



Working Paper 147/14

**PERSONAL INCOME TAX REFORMS: A GENETIC
ALGORITHM APPROACH**

**Matteo Morini
Simone Pellegrino**

Personal Income Tax Reforms: a Genetic Algorithm Approach

November 29th, 2014

Matteo Morini

ENS Lyon, Institut Rhône-Alpin des Systèmes Complexes (IXXI), Lyon (FR)
University of Torino, Department of Economics and Statistics, Torino (IT)

Simone Pellegrino

University of Torino, Department of Economics and Statistics, Torino (IT)

Abstract

Given a settled reduction in the present level of tax revenue, and by exploring a very large combinatorial space of tax structures, in this paper we employ a genetic algorithm in order to determine the ‘best’ structure of a real world personal income tax that allows for the maximization of the redistributive effect of the tax, while preventing all taxpayers being worse off than with the present tax structure. We take Italy as a case study.

JEL Codes: C63, C81, H23, H24

Keywords: Personal Income Taxation, Genetic Algorithms, Micro-simulation Models, Reynolds-Smolensky Index, Tax Reforms

1. Introduction¹

Personal income tax (hereafter, PIT) around the world is characterized by several parameters that define its structure: marginal tax rates, upper limits of the thresholds, allowances and deductions, as well as tax credits. Applied to the distribution of income observed in a specific country, the PIT structure of that country determines a given tax revenue and a given redistributive effect.

Starting from this position, a Government may want to cut down PIT revenue in order to increase the purchasing power of taxpayers; conversely, it may want to increase the redistributive effect of the tax leaving the tax revenue unchanged; finally, it may want to increase tax revenue by letting the richest taxpayers face all the tax increase.

In order to achieve one of these specific targets, how should the whole PIT structure change? For several reasons, it is not the case that policy makers consider this question when thinking of a PIT reform. The PIT structure observed in a country is indeed the result of several and partial adjustments that have occurred over the past years and, given a revenue constraint, whether those tax reforms were aimed at achieving the best way to obtain the specific target is debatable.

With respect to these arguments, Italy is perfect case study: the Italian PIT is very complicated and its structure incorporates more than thirty parameters. Moreover, in order to increase the purchasing power of ‘poor’ PIT taxpayers, as well as taxpayers belonging to the ‘middle class’ (a proxy of the redistributive effect maximization), the Italian Government recently reduced the PIT revenue by about 9.31 billion euros² by introducing a cash transfer of 80 euros per month, only for employees with a PIT gross income in the range of 8-26 thousand euros (about 10.9 million taxpayers).

Then two questions arise. Is this tax cut allocation the best one the Government could have considered? Or, given this settled amount of the tax cut, which is the best way to reform the whole PIT structure in order to achieve the highest redistributive effect, whilst having no taxpayers being worse off, with respect to the present tax structure?

¹ This is a revised version of the working paper “Personal Income Tax Reforms: a Genetic Algorithm Approach” by Matteo Morini and Simone Pellegrino. Some of the results have been updated to reflect better results provided by the algorithm. An earlier version (September 2014) of this working paper can be found at <https://ideas.repec.org/p/tur/wpapnw/026.html>. The paper was presented at the XXVI Annual Conference of the Italian Association of Public Economics, Pavia, September 2014. We would like to thank Tito Boeri, Paola Profeta, Federico Revelli and Ivica Urban for their useful comments that helped us improve the paper.

² According to the official statistics made available by the Government, the tax cut amounts to 9.5 billion euros. Here we consider the tax cut resulting from the micro-simulation model employed in this work.

The solution of this problem can face an equity-efficiency trade-off: in order for the redistributive effect to be the highest, the efficiency of the tax, (i.e. the level of the effective marginal tax rates), can worsen. As an example, in this paper we mainly focus on the equity side of the problem. This does not imply that we forget about the efficiency side; we suggest a few constraints to the allowable parameters of tax structure, in order not to arrive at both trivial and inefficient solutions.

To answer the above questions, we rely on a static micro-simulation model written in STATA (technical details are available in Pellegrino *et al.* (2011)) that employs, as input data, those provided by the Bank of Italy in its Survey on Household Income and Wealth, published in 2012 with regard to the 2010 fiscal year. The results of this micro-simulation model are very close to the official statistics made available by the Italian Ministry of Economy and Finance for the 2010 fiscal year; as a consequence, this instrument is suitable for the type of empirical analysis we propose.

The static micro-simulation model had to be re-implemented in a more versatile way, in order for the kind of analysis we are interested in to be feasible. We rewrote the modules of the micro-simulation model evaluating the Italian PIT in Python, a very mainstream language that conveniently allows for the use of parallel computing techniques, distributed across multiple nodes. Python also offers an excellent compromise between agility in programming – providing the developers with several libraries optimized for numerical calculations – and computational performance.

We then employed a genetic algorithm; that is, a search heuristic inspired by natural selection, well suited to the identification of the most promising solution to the problem under consideration. We were interested in coming up with a reasonable tax structure that inhibits both trivial and inefficient solutions. The genetic algorithm then had to be provided with a few specific constraints that had to be obeyed, in terms of some parameters of the tax structure. If this were not the case, problematic solutions would appear. For example, having to find the highest redistributive effect with no constraints at all, the genetic algorithm would certainly impose excessively high marginal tax rates on higher income earners and a zero marginal tax rate on too many of the poorest taxpayers; as a result, a polarization of tax rates and bandwidths of thresholds would appear and the tax revenue would consequently be too high. Or the genetic algorithm would disproportionately favour high levels of some peculiar tax credits, simply

because they are enjoyed by a small group of taxpayers, resulting in a negligible impact on the tax revenue, but in awkward preferential treatment for some income groups.

In order to avoid these unpleasant outcomes, we imposed two constraints.

First, we impose a condition that no taxpayers should be worse off as a result of the tax reform; that is, all taxpayers must pay a lower (or, at most, equal) amount of taxes than the present one. Therefore, since the Italian personal income tax does not allow for negative income taxation, and as we are looking for a tax reform with no losing taxpayers, we let the 'no tax' area be greater, or at least equal, to the present one. We also require the highest marginal tax rate, as well as the lower limit of the top threshold, to be lower, or at most equal, to the present values.

Second, we keep the rank applied by the present tax structure to certain kinds of tax credits unchanged: for example, the present tax credit applied to employees is greater than the one applied to pensioners, and the one applied to pensioners is greater than that which is applied to self-employed taxpayers; similarly, the tax credit for tenants is greater with regard to younger ones.

In doing so, we do not allow the genetic algorithm to run free with too 'imaginative' solutions.

Then, we have to define our target: we are interested in obtaining the highest possible redistributive effect of the tax. To measure it, we refer to the Reynolds-Smolensky index, given by the difference between the Gini coefficient for the pre-tax income distribution and the corresponding concentration coefficient for the post-tax distribution. Instead of the equivalent household gross and net income distributions, we refer to the taxpayers' ones. The main reasons are twofold: the cash transfer introduced by the Italian Government favours taxpayers; this is the first exercise that employs a genetic algorithm for a tax system optimization, so that by observing the composition of the tax cut by income classes, we can ensure that our result is among the 'best'. By referring to the equivalent household income distribution, this check could be much harder to assess.

We then let the genetic algorithm set up a 'population' of 200 different tax structures, each of them composed of 33 different parameters defining the present structure of the Italian personal income tax, and we let it evolve for 10 thousands of generations until the desired result was achieved.

Even if by employing the ‘best’ tax structure no taxpayer is worse off, its actual applicability could face political resistance, since all parameters of the tax change and, consequently, taxpayers could hardly believe that no one is worse off. We do not discuss these political economy inconveniences. Finally, it has to be noted that here we also do not consider taxpayers’ responses to the new parameters of the tax structures: it is a ‘short run’ solution that can help policy makers when they think of a PIT reform. In order to consider taxpayers’ responses, agent-based models could be employed. This is the baseline of our further research.

The structure of the paper is as follows. Section 2 describes in greater detail the 2010 structure of the Italian personal income tax, the baseline for our analysis. Section 3 briefly presents how tax progressivity, and the redistributive effect exerted by the tax, can be measured. Section 4 shows the data and peculiarities of the static micro-simulation model employed for simulations. Section 5 first describes how genetic algorithms work and then presents the implementation used in this work. Section 6 shows the results, whilst section 7 offers a conclusion.

2. The Personal Income Tax in the 2010 Fiscal Year: Technical Details

Let x_i be the personal gross income of taxpayer i ($i = 1, 2, \dots, n$). The 2010 Italian tax law considers two different kinds of deductions: d_i^1 is deduction for the main residence cadastral income; d_i^2 is the sum of deductions for social security contributions and alimonies as well as donations. The taxable income y_i is evaluated as:

$$y_i = \begin{cases} x_i - d_i^1 - d_i^2 & \text{if } d_i^1 + d_i^2 < x_i \\ 0 & \text{if } d_i^1 + d_i^2 \geq x_i \end{cases} \quad (1)$$

From 2007 onwards the rate schedule $S(y_i)$ contemplates 5 thresholds as reported in Table 1.

TABLE 1 AROUND HERE

The upper limits $UL_j = LL_{j+1}$ ($j=1,2,3,4$) of thresholds are 15, 28, 55, 75 thousand euros, being the first lower limit $LL_1 = 0$; tax rates t_j range between 23 and 43 percent.

By applying the rate schedule to the tax base the gross tax liability GT_i is obtained.

In order to determine the net tax liability T_i , tax law admits three distinct kinds of effective tax credits. They are: tax credits for earned income $c_i^1(x_i^{MR})$, tax credits for dependent individuals within the household $c_i^2(x_i^{MR})$, and tax credits for items of expenditure c_i^3 , where $x_i^{MR} = x_i - d_i^1$.

The net tax liability T_i is then evaluated as:

$$T_i = \begin{cases} GT_i - c_i^1(x_i^{MR}) - c_i^2(x_i^{MR}) - c_i^3 & \text{if } GT_i > c_i^1(x_i^{MR}) + c_i^2(x_i^{MR}) + c_i^3 \\ 0 & \text{if } GT_i \leq c_i^1(x_i^{MR}) + c_i^2(x_i^{MR}) + c_i^3 \end{cases} \quad (2)$$

In what follows we do not consider regional and municipal surtaxes and then we evaluate taxpayer i 's net income as $z_i = x_i - T_i$.

Focusing on tax credits for employees and pensioners as well as self-employed,

$$c_i^1(x_i^{MR}) = \begin{cases} t_1 m_r & \text{if } x_i^{MR} \leq m_r \\ (t_1 m_r - a_r) + a_r \left(\frac{LL_2 - x_i^{MR}}{LL_2 - m_r} \right) & \text{if } m_r < x_i^{MR} \leq LL_2 \\ (t_1 m_r - a_r) \left(\frac{LL_4 - x_i^{MR}}{LL_4 - LL_2} \right) + b & \text{if } LL_2 < x_i^{MR} \leq LL_4 \\ 0 & \text{if } x_i^{MR} > LL_4 \end{cases} \quad (3)$$

where t_1 is the lowest marginal tax rate (23 percent); m_r with $r = (1,2,3,4)$ (the level of x_i^{MR} below which taxpayer has a nil net tax liability) is equal to 8,000 euros for employees (m_1), 7,500 for pensioners younger than 75 (m_2), 7,750 for pensioners older than 75 (m_3), 4,800 for the self-employed (m_4), and zero for non-working taxpayers; a_r is equal to 502 euros for employees (a_1), 470 for pensioners younger than 75 (a_2), 486 for pensioners older than 75 (a_3), zero for self-employed (a_4); b , that ranges from 10 to 40 euros in the bandwidth 23-28 thousand euros, is applied only to employees (as discussed later, we always set $b = 0$ in simulations). Non-working taxpayers have no

tax credit for earned incomes. Finally, this tax credit decreases from zero to m_4 , and from m_4 to LL_4 only for self-employed taxpayers.

Four different tax credits for type of relationship are allowed: tax credit for dependent children $c_i^{2H}(x_i^{MR})$, further tax credit for households with more than three children c_i^{2HF} , tax credit for dependent spouse $c_i^{2S}(x_i^{MR})$, and tax credit for other household components $c_i^{2O}(x_i^{MR})$. The overall value for $c_i^2(x_i^{MR})$ is then $c_i^2(x_i^{MR}) = c_i^{2H}(x_i^{MR}) + c_i^{2HF} + c_i^{2S}(x_i^{MR}) + c_i^{2O}(x_i^{MR})$. In particular,

$$c_i^{2H}(x_i^{MR}) = \begin{cases} c_i^{2Hp} \frac{q + (f-1)e - x_i^{MR}}{q + (f-1)e} & \text{if } 0 < x_i^{MR} \leq q + (f-1)e \\ 0 & \text{if } x_i^{MR} > q + (f-1)e \end{cases} \quad (4)$$

where $f = \sum_{l=1}^4 f_l$ is the overall number of dependent children; f_1 is the number of dependent children older than 3 years if the dependent children within the household are 3 or less; f_2 is the number of dependent children younger than 3 years if the dependent children within the household are 3 or less; f_3 is the number of dependent children older than 3 years if the dependent children within the household are more than 3; f_4 is the number of dependent children younger than 3 years if the dependent children within the household are more than 3; e is equal to 15,000 euros; q is equal to 95,000; $c_i^{2Hp} = \sum_{l=1}^4 f_l c_i^{2Hpl}$; the present values for the potential tax credits are: $c_i^{2Hp1} = 800$, $c_i^{2Hp2} = 900$, $c_i^{2Hp3} = 1,000$, $c_i^{2Hp4} = 1,100$ euros.

Moreover, whenever $c_i^{2H}(x_i^{MR}) > 0$ and the dependent children within the households are more than 3 the tax law admits a further tax credit c_i^{2HF} equal to 1,200 euros for all beneficiaries. The tax credits for dependent children have to be split between spouses whenever both of them have a positive gross income. Finally,

$$c_i^{2S}(x_i^{MR}) = \begin{cases} c_i^{2Sp} - u \frac{x_i^{MR}}{LL_2} & \text{if } x_i^{MR} \leq LL_2 \\ c_i^{2Sp} - u & \text{if } LL_2 < x_i^{MR} \leq w \\ (c_i^{2Sp} - u) \frac{k - x_i^{MR}}{k - w} & \text{if } w < x_i^{MR} \leq k \\ 0 & \text{if } x_i^{MR} > k \end{cases} \quad (5)$$

and

$$c_i^{2O}(x_i^{MR}) = \begin{cases} c_i^{2Op} \frac{k - x_i^{MR}}{k} & \text{if } x_i^{MR} \leq k \\ 0 & \text{if } x_i^{MR} > k \end{cases} \quad (6)$$

where u is equal to 110 euros, w is equal to 40 thousand euros, k is equal to 80 thousand euros, c_i^{2Sp} is equal to 800 euros and c_i^{2Op} is equal to 750 euros. The present tax code considers higher values than $c_i^{2Sp} - u$ in the income range 29,000-35,200 euros. Instead of 690 euros, in this income range values ranging from 700 to 720 euros are applied. We do not consider these differences in simulations, always letting $c_i^{2Sp} - u$ be equal to a same value.

Tax credits for items of expenditures c_i^3 can be classified in two groups according to the percentage of the expense the tax law admits as a tax credit. There are expenses that allow a tax credit of 19 percent and 36 percent, respectively.³ The 19 percent tax credits (we label this variable *expenditure*₁) are very large, 19 different cases, such as expenses for health care, mortgage interests, etc.; 36 percent tax credits (*expenditure*₂) are allowed for home restructuring-related expenses. All together, tax law admits 30 different tax credits for items of expenditure. Finally, tax law admits a tax credit for tenants; it is 300 euros if $x_i \leq 15,494$ (we label this variable *tenants*₁); 150 if $15,494 < x_i \leq 30,987$ (*tenants*₂); 992 euros if $x_i \leq 15,494$ and if the taxpayers are younger than 30 (*tenants*₃).

³ The tax code considers also a 55 percent tax credit for interventions for energy saving and a 20 percent tax credit for purchasing of a washing machine. Because of the low number of taxpayers interested in these two kinds of tax credits, we did not considered them in the micro-simulation model.

3. Distribution of Income and Personal Income Tax Progressivity

Let x_1, x_2, \dots, x_n be the pre-tax income levels associated to n income units. The corresponding post-tax income levels and tax levels are z_1, z_2, \dots, z_n and T_1, T_2, \dots, T_n , respectively. We denote the pre-tax and the post-tax income distribution as well as the tax distribution by X, Z and T , respectively.

As is well known, inequality among pre- and post-tax income levels as well as tax levels can be evaluated by the Gini coefficient. Let G_X, G_Z and G_T be the corresponding Gini coefficient for pre-tax income, post-tax incomes and taxes, respectively. Then,

$$G_\varepsilon = \frac{2\text{cov}[\varepsilon, F(\varepsilon)]}{\mu_\varepsilon} \quad (7)$$

where $\varepsilon = X, Z, T$, μ_ε is the average value for pre-tax and post-tax incomes and taxes, cov represents the covariance and $F(\varepsilon)$ is the cumulative distribution function.

After the tax, it is not guaranteed that post-tax ordering be equal to the pre-tax income one. Indeed, it is most likely that these two orderings differ because of the re-ranking due to the tax. Therefore, the inequality of Z and T can be evaluated once these distributions are ordered according to the corresponding pre-tax incomes, ranked in a non-decreasing order. For what concerns post-tax incomes and taxes, the corresponding concentration coefficient can then be evaluated as follows:

$$C_\varepsilon = \frac{2\text{cov}[\varepsilon, F(X)]}{\mu_\varepsilon} \quad (8)$$

Progressive taxation produces two different effects on the distribution of pre-tax incomes: post-tax income inequality is lower than that measured on pre-tax income distribution, whilst tax inequality is greater. The first effect is known as the redistributive effect of the tax and the second one as departure from proportionality of the progressive taxation (Lambert, 2001). The overall redistributive effect of the tax RE can be evaluated as

$$RE = G_X - G_Z = (G_X - C_Z) - (G_Z - C_Z) = RS - R^{APK} \quad (9)$$

where $RS = G_X - C_Z$ is the Reynolds-Smolensky index, whilst $R^{APK} = G_Z - C_Z$ is the Atkinson-Plotnick-Kakwani index. The more the tax is progressive, the greater RE and RS ; the more the tax causes re-ranking, the greater the negative contribution of re-

ranking to the overall redistributive effect. Note that if the tax does not cause re-ranking ($R^{APK} = 0$), then $RE = RS$.

The departure from proportionality of the progressive taxation can instead be evaluated by the Kakwani index $K = C_T - G_x$. The Kakwani and the Reynolds-Smolensky indexes are linked by the overall average tax rate θ , namely

$$\theta = \frac{\sum_{i=1}^n T_i}{\sum_{i=1}^n x_i} \quad (10)$$

As a consequence, $RS = \frac{\theta}{1-\theta} K$. This formula tells us that the Reynolds-Smolensky

index has two determinants: the overall average tax rate and the Kakwani index.

In what follows we focus only on the Reynolds-Smolensky index and then we are interested in finding the ‘best’ tax structure able to determine a given tax revenue (smaller than the present one) and to yield to the greatest RS while getting no loser taxpayers.

Since we impose a reduction of the tax revenue, note that the value of θ will be smaller than the present one. Note also that the simulated K will be greater than the present one in order for RS to be the highest.

4. The Data and the Static Micro-simulation Model

The micro-simulation model used in this work estimated the most important taxes and contributions which characterise the Italian fiscal system. Here we employ the micro-simulation model module concerning the PIT. It considers as input data those provided by the Bank of Italy in its 2012 Survey on Household Income and Wealth (hereafter, BI-SHIW). The survey contains information on household income and wealth in the year 2010, covering 7,951 households and 19,836 individuals (Bank of Italy, 2012). The sample is representative of the Italian population, composed of about 24 million households and 60 million individuals.

The BI-SHIW provides information only on each individual's disposable income, which considers items of income that are taxed within the PIT or that can be exempt from the tax, as well as can be taxed under a separate regime. Therefore, the micro-simulation model first distinguishes all incomes included in the PIT taxable income definition, incomes exempt from any taxes and incomes taxed under a separate regime. Then the PIT gross income distribution is evaluated, starting from the PIT net income distribution. The transition from the post- to the pre-tax personal income of each individual has been computed by applying the algorithm proposed by Immervoll and O'Donoghue (2001).

Using original sample weights, the grossing-up procedure simply proportions the sum of individuals' sample weights to the dimension of the population as estimated by the National Statistical Office (ISTAT). Then the grossed-up number of PIT taxpayers has been obtained by considering individuals with a positive gross income within the sample (13,791 taxpayers, corresponding to about 40 million in the population).

Considering the income units, results concerning the PIT gross income distribution are very close to the Ministry of Finance's (2011) official statistics, both considering the gross income distribution by income classes and the composition of PIT income units by work status, as well as by their mean gross income. In addition, the overall tax revenue resulting from the micro-simulation model (148.75 billion euros) is very close to that showed in the official statistics. Considering all individual taxpayers, Figure 1 compares the frequency density function obtained with the micro-simulation model and the one obtained using the Ministry of Finance's official data by income classes. Similar pictures emerge considering the frequency density function for pensioners and employees, as well as the self-employed.

FIGURE 1 AROUND HERE

The column 'Present value' of Table 2 shows the inequality indices for individual taxpayers in the 2010 fiscal year, which is our reference situation for the Reynolds-Smolensky index maximization. The Gini coefficient for the gross income distribution is 0.44338, whilst that for the net income distribution is 0.39138. The overall redistributive effect *RE* is 0.05200. The concentration coefficient for the net income

distribution is 0.39076, whilst that on the net tax liability distribution is 0.67215; therefore, the Reynolds-Smolensky RS index is equal to 0.05262 and the Kakwani index K is 0.22877. The overall average tax rate is 0.18699, whilst the Atkinson-Plotnick-Kakwani R^{APK} index is equal to 0.00062.

TABLE 2 AROUND HERE

Table 2 (column ‘With cash transfer’) also shows the overall redistributive impact of PIT joined with the cash transfer Φ_i . Φ_i favours only employees with a PIT gross income in the range of 8-26 thousand euros (about 10.9 million taxpayers), as follows:

$$\Phi_i = \begin{cases} 960 & \text{if } m_1 < x_i \leq 24,000 \ \& \ GT_i - c_i^1(x_i^{MR}) > 0 \\ 960 \frac{26,000 - x_i}{2,000} & \text{if } 24,000 < x_i \leq 26,000 \ \& \ GT_i - c_i^1(x_i^{MR}) > 0 \\ 0 & \text{if } x_i > 26,000 \ \& \ GT_i - c_i^1(x_i^{MR}) > 0 \end{cases}$$

It does not modify the PIT structure at all. Beneficiaries obtain Φ_i , and continue to pay the same amount of PIT net tax liability T_i . The net effect $z_i^* = y_i - T_i + \Phi_i$ is an increase in these taxpayers’ disposable income by 960 euros per year, in the income range 8-24 thousands euros and a decreasing amount up to 26 thousand euros.⁴ Note also that if $T_i < \Phi_i$ beneficiaries obtain a subsidy. Taxpayers other than employees, on the contrary, do not gain from this cash transfer. The cost of this measure is 9.31 billion euros.

Even if only one taxpayer out of four obtains the cash transfer, the redistributive effect of the PIT joined with the transfer considerably improves: the Reynolds-Smolensky RS index increases by 8.5 per cent; on the contrary, the overall redistributive effect RE increases only by 7.9 per cent, since the Atkinson-Plotnik-Kakwani R^{APK} index worsens by about 53 per cent. On the efficiency side, the effective marginal tax rate resulting in the income range of 24-26 thousand euros increases by up to 80 per cent;

⁴ There is a political explanation behind this choice. The Government announced that it would reduce PIT tax liability by 80 euros per month for all taxpayers. Of course, such an announcement proved to be too expensive; moreover, it would have been very difficult to reach through a PIT structure reform, since it is hard to reform such a complex tax structure and, in the meantime, ensure an equal tax reduction for all taxpayers. In order to at least partially meet its commitment, the Government decided to apply the 80 euros pledge only to a group of taxpayers, given the revenue constraint of 9.31 billion euros. In order not to modify the PIT structure, it chose the cash transfer instrument.

for all other income levels, on the contrary, the effective marginal tax rates do not change with respect to the present tax structure.

By employing a genetic algorithm, in the next sections, we show that a more equity-oriented and more efficient reform is possible. From the methodological point of view, the specific measure employed by the Italian Government is not particularly interesting; it refers only to employees and then, a few parameters of the tax could be simultaneously changed in a genetic algorithm framework, i.e. parameters defining the structure of the tax credit for employees. We discuss an overall reform, which considers all the parameters of the tax are able to be trimmed. Since we split the tax cut between all taxpayers, both the average reduction of the net tax liability for all taxpayers and the value of R^{APK} are smaller than that guaranteed to employees by the Government's cash transfer.

5. Genetic Algorithm

5.1. General Overview

Genetic algorithms (henceforth, GAs) are a heuristic search which belong to the field of evolutionary algorithms, a subfield of artificial intelligence. Since their inception (Holland, 1975), GAs found a wealth of applications in the most varied research disciplines, beyond computational science, mathematics, physics, bioinformatics, etc. Applications in economics also exist, broadly including game theory, finance related works, schedule optimization and whenever some sort of learning mechanism is needed; historically, the first attempt at employing GAs in economics is due to Miller's (1986) research on adaptive behaviour.

To our knowledge, no previous attempts at employing GAs for tax system optimization exist, to date.

The huge solutions search space, which is the aftermath of the combinatorial effect of very many parameters, posing a serious challenge to traditional optimization techniques; brute force methods are out of the question, just like iterative methods (cfr. Newton's); GAs appear as an obviously appropriate choice.

Candidate solutions, which in the GA are internally represented as ‘individuals’ (each of them is characterized by her own ‘genome’, a vector of ‘chromosomes’)⁵, are generated as an initial ‘population’ at random. Evolutionary operators iteratively select, cross-breed and mutate the best (most ‘fit’, according to an objective function called ‘fitness function’) ‘individuals’, in order to produce an offspring of ‘individuals’ – the subsequent ‘generation’ – that will enter a new reproduction step. The ‘individuals’ average fitness increases after every generation, until a satisfactory solution is found. The stopping criterion normally employed is related to the ‘population’ homogeneity: as the search process becomes closer to an optimum, the ‘individuals’ become more and more similar among them.

The GA implementation employed in this work is based on Python’s open-source Pyevolve library (Perone, 2009); the ‘population’ selection mechanism across generations is the standard roulette wheel (fitness proportional) with elitism selection (the very best individuals in each generation are kept unaltered and carried over across successive generations), while the evolutionary crossover operator is a standard one-point. A low mutation rate value and a high crossover rate have also been utilized, in order to let the search process converge reasonably quickly on solutions, whilst maintaining the ability to escape local maxima. The “population size vs. number of generations” trade-off has been tackled and solved, favouring a small population vs. numerous evolutionary steps.

Detailing the trimming of the GAs technical parameters is out of the scope of this work; suffice it to say, as agreed upon by a vast literature, it is an ad-hoc process, to be performed mostly by trial and error, on every specific search domain.

5.2. The Structure of the Genetic Algorithm We Employ

As the starting point, we let the GA set up a ‘population’ of 200 different tax structures (‘individuals’ from the GA point of view) and then we let it evolve them for 10,000 ‘generations’. Consequently, the GA has to evaluate as much as 2 million candidate solutions, applying all these different tax structures to the same pre-tax income

⁵ In the early GA implementations (BCGA, binary-coded genetic algorithms), the solutions space had to be coded in binary numbers; RCGA, real-coded genetic algorithms, allow working with variables in continuous domains; cfr. Herrera et al. (1998).

distribution, composed of 13,791 taxpayers. We set the crossover rate as equal to 0.75, as well as the mutation rate as equal to 0.05 and the elitism selection equal to 7.

For each tax structure and each taxpayer, the GA computes all the relevant tax variables in the transition from the pre- to the post-tax income. For each tax structure, it then computes the overall tax revenue, the share of loser taxpayers by considering each taxpayer's actual net tax liability, the average loss for the loser taxpayers, as well as the Reynolds-Smolensky *RS* index (the four parts of the objective fitness function to maximize, see below); it then saves all these resulting values, in addition to all the parameters of the tax, on a dump file. We employ a computer powerful enough to evaluate 2 million runs in about 4 days; the duration of our average run is then 0.18 seconds.

The GA has to maximize a fitness function. We employed

$$fitness = e^{\frac{\alpha RS - \beta \Delta - \delta \Pi - \lambda \Omega}{10}} \quad (11)$$

where Δ is the percentage deviation of the computed tax revenue of each run from the target one (139.44 billion euros, 9.31 billion euros less than the present tax revenue), whilst Π is the share of taxpayers losing with the simulated tax structure, Ω is the average loss (in euros) for the loser taxpayers, α , β , δ and λ are all positive parameters. We fix $\alpha = 575$, $\beta = 53$, $\delta = 20$ and $\lambda = 150^{-1}$. We made several attempts, with different parameters and different functional forms; this fitness function has proved to be the most effective.

The exponential form of the fitness function helps the GA to always converge to the 'best' solution, since generation after generation, more-than-linearly high scores are assigned to the most promising candidate solutions.

The first term of the exponent shows the part of the fitness function depending on *RS*. We highly favour *RS* with respect to Δ , Π , and Ω , since we are interested in obtaining the highest Reynolds-Smolensky index. The second and third, as well as the fourth terms of the exponent of the fitness function, show the 'penalties' we impose on the fitness value when Δ , Π and Ω became too large. As a result, the smaller Δ , Π and Ω , the more the fitness value increases. The parameter λ is crucial for the convergence of the GA: there are combinatorial spaces where *RS* increases even if Π increases; a low value for λ allows the GA to attribute low scores to those candidate solutions.

When the GA has evolved for a reasonable number of generations, we obtain $\Delta \approx 0$, $\Pi \approx 0$, the smallest Ω , and the highest RS . The parameters of the ‘best’ structure of the tax can then be observed.

Figure 2 shows the maximum value of the fitness function applied in this work for each generation. As can be noted, the maximum value tends to increase generation after generation; top 1 per thousand highest values are reached starting from generation 2,500. The highest fitness value is 25.47, and it is reached in generation 8,914.

FIGURE 2 AROUND HERE

Each candidate tax structure is characterized by 33 different parameters (see Table 3), each of them related to a specific parameter of the Italian PIT structure described in Section 2. In order to evaluate each of the 33 parameters, the GA couples 36 ‘chromosomes’ that are values ranging from zero to 1.

We now turn to describing how we let the GA trim each ‘chromosome’.

First of all, we let the GA choose five marginal tax rates as in the actual tax code. Given the constraints we impose, we know that the top marginal tax rate cannot be too much higher than the present one (since we impose a ‘no loser’ taxpayers’ constraint); moreover, we do not want it to be higher than the present value. Conversely, we do not know the minimum allowable value of the bottom marginal tax rate; we set the lowest marginal tax rate at not lower than 10 per cent (being the present value equal to 23 per cent).

In particular, the GA randomly sets a group of six chromosomes serving for the definition of the five marginal tax rates. It then adds them up in order to obtain a

normalization value as follows: $norm_t = \frac{0.43 - 0.10}{\sum_{\varphi=1}^6 chromosome_{\varphi}}$.

The GA finally chooses the five tax rates t_{λ} with $\lambda = (1, 2, 3, 4, 5)$ as follows:

$$t_{\lambda} = 0.10 + \sum_{\varphi=1}^{\lambda} chromosome_{\varphi} * norm_t.$$

We then set a second group of 5 chromosomes (7-11) defining the four upper limits $UL_j = LL_{j+1}$ of the thresholds, being $LL_1 = 0$ by definition and $UL_j < UL_{j+1}$. We applied an empirical strategy similar to that employed for the definition of the marginal tax rates and we impose $UL_1 \geq 15,001$ (see on for the choice of this value) and we let the highest value of UL_4 be 75 thousand euros (as in the present tax structure). As a consequence,

$$norm_{UL} = \frac{75,000 - 15,001}{\sum_{\varphi=7}^{11} chromosome_{\varphi}}$$

and, for $j = 1, 2, 3, 4$,

$$UL_j = 15,001 + \sum_{\varphi=1}^j chromosome_{\varphi} * norm_{UL}.$$

Afterwards, we define 25 chromosomes related to the tax credits' structure. Starting at the tax credit for employees (Equation (3)), we let the GA choose the no tax area m_1 applied to employees, that is the limit of pre-tax income below which these taxpayers face a zero net tax liability, between 8 thousand euros (the present value) and $\psi * UL_1$:

$m_1 = 8,000 + chromosome_{12} * (\psi * UL_1 - 8,000)$. We set $\psi = \frac{8}{15}$, equal to the present value. As a consequence, UL_1 has to be set greater than or, at most equal, to 15 thousand euros being $\frac{8}{15} * 15,000 = 8,000 = m_1$.

Note that the choice of the parameter ψ influences the equity-efficiency trade-off. The higher the value chosen by the GA, the more likely the slope of the effective marginal tax credit in the income level $m_r - LL_2$ is high; therefore, the higher the effective marginal tax rate for this income bandwidth.⁶ The constraint we impose on the share of loser taxpayers lets the GA chose the highest admissible value for m_1 and at the same time keep under control the level of the effective marginal tax rates (see on).

The GA then choose m_2, m_3, m_4 as follows:

⁶ If we let the GA to run free with respect to this parameter, for example by letting the GA choose it up to 1, a lower than 1 value would be chosen (given the constraints we impose), a higher redistributive effect would be obtained, as well as a confiscatory effective tax rate for incomes just above the 'no tax' area. The reason is clear: *ceteris paribus*, having to maximize the Reynolds-Smolensky index given a revenue constraint, the larger the share of taxpayers with a nil net tax liability, the higher the Reynolds-Smolensky index (and the higher the tax rates and narrower the upper limit of each threshold in order to obtain the target tax revenue). As a consequence, the larger the 'no tax' area, the narrower the income bandwidth between the 'no tax' area and the upper limit of the bottom threshold, and then the more sharply the subsequent reduction of the effective tax credit.

$$\begin{aligned}
m_2 &= (0.9375 * m_1) - (\text{chromosome}_{13} * 0.2 * 0.9375 * m_1) \\
m_3 &= m_2 + \text{chromosome}_{14} * (0.96875 * m_1 - m_2) \\
m_4 &= 0.6 * m_1 - \text{chromosome}_{15} * 0.6 * m_1
\end{aligned}$$

where 0.2 is arbitrarily chosen, whilst $0.9375 = \frac{m_2}{m_1}$, $0.96875 = \frac{m_3}{m_1}$, and $0.6 = \frac{m_4}{m_1}$ are

the corresponding values according to the present tax structure. In so doing, we let the GA choose m_r with $r = (1, 2, 3, 4)$ in a large combinatorial space preserving the present rank of m_r .

Having the GA chosen t_1 and m_r , note that the potential tax credits $t_1 m_r$ are automatically defined. We then let the GA choose also the parameters a_r with $r = (1, 2, 3)$ in the range $(0 - t_1 m_r)$ by defining chromosomes 16, 17, and 18 (as described in Section 2, parameter a_4 is equal to zero and we keep it unchanged). In so doing, a very large combination of tax credits for earned incomes is allowable.

Finally, we always set the parameter b equal to zero. At present, it is applied only to employees and its values range from 10 to 40 euros for levels of x_i^{MR} belonging to the threshold 23-28 thousand euros. We prefer this parameter to be fixed at zero since, if it were positive, it would not let the tax credit under discussion be a continuous function for all levels of x_i^{MR} .

Finally, note that we do not let the effective tax credits for earned income be piecewise decreasing with respect to limits, others than those observed in the rate schedule. If $c_i^1(x_i^{MR})$ were piecewise decreasing with respect to other thresholds, the number and the level of effective marginal tax rates would not be under control, leading to unpleasant and inefficient outcomes.

We continue by defining specific chromosomes, in order to set the combinatory space for the three tax credits for dependent individuals within the household: $c_i^{2H}(x_i^{MR})$, $c_i^{2S}(x_i^{MR})$ and $c_i^{2O}(x_i^{MR})$.

Starting with the tax credits for dependent children $c_i^{2H}(x_i^{MR})$, we let the GA choose the potential level of the tax credits c_i^{2Hpl} in the range 600-3,000 euros (being the present values ranging between 800 and 1,100 euros). Similar to the choice of the tax rates and the

upper limits of the thresholds, we then define 5 different chromosomes (from 19 to 23) to set the 4 kinds of tax credits for dependent children. In so doing, we set specific constraints, in order to let the potential tax credit be higher for households with more than 3 children and lower for those with fewer than 3 children, as well as higher for children aged 3 or less and lower for a child aged more than 3. Then we introduce chromosomes 24 and 25 for the choice of parameters q and e . We let the GA choose q between 50 and 130 thousand euros (being the present value equal to 95 thousand), and the parameter e between zero and 40 thousand euros (being the present value equal to 15 thousand). Note that these are very large ranges, so that the GA can choose an extremely large set of combinations. Finally, the GA chooses chromosome 26 in order to set c_i^{2HF} between zero and 1,500 euros (being the actual value 1,200 euros).

Turning to the effective tax credit for the spouse $c_i^{2S}(x_i^{MR})$, we generate chromosome 27 in order for c_i^{2Sp} to range between 500 and 2 thousand euros (being the present value equal to 800 euros), and a further chromosome 28 in order for the parameter u to range between zero and c_i^{2Sp} . Chromosomes 27 and 28 let the effective tax credit for the spouse be a non-increasing function with respect to x_i^{MR} and let the GA choose among a very large combination of structures for this tax credit.

Looking at Equation (5), this effective tax credit $c_i^{2S}(x_i^{MR})$ is piecewise linearly decreasing with respect to three thresholds: from zero to LL_2 , from LL_2 to $LL_3 < w < LL_4$, and from w to $k > LL_5$. In order to define w and k , we introduced chromosomes 29 and 30 as follows:

$$w = LL_2 + chromosome_{29} * (LL_4 - LL_2)$$

$$k = w + chromosome_{30} * (80,000 - w)$$

Finally, concerning the tax credit for other dependent individuals within the household, we introduce chromosome 31 in order for c_i^{2Op} to range between $0.75 * c_i^{2Hp1}$ and $0.95 * c_i^{2Hp1}$, and we impose $c_i^{2O}(x_i^{MR})$ to be linearly decreasing between zero and k , and to be zero if $x_i^{MR} > k$.

Afterwards, we let the GA choose chromosomes 32, 33 and 34 in order to set tax credits for tenants:

$$tenants_1 = chromosome_{32} * 1,000$$

$$tenants_2 = chromosome_{33} * tenants_1$$

$$tenants_3 = tenants_1 + chromosome_{34} * (tenants_1 * 2)$$

According to the present tax code, these tax credits are applied to two income thresholds: 15,494 and 30,987 euros; we consider this aspect by letting them change according to UL_1 and UL_2 .

Finally, the GA chooses chromosomes 35 and 36 in order to set the percentage of the expenses the tax law admits as further tax credits for items of expenditure. We let it chose $expenditure_1$ between 0 and 40 per cent (at present equal to 19), and $expenditure_2$ between 20 and 70 per cent (at present equal to 36).

6. Results

Table 3 shows all the parameters of the ‘best’ tax structure able to maximize the Reynolds-Smolensky RS index, given that the tax revenue is 9.31 billion euros lower than the present one and no taxpayers have to be worse off due to the tax reform.

As can be noted, the bottom marginal tax rate t_1 significantly decreases from 23 to 19.42 per cent; this reduction lowers the gross tax liability, not only for the poorest taxpayers but also for all the other taxpayers. t_4 decreases from 41 to 39.95 per cent, whilst t_5 remains unchanged as expected. The other two marginal tax rates increase: t_2 from 27 to 27.96 per cent, and t_3 from 38 to 39.95 per cent. Note that $t_3 = t_4$: according to the GA, the highest redistributive effect can be obtained with 4 instead of 5 thresholds. In terms of the bandwidth of the thresholds, the first one broadens from 0-15,000 to 0-22,871 euros, whilst the second narrows from 15,000-28,000 to 22,871-26,454 euros. Then the original third and fourth thresholds unite and the after-reform third threshold goes from 26,454 to 57,158 euros. Finally, the top marginal tax rate is applied to incomes above 57,158 euros instead of 75 thousand euros.

TABLE 3 AROUND HERE

The ‘no tax’ area enlarges for all the four kinds of taxpayer: m_1 increases from 8,000 to 11,028 euros, m_2 from 7,500 to 10,339, m_3 from 7,750 to 10,684, whilst m_4 rises from 4,800 to 5,116. Note that $\frac{m_1}{UL_1} = 0.48220$ is not set to its maximum value.

The parameters defining the shape of the effective tax credits for earned income considerably increase: a_1 from 502 to 2,142 euros, a_2 from 470 to 2,008, and a_3 from 486 to 2,075 euros. Note that in all cases the GA sets $a_r = t_1 m_r$. This means that the shape of the tax credits for earned incomes changes with respect to those originally observed.

Since we impose the tax credit $c_i^1(x_i^{MR})$ piecewise decreasing with respect to LL_2 and LL_4 , note also that after the tax reform, it is positive up to LL_2 . As an example, Figure 3 compares the effective tax credits for employees before and after the tax reform. After the reform, the slope of the effective tax credit is higher (in absolute value) in the income range $(m_1 - LL_2)$. This shape also affects the level of the effective marginal tax rates (EMTR) in this income bandwidth, which increase with respect to the ones observed before the tax reform (Figure 4). Therefore, an equity-efficiency trade-off emerges: in order for the Reynolds-Smolensky index to be the highest, we have to agree to higher effective marginal tax rates.

FIGURE 3 AROUND HERE

FIGURE 4 AROUND HERE

The shape of the tax credit for a spouse $c_i^{2S}(x_i^{MR})$ is very similar to the present one (Figure 5). The potential tax credit is a little bit higher than before, and the effective one is higher in the income range 0- w ; it becomes zero above 80 thousand euros.

In terms of the tax credits for dependent children, the two tax credits for children if the dependent children within the household are 3 or less are very similar to those observed before the tax reform: c_i^{2Hp1} decreases from 800 to 788 euros, whilst c_i^{2Hp2} increases from

900 to 907. The other two tax credits are significantly higher: c_i^{2Hp3} increases from 1,000 to 1,687, whilst c_i^{2Hp4} from 1,100 to 2,265 euros.

FIGURE 5 AROUND HERE

Note also that since c_i^{2Hp3} and c_i^{2Hp4} are so much higher than before, c_i^{2HF} can be set equal to zero. The income limits above which this tax credit becomes zero, also change: q is equal to 86,803, whilst e is equal to 25,515. A similar picture emerges when considering the tax credit for other dependent individuals within a household: the potential tax credit is a little lower than the present value (748 euros), and is positive for income below 80 thousand euros.

Focusing on the remaining parameters of the tax, the tax credit for tenants with gross income below UL_1 is approximately the same (301 euros against 300); the tax credit for tenants with income in the range $(UL_1 - UL_2)$ is set to 209 euros instead of 150, whilst the tax credit for younger tenants is lower (589 euros instead of 992).

Finally, the percentages of expenses the tax law admits as a tax credit also remain relatively unchanged: 18.54 per cent instead of 19 per cent and 41.17 per cent instead of 36 per cent.

Very few taxpayers are worse off as a result of this tax reform (1.1 per cent and, on average, they lose 43 euros per year), whilst 24.5 per cent are unaffected (we consider as unaffected the taxpayers for whom the absolute value of the computed net tax liability differs from the present one by, at most, one euro). The remaining 74.4 per cent of taxpayers gain from the reform. Looking at the composition of the tax cut in terms of income classes (Table 4), 85.6 per cent of the tax cut favours taxpayers in the income range 8-28 thousands euro, whilst 2.5 per cent favours taxpayers with lower incomes. This is due to the fact that the Italian personal income tax system does not admit negative income taxation; therefore, taxpayers with a nil net tax liability (almost all taxpayers with income lower than 8 thousand euros) are not affected by the tax reform. If the Italian PIT allowed negative income taxation, the tax reform would show a different distribution in terms of the tax cut among income classes: in particular, there would be lower gains for the top income earners and higher gains for the bottom ones.

TABLE 4 AROUND HERE

Only 9 per cent of the tax cut favours taxpayers with incomes in the range 28-55 thousand euros, whilst the remaining 3.1 per cent favours richer taxpayers. It can be observed that the RS could be higher, were the (low) gains of the richer taxpayers transferred to the poorest ones. Given the structure of ‘chromosomes’ described in subsection 5.2, this is not possible, or at most, not likely, since the GA has to balance the effects on RS , Δ , Π and Ω due to 33 parameters.

Finally, Table 5 compares the inequality indexes for taxpayers according to the present tax structure, and those obtained by applying the new structure of the tax to the same pre-tax income distribution. As can be noted, RS is 8.3 per cent higher than the present value: since the overall average tax rate decreases from 18.70 per cent to 17.53 per cent, the Kakwani index increases by 17.2 per cent. Note also that this tax reform positively affects R^{APK} .

TABLE 5 AROUND HERE

7. Concluding Remarks

In this paper, we propose a new methodology to implement a personal income tax reform. In particular, given a settled tax cut decided upon by the Government, (note that a similar strategy can be applied if the tax revenue increases), we show how a genetic algorithm can be employed, in order to find out the values of all parameters defining the structure of the personal income tax able to satisfy a specific target. Our methodology can be applied to any other specific target; as an example, in this work our target is the maximization of the redistributive effect of the tax, while preventing all taxpayers being worse off with respect to the present tax structure. We apply this methodology to the Italian personal income taxation system for two reasons: the tax structure is quite complicated, and recently the Government decided to reduce tax revenue by about 9.31 billion euro starting from 2015. The aim of this tax cut is to increase the purchasing

power of 'poor' taxpayers and taxpayers belonging to the 'middle class', and the instrument is the introduction of a cash transfer (not related to the structure of the personal income tax) only for employees with gross incomes in the range 8-26 thousand euros (in order for the yearly gain to be about one thousand euros), whilst all other kinds of taxpayer are not affected by this money transfer. Here we show that a better and more equity-oriented reform is possible. This methodology allows a short run reform, and can help policy makers when they think of a tax reform.

References

- Bank of Italy (2012). Household Income and Wealth in 2010, Supplements to the Statistical Bulletin, XXII (New Series), No. 6.
- Holland, J. H. (1992). *Adaptation in Natural and Artificial Systems: an Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, MIT Press, Cambridge, MA; first edition (1975). University of Michigan Press, Ann Arbor, MI.
- Herrera F., Lozano M. and Verdegay J.L. (1998). Tackling Real-Coded Genetic Algorithms: Operators and Tools for Behavioural Analysis, *Artificial Intelligence Review*, Vol. 12(4), pp. 265-319.
- Immervoll, H. and O'Donoghue, C. (2001). Imputation of Gross Amounts from Net Incomes in Households Surveys. An application using EUROMOD, Working paper EM1, EUROMOD.
- Lambert, P. J. (2001). *The Distribution and Redistribution of Income*, Third edition, Manchester University Press, Manchester and New York.
- Miller, J. H. (1986). *A Genetic Model of Adaptive Economic Behavior*, University of Michigan Working Paper, Ann Arbor, MI.
- Ministry of Finance, Department of Finance (2011). *Statistical Reports*.
- Pellegrino, S., Piacenza, M. and Turati G. (2011). Developing a Static Microsimulation Model for the Analysis of Housing Taxation in Italy, *The International Journal of Microsimulation*, Vol. 4(2), pp. 73-85.
- Perone, C. S. (2009). Pyevolve: a Python Open-Source Framework for Genetic Algorithms, *ACM SIGEVolution Newsletter*, vol. 4, n. 1, pp. 12-20, ACM, New York.

Figure 1: Frequency density function for all individual taxpayers

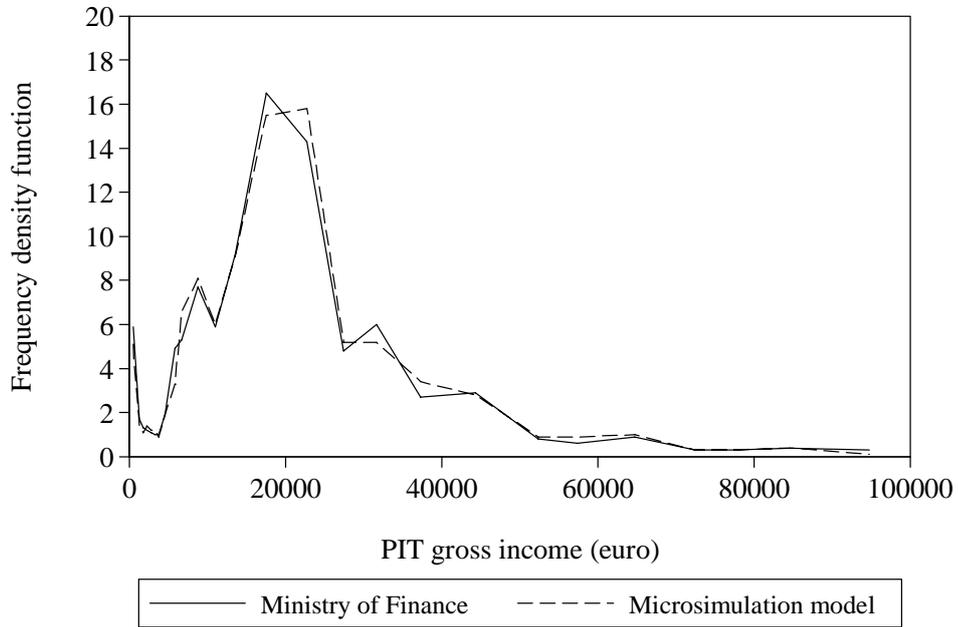


Figure 2: Maximum values of the fitness function

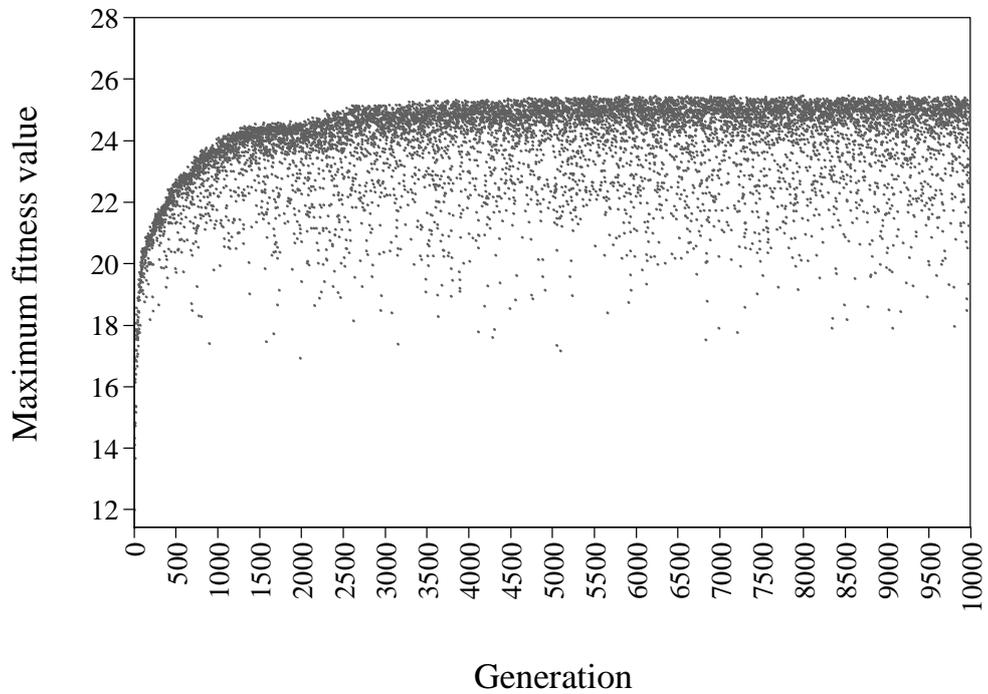


Figure 3: The effective tax credit for an employee

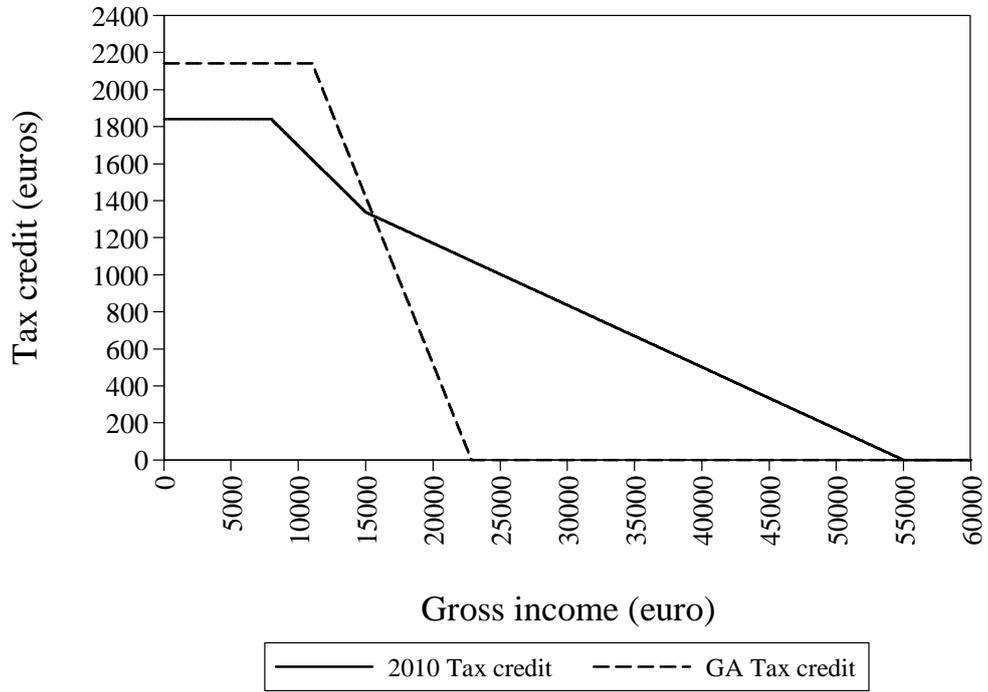


Figure 4: The effective marginal tax rates for a celibate employee without children

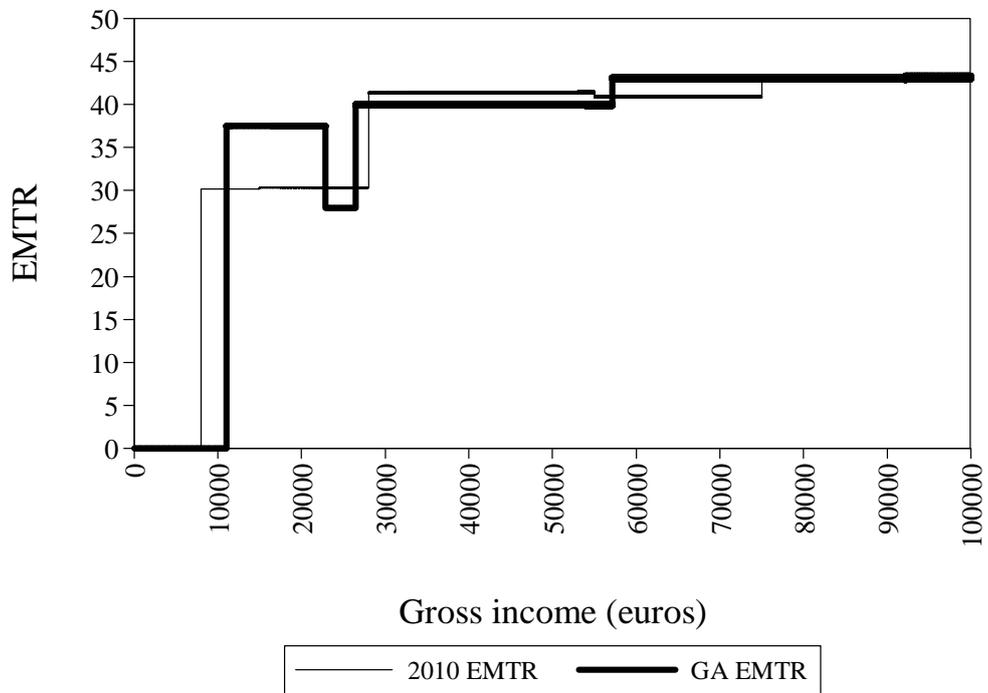


Figure 5: The effective tax credit for a spouse

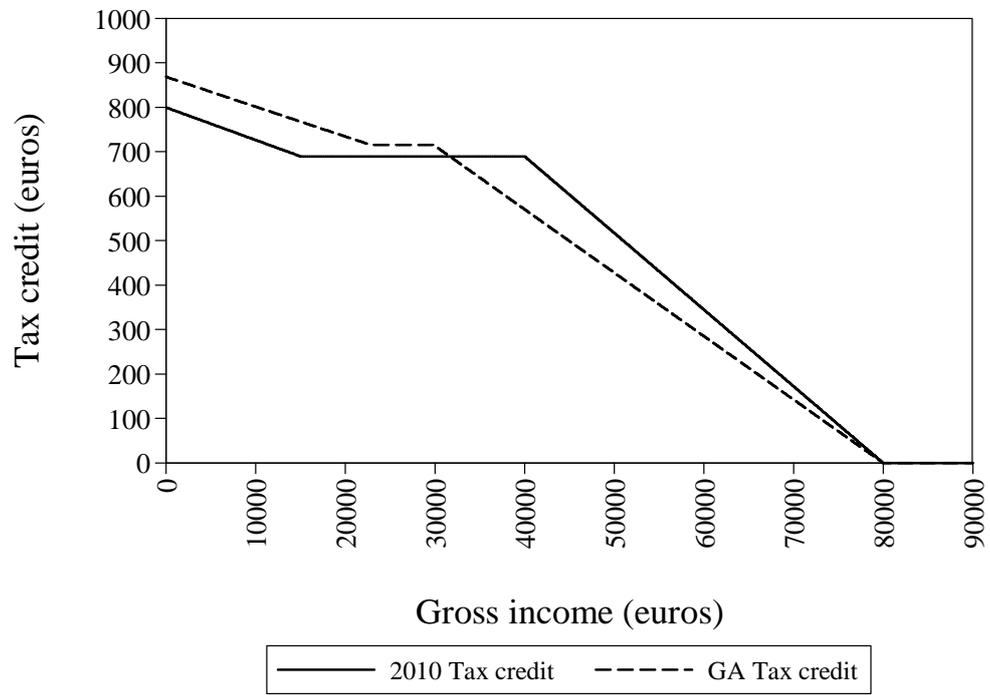


Table 1: Rate schedule

Taxable income (euros)			
Threshold (j)	Lower limit (LL)	Upper limit (UL)	Tax rate (%) (t)
1	0	15,000	23
2	15,000	28,000	27
3	28,000	55,000	38
4	55,000	75,000	41
5	75,000	-	43

Source: Italian Tax Code.

Table 2: Inequality indexes for taxpayers

Index	Present value	With cash transfer	Absolute difference	Percentage difference
Gini coefficient for the gross income	0.44338	0.44338	0.00000	0.0
Gini coefficient for the net income	0.39138	0.38727	-0.00412	-1.1
Concentration coefficient for the net income	0.39076	0.38632	-0.00445	-1.1
Gini coefficient for the net tax liability	0.68150	0.73255	0.05104	7.5
Concentration coefficient for the net tax liability	0.67215	0.71187	0.03972	5.9
Redistributive effect	0.05200	0.05611	0.00412	7.9
Reynolds-Smolensky index	0.05262	0.05707	0.00445	8.5
Kakwani index	0.22877	0.26849	0.03972	17.4
Atkinson-Plotnik-Kakwani index	0.00062	0.00095	0.00033	53.3
Average tax rate (%)	18.70	17.53	-1.17	-6.3

Source: Own elaborations based on BI-SHIW.

Table 3: Present and computed parameters of the tax

Parameters	Present value	Best value
t_1	0.23	0.19421
t_2	0.27	0.27955
t_3	0.38	0.39952
t_4	0.41	0.39952
t_5	0.43	0.43000
UL_1	15,000	22,871.01
UL_2	28,000	26,453.94
UL_3	55,000	44,296.30
UL_4	75,000	57,157.63
m_1	8,000	11,028.41
m_2	7,500	10,339.13
m_3	7,750	10,683.77
m_4	4,800	5,115.92
a_1	502	2,141.85
a_2	470	2,007.98
a_3	486	2,074.91
c_i^{2Sp}	800	868.82
u	110	153.46
w	40,000	29,870.12
k	80,000	80,000.00
c_i^{2Op}	750	748.18
c_i^{2Hp1}	800	787.56
c_i^{2Hp2}	900	907.11
c_i^{2Hp3}	1,000	1,687.32
c_i^{2Hp4}	1,100	2,265.06
q	95,000	86,802.79
e	15,000	25,514.69
c_i^{2HF}	1,200	0.00
$tenants_1$	300	301.47
$tenants_2$	150	208.55
$tenants_3$	992	588.65
$expenditures_1$	0.19	0.18538
$expenditures_2$	0.36	0.41171

Source: Own elaborations based on BI-SHIW.

Table 4: The composition of the tax cut by income classes

Income class (thousand euros)	Composition of the tax cut (%)	Winner (%)	Indifferent (%)	Loser (%)	Total (%)	Average win (euros)	Average loss (euros)
0-8	2.5	6.1	19.3	0.0	25.3	91	0
8-15	40.5	17.3	4.7	0.0	21.9	542	0
15-28	45.1	34.4	0.4	0.5	35.3	294	52
28-55	9.0	13.0	0.1	0.5	13.5	158	39
55-75	2.8	2.2	0.0	0.0	2.2	290	41
above 75	0.3	1.5	0.0	0.1	1.7	42	14
Total	100.0	74.4	24.5	1.1	100.0	307	43

Source: Own elaborations based on BI-SHIW.

Table 5: Inequality indexes for taxpayers

Index	Present value	Best tax structure	Absolute difference	Percentage difference
Gini coefficient for the gross income	0.44338	0.44338	0.00000	0.0
Gini coefficient for the net income	0.39138	0.38696	-0.00442	-1.1
Concentration coefficient for the net income	0.39076	0.38639	-0.00437	-1.1
Gini coefficient for the net tax liability	0.68150	0.72020	0.03870	5.7
Concentration coefficient for the net tax liability	0.67215	0.71150	0.03935	5.9
Redistributive effect	0.05200	0.05642	0.00442	8.5
Reynolds-Smolensky index	0.05262	0.05699	0.00437	8.3
Kakwani index	0.22877	0.26812	0.03935	17.2
Atkinson-Plotnik-Kakwani index	0.00062	0.00057	-0.00005	-8.3
Average tax rate (%)	18.70	17.53	-1.17	-6.3

Source: Own elaborations based on BI-SHIW.

Our papers can be downloaded at:

<http://cerp.carloalberto.org/en/publications>

CeRP Working Paper Series

N° 147/14	Matteo Morini Simone Pellegrino	Personal Income Tax Reforms: a Genetic Algorithm Approach
N° 146/14	Mariacristina Rossi Eva Sierminska	Single again? Asset and portfolio changes due to widowhood shock
N° 145/14	Johannes G. Hoogeveen Mariacristina Rossi Dario Sansone	Drivers of performance in primary education in Togo
N° 144/14	Elsa Fornero	Economic-financial literacy and (sustainable) pension reforms: why the former is a key ingredient for the latter
N° 143/14	Kees de Vaan Daniele Fano Heriaht Mens Giovanna Nicodano	A Reporting Standard for Defined Contribution Pension Plans
N° 142/14	Elisabetta Cagna Giulio Casuccio	Equally-weighted Risk Contribution Portfolios: an empirical study using expected shortfall
N° 141/14	Massimo Baldini Costanza Torricelli Maria Cesira Urzi Brancati	Family ties: occupational responses to cope with a household income shock
N° 140/14	Cecilia Boggio Elsa Fornero Henriette Prast Jose Sanders	Seven Ways to Knit Your Portfolio: Is Investor Communication Neutral?
N° 139/14	Laura Bianchini Margherita Borella	Cognitive Functioning and Retirement in Europe
N° 138/13	Claudio Morana	Insights on the global macro-finance interface: Structural sources of risk factors fluctuations and the cross-section of expected stock returns
N° 137/13	Claudio Morana	New Insights on the US OIS Spreads Term Structure During the Recent Financial Turmoil
N° 136/13	Anna Lo Prete	Inequality and the finance you know: does economic literacy matter?
N° 135/13	Rik Dillingh Henriette Prast Mariacristina Rossi Cesira Urzi Brancati	The psychology and economics of reverse mortgage attitudes: evidence from the Netherlands
N° 134/13	Annamaria Lusardi Olivia S. Mitchell	The Economic Importance of Financial Literacy: Theory and Evidence
N° 133/13	Annamaria Lusardi Pierre-Carl Michaud Olivia S. Mitchell	Optimal Financial Knowledge and Wealth Inequality
N° 132/13	Riccardo Calcagno Sonia Falconieri	Competition and dynamics of takeover contests
N° 131/13	Riccardo Calcagno Maria Cesira Urzi Brancati	Do more financially literate households invest less in housing? Evidence from Italy

N° 130/12	Maela Giofr�	Financial Education, Investor Protection and International Portfolio Diversification
N° 129/12	Michele Belloni Rob Alessie Adriaan Kalwij Chiara Marinacci	Lifetime Income and Old Age Mortality Risk in Italy over Two Decades
N° 128/12	Fabio Cesare Bagliano Claudio Morana	Determinants of US Financial Fragility Conditions
N° 127/12	Mariacristina Rossi Serena Trucchi	Liquidity Constraints and Labor Supply
N° 126/11	Margherita Borella Flavia Coda Moscarola Mariacristina Rossi	(Un)expected retirement and the consumption puzzle
N° 125/11	Carolina Fugazza	Tracking the Italian employees' TFR over their working life careers
N° 124/11	Agnese Romiti Mariacristina Rossi	Should we Retire Earlier in order to Look After our Parents? The Role of immigrants
N° 123/11	Elsa Fornero Maria Cristina Rossi Maria Cesira Urzi Brancati	Explaining why, right or wrong, (Italian) households do not like reverse mortgages
N° 122/11	Serena Trucchi	How credit markets affect homeownership: an explanation based on differences between Italian regions
N° 121/11	Elsa Fornero Chiara Monticone Serena Trucchi	The effect of financial literacy on mortgage choices
N° 120/11	Giovanni Mastrobuoni Filippo Taddei	Age Before Beauty? Productivity and Work vs. Seniority and Early Retirement
N° 119/11	Maarten van Rooij Annamaria Lusardi Rob Alessie	Financial Literacy, Retirement Planning, and Household Wealth
N° 118/11	Luca Beltrametti Matteo Della Valle	Does the implicit pension debt mean anything after all?
N° 117/11	Riccardo Calcagno Chiara Monticone	Financial Literacy and the Demand for Financial Advice
N° 116/11	Annamaria Lusardi Daniel Schneider Peter Tufano	Financially Fragile Households: Evidence and Implications
N° 115/11	Adele Atkinson Flore-Anne Messy	Assessing financial literacy in 12 countries: an OECD Pilot Exercise
N° 114/11	Leora Klapper Georgios A. Panos	Financial Literacy and Retirement Planning in View of a Growing Youth Demographic: The Russian Case
N° 113/11	Diana Crossan David Feslier Roger Hurnard	Financial Literacy and Retirement Planning in New Zealand
N° 112/11	Johan Almenberg Jenny S�ve-S�derbergh	Financial Literacy and Retirement Planning in Sweden
N° 111/11	Elsa Fornero Chiara Monticone	Financial Literacy and Pension Plan Participation in Italy

N° 110/11	Rob Alessie Maarten Van Rooij Annamaria Lusardi	Financial Literacy, Retirement Preparation and Pension Expectations in the Netherlands
N° 109/11	Tabea Bucher-Koenen Annamaria Lusardi	Financial Literacy and Retirement Planning in Germany
N° 108/11	Shizuka Sekita	Financial Literacy and Retirement Planning in Japan
N° 107/11	Annamaria Lusardi Olivia S. Mitchell	Financial Literacy and Retirement Planning in the United States
N° 106/11	Annamaria Lusardi Olivia S. Mitchell	Financial Literacy Around the World: An Overview
N° 105/11	Agnese Romiti	Immigrants-Natives Complementarities in Production: Evidence from Italy
N° 104/11	Ambrogio Rinaldi	Pension awareness and nation-wide auto-enrolment: the Italian experience
N° 103/10	Fabio Bagliano Claudio Morana	The Great Recession: US dynamics and spillovers to the world economy
N° 102/10	Nuno Cassola Claudio Morana	The 2007-? financial crisis: a money market perspective
N° 101/10	Tetyana Dubovyk	Macroeconomic Aspects of Italian Pension Reforms of 1990s
N° 100/10	Laura Piatti Giuseppe Rocco	L'educazione e la comunicazione previdenziale - Il caso italiano
N° 99/10	Fabio Bagliano Claudio Morana	The effects of US economic and financial crises on euro area convergence
N° 98/10	Annamaria Lusardi Daniel Schneider Peter Tufano	The Economic Crisis and Medical Care Usage
N° 97/10	Carlo Maccheroni Tiziana Barugola	E se l'aspettativa di vita continuasse la sua crescita? Alcune ipotesi per le generazioni italiane 1950-2005
N° 96/10	Riccardo Calcagno Mariacristina Rossi	Portfolio Choice and Precautionary Savings
N° 95/10	Flavia Coda Moscarola Elsa Fornero Mariacristina Rossi	Parents/children "deals": Inter-Vivos Transfers and Living Proximity
N° 94/10	John A. List Sally Sadoff Mathis Wagner	So you want to run an experiment, now what? Some Simple Rules of Thumb for Optimal Experimental Design
N° 93/10	Mathis Wagner	The Heterogeneous Labor Market Effects of Immigration
N° 92/10	Rob Alessie Michele Belloni	Retirement choices in Italy: what an option value model tells us
N° 91/09	Annamaria Lusardi Olivia S. Mitchell Vilisa Curto	Financial Literacy among the Young: Evidence and Implications for Consumer Policy
N° 90/09	Annamaria Lusardi Olivia S. Mitchell	How Ordinary Consumers Make Complex Economic Decisions: Financial Literacy and Retirement Readiness
N° 89/09	Elena Vigna	Mean-variance inefficiency of CRRA and CARA utility functions for portfolio selection in defined contribution pension schemes

N° 88/09	Maela Giofré	Convergence of EMU Equity Portfolios
N° 87/09	Elsa Fornero Annamaria Lusardi Chiara Monticone	Adequacy of Saving for Old Age in Europe
N° 86/09	Margherita Borella Flavia Coda Moscarola	Microsimulation of Pension Reforms: Behavioural versus Nonbehavioural Approach
N° 85/09	Cathal O'Donoghue John Lennon Stephen Hynes	The Life-Cycle Income Analysis Model (LIAM): A Study of a Flexible Dynamic Microsimulation Modelling Computing Framework
N° 84/09	Luca Spataro	Il sistema previdenziale italiano dallo shock petrolifero del 1973 al Trattato di Maastricht del 1993
N° 83/09	Annamaria Lusardi Peter Tufano	Debt Literacy, Financial Experiences, and Overindebtedness
N° 82/09	Carolina Fugazza Massimo Guidolin Giovanna Nicodano	Time and Risk Diversification in Real Estate Investments: Assessing the Ex Post Economic Value
N° 81/09	Fabio Bagliano Claudio Morana	Permanent and Transitory Dynamics in House Prices and Consumption: Cross-Country Evidence
N° 80/08	Claudio Campanale	Learning, Ambiguity and Life-Cycle Portfolio Allocation
N° 79/08	Annamaria Lusardi	Increasing the Effectiveness of Financial Education in the Workplace
N° 78/08	Margherita Borella Giovanna Segre	Le pensioni dei lavoratori parasubordinati: prospettive dopo un decennio di gestione separata
N° 77/08	Giovanni Guazzarotti Pietro Tommasino	The Annuity Market in an Evolving Pension System: Lessons from Italy
N° 76/08	Riccardo Calcagno Elsa Fornero Mariacristina Rossi	The Effect of House Prices on Household Saving: The Case of Italy
N° 75/08	Harold Alderman Johannes Hoogeveen Mariacristina Rossi	Preschool Nutrition and Subsequent Schooling Attainment: Longitudinal Evidence from Tanzania
N° 74/08	Maela Giofré	Information Asymmetries and Foreign Equity Portfolios: Households versus Financial Investors
N° 73/08	Michele Belloni Rob Alessie	The Importance of Financial Incentives on Retirement Choices: New Evidence for Italy
N° 72/08	Annamaria Lusardi Olivia Mitchell	Planning and Financial Literacy: How Do Women Fare?
N° 71/07	Flavia Coda Moscarola	Women participation and caring decisions: do different institutional frameworks matter? A comparison between Italy and The Netherlands
N° 70/07	Radha Iyengar Giovanni Mastrobuoni	The Political Economy of the Disability Insurance. Theory and Evidence of Gubernatorial Learning from Social Security Administration Monitoring
N° 69/07	Carolina Fugazza Massimo Guidolin Giovanna Nicodano	Investing in Mixed Asset Portfolios: the Ex-Post Performance
N° 68/07	Massimo Guidolin Giovanna Nicodano	Small Caps in International Diversified Portfolios

N° 67/07	Carolina Fugazza Maela Giofré Giovanna Nicodano	International Diversification and Labor Income Risk
N° 66/07	Maarten van Rooij Annamaria Lusardi Rob Alessie	Financial Literacy and Stock Market Participation
N° 65/07	Annamaria Lusardi	Household Saving Behavior: The Role of Literacy, Information and Financial Education Programs (Updated version June 08: "Financial Literacy: An Essential Tool for Informed Consumer Choice?")
N° 64/07	Carlo Casarosa Luca Spataro	Rate of Growth of Population, Saving and Wealth in the Basic Life-cycle Model when the Household is the Decision Unit
N° 63/07	Claudio Campanale	Life-Cycle Portfolio Choice: The Role of Heterogeneous Under-Diversification
N° 62/07	Margherita Borella Elsa Fornero Mariacristina Rossi	Does Consumption Respond to Predicted Increases in Cash-on-hand Availability? Evidence from the Italian "Severance Pay"
N° 61/07	Irina Kovrova	Effects of the Introduction of a Funded Pillar on the Russian Household Savings: Evidence from the 2002 Pension Reform
N° 60/07	Riccardo Cesari Giuseppe Grande Fabio Panetta	La Previdenza Complementare in Italia: Caratteristiche, Sviluppo e Opportunità per i Lavoratori
N° 59/07	Riccardo Calcagno Roman Kraeusl Chiara Monticone	An Analysis of the Effects of the Severance Pay Reform on Credit to Italian SMEs
N° 58/07	Elisa Luciano Jaap Spreeuw Elena Vigna	Modelling Stochastic Mortality for Dependent Lives
N° 57/07	Giovanni Mastrobuoni Matthew Weinberg	Heterogeneity in Intra-Monthly Consumption. Patterns, Self-Control, and Savings at Retirement
N° 56/07	John A. Turner Satyendra Verma	Why Some Workers Don't Take 401(k) Plan Offers: Inertia versus Economics
N° 55/06	Antonio Abatemarco	On the Measurement of Intra-Generational Lifetime Redistribution in Pension Systems
N° 54/06	Annamaria Lusardi Olivia S. Mitchell	Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth
N° 53/06	Giovanni Mastrobuoni	Labor Supply Effects of the Recent Social Security Benefit Cuts: Empirical Estimates Using Cohort Discontinuities
N° 52/06	Luigi Guiso Tullio Jappelli	Information Acquisition and Portfolio Performance
N° 51/06	Giovanni Mastrobuoni	The Social Security Earnings Test Removal. Money Saved or Money Spent by the Trust Fund?
N° 50/06	Andrea Buffa Chiara Monticone	Do European Pension Reforms Improve the Adequacy of Saving?
N° 49/06	Mariacristina Rossi	Examining the Interaction between Saving and Contributions to Personal Pension Plans. Evidence from the BHPS
N° 48/06	Onorato Castellino Elsa Fornero	Public Policy and the Transition to Private Pension Provision in the United States and Europe

N° 47/06	Michele Belloni Carlo Maccheroni	Actuarial Neutrality when Longevity Increases: An Application to the Italian Pension System
N° 46/05	Annamaria Lusardi Olivia S. Mitchell	Financial Literacy and Planning: Implications for Retirement Wellbeing
N° 45/05	Claudio Campanale	Increasing Returns to Savings and Wealth Inequality
N° 44/05	Henrik Cronqvist	Advertising and Portfolio Choice
N° 43/05	John Beshears James J. Choi David Laibson Brigitte C. Madrian	The Importance of Default Options for Retirement Saving Outcomes: Evidence from the United States
N° 42/05	Margherita Borella Flavia Coda Moscarola	Distributive Properties of Pensions Systems: a Simulation of the Italian Transition from Defined Benefit to Defined Contribution
N° 41/05	Massimo Guidolin Giovanna Nicodano	Small Caps in International Equity Portfolios: The Effects of Variance Risk.
N° 40/05	Carolina Fugazza Massimo Guidolin Giovanna Nicodano	Investing for the Long-Run in European Real Estate. Does Predictability Matter?
N° 39/05	Anna Rita Bacinello	Modelling the Surrender Conditions in Equity-Linked Life Insurance
N° 38/05	Carolina Fugazza Federica Teppa	An Empirical Assessment of the Italian Severance Payment (TFR)
N° 37/04	Jay Ginn	Actuarial Fairness or Social Justice? A Gender Perspective on Redistribution in Pension Systems
N° 36/04	Laurence J. Kotlikoff	Pensions Systems and the Intergenerational Distribution of Resources
N° 35/04	Monika Bütler Olivia Huguenin Federica Teppa	What Triggers Early Retirement. Results from Swiss Pension Funds
N° 34/04	Chourouk Houssi	Le Vieillissement Démographique : Problématique des Régimes de Pension en Tunisie
N° 33/04	Elsa Fornero Carolina Fugazza Giacomo Ponzetto	A Comparative Analysis of the Costs of Italian Individual Pension Plans
N° 32/04	Angelo Marano Paolo Sestito	Older Workers and Pensioners: the Challenge of Ageing on the Italian Public Pension System and Labour Market
N° 31/03	Giacomo Ponzetto	Risk Aversion and the Utility of Annuities
N° 30/03	Bas Arts Elena Vigna	A Switch Criterion for Defined Contribution Pension Schemes
N° 29/02	Marco Taboga	The Realized Equity Premium has been Higher than Expected: Further Evidence
N° 28/02	Luca Spataro	New Tools in Micromodeling Retirement Decisions: Overview and Applications to the Italian Case
N° 27/02	Reinhold Schnabel	Annuities in Germany before and after the Pension Reform of 2001
N° 26/02	E. Philip Davis	Issues in the Regulation of Annuities Markets
N° 25/02	Edmund Cannon Ian Tonks	The Behaviour of UK Annuity Prices from 1972 to the Present

N° 24/02	Laura Ballotta Steven Haberman	Valuation of Guaranteed Annuity Conversion Options
N° 23/02	Ermanno Pitacco	Longevity Risk in Living Benefits
N° 22/02	Chris Soares Mark Warshawsky	Annuity Risk: Volatility and Inflation Exposure in Payments from Immediate Life Annuities
N° 21/02	Olivia S. Mitchell David McCarthy	Annuities for an Ageing World
N° 20/02	Mauro Mastrogiacomo	Dual Retirement in Italy and Expectations
N° 19/02	Paolo Battocchio Francesco Menoncin	Optimal Portfolio Strategies with Stochastic Wage Income and Inflation: The Case of a Defined Contribution Pension Plan
N° 18/02	Francesco Daveri	Labor Taxes and Unemployment: a Survey of the Aggregate Evidence
N° 17/02	Richard Disney and Sarah Smith	The Labour Supply Effect of the Abolition of the Earnings Rule for Older Workers in the United Kingdom
N° 16/01	Estelle James and Xue Song	Annuities Markets Around the World: Money's Worth and Risk Intermediation
N° 15/01	Estelle James	How Can China Solve its Old Age Security Problem? The Interaction Between Pension, SOE and Financial Market Reform
N° 14/01	Thomas H. Noe	Investor Activism and Financial Market Structure
N° 13/01	Michela Scatigna	Institutional Investors, Corporate Governance and Pension Funds
N° 12/01	Roberta Romano	Less is More: Making Shareholder Activism a Valuable Mechanism of Corporate Governance
N° 11/01	Mara Faccio and Ameziane Lasfer	Institutional Shareholders and Corporate Governance: The Case of UK Pension Funds
N° 10/01	Vincenzo Andrietti and Vincent Hildebrand	Pension Portability and Labour Mobility in the United States. New Evidence from the SIPP Data
N° 9/01	Hans Blommestein	Ageing, Pension Reform, and Financial Market Implications in the OECD Area
N° 8/01	Margherita Borella	Social Security Systems and the Distribution of Income: an Application to the Italian Case
N° 7/01	Margherita Borella	The Error Structure of Earnings: an Analysis on Italian Longitudinal Data
N° 6/01	Flavia Coda Moscarola	The Effects of Immigration Inflows on the Sustainability of the Italian Welfare State
N° 5/01	Vincenzo Andrietti	Occupational Pensions and Interfirm Job Mobility in the European Union. Evidence from the ECHP Survey
N° 4/01	Peter Diamond	Towards an Optimal Social Security Design
N° 3/00	Emanuele Baldacci Luca Inglese	Le caratteristiche socio economiche dei pensionati in Italia. Analisi della distribuzione dei redditi da pensione (only available in the Italian version)
N° 2/00	Pier Marco Ferraresi Elsa Fornero	Social Security Transition in Italy: Costs, Distorsions and (some) Possible Correction
N° 1/00	Guido Menzio	Opting Out of Social Security over the Life Cycle