MAGYAR TUDOMÁNYOS AKADÉMIA
Közgazdaság- és Regionális Tudományi Kutatóközpont
MTA KRTK KTI
Centre for Economic and Regional Studies
HUNGARIAN ACADEMY OF SCIENCES

MŰHELYTANULMÁNYOK DISCUSSION PAPERS

MT-DP – 2013/24

A family of simple paternalistic transfer models

ANDRÁS SIMONOVITS

INSTITUTE OF ECONOMICS, CENTRE FOR ECONOMIC AND REGIONAL STUDIES,
HUNGARIAN ACADEMY OF SCIENCES BUDAPEST, 2013

Discussion papers
MT-DP – 2013/24

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

KTI/IE Discussion Papers are circulated to promote discussion and provoque comments.
Any references to discussion papers should clearly state that the paper is preliminary.
Materials published in this series may subject to further publication.

A family of simple paternalistic transfer models

Author:

András Simonovits
research advisor
Institute of Economics
Centre for Economic and Regional Studies
Hungarian Academy of Sciences
also Mathematical Institute, Budapest University of Technology,
and Department of Economics, CEU
email:simonovits.andras@krtk.mta.hu

August 2013

ISBN 978-615-5243-81-3
ISSN 1785 377X

A family of simple paternalistic transfer models

András Simonovits

Abstract

A general framework is analyzed which contains several special transfer (tax and pension) models. In our static two-overlapping-generation framework, every individual works in the first stage of the adult age, while is retired in the second. The government operates a balanced linear transfer system, sometimes with caps. In the models, the individuals may optimize their situation in various ways: contributing to voluntary pension, restraining labor supply and underreporting wages. Individuals are typically short-sighted, therefore they choose paternalistically suboptimal decisions. The models provide useful information on the socially optimal paternalistic transfer system.

Keywords: tax systems, pension systems, pension models, overlapping generations, paternalism

JEL classification: H24, I31, J22, J26

Acknowledgements:

I am indebted to the support of the Hungarian Science Research Foundation's project K 67853. I express my general gratitude to the help of those who helped me writing the papers underlying the present paper. Erzsébet Kovács and Balázs Muraközy's useful comments are also acknowledged.

Egyszerű paternalista transzfermodellek családja

Simonovits András

Összefoglaló

Olyan általános keretet elemzünk, amely néhány speciális transzfer (adó- és nyugdíj-) modellt tartalmaz. Statikus, két együtt élő nemzedékes keretünkben minden egyén dolgozik az első időszakban, és nyugalomban van a másodikban. A kormányzat kiegyensúlyozott lineáris transzferrendszert működtet, néha plafonokkal. Modelljeinkben az egyének helyzetüket különféleképp optimalizálhatják: részt vesznek egy önkéntes nyugdíjrendszerben, visszafogják munkakínálatuk és eltitkolják keresetük egy részét. Az egyének rendszerint rövidlátók, ezért paternalista szempontból szuboptimális döntéseket hoznak. A modellek hasznos információkat adnak a társadalmilag optimális paternalista transzferrendszerekről.

Tárgyszavak: adórendszerek, nyugdíjrendszerek, nyugdíjmodellek, együtt élő nemzedékek, paternalizmus

JEL kód: H24, I31, J22, J26

A family of simple paternalistic transfer models*

by
András Simonovits

Institute of Economics, CERS, Hungarian Academy of Sciences
also Mathematical Institute, Budapest University of Technology,
and Department of Economics, CEU
Budapest, Budaörsi út 45, Hungary
simonovits.andras@krtk.mta.hu
July 20, 2013

Abstract

A general framework is analyzed which contains several special transfer (tax and pension) models. In our static *two-overlapping-generation* framework, every individual works in the first stage of the adult age, while is retired in the second. The government operates a balanced linear transfer system, sometimes with caps. In the models, the individuals may optimize their situation in various ways: contributing to voluntary pension, restraining labor supply and underreporting wages. Individuals are typically short-sighted, therefore they choose paternalistically suboptimal decisions. The models provide useful information on the socially optimal paternalistic transfer system.

Keywords: tax systems, pension systems, pension models, overlapping generations, paternalism

JEL numbers: H24, I31, J22, J26

* I am indebted to the support of the Hungarian Science Research Foundation's project K 67853. I express my general gratitude to the help of those who helped me writing the papers underlying the present paper. Erzsébet Kovács and Balázs Muraközy's useful comments are also acknowledged.

1. Introduction

I have recently analyzed several simple transfer (i.e. tax-and-pension) models (sometimes written with coauthors) almost independently of each other. Simple means here utmost simplicity: everything is neglected what can be neglected. Having completed their analyses, the joint features of the models became much clearer. (i) Though they follow the logic of the Overlapping Generations model (cf. Weil, 2008), they are basically static, i.e. they only consider a population consisting of people born in the same time period (year). (ii) They assume away the changes in wages and pensions occurring during the two life stages. (iii) The population is heterogeneous, the types' earnings or discount factors are different. (iv) The rules of the transfer system strongly influence individual behavior, based on maximization of discounted utility: labor supply, saving at young age, report on earnings etc. (v) The *paternalistic* government determines the transfer rules by maximizing an *undiscounted* social welfare function.[1]

Of course, these models were not the first ones of this type. Nevertheless, I give a survey of them. For reference points, I single out three such models which preceded mines: Feldstein (1985), (1987) (twins) and Cremer, De Donder, Maldonaldo and Pestieau (2008).[2] I only mention few related but complex models which have dynamic and detailed age-structures: Auerbach and Kotlikoff (1987), Fehr and Habermann (2008), Sefton, van de Ven and Weale (2008).[3] It is of interest that the complex models generally neglect myopia (for an exception, Fehr and Kindermann, 2010). Note that all these models overlooked the complex formulation of *mechanism design* invented by Mirrlees (1971) and treated in an introductory way by Diamond (2003) in a two-generation framework.

Table 1 surveys the types of assumptions of the foregoing models. The order of discussion follows the order of complexity. For the sake of clarity, I will refer to my models and the underlying assumptions by keywords. To avoid any confusion, however, I give here their resolutions.

I start with the models' *abbreviations*. Voluntary pension: Simonovits (2009a) and (2011b), cap on contributions: Simonovits (2012a) and (2013), labor supply: Simonovits (2012d), wage reporting: Simonovits (2009b), pension credit: Simonovits (2012c). Continuing with the *assumptions*; labor supply: labor supply is flexible; wage report: the reported earnings are endogenous; saving: there exist positive private life-cycle savings (with or without tax favors), nonlinearity: mostly piecewise linearity; tax: in addition to pension contribution, there is a personal income tax. Sign + stands for a property present or endogenous; sign – stands for a property missing or exogenous. (In contrast to some authors, negative savings are excluded!)

---

[1] For the *pros* and *cons* of paternalism, see e.g. Sunstein and Thaler (2003) and Epstein (2007), respectively.

[2] Cremer and Pestieau (2011) surveyed the whole field.

[3] Andersen (2012) is an interesting crossover being a simple dynamic model, investigating the question of more saving or later retirement is better against population aging. Varian (1980) is a simple static model covering the issue of social insurance.

Table 1. *Models and assumptions*

Assumptions / Models	Labor supply	Wage report	Saving	Nonlin-earity	Tax
Feldstein	–	–	+	+	–
Cremer et al.	+	–	+	–	–
Voluntary pension	–	–	+	+	+
Cap on contribution	–	–	+	+	–
Labor supply	+	–	–	–	+
Report	–	+	+	–	–
Pension credit	+	–	+	+	–

At surveying the results, I pay little attention to the specific issues, rather I concentrate on the similarities. Nevertheless, it is worth summarizing in a nutshell the messages of the specific models and sometimes illustrating their weights by Hungarian or international data.

a) The *voluntary pension* system complements or supplements a mandatory one. Taking into account the arising tax expenditures, the voluntary system is not as attractive as is frequently depicted. Among others, Engen, Gale and Scholz (1996); Baily and Kirkegard (2009, 452–461) emphasize the illusory effects of savings incentives.[4]

b) The *cap on the pension contribution base* is quite neglected in the literature. Barr and Diamond (2008, p. 63) list two functions: (i) In a proportional system, the government has no mandate to force high old-age consumption on high-earners. (ii) In a progressive system, the cap limits the redistribution from the rich to the poor. Others discuss a third function: to limit the perverse redistribution from the poor to the rich caused by the strong correlation between lifetime earning and life expectancy. We add a fourth one: the cap reduces the effective contribution rate of those who earn above the cap and thus making room for the presumably more efficient private savings.[5]

c) Designing an optimal transfer system, the *elasticity of labor supply* plays a key role. First, it limits the scope of redistribution; second, at least in our model, ensuring old-age living standard has a priority over the earning redistribution. (Disney (2004) thoroughly analyzes the redistribution in various public pension systems.) Nevertheless, the impact of elasticity of labor supply is a hotly debated issue. We only mention that the Czech Republic, which is quite similar to Hungary in many respects, is able to operate a strongly redistributive pension system without having significant other pillars (see also underreported earnings below).

[4] For example, in Hungary, in 2009 the voluntary contributions—mostly paid by the well-to-do—amounted only to 0.3 percent of the GDP but requested a 50 percent matching rate. (For comparison, the mandatory pension expenditures were about 10 percent of the GDP in 2009.)

[5] The example of Hungary is illuminating: Between 1992 and 2012 the relative value of the cap oscillated between a high 3 and a low 1.6. Since 2013, there is no cap and the benefit has become unlimited as well, creating unwanted tensions and reducing the long-run welfare around 0.7 percent in our model.

d) In the *wage reporting* model, the unobservable *tax morale* explains the variation of reported income regardless of the tax rate and the probability of audits and fees (Frey and Weck-Hannemann, 1984).[6] Our discussion extends the validity of the tax model (e.g. Simonovits, 2011a) to the pension system.[7] The basic insight is intuitive: the higher the tax morale, the more redistributive the socially optimal transfer system.

e) Rather than choosing between two rigid schemes, namely the flat and the means-tested ones, a lot of countries have been applying the *pension credit* system. Sefton et al. model the corresponding British system. For example, raising the value of the basic benefits and withdrawing 50 percent of the proportional benefits from the basic pension, the poorest pensioners' standard of living could be increased without burdening the budget with a uniform basic pension.

Finally we must mention the missing topic of *mandatory private* pension systems. For a number of works (e.g. World Bank, 1994), this is *the most important* issue of the pension reforms. Though the early work of Barr (1979) had already exposed the theoretical flaws of this approach, only the recent international financial crisis demonstrated the practical problems of such a system and of the transition leading to it.[8]

Having listed these models, a natural question arises: Why do I *not* unify all these simple models into one? My answer is as follows: The unification would only complicate the model without bringing us too much closer to the reality; dynamics would still be missing, the age-dependence of earnings and of pensions would still be neglected.

The negation of dynamics is especially hurting: a) the contradiction between population aging and productivity rise is left out; b) the adverse impact of erratically changing government policies (e.g. in Hungary as well as in Great Britain) on the incentive schemes is overlooked; and c) it is impossible to model the transition from one system into another within our two-generation framework. Finally we allude to an alternative of the welfare approach, namely the political economy (cf. Casamatta, Cremer and Pestieau, 2000).

The structure of the remaining part is as follows: Section 2 analyzes three introductory examples: (i) the limits on redistribution with flexible labor supply without old-age consumption; (ii) the relation of the insufficient private saving and the mandatory pension system of a representative agent; (iii) the practical impossibility of full intra- and intergenerational redistribution. Section 3 presents the general framework. Section 4 outlines the specific pension models mentioned above. Section 5 returns to the basic specifications in the original papers: discounted Leontief utility function and the continuous Pareto earnings distribution, applied to the calculation of the optimal contribution rate and of the cap; and the pension credit. (An additional gain is that comparing Sections 4 and 5 underlines the excessive sensitivity of our results to apparently unessential details.) Section 6 draws the conclusions. At the end of the paper a list contains the notations.

---

[6] An alternative approach is the elasticity of taxable income, for the latest survey, see Saez, Slemrod and Giertz (2012).

[7] As a *detour*, let us call the reader's attention to two related works: Garay, Simonovits and Tóth (2011) analyzed the dynamics of tax declaration with the traditional methods, while Méder, Simonovits and Vincze (2012) did the same with the new tools of agent-based modeling.

[8] As an afterthought, I mention van Groezen, Leers and Meijdam (2003) whose static protomodel considers the problem of optimal saving and fertility in a similar framework.

2. Introductory examples

As an introduction into models of the transfer systems, we consider the three *simplest* paternalistic models: one for income redistribution, another for pension and a third for their unification.

2.1. Income redistribution with flexible labor supply

We consider a static model (cf. Simonovits, 2012b) for income redistribution with more than one type. To make the model as simple as possible, we confine our attention to the personal income tax and neglect any other tax or pension. Let w be the *total wage* (or full compensation) *rate* of a certain type of workers, T be his annual (or life) time limit, l be his corresponding labor supply, $0 < l < T$, then wl is his earning. Let θ be the personal income tax rate (representing all other taxes as well), $0 \leq \theta \leq 1$ and γ be the basic income. His consumption is equal to

$$c = \gamma + (1 - \theta)wl.$$

We turn to the determination of the optimal labor supply. The simplest utility function is copied from Doerrenberg, Duncan, Fuest and Peichl (2012):

$$U(w, c, l) = c + \frac{1}{2}wl^2.$$

Substituting the formula of the consumption into the utility function, we arrive at the *reduced* utility function with parameter w and variable l:

$$U^*(w, l) = \gamma + (1 - \theta)wl - \frac{1}{2}wl^2.$$

Take the value of the basic income as given. Due to the strict concavity of U^*, the optimal labor supply is determined as a root of the marginal utility–labor supply function:

$$0 = U_l^{*'}(w, l) = (1 - \theta)w - wl.$$

We have arrived to

Lemma 1. *The optimal labor supply and the utility function are equal to*

$$l^* = 1 - \theta \quad \text{and} \quad U^*(w, l^*) = \gamma + \frac{1}{2}w(1 - \theta)^2.$$

Ceteris paribus, the optimal labor supply is independent of the wage rate and of the basic income and is a decreasing function of the tax rate.

Having determined the individual optima, we turn to the macrorelations. Denoting expectation over the wage distribution by $\mathbf{E}$, and choosing an appropriate unit of measurement, the average wage rate can be taken unity: $\mathbf{E}(w) = 1$. Then the tax balance is given by

$$\gamma = \theta\mathbf{E}(wl^*) = \theta(1 - \theta).$$

To determine the socially optimal tax system, we shall use the simplest social welfare function, namely the Rawlsian one which is the minimum of the individual utility functions: $V(\theta) = \min_w U^*(w, l^*)$. Denoting the lowest wage rate by $w_L \in (0, 1)$, it is easy to see that

$$V(\theta) = \theta(1-\theta) + \frac{1}{2}w_L(1-\theta)^2 = \frac{1}{2}w_L + (1-w_L)\theta - \left(1 - \frac{1}{2}w_L\right)\theta^2.$$

Since the socially optimal tax rate satisfies the necessary first-order condition: $0 = V'(\theta^*) = 1 - w_L - \theta(2 - w_L)$, we have

Theorem 1. *The Rawlsian optimal tax rate is equal to*

$$\theta^* = \frac{1-w_L}{2-w_L} < \frac{1}{2}.$$

It is not enough to determine the social optimum, we must evaluate the efficiency gain having an optimal tax system with respect to having a suboptimal one or no tax system. Since the numerical value of V has no direct meaning, we define the *relative efficiency* e of θ^* with respect to $\theta = 0$ as follows. Multiplying the earnings by a positive scalar e such that the social welfare of the no tax system is the same as the social welfare of the optimal tax with the original wages. Denoting the dependence of the welfare on the (average) wage e, we have the following definition for efficiency:

$$V[e, 0] = V[1, \theta^*].$$

By simple calculations,

$$V[e, 0] = \frac{1}{2}ew_L \quad \text{and} \quad V[1, \theta^*] = \frac{1}{2(2-w_L)},$$

hence the relative efficiency is equal to

$$e = \frac{1}{w_L(2-w_L)} > 1.$$

This is as expected: the lower the minimal wage rate, the higher the relative efficiency of the optimal tax system to the no tax.

We have paid a high price for getting an explicit formula: (i) the utilitarian social welfare function $\mathbf{E}U^*$ would yield zero optimal tax rate; (ii) total redistribution ($\theta = 1$) would imply zero labor supply ($l^* = 0$); (iii) in the two-period model of the next subsection, the linear consumption utility would make perfect substitution between the young- and the old-age consumption, undermining the pension problem. As a way out, in Section 4, we shall consider the appropriate generalization of the Cobb–Douglas-utility functions of the form

$$U(w, c, l) = \log c + \xi \log[w(T - l)],$$

where $T \geq 1$ is the time limit and $\xi \geq 0$ is the relative labor disutility coefficient.

2.2. Paternalistic pension system

Following Feldstein (1985, first case), the whole society is represented by a single individual, living for two periods. Let c and d be the young- and old-age consumption, respectively, and δ be the discount factor, $0 \leq \delta \leq 1$. The individual has a Cobb–Douglas lifetime utility function

$$U(c, d) = \log c + \delta \log d.$$

The individual only works in the first period, his earning is 1 and saves s. In the second period he enjoys the fruits of his savings, ρs, where $\rho \geq 1$ is the interest factor. Denoting the dependence on the discount factor by a subindex δ, the reduced utility function is

$$U^*_\delta(s) = \log(1 - s) + \delta \log(\rho s).$$

To maximize this utility, he chooses such a saving at which the marginal lifetime utility (denoted by prime) becomes zero:

$$U^{*\prime}_\delta(s) = -\frac{1}{1-s} + \frac{\delta\rho}{\rho s} = 0, \quad \text{i.e.} \quad s_\delta = \frac{1}{1+\delta^{-1}} \leq \frac{1}{2}.$$

Note that here the Cobb–Douglas utility function is degenerated, because the optimal saving is independent of the interest factor.

Then the optimal consumption pair are equal to

$$c_\delta = \frac{\delta^{-1}}{1+\delta^{-1}} \quad \text{and} \quad d_\delta = \frac{\rho}{1+\delta^{-1}}.$$

We specify the old-age consumption for two special interest factors ($0 < \delta < 1$):

$$d_\delta = \frac{1}{1+\delta^{-1}} < c_\delta \quad \text{if} \quad \rho = 1$$

and

$$d_\delta = \frac{\delta^{-1}}{1+\delta^{-1}} = c_\delta \quad \text{if} \quad \rho = \delta^{-1}.$$

Let us assume that $1 < \rho < 1/\delta$ and the paternalistic government does not accept the shortsightedness, i.e. undersaving. Then it replaces the private saving s_δ by a paternalistic mandatory retirement contribution τ yielding old-age benefit τ. There is no discounting in the paternalistic utility function (therefore δ is replaced by 1 in the subindex):

$$V_1[\tau] = \log(1 - \tau) + \log \tau.$$

The socially optimal contribution rate now is determined by

$$V_1'[\tau] = -\frac{1}{1-\tau} + \frac{1}{\tau} = 0, \quad \text{i.e.} \quad \tau_1 = \frac{1}{2}.$$

It is evident that in this model, the pension system achieves a higher saving than the private scheme does: $\tau_1 > s_\delta$. But the relevant question sounds as follows: for what parameter values (δ, ρ) is the pension system superior to the saving scheme in the paternalistic framework? The paternalistic welfare values of the two schemes are respectively

$$U^*_1(s_\delta) = \log(1 - s_\delta) + \log(\rho s_\delta) \quad \text{and} \quad V_1[\tau_1] = \log(1 - \tau_1) + \log \tau_1.$$

It is relatively straightforward to verify that $U^*_1(s_\delta)$ is an increasing function of δ. Then a simple calculation yields

Theorem 2. *For any discount factor $0 < \delta < 1$, there exists a critical interest factor ρ_δ, $1 < \rho_\delta < 1/\delta$ such that the private scheme is better/worse than the public one if the interest factor is above/below the critical value:*

$$U_1^*(s_\delta) < V_1[\tau_1] \quad \text{if} \quad \rho < \rho_\delta$$

and

$$U_1^*(s_\delta) > V_1[\tau_1] \quad \text{if} \quad \rho > \rho_\delta.$$

In words: in the representative agent's saving model, *the introduction of a mandatory pension system can be justified if and only if the efficiency of the private system is limited with respect to the shortsightedness.* Evidently, the higher the discount factor, the lower the critical interest factor below which paternalism is needed.

Notwithstanding all deficiencies of the model, it is worth demonstrating numerically how the critical value depends on the discount factor. Considering 30-year-long life-stages, our illustration applies only *annual* factors, derived from

$$\delta(1) = \delta^{1/30} \quad \text{and} \quad \rho(1) = \rho^{1/30}.$$

For example, in the first line of Table 2, the annual discount rate of 5 percent defines a critical interest rate of 1.8 percent, and the ratio of old- to young-age consumption is equal to 0.37. But raising the discount rate, the critical interest rate quickly converges to zero and the consumption ratio slowly approximates 1.

Table 2. *Autonomy and paternalism, annual factors*

Discount factor $\delta(1)$	Critical interest factor $\rho_{\delta(1)}(1)$	Young-age consumption c_δ^*	Old-age consumption d_δ^*
0.95	1.018	0.823	0.304
0.96	1.012	0.773	0.324
0.97	1.007	0.714	0.350
0.98	1.003	0.647	0.387
0.99	1.001	0.575	0.435
1.00	1.000	0.500	0.500

Remark. In the socially optimal mandatory pension system, $c_1 = 0.5$ and $d_1 = 0.5$.

2.3. Total intra- and intergenerational redistribution

At the end of this Section, unifying and simplifying the previous two models, we outline the model of Marxian communism: for inflexible labor supply and zero interest rate, full income redistribution is socially optimal. We do not need formulas: all incomes are taxed away and everybody receives the same transfer; the workers as tax returns and the pensioners as basic benefit.

These models are extremely simple, because they neglect at least one of the following complications: the intertemporality/heterogeneity of the population, the flexibility of labor supply and the possibility of underreporting earnings. Before entering the discussion of these problems, we outline a general framework, assuming individuals with differing wage rates, discount factors etc.

3. General model

We distinguish the micro and the macro sides. The primitive quantities appearing in the model are generally positive (possibly zeros) and the functions are smooth. We assume that the population as well as the economy is stationary, moreover, there is no inflation.

3.1. Micromodel

Let vector $\mathbf{p}$ of dimension $n+1$ denote the *individual characteristics* of a person (e.g. his wage rate, discount factor etc.). Let vector q of dimension N denote the *characteristics of the government transfer system* (e.g. contribution rate, accrual rate etc.). The employee works for a unit period and then is retired for another period, with length μ, $0 < \mu \leq 1$. His young-age consumption is c, his per-period old-age consumption is d. (Unlike in Subsection 2.2, at this stage we depart from the widespread but superfluous practice of identifying the lengths of the two periods in static models.) These variables depend on his *decision* vector x of dimension k. In the special models, the decisions are savings (including tax favored ones), labor supply and reported earnings.

Since one of the main reasons of having a mandatory retirement system is the individuals' shortsightedness, we single out the discount factor δ of the general characters, having thus $\mathbf{p} = (p, \delta)$, where p is an n-vector. Then it is worthwhile distinguishing the per-period young- and old-age utility functions $u_1(p, q, x, c)$ and $u_2(p, q, x, d)$. The *lifetime utility function* is

$$U(p, \delta, q, x, c, d) = u_1(p, q, x, c) + \mu\delta u_2(p, q, x, d).$$

In our abstract model, we have the following connection between the decision x and the feasible consumption pair:

$$c = c(p, \delta, q, x) \quad \text{and} \quad d = d(p, \delta, q, x).$$

For example, since the government operates with a contribution rate τ, out of the full wage w only the net wage $(1-\tau)w$ flows into the young-age consumption. This can be further diminished by taking into account the restrained labor supply $l(< 1)$ to $(1-\tau)wl$. The old-age consumption, in turn, primarily depends on the pension benefits: $\gamma + \beta wl$. In addition, if the government subsidizes a voluntary pension scheme through a matching rate φ and finances the scheme by an earmarked tax with rate θ, then the subsidized saving r, the unsubsidized saving s and the earmarked tax θwl also reduce young-age consumption by $r + s + \theta wl$ and increase the old-age consumption by $\mu^{-1}\rho[s + (1+\varphi)r]$, where ρ is the interest factor. In the more sophisticated models, for given consumption, the young-age per-period utility is lower than its old-age counterpart, the difference being the labor disutility $h(1-l)$.

Here we introduce the composite $n+1+N$-dimensional vector of individual and the government characteristics $P = (p, \delta, q)$. Substituting the two consumption functions into the lifetime utility function yields the *reduced utility function*:

$$U^*(P, x) = U(P, x, c(P, x), d(P, x)).$$

Following the neoclassical theory, we assume that every individual takes the government transfer system as given and maximizes his reduced utility function with respect to x:

$$U^*(P,x) \to \max .$$

Then we have

Theorem 3. *a) For the locally optimal decision function $x(P)$, condition*

$$U^{*\prime}_x(P,x) = 0$$

holds.

b) Assume that the k-order square matrix $U^{*\prime\prime}_{xx}$ is strictly negative definite, then the matrix of rates of marginal change of $x(P)$ is

$$x'(P) = [U^{*\prime\prime}_{xx}]^{-1} U^{*\prime\prime}_{xP}.$$

We shall need the balance function of the individuals, denoted by $T(P,x)$. This determines the balance of the contributions and the benefits of type (p,δ) at the transfer system q.

The *neutral* system, where the balance is identically zero: $T(P,x) \equiv 0$, is theoretically interesting. Then the socially suboptimal decision $x_{\mathrm{N}}(P)$ satisfies $L'_x = 0$, where the individual Lagrange-function is

$$L(P,x) = U^*(P,x) - \lambda(P)T(P,x)$$

and $\lambda(P)$ is the appropriate Lagrange multiplier.

3.2. Macromodel

Until now we have only considered a single type with given individual characteristics p and δ and government characteristics q. Now we introduce the heterogeneity of types. Let $\mathbf{E}$ be the operator of expectation of the individual types (typically of wages and discount factors). Then the balance of the transfer system can be formulated as

$$\mathbf{E}T(P,x(P)) = 0.$$

Note that this is already a general equilibrium model, where individual decisions depend on the transfer system and the balance of the latter also depends on the former.

In certain models, there is only a single transfer system (namely, the public pension) but in others, the tax system also appears. It is a question of choice to have either a unified or separated balances.

In our framework, the government maximizes a *paternalistic* social welfare function, where the individual discount factors are replaced by 1 and the transfer system is balanced. Also, the individual utility functions can be transformed by an increasing and concave function ψ, making the social welfare function *progressive*:

$$V(q) = \mathbf{E}\psi[U^*(p,1,q,x(p,\delta,q))].$$

If $\psi(U) \equiv U$, then we receive pure utilitarianism.

Then the government chooses a vector of transfer characteristics q^* such that $V(q)$ is maximal with respect to the balance equation(s). Again considering a local maximum, and relying on the method of Lagrange multiplier, we have

Theorem 4. *A necessary first-order condition for a social optimum is*

$$\mathbf{L}'_q = \mathbf{E}\left[\psi'(U^*)\frac{dU^*}{dq}(p,1,q,x(p,\delta,q))\right] - \nu\mathbf{E}\frac{dT}{dq}(p,\delta,q,x(p,\delta,q)) = 0$$

and $\mathbf{E}T(p,\delta,q,x(p,\delta,q)) = 0$.

Remarks. 1. It is clear that for an N-vector q we have N independent scalar equations and ν is also an N-vector.

2. Note the contradiction between individual myopia and paternalism: V has parameter values 1 and δ at the same time.

Comparing the social welfare provided by two transfer systems A and B, we cannot rely on the numerical values of the social optima, since the utility functions' scales are also arbitrary. As in Subsection 2.1 above, we shall use an indirect method for comparison. Let e be a positive scalar and $V_X(e)$ be the social welfare assured by system X when all wages (rates) are uniformly multiplied by e. Then the efficiency of system A is e times that of B if $V_B(e) = V_A(1)$.

What problems can be investigated with the help of such a family of models? A skeptic might say that such stylized models are useless. Luckily there are such or even more stylized models—or families of models—which give rise to strong conclusions. At least such models are able refute certain qualitatively wrong ideas. For example, neutral systems are typically socially inferior to redistributive ones. On the other hand, excessive redistribution may undermine the efficiency of transfers so much that even the preferred types suffer from redistribution. In Sections 4 and 5 only specific models will be analyzed.

4. Specific models

In this Section we outline the specific models, transformed into an approximately common form. Because of the lack of illuminating analytical results, we must often rely on numerical illustrations, and generally we rest satisfied with two types: denoted by L(ow) and H(igh). We have wages $w_L < w_H$, the types' shares are f_L and f_H with $f_L + f_H = 1$. The average wage rate is 1, implying $w_H = (1 - f_L w_L)/f_H$. (Note, however, the continuous wage distribution in Subsection 5.2 below). In the two-overlapping-generation models, numerically we work with $\mu = 1/2$, i.e. the retirement period's length is half of the working period.

Just as a warming-up, we return to the socially optimal contribution rate (of Subsection 2.2), when saving is zero and μ ceases to be 1. Now

$$V_1[\tau] = \log(1-\tau) + \mu\log(\mu^{-1}\tau).$$

Therefore the socially optimal contribution rate is determined by

$$V_1'[\tau] = -\frac{1}{1-\tau} + \frac{\mu}{\tau} = 0, \quad \text{i.e.} \quad \bar{\tau} = \frac{1}{1+\mu^{-1}}.$$

If $\tau > \bar{\tau}$, then $d > c$, excess to be excluded. We shall call $\bar{\tau}$ *maximal*, corresponding to the perfectly far-sighted worker's optimal saving rate. The corresponding net replacement rate is unitary: $c = d$.

4.1. Voluntary pension system

We start with the simplest special model, namely the interaction of mandatory and voluntary pension systems. The basic idea is simple: in economics, it is a commonplace that voluntary participation is generally better than mandatory one. But we have already shown in Subsection 2.2 that in the case of life-cycle saving of myopic workers, some coercion may be needed. One hopes that this can be minimized, for example, by introducing voluntary pensions. Nevertheless, if one takes into account the tax expenditure on tax-favored voluntary pension systems, then the attractiveness of this form is greatly diminished. Perhaps the existence of a cap on voluntary pension contributions already signals this effect.

As outlined in the introduction to the general framework, there is a worker earning wage w who pays a mandatory pension contribution τw and will receive a proportional benefit $b = \mu^{-1}\tau w$ when retired. In addition to this mandatory transfer, there is a possibility to contribute r to a voluntary system, matched by a factor φ by the government up to the cap $\bar{r}$. If the worker wants to save even more, he can save s without getting any matching. The factor of return to these private savings is ρ, i.e. the old-age consumption becomes

$$d = \mu^{-1}[\tau w + \rho\{(1+\varphi)r + s\}].$$

Note, however, that this matching must be covered by proportional earmarked taxes, θw. Therefore the worker's consumption reduces to

$$c = w - (\tau + \theta)w - r - s.$$

In this Section we neglect the real interest rate: $\rho = 1$ and work with CRRA rather than Cobb–Douglas-utility function: $u(c) = \sigma^{-1}c^{\sigma}$, where $\sigma < 0$. (For $\sigma = 0$, we would obtain the log-utility function! Note that in Subsection 2.2, the optimal saving s_δ was independent of the interest factor ρ, showing degeneracy!) Then for positive savings, the optimal consumption pair satisfy

$$d = cD(\delta, \varphi), \quad \text{where} \quad D(\delta, \varphi) = [\delta(1+\varphi)]^{1/(1-\sigma)}.$$

For zero saving, $d > cD(\delta, \varphi)$.

We assume that the lower paid have a lower discount factor than the higher paid have:

$$0 \le \delta_L < \delta_H \le 1.$$

We have two balance equations.
Pension balance:

$$\mu\beta\mathbf{E}w = \tau\mathbf{E}w, \quad \text{i.e.} \quad \tau = \mu\beta.$$

Earmarked tax balance:

$$\varphi\mathbf{E}r = \theta\mathbf{E}w.$$

It is evident that no rational worker would save normally before exhausting the possibilities of voluntary contributions in $[0, \bar{r}]$.

We compare three idealized systems: (i) the *purely mandatory* system, (ii) the *asymmetric* system, where only the far-sighted participate and (iii) the *symmetric* system, where both types participate proportionally to their wages. We shall see that the symmetric system is socially superior to both the pure mandatory system and the asymmetric one, which in turn are roughly equivalent.

To save space, here we only discuss the asymmetric voluntary system, and confine attention to a single case from Theorem 3 of Simonovits (2009). As a preparation, let us denote δ^{o} the government implicit discount factor, implying the mandatory contribution rate τ, as an individual optimum ($\tau = s_{\delta^{\mathrm{o}}}$). Then let φ_L be that positive scalar, for which $\delta_L(1+\varphi_L) = \delta^{\mathrm{o}}$, i.e. that matching rate at which type L is just without a voluntary contribution.

Theorem 5. *If the matching rate is low enough, $0 < \varphi \le \varphi_L$ and the cap on voluntary contribution is high enough:*

$$\bar{r} > r_H(\varphi) = \frac{[D(\delta_H,\varphi)(1-\tau) - \mu^{-1}\tau]w_H}{D(\delta_H,\varphi)(1+f_H\varphi w_H) + (1+\varphi)\mu^{-1}},$$

then type H's optimal contribution is equal to $r_H(\varphi)$, while $s_H = 0$.

To illustrate our findings numerically, we choose $\sigma = -1$, $\delta_L = 0.15$; $\delta_H = 0.5$ (annually $\delta_L(1) = 0.939$; $\delta_H(1) = 0.977$). The government contribution rate $\tau = 0.183$ corresponds to a medium discount factor $\delta^{\mathrm{o}} = 0.2$ (annually $\delta^{\mathrm{o}}(1) = 0.948$). To have simple algebra, the cap in each voluntary system is to be equal to H's voluntary contribution (italicized in Table 3). To amplify the effects, we choose a rather unequal wage distribution: $f_L = 2/3$ and $f_H = 1/3$ with $w_L = 0.5$ and $w_H = 2$. Table 3 presents a numerical illustration.

Table 3. *Comparison of combinations of mandatory and voluntary pensions*

Earning w_i	Voluntary contribution r_i	Traditional saving s_i^{o}	Worker consumption c_i^{o}	Pensioner consumption d_i^{o}	Voluntary transfer T_i^{o}	Relative efficiency e
Pure mandatory system ($\varphi = 0$)						1
0.5	0	0	0.409	0.183	0	
2.0	0	0.157	1.478	1.045	0	
Asymmetric voluntary system ($\varphi = 1/3$)						0.989
0.5	0	0	0.399	0.183	–0.009	
2.0	*0.165*	0	1.433	1.170	0.018	
Symmetric voluntary system ($\varphi = 1$)						1.038
0.5	0.006	0	0.393	0.215	0	
2.0	*0.023*	0.092	1.478	1.045	0	

(i) The pure mandatory system is only displayed as a benchmark, with relative efficiency 1. Note the unacceptably low old-age consumption of the myope: $d_L^{\mathrm{o}} = 0.183$.

(ii) The mandatory pillar supplemented with an asymmetric voluntary pillar only makes things a little bit worse because the matching rate is too low to help the myope ($\varphi = 1/3$) and the ceiling on voluntary contribution is high enough ($\bar{r} = 0.165$) to allow the far-sighted to appropriate all the benefits. The earmarked tax rate creates a net transfer from the myopes to the far-sighted. The young-age consumption of the former slightly diminishes, just to help raise the far-sighted's old-age consumptions. The pure mandatory pillar can achieve the same social welfare as the asymmetric voluntary system with 1.1 percent lower wages.

(iii) The mandatory pillar supplemented with a symmetric voluntary pillar redresses the injustice: the matching rate is raised to 1, while the ceiling is drastically lowered to 0.023. The myope's old-age consumption rises now from $d_L^{\mathrm{o}} = 0.183$ to 0.215, the young-age consumption drops from $c_L^{\mathrm{o}} = 0.409$ to 0.393, while c_H^{o} and d_H^{o} remain invariant. The pure mandatory pillar can only achieve the welfare of the symmetric voluntary system by increasing wages uniformly by 3.8 percent. Note, however, that this is the same as a $1.2 + 1.2 = 2.4$ percent point increase of the mandatory contribution rate!

4.2. Cap on the contribution base

In most countries, workers do not pay pension contributions after the earning above the so-called contribution base cap, shortly *cap*. In the Introduction, we promised to model a neglected function of the cap: it effectively limits the ratio of the contribution to the earnings, reducing the presumed adverse efficiency impact of mandatory saving: this is the basis of the present model. We shall see that in our specification, the cap has hardly any welfare impact in the relevant interval but is socially superior to the capless case.

Our starting point is the cap $\bar{w}$ and the covered total wage $\hat{w} = \min(w, \bar{w})$. Then the contribution is equal to $\tau\hat{w}$. We introduce the covered net wage and the corresponding benefit:

$$\hat{v} = (1 - \tau)\hat{w} \quad \text{and} \quad b = \beta\hat{v}.$$

We shall also need the *effective contribution rate*, defined as the ratio of contribution to wage:

$$\tilde{\tau} = \frac{\tau\bar{w}}{w} \quad \text{if} \quad w > \bar{w} \quad \text{and} \quad \tilde{\tau} = \tau \quad \text{if} \quad w \leq \bar{w}.$$

Obviously, for a wage higher than the cap, the effective contribution rate is lower than the statutory one: $w > \bar{w}$ implies $\tilde{\tau} < \tau$.

We only display the most important formulas. We should distinguish again two cases: the saving intention is either (i) nonnegative or (ii) negative.

To determine the optimal saving, we need the Cobb–Douglas-lifetime utility function

$$U(w, \delta, c, d) = \log c + \mu\delta \log d.$$

Inserting the consumption functions

$$c = w - \tau\hat{w} - s \quad \text{and} \quad d = \beta\hat{v} + \rho\mu^{-1}s,$$

we obtain the derived utility function

$$U^*(w, \delta, s) = \log(w - \tau\hat{w} - s) + \mu\delta \log(\beta\hat{v} + \rho\mu^{-1}s).$$

Introducing the cap-dependent net earning $w - \tau\bar{w}$ and taking the derivative by s and equating it to zero; yields

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

Hence we have the following result.

Theorem 6. *In a pension system with the cap-dependent net earning $w - \tau\bar{w}$, the optimal saving intention is*

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

If $\delta\rho v \geq b$, then $s^i \geq 0$, i.e. $s^{\mathrm{o}} = s^i$ and $d^{\mathrm{o}} = \delta c^{\mathrm{o}}$. Otherwise, if $\delta\rho v < b$, then $s^i < 0$, i.e. $s^{\mathrm{o}} = 0$ and $d^{\mathrm{o}} > \delta c^{\mathrm{o}}$.

In our paternalistic framework, the government maximizes the expected value of

$$U^*[w, 1] = \log(w - \tau\bar{w} - s^{\mathrm{o}}) + \mu \log(\beta\hat{v} + \rho\mu^{-1}s^{\mathrm{o}})$$

rather than $\mathbf{E}U^*[w, \delta]$. That is,

$$V(\tau, \bar{w}) = \mathbf{E}U^*[w, 1].$$

We confine our numerical illustrations to the 2-type case. Input data: The compounded discount factors for 30 years are $\delta_L = 0.1$; and $\delta_H = 0.6$ (annually $\delta_L(1) = \delta_L^{1/30} = 0.926..$ and $\delta_H(1) = \delta_H^{1/30} = 0.983..$); while the interest factor is $\rho = 2$ (annually $\rho(1) = 2^{1/30} = 1.0234..$). By theoretical and numerical reasoning, the socially optimal contribution rate $\tau^{\mathrm{o}} = 0.333$. Table 4 shows that the social welfare function increases with the cap until w_L and then it drops. The key to the solution is to be found in the consumption of the lower paid: $d_L^{\mathrm{o}}(\bar{w})$, which grows until $\bar{w}$ reaches w_L (italicized row), beyond which it only reduces the possibilities of the high-earners.

Table 4. *The impact of cap on the social welfare*

	Low earner		High earner		Social
	young-age	old-age	young-age	old-age	welfare
Cap		consumption			function
$\bar{w}$	c_L^{o}	d_L^{o}	c_H^{o}	d_H^{o}	V
0.0	0.476	0.095	1.538	1.846	–1.033
...					...
0.4	0.367	0.266	1.487	1.785	–0.881
0.5	*0.334*	*0.333*	*1.474*	*1.769*	*–0.874*
0.6	0.334	0.333	1.462	1.754	–0.878
...					...
2	0.334	0.333	1.282	1.539	–0.944

Remark. $\tau^* = 0.333$, $\bar{w}^* = 0.5$.

The wages should be uniformly increased by 11.1 and 4.6 percent in the no-pension or the no-cap pension system respectively to have as high social welfare as with the socially optimally capped pension system.

4.3. Redistribution with flexible labor supply

This is *the* classical model, where the two introductory models (Subsections 2.1–2.2) are modified, unified and the third (Subsection 2.3) is generalized.[9] Here the labor supply is flexible and is influenced by the personal income tax rate and the forced saving for old-age. To keep the model as simple as possible, there is no private saving and the personal income tax θwl is copied from Subsection 2.1 but the basic income γ is extended to the old. We recall some earlier notations and formulas: τ = contribution rate, θ = personal income tax rate, $t = \tau + \theta$ = transfer rate and $\bar{t} = 1 - t$ = the net-of-transfer rate. Furthermore, l = labor supply. The pension benefit in turn consists of two parts: flat benefit γ and the proportional benefit βwl, where β is the so-called *accrual rate* (or equivalently, marginal replacement rate) and wl is the total earning. Hence the total benefit is $b(wl) = \gamma + \beta wl$. We display the consumption pair:

$$c = \gamma + \bar{t}wl \quad \text{and} \quad d = \gamma + \beta wl.$$

Without proof, we give the necessary condition for individual optimality.

Theorem 7. *The necessary condition for the optimal labor supply is*

$$\frac{\bar{t}w}{\gamma + \bar{t}wl} - \frac{\xi}{T - l} + \frac{\delta\mu\beta w}{\gamma + \beta wl} = 0.$$

With rearrangement, a quadratic equation is obtained:

$$a_2 l^2 + a_1 l + a_0 = 0,$$

where

$$a_2 = -(1 + \xi + \delta\mu)\bar{t}\beta w^2,$$

$$a_1 = \bar{t}w(T\beta w - \gamma) - \xi w(\gamma\beta + \bar{t}\gamma) + \delta\mu\beta w(\bar{t}wT - \gamma),$$

and

$$a_0 = \bar{t}wT\gamma - \xi\gamma^2 + \delta\mu\beta w\gamma T.$$

We shall consider two special cases: (i) the proportional pension system (P) without redistribution: $\gamma = 0$ and (ii) the pure flat benefit system (F): $\beta = 0$. Here the optimal labor supplies are as follows:

$$l_0^P = \frac{T}{1 + \xi} < l^P = \frac{(1 + \delta\mu)T}{1 + \xi + \delta\mu} < T \quad \text{and} \quad 0 < l_w^F = \frac{T - \xi\gamma/(\bar{t}w)}{1 + \xi} < l_0^P,$$

where l_0^P is the labor supply of Feldstein's myopes ($\delta = 0$). The unified balance condition is as follows:

$$t\mathbf{E}(wl) = (1 + \mu)\gamma + \mu\beta\mathbf{E}(wl), \quad \text{i.e.} \quad \gamma = (1 + \mu)^{-1}(t - \mu\beta)\mathbf{E}(wl), \qquad t \geq \mu\beta.$$

We only illustrate the impact of wage rate inequality, i.e. reduce the ratio $\omega = w_H/w_L$ from 4 to 1. It is to be expected that the socially optimal transfer rate is

[9] Among the models surveyed, this is the closest to Cremer et al. (2008)!

dropping along, but only numerical calculations, displayed in Table 5, give the measures: namely the socially optimal transfer rate drops from 0.52 to 0.33 and through the rise of labor supply, the average wage rises from 0.74 to 0.91. But it is quite surprising that the optimal redistribution among workers disappears well before reaching wage rate equality, namely slightly below $\omega = 2$ (exactly at 1.8)! At the same time, the accrual rate β^* steeply rises and reaches the value of the optimal proportional system. We cite Table 4 from Simonovits (2012d).

Table 5. *The impact of earning inequality on the social optimum*

Ratio of wage rates ω	Low w_L	Transfer rate t^*	Basic income γ^*	Accrual rate β^*	Average wage $\mathbf{E}(wl^*)$	Low worker consumption c_L^*	Low pensioner consumption d_L^*
4	*0.5*	*0.52*	*0.158*	*0.40*	*0.741*	*0.281*	*0.260*
3	0.6	0.46	0.111	0.51	0.811	0.337	0.324
2	0.75	0.35	0.015	0.65	0.905	0.440	0.440
1	1	0.33	0.000	0.66	0.908	0.596	0.587

Remark. $\mu = 0.5$, $\delta_L = 0.4$ and $\delta_H = 0.7$.

4.4. Underreporting wages

One of the most important problems of public economics, especially during transition (from socialism to capitalism) is as follows: how to achieve income redistribution in such a way that the lower income people do not starve and the higher income people do not hide too much of their incomes? Using the tax morale, mentioned in the Introduction, endogenizes underreporting of labor incomes. Excluding the old-age, we had already been able to derive quite sharp analytical results in the generalization of Subsection 2.1 (cf. Simonovits, 2012b). But extending the analysis to the second stage, we have to be content with more modest results.

We keep the notations of Subsection 4.3, except for wl is replaced by *reported wage* v. We choose the simplest progressive tax system, where every worker pays a linear earning tax θv and receives a lump-sum return ε. Thus the *net tax* is equal to $\theta v - \varepsilon$. The pension benefit in turn consists of two parts: flat benefit α and the proportional benefit βv, where β is the so-called accrual rate. Hence the total benefit is $b(v) = \alpha + \beta v$. In addition to forced savings via the pension system, workers also can save privately, its value is denoted by s, there is no interest: $\rho = 1$. We display the consumption pair:

$$c = w - tv + \varepsilon - s \quad \text{and} \quad d = \alpha + \beta v + \mu^{-1} s.$$

In the spirit of the tax morale literature, we introduce the tax morale parameter m, influencing the subjective utility of the worker underreporting his true earning w:

$$z(w, m, v) = m(\log v - v/w).$$

Due to the linear correction term, $z(w, m, \cdot)$ has a maximum at $v = w$. In fact,

$$z'_v(w, m, v) = \frac{m}{v} - \frac{m}{w} > 0$$

if and only if $0 < v < w$, i.e. the overreporting individual is not rational.

The other terms of the lifetime utility function are standard and have been introduced in Section 3:

$$u_1(c) = \log c \quad \text{and} \quad u_2(d) = \delta \log d.$$

In sum, the total individual utility function is

$$U(w, m, \delta, c, d, v) = \log c + m(\log v - v/w) + \mu\delta \log d.$$

Executing the substitutions and equating the partial derivatives (with respect to wage reporting and saving) to zero, yield the individual optimum. Because savings cannot be negative, we must distinguish again between two cases: a) the saving intention is nonnegative or b) negative. In the former case, the saving intention is realized (slack credit constraint); in the latter case, the actual saving becomes zero (binding credit constraint). We only consider here the former case.

Theorem 8. *Under our assumptions, if the worker has a slack credit constraint, then the individually optimal young-age consumption c is the positive root of the following quadratic equation:*

$$h_2 c^2 + h_1 c + h_0 = 0, \quad \text{where} \quad t_1 = t - \mu\beta$$

and

$$h_2 = m(1+\mu\delta), \quad h_1 = t_1(1+\mu\delta)w + t_1 m w - m(w + \mu\alpha + \varepsilon), \quad h_0 = -t_1 w(w + \mu\alpha + \varepsilon).$$

This individually optimal worker's consumption defines the other components like reported wage v.

To get rid of the mathematical difficulties of the tax balance equation, we assume the existence of a positive net average tax $\bar{a} = \mathbf{E}(\theta v - \varepsilon)$. Furthermore, the individual utility is enhanced by the utility of public goods, modeled as

$$q(\bar{a}) = (1 + \mu)\kappa \log \bar{a},$$

where κ is the coefficient of relative utility of public goods to that of young-age consumption.

At this point, we must consider the pension balance equation:

$$\tau \mathbf{E} v = \mu \mathbf{E} b(v).$$

We arrived to a relatively simple general equilibrium model mentioned in Section 3.

We display two numerical calculations, where the tax morale is relatively weak and strong, respectively: a *dark grey* economy with $m = 0.5$ and a *light grey* economy with

$m = 2$. For the relative utility parameter $\kappa = 1/3$, we determine the socially optimal values of the transfer system.

Unlike in light grey economy, in the dark grey economy, the socially optimal basic income is always zero, therefore Table 6 has one column less than Table 7. Citing Table 4 in Simonovits (2009b) we see that the higher the flat component, the lower the average reported wage and consequently, the lower the social welfare, assuming that the government chooses socially optimal tax rates. Repeating our warning again: there is not much sense in differences between the values of social welfare functions, rather we must measure the decline by *consumption equivalence*. For example, let $e > 1$ be such a scalar that by multiplying both types' wages with it, the social welfare of a flat component $\alpha = 0.2$ becomes identical to that of the original proportional system. We obtain $e = 1.03$, i.e. the economy should have wages uniformly higher by 3 percent to achieve the original welfare of the proportional system in the foregoing flat system.

Table 6. *An optimal system with a given flat benefit, dark grey economy,* $m = 0.5$

Flat benefit α	Contribution rate τ^{o}	Tax rate θ^{o}	Marginal replacement rate β^{o}	Reported wage $\bar{v}^{(P)}$	Average net tax $\bar{a}$	Social welfare V
0.00	*0.22*	*0.28*	*0.440*	*0.547*	*0.102*	*–3.108*
0.04	0.35	0.29	0.620	0.500	0.097	–3.113
0.08	0.37	0.29	0.568	0.466	0.090	–3.124
0.12	0.36	0.29	0.447	0.439	0.085	–3.140
0.16	0.36	0.29	0.331	0.411	0.079	–3.161
0.20	0.35	0.29	0.181	0.385	0.074	–3.185

Turning to the light grey economy (Table 5 in Simonovits 2009b), the purely proportional system is less efficient than the optimally redistributive one. In our specification, the average reported wage is also a decreasing function of flat component, but its impact is weak enough that the proportional benefit disappear around $\alpha = 0.28$, before the system reached the optimum at $\alpha = 0.32$. In the purely proportional system, wages should be increased uniformly by 2.1 percent to achieve the same social welfare as in the social optimum with basic benefit $\alpha = 0.32$.

Table 7. *An optimal system with a given flat benefit, light grey economy, $m = 2$*

Flat benefit α	Contribution rate τ^{o}	Net tax rate θ^{o}	Basic income ε	Marginal replacement rate β^{o}	Reported wage $\bar{v}^{(F)}$	Average net tax $\bar{a}^{(F)}$	Social welfare $V^{(F)}$
0.00	0.26	0.40	0.10	0.520	0.732	0.128	–4.810
...							...
0.20	0.26	0.36	0.07	0.239	0.712	0.124	–4.773
0.24	0.21	0.27	0	0.094	0.735	0.132	–4.772
0.28	0.20	0.27	0	0.013	0.724	0.130	–4.767
0.32	*0.22*	*0.28*	*0*	*0*	*0.707*	*0.132*	*–4.755*
0.36	0.25	0.26	0	0	0.703	0.122	–4.760
0.40	0.28	0.24	0	0	0.699	0.112	–4.773

5. Original specifications

As was mentioned in the Introduction, in displaying the specific models in Section 4, to approach a common framework, we have frequently deviated from the original specifications. Here we present some original ones, which shed some light on the sensitivity of the models as well.

5.1. Discounted Leontief utility and cap

As is known, the additive discounted utility function, introduced by Samuelson (1937) and specified as CRRA, contains the discounted Cobb–Douglas and the *undiscounted* Leontief utility functions as special cases. But following the spirit of Shane, Loewenstein and O'Donoghue (2002), we gave up the additivity assumption, and to help analytical discussion of the contribution cap, originally we chose a specific utility function, called discounted Leontief utility function:

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

Note that this lifetime utility function is *not* age-additive, and surprisingly it is the young- rather than the old-age consumption which is discounted. Note, however, that for any positive saving, at the optimum, $\delta c = d$, i.e. the lower the discount factor, the lower the old-age consumption. An additional advantage of this specification lies in that for reasonable choice of the contribution rate, $c > d$, i.e. $U(w, \delta, c, d) = d$, yielding a simple social welfare function: $V = \mathbf{E}d$. Simonovits (2013) conjectured that the unitary replacement rate is socially optimal.

We only display the most important formulas. Again, we should distinguish two cases: the saving intention is (i) nonnegative or (ii) negative.

Ad (i). At the optimum, $\delta c^* = d^*$, i.e. $\delta v - \delta s = b + \mu^{-1}\rho s$ yielding the saving intention $s^i = (\delta v - b)/(\delta + \mu^{-1}\rho)$, which is realized of and only if $\delta \geq \delta_\tau = \mu^{-1}\tau\hat{w}/v$. Then $c_1^* = (\rho v + \mu b)/(\mu\delta + \rho)$.

Ad (ii). If $s^i < 0$, then $s_2^* = 0$, i.e. $c_2^* = v$ and $d_2^* = b$.

Here we confine our numerical illustrations to the 3-type case. Input data: $f_L = 0.7$; $f_E = 0.25$ and $f_H = 0.05$, wages: $w_L = 0.7$; $w_E = 1.3$ and $w_H = 1.7$. The compounded interest factors for 30 years are $\delta_L = 0.215$; $\delta_E = 0.545$ and $\delta_H = 0.97$; at annual level: $\delta_L(1) = 0.977$; $\delta_E(1) = 0.988$ and $\delta_H(1) = 0.999$. The socially optimal contribution rate $\tau^* = 0.333$ and the corresponding cap $\bar{w}^* = w_E$. Table 2 of Simonovits (2012a) is reproduced here.

Table 8. *The impact of cap on the social welfare, 3 types, Leontief*

Cap $\bar{w}$	Medium consumption d_E^o	High consumption d_H^o	Social welfare $V = \mathbf{E}d$
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. $\tau^* = 0.333$, $\bar{w}^* = 1.3$.

Note that raising the cap from 1 to the medium wage 1.3, the weighted increase in d_E^o is greater than the weighted decrease in d_H^o. Above w_E, the further rise in $\bar{w}$ only diminishes d_H^o, without raising d_E^o.

5.2. Pareto-distribution

Though in our survey, the earning distribution is generally discrete, with two or three types, now we present a more realistic, continuous earning distribution named after Wilfredo Pareto (cf. Diamond and Saez, 2011). The distribution function is given as

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

$\sigma > 1$ and $w_m > 0$ is the minimal earning. Normalizing the average wage to 1, the minimum wage is $w_m = (\sigma - 1)/\sigma$.

Selected numerical values are displayed in Table 9. Note that the lower 50 percent of earners have only 25 percent of the total earnings, while the upper 3 percent of earners have 17 percent.

Table 9. *Pareto-probabilities and covered wage shares as a function of the cap*

Earning cap $\bar{w}$	Probability $F(\bar{w})$	Covered wage share $\mathbf{P}(w < \bar{w})\mathbf{E}(w \mid w < \bar{w})$
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

Returning to the cap, we show the result obtained for Pareto-distribution and discounted Leontief utility function. We specify an appropriate discount factor–wage function

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

where δ_m and δ_M are the minimal and the maximal discount factors, respectively, while η measures the strength of the dependence on wage. Let $\delta_m = 0.25$ (annually 0.955, quite low); $\delta_M = 1$ and $\eta = 0.5$.

Simonovits (2013) investigated the optimal choice of the contribution rate without imposing a cap. For our parameterization, the maximal contribution rate is the social optimum: $\tau^* = \bar{\tau}$. According to Table 10, the social welfare achieved for the contribution rates lower than 0.17 is inferior to the no-pension's!

Table 10. *The welfare impact of the contribution rate without cap, Pareto*

Contribution rate τ	10× Social welfare $10U$	Expected saving $\mathbf{E}s$
0.000	4.35	0.109
0.100	4.14	0.053
0.120	4.10	0.043
0.170	4.34	0.024
0.200	4.64	0.016
0.300	6.07	0.002
0.333	6.67	0.000

Turning to the issue of optimal cap, Table 11 shows that the cap hardly influences the social welfare in the relevant interval [2.5,4.5], though its total lack (last row) diminishes the welfare by 0.7 percent.

Table 11. *The welfare impact of the cap, Pareto*

Cap $\bar{w}$	10× Social Welfare $10U$	Average saving $\mathbf{E}s$
2.5	6.690	0.017
3.0	6.704	0.015
3.5	6.709	0.013
4.0	*6.709*	*0.012*
4.5	6.708	0.010
5.0	6.705	0.009
...		...
∞	6.660	0.000

Remark. $\tau = 1/3$.

5.3. Pension credit

This model generalizes the model with flexible labor supply (Subsection 4.3) in two dimensions: it includes private saving and replaces the linear "flat plus" pension benefit by the nonlinear pension credit. This also generalizes the means-tested pension.

The basic pension is a universal transfer, regardless of the individual's other incomes. The means-tested benefit is only given to those whose other income fall below the minimum, and this insufficient income is topped up to the minimum. It can be seen that due to its size, the "basic plus" system may undermine incentives to contribute; while the means-tested system creates a poverty trap.[10] The pension credit is a clever combination of the two special systems: it pays full basic benefit for zero income and zero flat pension for sufficiently high income. In between, the basic income is smoothly reduced by a share of the other incomes.

To avoid repetition, here we only present the elementary part, without dealing with optimality. It is assumed that a given share, say ε ($0 \leq \varepsilon \leq 1$) of the proportional pension βv is deducted from the flat part. Denote by x_+ the positive part of x, i.e. $x_+ = x$ for $x \geq 0$ and $x_+ = 0$ for $x < 0$. Then the pension credit benefit is given by

$$b(v) = [\gamma - \varepsilon\beta v]_+ + \beta v.$$

It is worth rewriting this formula:

$$b(v) = \max[\gamma + (1-\varepsilon)\beta v, \beta v].$$

In words: the benefit is equal to the maximum of the following two quantities: a) the flat benefit + the reduced proportional benefit, and b) the proportional benefit.

We shall work with three types and the following shares: $f_L = 0.5$, $f_E = 0.3$ and $f_H = 0.2$ and fix β so that the flat benefit be half of the average benefit. Table 12

[10] Following Feldstein (1985), this was the starting point of Feldstein (1987), which initiated the whole research direction!

cites Table 2 from Simonovits (2012c). It shows that the flat system is very expensive ($\mathbf{E}b = 0.6$), while the means-tested system assures too low a benefit even to the average earner ($b_E = 0.3$). The pension credit is not too expensive ($\mathbf{E}b = 0.458$) and it assures an acceptable minimum benefit ($b_L = 0.375$).

Table 12. *Pension credit,* $\alpha = 0.3$

Proportion of credit $1-\varepsilon$	Full wage v_i	Proportional βv_i	Conditional pension $\alpha + (1-\varepsilon)\beta v_i$	Total b_i	Average $\bar{b}$
0	0.5	0.150	0.300	0.300	
	1	0.300	0.300	0.300	
	2.25	0.675	0.300	0.675	
means-tested					0.375
0.5	0.5	0.150	0.375	0.375	
	1	0.300	0.450	0.450	
	2.25	0.675	0.638	0.675	
pension credit					0.458
1	0.5	0.150	0.450	0.450	
	1	0.300	0.600	0.600	
	2.25	0.675	0.975	0.975	
flat					0.6

6. Conclusions

We have presented a general elementary framework within which many interesting special transfer (tax and pension) models can be investigated. The joint framework is the two-overlapping-generation model, where the individuals work in the first stage of their adult life and retire for the second. In every model of this family, the government operates a balanced linear or piecewise linear transfer systems, possibly with caps. The individuals optimize their lifetime utilities by choosing paying voluntary pension contributions, their reported earnings, reacting with increased savings to lowered caps, supplying more labor. The individuals are typically shortsighted, therefore left for themselves they would save too little, supply too little labor and hide a large part of their earnings. The paternalistic government tries to steer them in the good direction through incentives. Due to the linearity (or piecewise linearity) of the transfer rules, their optimal parameter values can easily determined. The analysis would be much more difficult if genuinely nonlinear transfer systems were considered.

List of symbols

Common symbols

w = total wage (cost)
L = low
T = labor time
τ = contribution rate
t = transfer rate
c = young-age consumption
b = pension
μ = length ratio
$u = U$ = utility function
f = relative frequency
β = accrual rate
δ = discount factor
H = high
l = labor supply
θ = tax rate
$\bar{t}$ = net-of-transfer rate
d = old-age consumption
s = saving (i superscript=intention)
ρ = interest factor
V = social welfare function
$\mathbf{E}$ = expected value

Income redistribution under flexible labor income supply

γ = basic income
ξ = relative utility of leisure

Framework

p = individual characteristics
e = relative efficiency
u_1 = young-age utility
T = transfer function
L = individual Lagrange-function
λ = individual Lagrange-multiplier
q = government characteristics
P = combined characteristics
u_2 = old-age utility
x = decision
$\mathbf{L}$ = social Lagrange-function
ν = social Lagrange-multiplier

Voluntary pension

r = voluntary contribution
α = matching rate
γ = generalized discount factor
$\bar{r}$ = cap on voluntary contribution
σ = CRRA-index

Cap on the contribution base

$\bar{w}$ = cap
$\hat{v}$ = covered net earning
$\hat{w}$ = covered total earning
v = net earning

Flexible labor supply

ω = wage rate ratio
F = flat system
γ = basic income
P = proportional system

Reported earning and redistribution

v = reported earning
z = morale utility
κ = coefficient of utility of public goods
α = basic income
m = tax morale
a = net tax
E = medium type
ε = basic pension

Pareto-distribution

f = density function
σ = density index
w_N = upper earning
δ_m = minimal δ
η = coefficient of discount factor
F = distribution function
w_m = minimum wage
w_K = representative earning
δ_M = maximal δ

Pension credit

ε = exemption rate
v=earning

References

Andersen, T. M. (2012): "Fiscal Sustainability and Demographics – Should we Save or Work More?", *Journal of Macroeconomics 34*, 264–280.

Auerbach, A. J. and Kotlikoff, L. J. (1987): *Dynamic Fiscal Policy, Cambridge*, Cambridge University Press.

Baily, M.N. and Kirkegaard, J.F. (2009): *US Pension Reform: Lessons from other Countries,* Petersons Books.

Barr, N. (1979): "Myth my Grandpa Taught Me", *Three Bank Review*, 124, 27–55.

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

Cremer, H.; De Donder, Ph.; Maldonaldo, D.and Pestieau, P. (2008): "Designing a Linear Pension Scheme with Forced Savings and Wage Heterogeneity", *Journal of Economic Surveys 22*, 213–233.

Cremer, H. and Pestieau, P. (2011): "Myopia, Redistribution and Pensions", *European Economic Review 55*, 165–175.

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

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.

Disney, R. (2004): "Are Contribution to Public Pension Programmes a Tax on Employment?", *Economic Policy 39*, 267–311.

Doerrenberg, Ph.; Duncan, D; Fuest, C. and Peichl, A. (2012): "Nice Guys Finish Last: Are People with Higher Tax Morale Taxed More Heavily?", IZA Discussion Paper 6275.

Engen, E. M., Gale, W. G. and Scholz, J. (1996): "The Illusory Effects of Savings Incentives on Saving", *Journal of Economic Perspectives 10:4*, 111–138.

Epstein, R.A. (2007): "The Neoclassical Economics of Consumer Contract", *New York University Law and Economics Working Papers.*

Fehr, H. and Habermann, C. (2008): "Risk Sharing and Efficiency Implications of Progressive Pension Arrangements", *Scandinavian Journal of Economics, 110*, 419–443.

Fehr, H. and Kindermann, F. (2010): "PensionFunding and Individual Accounts in Economies with Life Cyclers and Myopes", *CESifo Economic Studies 56(3)* 404–443.

Feldstein, M. S. (1985): "The Optimal Level of Social Benefits", *Quarterly Journal of Economics 100*, 302–320.

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

Frey, B. S. and Weck-Hannemann, H. (1984): "The Hidden Economy as an 'Unobserved' Variable", *European Economic Review 26*, 33–53.

van Groezen, B.; Leers, Th. and Meijdam, L. (2003): "Social Security and Endogenous Fertility: Pensions and Child Allowances as Siamese Twins", *Journal of Public Economics 87*, 233–251.

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

Saez, E.; Slemrod, J. and Giertz, S. H. (2012): "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review", *Journal of Economic Literature 50*, 3–50.

Samuelson, P. A. (1937): "A Note on Measurement," *Review of Economic Studies 4,* 156–161.

Sefton, J.; van de Ven, J. and Weale, M. (2008): "Means Testing Retirement Benefits: Fostering Equity or Discourageing Saving?", Economic Journal 118, 556–590.

Shane, F.; Loewenstein, G. and O'Donoghue, T. (2002): "Time Discounting and Time Preferences: A Critical Review," *Journal of Economic Literature 40,* 351–401.

Simonovits, A. (2009a): "When and how to Subsidize the Tax-favored Pension Accounts?", IE-HAS Working Paper, 2.

Simonovits, A. (2009b): "Underreported Earning and Age-Specific Income Redistribution in Post-Socialist Economies", IE-HAS Discussion Papers 27.

Simonovits, A. (2011a): Higher Tax Morale Implies a Higher Optimal Income Tax Rate, IE-HAS Working Paper 37.

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

Simonovits, A (2012a): "Optimal Cap on Pension Contributions", IE-CERS-HAS, Discussion Paper 8.

Simonovits, A. (2012b): "Does Higher Tax Morale Imply a Higher Optimal Labor Income Tax Rate", IE-CERS-HAS Working Paper, 18.

Simonovits, A. (2012c): "Flat or Means-tested Pensions? Pension Credit!", IE-CERS-HAS Working Paper 21.

Simonovits, A. (2012d): "Optimal Linear Redistributive Tax and Pension Systems with Flexible Labor Supply", IE-CERS-HAS Working Paper 33.

Simonovits, A. (2013): "The Optimal Replacement Rate", IE-CERS-HAS Working Paper.

Sunstein, C. R. and Thaler, R. H. (2003): "Libertarian Paternalism is not an Oximoron", *University of Chicago Law Review*, 70, 1159.

Varian, H. R. (1980): "Redistributive Taxation as Social Insurance", *Journal of Public Economics 14*, 49–68.

Weil, Ph. (2008): Overlapping Generations: The First Jubilee, *Journal of Economic Perspectives 22:4*, 115–134.

World Bank (1994): *Averting the Old-Age Crisis*, New York, N.Y., Oxford University Press.