The Static Microsimulation Model of the Italian Department of Finance: Structure and First Results Regarding Income and Housing Taxation

Published version:

DOI:10.3280/EP2015-002004

The Static Microsimulation Model of the Italian Department of Finance: Structure and First Results Regarding Income and Housing Taxation

By Fernando Di Nicola[*] Giorgio Mongelli[**] Simone Pellegrino[***]

ABSTRACT

In this paper we present a first attempt to develop a representative and flexible static tax-benefit microsimulation model, based on an exact match among a representative sample survey of the Italian income and living conditions in 2009-2010, provided by the Italian Institute of Statistics, and corresponding personal income tax returns, as well as cadastral data of the real estate properties of each individual, provided by the Department of Finance of the Italian Ministry of Economy and Finance on the same tax year. This static tax-benefit model can evaluate tax revenue and the redistributive impact of property and personal income taxation based on income types and levels actually declared with details of tax deductions. It should allow more reliable and detailed results compared to those based only on survey data and incomes declared to interviewers.

JEL: C81, H23, H24
Keywords: Microsimulation models, personal income taxation, property tax.

* Department of Finance, Ministry of Economy and Finance, Via dei Normanni 5, Roma (IT); Phone: +39 06 9383 6168; E-mail: fernando.dinicola@finanze.it.
** Department of Finance, Ministry of Economy and Finance, Via dei Normanni 5, Roma (IT); Phone: +39 06 9383 6510; E-mail: giorgiomongelli@gmail.com.
(Corresponding Author), Department of Economics and Statistics - University of Torino, C.so Unione Sovietica 218 bis 10134 Torino (IT); Phone: +39 011 670 6060; Corresponding author's e-mail: simone.pellegrino@unito.it.
We would like to thank two anonymous referees for their helpful comments that helped us improve the paper.

Economia Pubblica, numero 2, anno 2015

1. Introduction

In this paper we first present a static tax-benefit microsimulation model developed from 2012 by the research unit of the Department of Finance of the Italian Ministry of Economy and Finance (hereafter MEF) for study purposes, and then we show a first application.

The microsimulation model here discussed is based on both a representative sample survey (hereafter SILC Dataset) of the Italian incomes and living conditions in 2009-2010 made available by the Italian Institute of Statistics (hereafter ISTAT) and the corresponding administrative tax data (personal tax returns and cadastral data of real estate properties) of each individual within SILC Dataset on the same fiscal year made available by MEF.

Written in both STATA and SPSS, the model is primarily aimed at studying revenues and impacts of fiscal policies (personal income tax, housing taxation, family allowance, social contributions) with the redistributive impact at both individual and household level. It has further applications, such as the evaluation of global comprehensive personal income, the analysis of a broader range of tax measures, a quantification of tax erosion and a measure of the share of tax evasion derived by difference among tax return income and that declared by respondents in the survey sample.

The remainder of the paper is structured as follows. Section 2 explains the reasons that induced the MEF to develop its own microsimulation model. Section 3 describes the three datasets employed in developing our microsimulation model, and emphasizes strengths and weaknesses of each dataset. Section 4 presents the main structure of the microsimulation model, whilst Section 5 offers a comparison between the microsimulation model and the official statistics of the year to which the data refer (2009). Section 6 offers some applications of the microsimulation model on the 2013 tax year. Finally, Section 7 concludes.

2. Reasons and opportunities for an exact match approach

Several reasons induced the MEF to develop its own microsimulation model. First of all, although tax authorities have information on all tax returns filled in by Italian individual taxpayers, there is no way to extend such huge set of information for analysing the impact of taxes and transfers to the household level (as from family registry) using administrative data, since tax return forms are collected and observable only for the "fiscal family".[1] This is a strong

---

[1] Usually sample surveys consider a household as a group of individuals who live together. On the contrary, this classification is not possible when considering administrative

limitation of the administrative data, since the ability to observe socio-economic characteristics at household level is essential for analysing the effects of tax reforms.

On the other hand, even if several static tax-benefit microsimulation models based on sample surveys have been proposed in Italy during the past three decades (Marenzi, 1989; Bernardi et al., 1992, Di Biase et al., 1995; Rizzi, 1996, Brandolini, 1999, Proto, 1999, Baldini, 2000; D'Amuri and Fiorio, 2006, Maitino and Sciclone, 2008, Fiorio, 2008; Department of Treasury, 2010; Pellegrino et al., 2011, Baldini et al., 2015, among others), they encounter some difficulty for a good approximation of the real estate distribution among households as well as the actual impact of the tax system, based on incomes declared in tax return.

Two are the basic limits of this kind of models. First, they are based only on survey data and related declared incomes, so that by definition they cannot take advantages of both real level and the particularly detailed set of information collected in the tax return forms. As a consequence, their results are more reliable when focusing on a high share of taxpayers and/or the main features of a tax. Second, by employing a complex set of procedures and hypothesis, the survey-based models estimate each taxpayer's gross income (and net tax liability as well as social security contributions) starting from the net one stated by the interviewed.[2] They also face specific problems for neutralising over-reporting, under-reporting or misclassification of specific income issues or real estate properties (particularly under-reported in surveys, less than fifty per cent respect to cadastral data information; see Pellegrino et al., 2011, on this point).

On the contrary, by matching real tax return forms and real estate properties for each individual in SILC Dataset, the model here presented does not suffer of these kind of approximations, over-reporting or under-reporting, since the transition from the gross to the net income is made possible by algorithms replicating all the variables of each kind of income form filled in by each taxpayer.

Finally, in Italy, few attempts have been proposed for presenting gross and net individual income distribution starting from tax return forms. A first

---

data and tax return forms. This is due to the set of information collected by tax authorities. Let's take an example. Consider a household composed by father, mother and son; they live together. The father and the son are employees, whilst the mother is a housewife who earns no money. From a sample survey point of view they form a household. From the administrative data point of view these three individuals form two distinct fiscal families: the first fiscal family is composed by the spouses (the husband fills in the tax return form with a dependent spouse and zero dependent sons), whilst the second one is composed by the son himself.

[2] These models often simulate all the particularities of the tax system by adopting an adequate net-to-gross procedure, as proposed by Immervoll and O'Donoghue (2001), and they have to be based on the most reliable set of survey data (Ceriani et al., 2013).

attempt was made by Di Nicola and Monteduro (2004), who implemented a dataset for the Italian Personal Income Tax (hereafter PIT) by selecting a highly representative sample (above 350,000 taxpayers) among all individual PIT tax returns in the fiscal year 2000. More recently, Betti et al. (2011) proposed a net-to gross procedure for SILC Dataset income variables, and this partially employs tax return forms. It has to be noted that the SILC Dataset net-to gross procedure is not available for researchers, so that only the pre- and post-tax global income distributions can be observed. This is the most relevant limit for the SILC Dataset to be employed for some study purposes.

Baldini et al. (2006) presented a review of measurement errors in the reports of incomes and the comparisons of individual income in an interview with the individual income recorded in the tax form produced by the Ministry of Economy and Finance. The comparison was possible by an exact match among survey and fiscal data and, a relevant difference between the two source is shown.

In order to overcome these limitations, MEF decided to develop its own microsimulation model. The entire analysis has been managed by a basic principle of maximizing information within the three datasets we employed to define a broad and articulated picture of the Italian population and its tax situation, with particular emphasis on the socio-economic characteristics of Italian households. Finally, simulations were carried out in two steps. As all datasets we employed refer to the 2009 fiscal year. We first refer to this fiscal year as a benchmark analysis; then, population, incomes, fiscal rules and results are updated to the 2013 fiscal year.

These are important innovations, as up to now administrative data have been employed to study the redistributive effect of taxation on individual taxpayers, without a full representation of households and entire population. However, in Italy the “classic” microsimulation models, which also allow for analysis on households, are based, as observed, only on survey data, and are able to catch with approximations both declared income and paid taxes to tax authorities.

3. The Original Datasets and the Integrated Dataset Used for Simulations

3.1 Overall review

Input data combine three different datasets, and two of them are based only on administrative data. The integration of these databases has not occurred by making use of statistical criteria, but in a timely manner that couples information from multiple databases on the basis of an individual identifier.[3] In order to consider a representative sample of the Italian population, we make use of the 2010 SILC survey on income and living conditions provided by ISTAT on the 2009 fiscal year. This survey contains information on the demographic characteristics of Italian individuals and households, as well as their corresponding estimated gross and net incomes by type[4], including exemptions, as social pensions, subsidies, social safety nets, family allowances.

The survey covers 19,147 households and 47,551 individuals, and it is also representative at regional level. Whilst information on the main residences is fully registered, few data are available on second homes and buildings; moreover, only aggregate information by income sources is registered. SILC Dataset also shows the sample weight for each household (each individual within the household has the same household weight). First, we considered a grossing up procedure by applying the original sample weight. Next, as ISTAT sample weight were designed for demographic representativeness, we reassigned a sample weight to each individual in order for the grossing up procedure to give a better adherence with respect to that – referred to taxpayers – observed within the official statistics made available by MEF.

Using the sample selected by ISTAT (SILC Dataset), we were able to observe each individual of the sample together with their whole personal income tax return presented in 2010, with reference to the 2009 fiscal year (PIT Data), as well as all information on their real estate properties (Real Estate Data) provided by MEF, and always with reference to the same fiscal year.

A limited misalignment of some information concerning individuals we used in the surveys may be registered.[5] All mismatches were fixed by re-

---

[3] The exact match was made possible by employing the tax code observed for almost all individuals in each of the three datasets we employed. The tax code of each interviewer within the SILC Dataset was provided by ISTAT to the Department of Finance. Having identified the corresponding tax return forms, the Department of Finance supplied to us the three dataset containing an anonymous individual identifier for each respondent, instead of her tax code, in order to ensure anonymity of taxpayers.

[4] Note that the algorithm for the transition from pre- to post-tax incomes is not made available by ISTAT.

[5] Three examples can clarify this issue: a) within SILC Dataset some observations had a different age or sex compared to those inferred from the tax return (this was primarily due to a reshuffling of individuals within a household belonging to SILC Dataset); b) an individual owns her main residence starting from 2010: she correctly completed the answer within the SILC Dataset questionnaire, but she does not own the main residence according to PIT and Real Estate Data; c) an individual begins retirement in 2010, and now registers herself as pensioner and former employee within SILC Dataset, but she still appears to be an employee within PIT and Real Estate Data. Making up the integrated dataset, we considered all these anomalies by preserving as much detail as possible, as well as referring all information to the 2009 fiscal and income year.

examining any individuals showing such problems and changing the variables where necessary. However, we were not able to check all possible mismatches among our three datasets, but the integration helped to correct about one third of cases and to apply the specific tax rules.

Together, these three datasets give a broad set of information for all incomes and real estate properties of the Italian population. In order for the whole set of information to be fully implemented within the microsimulation model, several procedures and manipulations were carried out. These are explained in the following sub-sections.

3.2. IT-SILC – European Union statistics on income and living conditions in Italy (SILC Dataset)

SILC Dataset gathers socio-economic information of each individual within the sample by collecting answers to the completed questionnaires.[6] Its most relevant information, used and implemented for the microsimulation model, can be grouped in three categories: a) socio-economic characteristics of the household and each component; b) net income sources (employment and incomes); and c) real estate owned.

For category (a) we obtained information on region of residence, composition of household, personal data of each member of the household: sex, status in household, marital status, educational qualification, disabilities. For category (b) we observed employment status, job title, type of work, sector of activity, type of employer, number of hours worked, type of pensioner, income sources (income from employment, self-employment, pension, atypical work, unemployment and mobility benefit, scholarship, financial assets, real estate, other sources of income such as alimony, income support, regular and occasional gifts or cash).

Finally, for category (c) we obtained information on real estate owned by the household and individuals. Information on the main residence is accurate: tenure status and individual share of ownership of the main residence, year of construction, number of rooms, surface area of the dwelling in squared metres, rent payment, kind of landlord, imputed rent, amount and term of the mortgage, mortgage payment, market value of the home. Information on other real estate, on the contrary, is only partial: number of other dwellings, imputed rent, and rent received.

Most of the social and demographic characteristics available in the 2010 SILC Dataset refer to the year 2010, whilst all economic characteristics refer to

---

[6] For an exhaustive discussion on sample selection and a description of the whole set of variables see the SILC Dataset documentation.

the 2009 fiscal year, consistently with European EUROSTAT approach for EU-SILC national surveys.

3.3. Personal income tax returns (PIT Dataset)

Information on each individual's tax return is made available by MEF only for internal use. Here, we obtain approximately 3,000 variables, as all the qualitative and quantitative information of each tax return is available. As we will see later in greater detail, the most important variables concern the transition from pre- to post-tax income (tax deductions and allowances, taxable income, gross tax liability, tax credits, net tax liability, regional and municipal surtaxes), and sections of the tax return regarding dependent individuals, buildings owned, sources of incomes and items of expenditures.[7] It should be noted that one of the main tools of the microsimulation model is the evaluation of specific STATA and SPSS algorithms for replying all features of the tax returns.

But also availability of each kind of fiscally defined and declared income is relevant for modelling actual vs hypothetical tax and benefit system.

[7] More precisely, given the 47,551 individuals within the SILC Dataset, all records within PIT Dataset were matched: 31,082 of them have both an accurate and reliable individual identifier and a corresponding tax return registered by the tax authority; 14,355 individuals have a reliable individual identifier but no tax return (children and other persons without income, as well as persons not obliged to present the tax return, such as, under specific conditions, individuals with only income from buildings); for the residual 2,114 persons the individual identifier could not be observed within IT-SILC. Focusing on the type of tax return, 7,913 individuals presented the standard tax return (the so-called "Unico" tax return according to the Italian administrative system); 14,495 individuals presented a simplified tax return reserved to persons without self employed activity (the so-called "730" tax return); for 8,674 individuals, who earned only retirement income or employment, as well as coordinate collaboration incomes and, optionally, the cadastral income of the main residence, income information was directly communicated to the tax authorities by the employer (the "770" tax return); 928 taxpayers supposedly presented a tax return, but this could not be observed as the absence of a reliable individual identifier; and finally, 15,541 individuals were not obliged to present a tax return to the authorities (e.g. those aged less than 15 years old, as well as those with no source of income). As mentioned above, for a few taxpayers (2,114) within SILC Dataset, ISTAT did not coupled survey information and an individual identifier. Among these, 928 individuals completed the questionnaire registering income sources. Almost certainly, these individuals were income earners during the fiscal year 2009. However, due to the lack of an individual identifier, we were not able to verify if they were taxpayers for the tax authorities. Moreover, among individuals characterized by a regular individual identifier, some completed the questionnaire by stating they were income earners even if they did not complete the tax return. In order to consider these issues, we decided to define tax variables used in the microsimulation model according to dual criteria.

For individuals within SILC Dataset, with a correct corresponding tax return within PIT Dataset, we used information contained in the tax return. For individuals declaring taxable incomes within SILC Dataset, but without a tax return within PIT Dataset, in some cases[8] we accepted the income information within SILC Dataset and processed it as truthful tax information.

As a consequence of all these income integrations derived from the exact match among SILC and PIT Datasets, it was possible to separate fiscal declared from other incomes, catching the presence of a relevant shadow economy, also involving employees figures.

3.4. Land and building register (Real Estate Data)

Differently from other existing microsimulation models, the one here presented pays particular attention to buildings and their tax regime, and has to be considered supplementary compared to the PIT structure. Moreover, unlike the microsimulation models based only on survey data, a full set of administrative microdata on buildings was duly employed to apply the PIT and property tax regimes. As the survey data for buildings owned is only partial, and it is subject to strong underreporting and some approximations (e.g. it does not provide the municipal location of the building), this is essential for the estimation of municipality income and property taxes, as well as distributional analysis based on capital income and global actual income. The information we obtained within Real Estate Data are: the cadastral value of each building, type of use, share of ownership, municipality where the building is located, property tax due, and, most importantly, the cadastral category of each building.

Income tax return is indeed not the only source of administrative data for what concerns real estate properties owned by each taxpayer. The land and building register (the Italian so-called "catasto edilizio urbano") gives a set of complementary and control information for at least three reasons.

First, for taxpayers completing "730" tax return and "740" tax return (about two-thirds of overall PIT taxpayers), the information on properties owned is registered both within Datasets 2 and 3. As a consequence, Real Estate Data acts as a control set of information for properties declared within PIT Data. Moreover, Datasets 2 and 3 differ for two variables:[9] the number of days of

[8] Two types of individuals belong to this group: those declaring within SILC Dataset to be public employees or pensioners, regardless of their individual identifiers; and those characterized by a missing or incorrect individual identifier, and only earning other income sources declared within SILC Dataset, and not declared within PIT Data.

[9] Information within PIT Data also registers the "particular cases" according to PIT tax law. In these circumstances, a specific tax regime has to be applied to determine the income from buildings to be considered within the PIT base. Moreover, whenever the building

ownership of each building is registered only within PIT Data, whilst the cadastral category of each building is shown only within Real Estate Data. Both of these variables are relevant for full information on owned properties, as one of the main targets of the microsimulation model is also evaluation of the property tax.

Second, even if the situation has improved in recent years, about one million buildings registered within Real Estate Data are not declared by taxpayers within PIT Data (ghost buildings). As a consequence, by exploiting this difference we are able to assess a complete picture of tax evasion on buildings within the PIT.

Third, taxpayers whose gross income equals the cadastral income of the main residence and taxpayers whose only income (almost solely cadastral) is less than 500 euro from buildings, as well as taxpayers filling in the "770" tax return (about one third of overall PIT taxpayers), do not have to certify the ownership of the main residence within the tax return. In these cases, real properties owned are registered only in Real Estate Data.

The weak point of the land and building register (Real Estate Data) is in its updating. We obtained a complete picture of the 2009 tax year, but this register is not guaranteed to be up to date for all individuals. In order to consider this issue, if the information on buildings is missing within PIT and Real Estate Datasets, even if expected, we check for differences by using SILC Dataset. In general, if information was available from two or three sources, we adopted this information. Considering the traditional reticence to declare all properties owned, this kind of "multiple sources integration" allowed us to obtain by far more reliable picture of real estate owned by each individual.

For giving a reliable idea of real estate underreporting in Italian surveys, Table 1 shows rents or imputed rents estimation based only on SILC and on integrated datasets, distinguishing between dwellings and other buildings.

Two conflicting results arise. While for dwellings is confirmed the known tendency to slightly overestimate imputed rents (about 156 billion euros according to the SILC Dataset and about 147 according to the matched datasets), the opposite occurs for other buildings; the estimated rents by means of our matched dataset are more than double than those obtained by employing only the SILC Dataset. Our result is a significant improvement not only in representing real estate estimates, but also their distribution.

changes type of use within the tax year, PIT Data contains specific information for each period and each type of use.

Table 1 - Real or imputed rents from SILC vs matched dataset

	SILC	*Matched Datasets*	*Difference (%)*
Dwellings	155.7	146.5	-5.9
Other buildings	41.5	90.5	118.0

Note: Billion euros.
Source: Own elaborations.

4. The Structure of the Microsimulation Model

The microsimulation model is structured in seven stages; each stage is replicated for both the three types of tax returns and the residual group of taxpayers with no individual identifier in the SILC Dataset:

1) reading and reorganizing all variables within the three surveys;
2) evaluating the cadastral incomes and property tax base for each taxpayer;
3) evaluating the tax credits for dependent individuals;
4) evaluating the tax credits for earned incomes;
5) evaluating the tax credits for items of expenditures;
6) evaluating the gross and net tax liability and the whole transition from the pre- to the post-tax income for each individual, including regional and municipal surtaxes, apart from social contributions[10] and family allowances[11];
7) evaluating property tax for each building owned by each taxpayer.

The flow chart 1 can help to represent the logical and statistical steps of the model.

With approximately 3,000 variables in PIT and Real Estate Datasets, the first goal is to create a set of algorithms able to replicate the transition from gross to net PIT income for all taxpayers within SILC Dataset. As all and detailed information on tax returns is available, broader peculiarities of the Italian PIT legislation can be simulated (for example special determination of taxable income or single type of tax detraction, such as for interest on loans, maintenance of the house, health expenditure, etc.).[12]

[10] Depending on gross income, sector of activity, company size, age and other known parameters.

[11] Depending on all kind of incomes and the household structure.

[12] The detailed description of the structure of the Italian Personal Income Tax is out of the scope of this work. Suffice it to say, we were able to replicate all the peculiarities of the tax system, by translating in STATA and SPSS language the instructions for filling in the tax return forms made available by MEF. See Morini and Pellegrino (2014) for a presentation of the main characteristics of the Italian PIT structure we simulate in this work.

We built a specific set of algorithms for each type of tax return (“Unico”, “730” and “770”) for two reasons. First, the instructions for completing ordinary tax returns are sometimes different to those of the other two kinds of simplified tax returns. Second, how variables are collected and itemised by the tax authority differs according to the kind of tax return (“Unico”, “730” and “770”).

This precise procedure can be done for almost all taxpayers within the survey, as complete information for a set of tax returns is available. For a few taxpayers (928 out of 32,010), for whom the tax return is not available, we obtained information from Survey 1 and replicated these tax returns as precisely as possible.

At the end of this procedure, we had the necessary algorithms for replicating the PIT structure. Moreover, by observing the original gross and net PIT incomes reported by respondents in SILC Dataset, we could also compare these distributions and give a reliable estimation of the differences, one of which is a portion of PIT evasion.

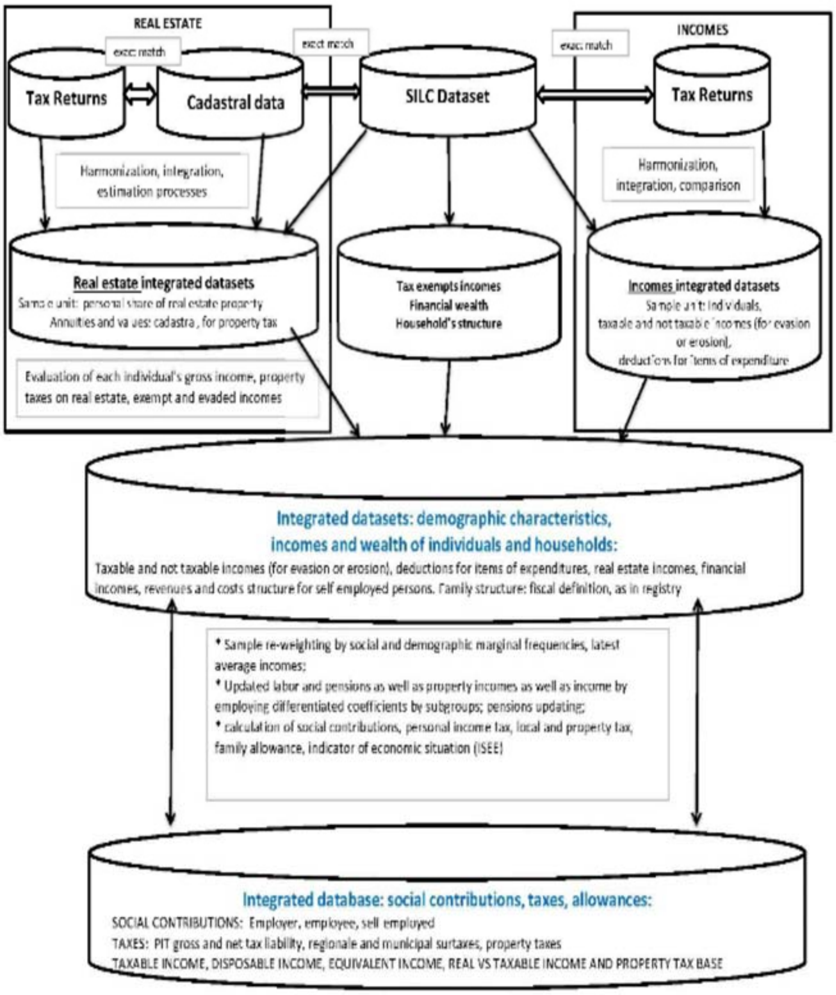
A similar strategy was employed for replicating a complete set of information regarding each individual’s property and for evaluating the corresponding property taxes. Again, for the few taxpayers for whom the individual identifier was not available, we completed our set of information by following the strategy described in the previous sections.

Stages two to six recalculate all structures of personal income tax, whilst stages two and seven recalculate the structure of the property tax.

Algorithms were written to obtain evaluated values of the microsimulation exactly equal (only rounding differences were tolerated) to the corresponding ones observed within the tax return. This was done for all taxpayers.

In so doing, we also corrected possible mistakes by taxpayers or business consultants when completing their tax return (e.g. an erroneous value of a tax credit for a spouse with respect to the taxpayer’s gross income, or an erroneous tax credit for a child claimed by the father as the child earns a higher income allowing the tax credit for the father; and so on). This has an important consequence: this model can take into account the true net tax liability as identified by the automatic audits carried out by the tax authority.

Flow Chart 1- Logical and statistical steps of the model



Source:Own elaborations.

5. Outputs with Respect to 2009 Fiscal Year

Having simulated peculiarities of the tax returns, we can now turn to some important results of the simulation, and to their comparison with the official statistics made available by MEF. We began by comparing the microsimulation model to the official statistics in the fiscal year 2009. As we employed a representative sample of the Italian population selected by ISTAT, the sample weight played an important role in an exhaustive replication of real taxes, even when employing a full set of real tax returns.

According to the official statistics made available by MEF for 2009 fiscal year, the overall number of PIT taxpayers was approximately 41.5 million, whilst the overall gross income amounted to approximately 783.3 billion euro.

By applying the original weighted samples of SILC Dataset, the number of overall taxpayers in the microsimulation model is 39.8 million, about 4.2 percent less than expected (that is, by the way, a good approximation), whilst the amount registered by the microsimulation model is 3 percent lower (751 billion euros).

By considering SILC and PIT Datasets, Table 2 compares aggregate original overall amount of gross income by source of income, and shows the percentage difference. It is of great evidence the underreporting of incomes, mainly due to tax evasion: the underreporting of self-employed incomes is about a half, whilst it is about 6 and 1 percent for labor and retirement income, respectively. The picture we obtain with our elaborations confirms previous estimation of tax evasion, e.g. those obtained by D'Amuri and Florio (2005).

Table 2 - A comparison between incomes within SILC and PIT Datasets

	SILC Dataset	*PIT Dataset*	*Difference (%)*
Employee income	446.6	420.2	-5.9
Retirement income	226.7	224.8	-0.8
Self-employed income	205.8	106.0	-48.5
Total	879.1	751.0	-14.6

Note: Billion euros.
Source: Own elaborations.

Turning to the overall number of taxpayers, and by employing original sample weights, we obtained a composition of taxpayers by source of income very close to the official fiscal statistics: about 40 million taxpayers against 41.5 million according to the official statistics.[13] The missing 1.5 million taxpayers can be due to the underreporting of income sources of a very small amount within SILC Dataset; however, these small incomes are registered within PIT and Real Estate Datasets.

Starting from this picture, we improve the goodness of fit for our model by changing the weights of the household sample. We employ the reweighting methodology proposed by ISTAT (2005) in its Genesees programme, and

[13] Without these manipulations, and using the original sample weight, we obtain only about 38.6 million taxpayers and consequently, a lower value for the overall gross income.

considered a large set of variables for the reweighting, basically all available household socio-economic characteristics, overall number of PIT taxpayers by gross income class, type of work. The main goal of the re-sampling strategy is twofold: a) to calibrate the PIT variable of the model with a higher degree of accuracy; b) to not modify the distribution of households by socio-economic characteristics. Table 3 presents the results comparing the overall number of individual PIT taxpayers and the overall amount for the most relevant tax variables by applying the re-weighted sample weights. As can be seen, the results are very close to the official statistics for all variables.

Table 3 - Number of PIT taxpayers and tax amounts - 2009 fiscal year

	Number of taxpayers (million)		*Amounts (billion euro)*	
Variables	*MEF*	*Model*	*MEF*	*Model*
Gross income	41.52	41.34	783.25	788.88
Taxable income	39.89	40.25	753.94	755.06
Gross tax liability	39.02	40.25	202.62	203.24
Net tax liability	31.01	30.53	146.49	146.66
Regional surtax	30.85	30.53	8.33	8.31
Municipal surtax	25.59	25.17	3.01	3.08

Source: Own elaborations.

The only relevant difference regards the number of taxpayers with a positive gross tax liability. There is a technical explanation for this difference. Current tax return instructions say that pensioners with a gross income of less than 7,500 euro, who own their main residence, have to show a nil gross tax liability even if it is positive. We prefer to keep an evaluated gross tax liability that does not affect the overall estimation of taxpayers who pay no net taxes and the overall amount of the gross tax liability.

Appendix A supplies several figures showing the distributions of taxpayers by gross income classes as well as the average values by income classes for several tax variables according to our microsimulation model and to the official statistics made available by MEF. In more details, we provides figures for gross income, taxable income, gross tax liability, net tax liability.

Given the absolute similarity for all tax variables in the transition from the pre- to the post-tax income, our microsimulation model can be successfully employed to estimate the effects on taxpayers of changes in the PIT structure, one of the main reasons why this model was developed by MEF.

6. A First Application of the Model: the 2013 Fiscal Year

Once the model was applied to the year of the first datasets, a further step was to update it to the 2013 fiscal year. In order to proceed in this direction, three phases had to be considered: a) to update the algorithms described in Section 3 according to the 2013 tax rules; b) to update the monetary values from 2009 to 2013; and c) to establish a new set of sample weights to take into account changes in the period 2009-2013 concerning both the population size and the composition of taxpayers.

With regards to phase a) several differences in changes to tax laws between 2009 and 2013 were simulated. Beginning in 2012, rents can be considered to be a part of PIT gross income or can be taxed with a proportional rate (21 or 19 percent according to the type of rent) according to the taxpayer's choice; income from not rented buildings are no longer taken into account in the definition of PIT gross income; a new set of tax credits for children has been introduced; tax credit for restructuring charges has increased from 36 to 50 percent; the group of 20 percent tax credits has been removed; a 3 per cent "solidarity contribution" (deductible from the gross income) for taxpayers earning more than 300,000 euro has been introduced; and finally, tax rates for regional and municipal surtaxes have changed.

Turning to phases b) and c), updating data was processed in two different steps. When the model was built only official statistics from the fiscal year 2011 were available. We updated all the relevant income variables from 2009 to 2013, according to reliable hypotheses connected to zero growth of GDP and the very low inflation rate. We applied different rules: retirement incomes had been updated following the rules of pension schemes in the period 2009-2013; labour incomes of private employees had been updated by 2 percent yearly; labour incomes of public employees had not been updated because of the wage freeze of the last years; income of self-employed taxpayers as well as deductions and tax credits for items of expenditures had been updated by 2 percent yearly, following the available statistics made available by National Accounts. Finally, cadastral income of buildings and deduction for donations had not been updated since they did not change. The re-weighting procedure was similarly employed, as cited in Section 4, whilst the updating of incomes was simulated by applying nominal changes observed for different types of taxpayers belonging to each income class, according to the MEF official statistics for 2009-2011, and to national accounts for 2011-2013.

6.1. Income taxes

In this section we present results regarding income taxation and consider both taxpayers and equivalent households. Equivalent income distribution was obtained by applying the equivalent scale determined by the squared root of the number of household members. As for the definition of income of each taxpayer, we consider the sum of two different sources of incomes: PIT gross income[14], and income from "minimal taxpayers" (the so called "regime dei contribuenti minimi").

6.1.1. Income taxes considering taxpayers

Table 4 shows the most important indices of inequality (see Appendix B for the definitions of inequality measures described in Tables 4 and 5) considering individual taxpayers. We evaluate different indices for the central government PIT and regional and municipal surtaxes.

The Gini coefficient for gross income is equal to 46.059, showing a high level of inequality.

Focusing on central government tax, the Gini coefficient for net incomes[15] is equal to 41.224, whilst the corresponding concentration coefficient is 41.132. The overall redistributive effect (the difference between the Gini coefficient of the pre-tax income distribution and the corresponding Gini coefficient of the post-tax income distribution) is equal to 4.835, whilst the Reynolds-Smolensky indices (the difference between the Gini coefficient of the pre-tax income distribution and the corresponding concentration coefficient of the post-tax income distribution) is 4.927.

Therefore, the Atkinson-Plotnik-Kakwani index (the difference between the Gini coefficient of the post-tax income distribution and the corresponding concentration coefficient of the post-tax income distribution, also equal to the difference between the Reynolds-Smolensky index and the overall redistributive effect of the tax) is equal to 0.092, and denotes the re-ranking exerted by the tax. The Gini coefficient for the tax is 68.194, whilst the concentration coefficient is 66.910. These values indicate a high concentration of the tax liability, since the tax is progressive. The Kakwani index, given by the difference between the concentration coefficient of the tax liability distribution and the Gini coefficient of the pre-tax income distribution, is equal to 20.851.

[14] For what concerns rents, they are included in the PIT definition of income up to 2011 fiscal year. Starting from the 2012 fiscal year, taxpayer can choose a separate regime (the so called "cedolare secca sui canoni di locazione"). In these cases we continue considering rents as a component of income.

[15] Here we consider only the net tax liability accruing to the central government; therefore, each taxpayer's net income is defined as the difference between the gross income and the central government tax (surtaxes are not considered).

Table 4 - Inequality indices for individual taxpayers

Index	*Value*
Gini coefficient for the gross income	46.059
Central government	
Gini coefficient for the net income	41.224
Concentration coefficient for the net income	41.132
Gini coefficient for the tax	68.194
Concentration coefficient for the tax	66.910
Average tax rate	19.113
Redistributive effect	4.835
Reynolds-Smolensky index	4.927
Kakwani index	20.851
Atkinson-Plotnick-Kakwani index	0.092
Regional government	
Gini coefficient for the tax	55.748
Concentration coefficient for the tax	53.656
Average tax rate	1.345
Municipal government	
Gini coefficient for the tax	62.309
Concentration coefficient for the tax	55.085
Average tax rate	0.501
Overall	
Gini coefficient for the net income	40.937
Concentration coefficient for the net income	40.83
Gini coefficient for the tax	66.961
Concentration coefficient for the tax	65.777
Average tax rate	20.959
Redistributive effect	5.122
Reynolds-Smolensky index	5.228
Kakwani index	19.718
Atkinson-Plotnick-Kakwani index	0.107

Source: Own elaborations.

Therefore, the degree of progression is very high, since the maximum admittable value of the Kakwani index is the difference between 1 and the Gini coefficient of the pre-tax income distribution.

The Kakwani and the Reynolds-Smolensky indexes are linked by the overall average tax rate (the ratio between the sum of the tax liability of each taxpayer and the sum of the pre-tax income of each taxpayer), which amounts to 19.113 percent.

As expected, considering overall taxation[16] we observe a greater average tax rate (20.959 percent), a greater Reynolds-Smolenky index (5.228), and a smaller Kakwani index (19.718).[17] The corresponding overall redistributive effect is 5.122, whilst the Atkinson-Plotnik-Kakwani index is equal to 0.107.

Very few scientific analysis present the most important personal income tax redistributive indexes with respect to taxpayers. The values we obtained are very similar to those according to the official statistics made available my MEF (2010) for the period 2001-2007 and Martone (2008); considering other microsimulation model based on survey data, our results are very similar to those obtained by Morini and Pellegrino (2014).

6.1.2. Personal income tax considering equivalent households

Similarly to Table 4, Table 5 shows the same inequality indices with regard to equivalent households. The Gini coefficient for the gross income is equal to 42.882.

Focusing on central government tax, the Gini coefficient for net incomes is equal to 38.297, whilst the corresponding concentration coefficient is 38.221. Therefore, the overall redistributive effect is equal to 4.585, whilst the corresponding Reynolds-Smolenky is 4.661. The Atkinson-Plotnik-Kakwani index is then equal to 0.076. The Gini coefficient for the tax is 63.361, whilst the concentration coefficient is 62.471, so that the Kakwani indices is 19.589. The observed average tax rate is 19.222 percent.

Considering overall taxation, the Gini coefficient for net income is 38.060, whilst the corresponding concentration coefficient is 37.970. The Gini coefficient for the tax is 62.129 and the concentration coefficient is 61.283. The overall average tax rate is 21.070 percent. We also observe that the Reynolds-Smolenky index is 4.912 and the Kakwani coefficient is 18.401. The overall redistributive effect is 4.822.

[16] Here we consider both the net tax liability accruing to the central government and the surtaxes accruing to local governments; therefore, each taxpayer's net income is defined as the difference between the gross income and the sum of central government tax and regional as well as municipal surtaxes.

[17] See Baldini *et al.* (2003) for the proof of these results.

Table 5 - Inequality indices for equivalent households

	Indexes
Gini coefficient for the gross income	42.882
Central government	
Gini coefficient for the net income	38.297
Concentration coefficient for the net income	38.221
Gini coefficient for the tax	63.361
Concentration coefficient for the tax	62.471
Average tax rate	19.222
Redistributive effect	4.585
Reynolds-Smolensky index	4.661
Kakwani index	19.589
Atkinson-Plotnick-Kakwani index	0.076
Regional government	
Gini coefficient for the tax	50.43
Concentration coefficient for the tax	48.537
Average tax rate	1.347
Municipal government	
Gini coefficient for the tax	57.599
Concentration coefficient for the tax	49.973
Average tax rate	0.501
Overall	
Gini coefficient for the net income	38.06
Concentration coefficient for the net income	37.97
Gini coefficient for the tax	62.129
Concentration coefficient for the tax	61.283
Average tax rate	21.07
Redistributive effect	4.822
Reynolds-Smolensky index	4.912
Kakwani index	18.401
Atkinson-Plotnick-Kakwani index	0.09

Source: Own elaborations.

Table 6 shows the composition of gross income, net income and tax by deciles of equivalent households; the Table also shows the cut points of equivalent incomes at each decile. Before tax, the first ten percent of households receive only 0.8 percent of overall gross income, whilst the top ten percent of households receive 32.1 percent. Because of the tax, these proportions are compressed: the first ten percent of households receives 1 percent of the post-tax incomes, whilst the top ten percent of households receives 28.3 percent. Note the first eighty percent of households receive a greater share of net income with respect to the pre-tax situation. Note also that the top ten percent of households pay a little less than half of the overall net tax liability.

Table 6 - Composition of incomes and tax by deciles

Deciles	*Cut points of equivalent incomes*	*Gross income*	*Net tax liability*	*Net income*
1	5,452	0.8	0.1	1
2	9,140	3.6	0.4	4.4
3	11,898	4.7	1.8	5.5
4	14,863	6	3.5	6.7
5	17,802	7.3	5.1	7.9
6	21,014	8.6	6.8	9.1
7	24,951	9.9	8.5	10.3
8	30,253	12	11.3	12.1
9	39,656	15	16	14.7
10	1,659,799	32.1	46.5	28.3
Total	-	100	100	100

Note: All households within the integrated Datasets have been considered.

Source: Own elaborations.

Finally, Table 7 shows the Pfähler (1990) decomposition of the Reynolds-Smolensky index (see Appendix C for details). Considering both individual taxpayers and equivalent households, about 60 percent of the Reynolds-Smolensky index is due to tax credits (and to dependent individuals and earned income ones in particular), whilst the remaining 40 percent is due to the tax rate schedule. As can be noted, allowances do not contribute to the overall Reynolds-Smolensky index.

Table 7 - The Pfähler decomposition

	Tax allowances	*Rate schedule*	*Tax credits*	*Total*
Taxpayers	-0.4	41.7	58.6	100
Equivalent Households	-0.7	40.7	60.0	100

Source: Own elaborations.

6.2. Distribution of household wealth (real estate)

In order to fully show the results of our microsimulation model, we discuss some specific statistics on buildings owned by households. Specifically, in what follows we consider equivalent values for the tax base of the Italian property tax (hereafter IMU). The Gini coefficient of the IMU tax base on main residence is equal to 55.843, whilst on other buildings it is equal to

84.308. Considering all buildings together, the Gini coefficient is equal to 60.159. More than 70 percent of households own the main residence, whilst only one household out of three own at least one building. As a consequence, Gini coefficients show a high concentration of wealth among households, a concentration higher than those observed for incomes.

If we evaluate the corresponding concentration coefficients of the IMU tax base once households are ordered according to their PIT income, this picture is resized: the concentration coefficient of the IMU tax base on main residence decreases to 22.860, whilst that on other buildings to 40.902. Considering all buildings together, the Gini coefficient is equal to 30.104. The high share of households who own at least one building can explain these differences.

To sum up, Italian households appear to be “cash-poor” (about 2 taxpayers out of three declare a PIT income lower than 20 thousand euros) and “asset-rich”. This dichotomy should be carefully considered when thinking of restructuring the whole tax system in order to both reduce inequality and boost economic growth: recent tax reforms in most European countries proposed a tax shift from direct to indirect as well as wealth taxation; given the budget constraints, Italy increased indirect and wealth taxation, whilst postponing the reduction of income taxation. On the redistributive point of view, such tax shift can cause ambiguous effect, since wealth taxation turns out to be progressive with respect to income whenever the ratio between wealth and income is increasing with income. Further statistics on wealth and income distributions can then help understanding this point.

Table 8 shows the share of households owning at least one building by deciles of equivalent PIT income. In particular, Table 8 shows three columns. The first concerns the property of the main residence; the second, the property of other buildings (not only dwellings, but also every other type of building); and the third considers all buildings together. As can be noted, the higher the decile, the higher the share of household ownership for all the three columns. About 71 percent of households are owner-occupiers, but only 45 percent of households belong to the bottom ten percent, and more than 80 percent of households belong to the top three deciles. A different picture emerges regarding buildings other than the main residence. About one quarter of households belonging to the bottom ten percent who own at least one dwelling, whilst this proportion rises to about 72 percent for households in the top ten percent.

Table 8 - Share of households with positive IMU tax base by deciles of equivalent households PIT income

Deciles	*IMU tax base Main residence*	*IMU tax base Other buildings*	*IMU tax base All buildings*
1	45.0	24.0	50.3
2	58.4	28.9	63.4
3	64.8	35.7	70.7
4	66.8	37.8	72.2
5	69.5	39.3	74.8
6	76.1	43.5	81.7
7	79.5	44.9	84.2
8	81.6	50.5	88.0
9	85.4	57.6	91.2
10	88.0	71.7	95.5
Total	71.1	43.2	76.8

Source: Own elaborations.

More surprising are the results shown in Table 9. Here, we show the percentage ratio between the value of the IMU tax base and the amount of PIT gross income.

Table 9 - Incidence of the IMU tax base on PIT income (%)

Deciles	*IMU tax base Main residence*	*IMU tax base Other buildings*	*IMU tax base All buildings*
1	1,339.4	641.6	1,981.0
2	409.4	173.9	583.2
3	332.0	169	501.0
4	275.7	143.6	419.3
5	257.4	129.0	386.5
6	264.7	117.8	382.5
7	238.0	112.4	350.4
8	200.1	126.5	326.6
9	189.6	133.5	323.2
10	143.1	169.4	312.5
Total	221.6	148.5	370.1

Source: Own elaborations.

As can be seen, Italian households have a high value of wealth with respect to their annual income: 221.6 percent for the main residence, 148.5 percent for buildings other than their main residence, and about 370 percent for all buildings together. Moreover, it has to be considered that the IMU tax base is, on average, about a half of the corresponding market value of the buildings. Even if the ratio between wealth and income is not increasing with income, these ratios are very high for all households, and incredibly for those belonging to the bottom ten percent (1,339.4, 641.6 and 1,981 percent, respectively) compared to those registered to the top ten percent (143.1, 169.4 and 312.5 percent, respectively).

7. Concluding Remarks

In this paper we presented and discussed a first attempt to develop a microsimulation model for Italy aimed to well simulate detailed tax and benefit rules and their possible change, preserving distributional information. On the one hand it considers a representative sample of the Italian population made available by the Italian National Institute of Statistics; on the other, it matches all income units from this survey with one to one information contained within the administrative data of both personal income tax returns and the building register, provided by the Department of Finance of the Italian Ministry of Economy and Finance. Given this integrated amount of information, we discuss the strengths and weaknesses of these datasets, and then explain in greater detail each stage to obtain our microsimulation model. Finally, we present some results of this model, used by the research unit of the Department of Finance of the Italian Ministry of Economy and Finance for study and support purposes.

APPENDIX A

Figure 1 - Frequency density function for gross income

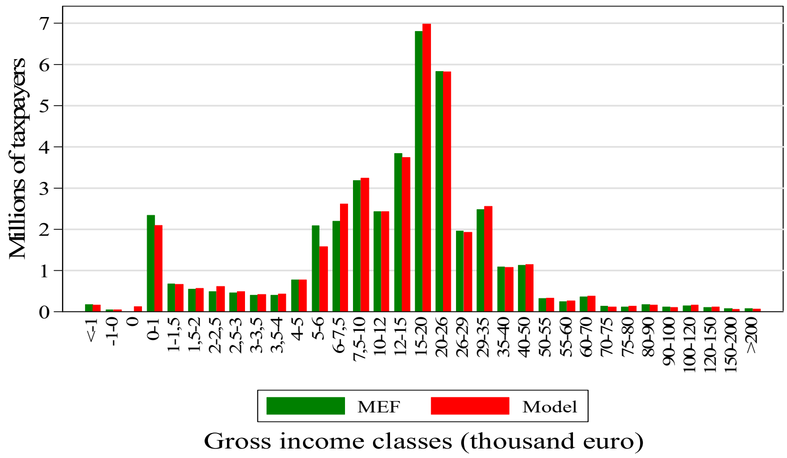


Figure 2a - Distribution of taxpayers by gross income classes - Taxable income

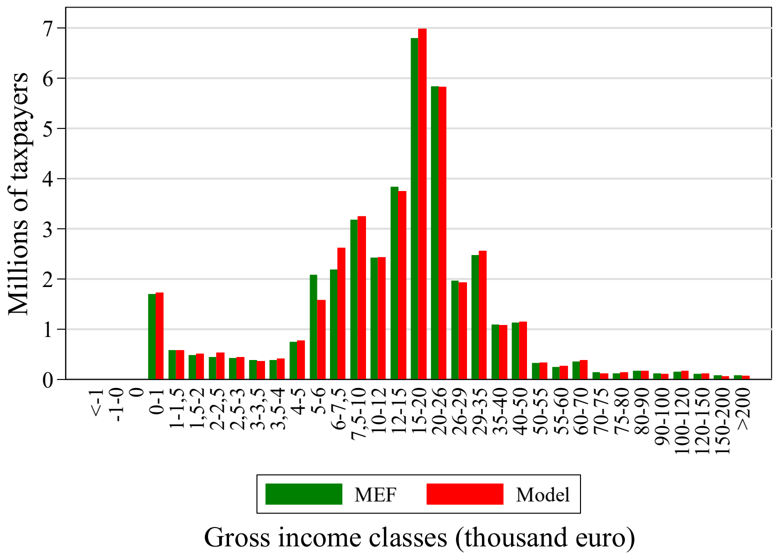


Figure 2b - Average value for taxable income by gross income classes

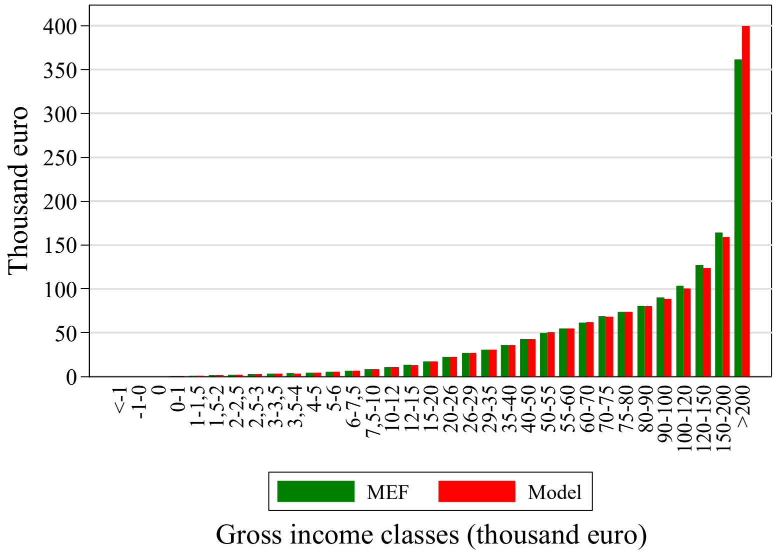


Figure 3a - Distribution of taxpayers by gross income classes - Gross tax liability

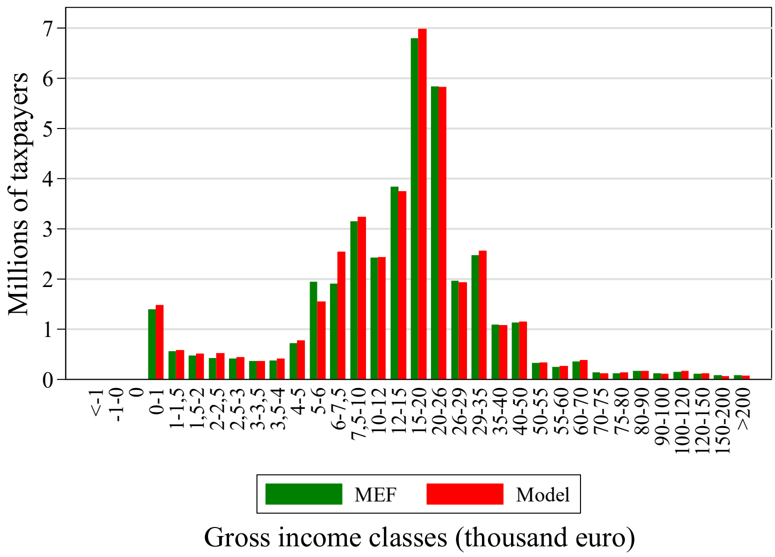


Figure 3b - Average value for gross tax liability by gross income classes

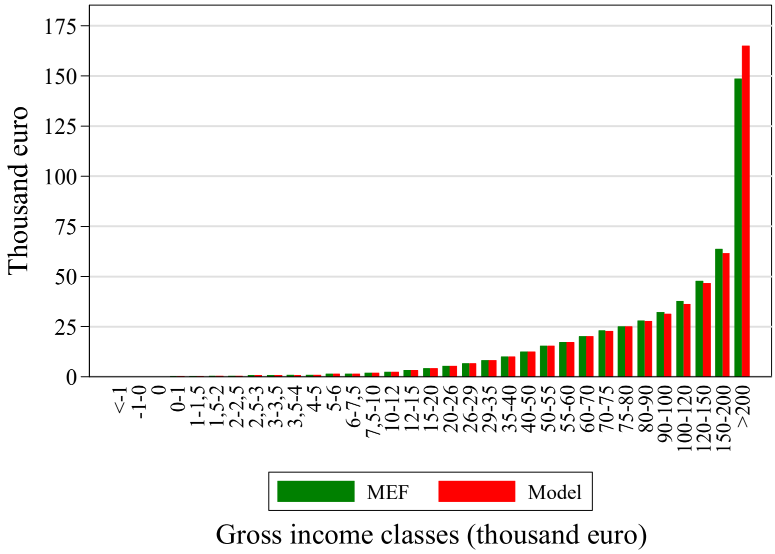


Figure 4a - Distribution of taxpayers by gross income classes - Net tax liability

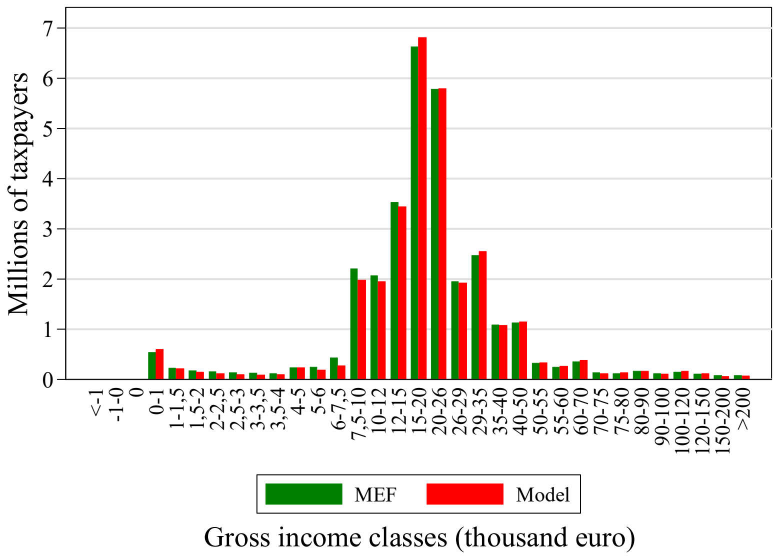
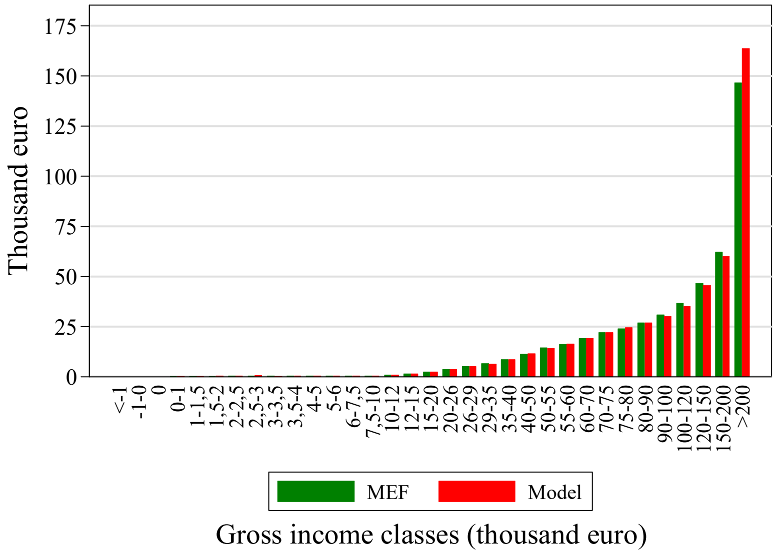


Figure 4b - Average value for net tax liability by gross income classes



APPENDIX B

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

Inequality among pre- and post-tax income levels as well as tax levels can be evaluated by the Gini coefficient, which ranges from zero to 1. Let G_X, G_Z and G_T, be the corresponding Gini coefficient for pre-tax income, post-tax incomes and taxes, respectively. Then,

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

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. 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 coefficients can then be evaluated as follows:

$$C_\varepsilon = \frac{2\operatorname{cov}[\varepsilon, F(X)]}{\mu_\varepsilon}.$$

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 is equal to $RE = G_X - G_Z = (G_X - C_Z) - (G_Z - C_Z) = RS - R^{APK}$, 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. 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, namely

$$\theta = \frac{\sum_{i=1}^{n} T_i}{\sum_{i=1}^{n} x_i}$$

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.

APPENDIX C

According to Pfähler (1990) and Lambert (2001), it is possible to 'decompose' the RS index in the contributions due to deductions, the rate schedule and the ax credits. Here we follow the methodology summarised by Urban (2006). The 'share' of RS and K due to the deductions can be evaluated as follows:

$$RS^{D} = \left(G_X - C_Y\right) = \frac{\delta}{1-\delta}\left(C_D - G_X\right)$$

where

$$\delta = \frac{\sum_{i=1}^{n} d_i}{\sum_{i=1}^{n} x_i}$$

and $K^D = C_D - G_X$, and Y is the taxable income.

The 'share' of RS and K due to the marginal tax rate schedule can be evaluated as follows:

$$RS^{RATE} = \left(C_Y - C_{Y-GT}\right) = \frac{\eta}{1-\delta-\eta}\left(C_{GT} - C_Y\right)$$

where

$$\eta = \frac{\sum_{i=1}^{n} GT_i}{\sum_{i=1}^{n} x_i}$$

and $K^{RATE} = C_{GT} - C_Y$.

Finally, the 'share' of RS and K due to tax credits can be evaluated as follows:

$$RS^C = \left(C_{X+C} - G_X\right) = \frac{\lambda}{1+\lambda}\left(C_C - G_X\right)$$

where

$$\lambda = \frac{\sum_{i=1}^{n} c_i}{\sum_{i=1}^{n} x_i}$$

and $K^C = C_C - G_X$.

Hence, the overall RS can be decomposed as:

$$RS = \begin{cases} \dfrac{-\eta}{1-\eta+\lambda}\dfrac{1-\delta}{\psi}RS^D + \\ +\dfrac{\eta}{1-\eta+\lambda}\dfrac{\psi-\eta}{\eta}RS^{RATE} + \\ +\dfrac{1+\lambda}{\eta-\lambda-1}RS^C \end{cases}$$

where

$$\psi = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i}.$$

References

Baldini M. (2000). MAPP98: un modello di analisi delle politiche pubbliche. *Dipartimento di Economia Politica, Materiali di discussione*, 331.

Baldini M., Bosi P., Guerra M.C., Silvestri P. (2003). Distribuzione del reddito e politiche fiscali in un contesto locale: il caso del comune di Modena. *XV Conference of the Italian Association of Public Economics*, Pavia.

Baldini M., Bosi P., Lalla M. (2009). Errors and underreporting of tax form's and interview report's income in the survey on economic and social conditions of families at Modena. *Department of Political Economy, University of Modena and Reggio Emilia*, Modena.

Baldini M, Giarda E., Olivieri A. (2015) A Tax-benefit Microsimulation Model for Italy: a Partial Evaluation of Fiscal Consolidation in the Period 2011-2014, *Prometeia*, Nota di lavoro n. 1.

Bernardi L., Marenzi A., Pozzi F. (1992), L'analisi di microsimulazione delle imposte dirette e dei contributi sociali a carico delle famiglie: modello e risultati. In: Bosi P. e Lugaresi S., a cura di, *Bilancio pubblico e redistribuzione*. Bologna: Il Mulino.

Betti G., Donatiello G., Verma V. (2011). The Siena Microsimulation Model (sm2) for Net-gross Convertion of EU-SILC Income Variables. *The International Journal of Microsimulation*, 4: 35-53.

Brandolini A. (1999). The Distribution of Personal Income in Post-War Italy: Source Description, Data Quality, and the Time Pattern of Income Inequality. *Temi di discussione del Servizio Studi*, n. 350, Banca d'Italia, Roma.

Ceriani L., Fiorio C.V., Gigliarano C. (2013). The Importance of Choosing the Data Set for Tax-benefit Analysis. *The International Journal of Microsimulation*, 6: 86-121.

Di Biase R., Di Marco M., Di Nicola F., Proto G. (1995). ITAXMOD. A Microsimulation Model of the Italian Personal Income Tax and Social Security Contributions'. *Documenti di Lavoro ISPE*, 16/95.

Department of Finance of the Italian Ministry of Economy and Finance (2010). *Misure globali della progressività, dell'incidenza e della redistribuzione dell'imposta - Irpef anni d'imposta 2001-2007*.

Di Nicola F., Monteduro M.T. (2004). Un modello di microsimulazione delle imposte sulle persone fisiche, *SECIT*.

Fiorio C.V. (2008). Analisi dell'IRPEF 2007 vs sistemi di imposizione alternativi utilizzando TABEITA04. *Libro Bianco*, Scuola Superiore dell'Economia e delle Finanze, Ministero dell'Economia e delle Finanze, Tributi, Supplemento 1.

Fiorio C.V., D'Amuri F. (2005). Workers' Tax Evasion in Italy. *Giornale degli Economisti ed Annali d'Economia*, 64(2/3): 247-270.

Immervoll H., O'Donoghue C (2001). Imputation of Gross Amounts from Net Incomes in Households Surveys. An Application Using EUROMOD, *EUROMOD*, Working paper EM1.

ISTAT (2005). GENESEES v.3 - Funzione Riponderazione. Manuale utente e aspetti metodologici.

Lambert P.J. (2001). The Distribution and Redistribution of Income. Manchester and New York: Manchester University Press, Third edition.

Marenzi A. (1989). La distribuzione del carico fiscale in Italia: un modello di microsimulazione. *Dipartimento di Economia Pubblica e Territoriale*, Pavia.

Dipartimento del Tesoro – Ministero dell'Economia e delle Finanze (2010). ITaxSim – Italian Taxation Simulation Model, Roma.

D'Amuri F. and Fiorio C.V. (2006). Grossing-up and Validation Issues in an Italian Tax-benefit Microsimulation Model. *Econpubblica Università Bocconi*, Milano, WP No. 117.

Maitino M.L. and Sciclone N. (2008). Il modello di microsimulazione multiregionale dell'IRPET MICROREG, *Società italiana di economia pubblica*, Pavia, WP No. 604.

Martone A. (2008). Dieci anni di riforme a confronto. *Libro Bianco*, Scuola Superiore dell'Economia e delle Finanze, Ministero dell'Economia e delle Finanze, Tributi, Supplemento 1.

Morini M., Pellegrino S. (2014). Personal Income Tax Reforms: a Genetic Algorithm Approach, *Department of Economics and Statistics*, University of Torino, Working Paper Series, 26.

Pellegrino S., Piacenza M., Turati G. (2011). Developing a Static Microsimulation Model for the Analysis of Housing Taxation in Italy. *The International Journal of Microsimulation*, 4: 73-85.

Pfähler W. (1990). Redistributive Effect of Income Taxation: Decomposing Tax Base and Tax Rates Effects. *Bulletin of Economic Research*, 42(2):121-129. DOI:10.1111/j.1467-8586.1990.tb00295.x

Proto G. (1999). Il modello di microsimulazione MASTRICT: struttura e risultati, *Rivista di statistica ufficiale*, ISTAT.

Rizzi D. (1996). TBM. Un modello statico di microsimulazione Tax-Benefit per l'Italia, *Dipartimento di scienze economiche, Università Ca' Foscari di Venezia*.

Urban I. (2006). Progressivity of Personal Income Tax in Croatia: Decomposition of Tax Base and Rate Effects. *Financial Theory and Practice*, 30(3): 207-231.