of the annual interest factor  $\rho[1]$  for a fixed elasticity, again  $\xi = 0.2$ . As the interest factor rises from 1.02 to 1.08 in Table 6, the socially optimal contribution rate drops slowly but then suddenly from 0.333 to 0.30. The decrease in the corresponding cap starts immediately: from 1.72 to 1.19 and then to the minimum wage: 0.5. This is the quantitative reflection of another of our qualitative ideas: the more efficient the private saving, the less is the need for mandatory saving. Unlike Table 5, in this case, the contribution rate and the cap drop together. The efficiency of the capped system declines from 1.72 to 1.08.

**Table 6.** *The impact of interest factor on the socially optimal cap*

| Annual interest factor<br>$\rho$ | Contribution rate<br>$\tau$ | Optimal cap<br>$\bar{w}_C$ | Relative efficiency<br>$\varepsilon$ | Expected saving<br>$\mathbf{E}s_C$ |
|----------------------------------|-----------------------------|----------------------------|--------------------------------------|------------------------------------|
| 1.02                             | 0.333                       | 1.72                       | 1.720                                | 0.033                              |
| 1.03                             | 0.333                       | 1.19                       | 1.566                                | 0.041                              |
| 1.04                             | 0.333                       | 0.92                       | 1.433                                | 0.048                              |
| 1.05                             | 0.333                       | 0.72                       | 1.320                                | 0.055                              |
| 1.06                             | 0.332                       | 0.59                       | 1.224                                | 0.060                              |
| 1.07                             | 0.330                       | 0.51                       | 1.145                                | 0.064                              |
| 1.08                             | 0.300                       | 0.50                       | 1.080                                | 0.067                              |

Remark.

$\xi = 0.2$ .



It is an unfortunate feature of our numerical experiments that either the capless suboptimal contribution rate is almost myopic or the optimal cap is close to the minimum wage. To display the full power of our Conjecture we would need an example, where the capless suboptimum is far from the myopic rate:  $\tau_N \ll \bar{\tau}$  and the optimal cap is far from the minimum wage:  $\bar{w}_C \gg w_m$ .

We have not yet commented the changes in the expected saving as a result of imposing a cap (last columns of Tables 3 vs. 5 and Tables 4 vs. 6). Typically the socially optimal expected saving is higher than the suboptimal one:  $\mathbf{E}s_C > \mathbf{E}s_N$ .

In summary, some results (especially on the optimal contribution rate) are quite robust, while other results (notably on the socially optimal cap) are very sensitive to the key parameters of the model, namely to the discount factor–wage elasticity ( $\xi$ ) and the interest factor ( $\rho$ ).

#### 4 Conclusions

We have constructed a very simple model of the proportional pension system to analyze the impact of the socially optimal contribution rate and especially of the contribution (base) cap on the social welfare (and private savings). We have concentrated on the contradiction between the needs of low-earning and myopic workers and of high-earning and far-sighted ones: the former need a high contribution rate to make up for their low saving intentions; the latter need a low contribution rate to make room for their high and efficient saving. For low enough wage-elasticities of discount factor and interest factors, the socially optimal contribution rate in a capless system is close to the myopic one, approaching the old-age consumption of the lower-paid shorter-sighted to their young-age one. For other cases, the scissors between young-age and old-age consumptions of the high-earners grow wide. A politically convenient compromise is the introduction of an appropriate cap on the contribution: a well-chosen cap does not diminish the contribution as well as the utility 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 interest factors. In fact, the expected life span and the interest factor rise with lifetime wages (cf. Borck, 2007). 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, which have been studied with other, related elementary models (Cremer et al., (2008) and Simonovits (2009)).

#### References

1. Barr, N. and Diamond, P. (2008): *Reforming Pensions: Principles and Policy Choices*, Oxford, Oxford University Press.
2. Becker, G. and Mulligan, C. B. (1997): “On the Optimal Determination of Time Preferences”, *Quarterly Journal of Economics* 112, 729–758.
3. Borck, R. (2007): “On the Choice of Public Pensions when Income and Life Expectancy are Correlated”, *Journal of Public Economic Theory* 9, 711–725.

4. Cremer, H.; De Donder, Ph.; Maldonado, D. and Pestieau, P. (2008): “Designing a Linear Pension Scheme with Forced Savings and Wage Heterogeneity”, *International Taxation and Public Finance* 15, 547–562.
5. Cremer, H. and Pestieau, P. (2011): “Myopia, Redistribution and Pensions”, *European Economic Review* 55, 165–175.
6. Diamond, P. A. and Orszag, P. (2004): *Saving Social Security: A Balanced Approach*, Washington, D.C., Brookings Institution.
7. 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.
8. Disney, R. (2004) “Are Contributions to Public Pension Programmes a Tax on Employment?” *Economic Policy* 39, 267–311.
9. 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.
10. Fehr, H.; Kallweit, M. and Kindermann, F. (2013): “Should Pensions be Progressive?” *European Economic Review* 63, 94–116.
11. Feldstein, M. S. (1985): “The Optimal Level of Social Benefits”, *Quarterly Journal of Economics* 100, 302–320.
12. Feldstein, M. S. (1987): “Should Social Security Means be Tested?”, *Journal of Political Economy* 95, 468–484.
13. Feldstein, M. (2005): “Structural Reform of Social Security”, *Journal of Economic Perspectives* 19:1, 33–55.
14. Findley, T. S. and Caliendo, F. N. (2009): Short Horizons, Time Inconsistency, and the Optimal Social Security, *International Journal of Public Finance*, 16, 487–513.
15. Laibson, D. (1997): “Golden Eggs and Hyperbolic Discounting”, *Quarterly Journal of Economics* 112, 443–477.
16. Lovell, M. C. (2009) “Social Security’s Five OASI Inflation Indexing Problems”, *Economics: The Open Access, Open Assessed E-Journal*, 140.
17. Pestieau, P. and Ponthiere, G. (2014): “Policy Implications for Changing Longevity”, *CESifo Economic Studies*, 60, 178–212.
18. Samuelson, P. A. (1975): “Optimum Social Security in a Life-Cycle Growth Model”, *International Economic Review* 16 539–544.
19. Simonovits, A. (2009): “Underreported Earning and Age-Specific Income Redistribution in Post-Socialist Economies”, IE-HAS Discussion Papers 27.
20. Simonovits, A. (2011): “When are Voluntary Pensions Indifferent?”, *Economics Letters* 111, 153–155.
21. Simonovits, A. (2012): “Optimal Cap on Pension Contributions”, IE-CERS–HAS Discussion Papers 8. Budapest.
22. Tenhunen, S. and Tuomala, M. (2010): “On Optimal Lifetime Redistribution Policy”, *Journal of Public Economic Theory* 12, 171–198.
23. Valdes-Prieto, S. and Schwarzhupt, U. (2011): “Optimal Compulsion when Behavioural Biases Vary and the State Errs”, CESifo Working Paper 3316.
24. World Bank (1994): *Averting Old-age Crisis*, Washington, DC, World Bank.