Dinamiche e persistenze nel mercato del lavoro italiano ed effetti di politiche
(basi di dati, misura, analisi)

Progetto di ricerca cofinanziato dal MIUR
(Ministero dell'Istruzione, dell'Università e della Ricerca) – Assegnazione: 2001
Coordinatore: Ugo Trivellato

Personal income in the gross and net forms:
applications of the Siena micro-simulation model (SM2)


* Dipartimento di Metodi Quantitativi, Univ. di Siena
** International Social Research Ltd e Univ. di Siena

Working Paper n. 54 settembre 2003

Unità locali del progetto:

Dip. di Economia “S. Cognetti De Martiis”, Univ. di Torino (coord. Bruno Contini)
Dip. di Scienze Economiche, Univ. “Ca’ Foscari” di Venezia (coord. Giuseppe Tattara)
Dip. di Metodi Quantitativi, Univ. di Siena (coord. Achille Lemmi)
Dip. di Scienze Statistiche, Univ. di Padova (coord. Ugo Trivellato)
Dip. di Politiche Pubbliche e Scelte Collettive, Univ. del Piemonte Orientale (coord. Alberto Martini)

Dip. di Scienze Statistiche
via C. Battisti 241-243, 35121 Padova
1 The Siena Micro-Simulation Model (SM2)

Income of households is made up of diverse components received by multiple individuals. Its elements may be compiled from different types of sources, which may differ in concepts and definitions and may not refer to exactly the same reference time. The different sources may be subject to differing patterns of response and recording errors, sampling errors, inconsistencies and incompleteness etc. This paper is not concerned with such conceptual and measurement issues, but with the following additional important problem.

Income can be recorded in various forms - such as gross, or net of taxes and/or other retentions at source, or as the final amounts actually received - differently for different components and for different income earners in the household. Aggregating these elements of income into the household's total income and its main components requires not only that information is available on all the elements, but also that it is in a homogeneous form to permit aggregation. The form must also be the same for all households to permit aggregation to the sample or the population. Furthermore, the same information in more than one forms is often required to meet different analytical objectives. For instance, for poverty and social exclusion research it is necessary to have information on total household disposable income. Total disposable income means gross income less income tax, regular taxes on wealth, employees', self employed, unemployed and other compulsory social insurance contributions, employers' social insurance contributions and inter household transfers paid. To study the effect on income distribution, the breakdown of disposable household income into old-age and survivors' benefits and other social transfers is needed. On the other hand, for the study of redistribution effect of taxes, for micro-simulation and many other research and policy purposes information is also required on gross income and its detailed breakdown by component and individual income recipient.

Different forms (gross, disposable or net, etc.) are related through extremely complex national fiscal systems, often with sub-national variations as well. This complexity has many aspects. (i) The relationships or rules vary by income component and according to characteristics and circumstances of the income recipient, in great detail and with many special cases. Some components may be exempt from tax and other deductions, while others may be subject to either or both, fully or in part. (ii) The rules may apply to different types of units, to individual persons, whole households, or some other intermediate 'tax units' within households. (iii) Different aggregations across income components may be involved in the application of the rules: some components may be treated by themselves, while others pooled together. (iv) How income is received can vary: it may for instance be received after retentions at source, or received gross to be taxed later. (v) What form it is reported in may vary from one component and recipient to another in the same data set. (vi) Who receives income may vary: while most income is received by individuals, some parts (e.g. certain transfers) may actually pertain to the whole household. (vii) When the income is received or deductions made and the period to which relate may differ. All this complexity is increased where individuals have a choice among alternative rule-sets to be applied to their particular case. We hardly need to mention discrepancies between rules and their actual application: individuals failing to claim benefits, other transfers and reimbursements to which they are entitled; or not paying taxes and deductions which are due.

Various micro-simulation models have been developed to simulate taxes, social insurance contributions, benefits and other transfers received to affect the transformation between
gross and net forms of income, mediated through complexities of the national fiscal systems. Important examples are Euromod and similar national micro-simulation models. A main objective of these models is to provide, on the basis of specific micro-datasets incorporated into the system, a comprehensive facility for simulation of the effect of varying parameters of the tax-benefit system on the income received by various segments of the population. Simulation of taxes and benefits under different regimes (fiscal policy options) forms the output of the system.

Our orientation in developing the Siena Micro-Simulation Model (SM2) described in this paper is somewhat different, though this new system shares much in objectives and methods with existing micro-simulation models. At the outset, SM2 has been designed for multi-country application, to be a flexible tool which is portable to the maximum extent possible across (at least the European) countries despite great differences in their fiscal systems. The immediate context for the development of SM2 has been the requirements of EU-SILC (EU Statistics on Income and Living Conditions). EU-SILC is a statistical source, being developed by European Commission (Eurostat) and implemented by all EU and also many other European countries, for the generation of comparable and detailed information on living conditions and income of households and persons. The central issue to be addressed is that, while the source, type and form of input (collected) information varies across and even within countries, the output required at the European level has to be comparable and standardised. Furthermore, while the information which can be collected is limited to particular forms by limitations of the sources providing it, it is required in both net and gross forms for diverse academic and policy research. We see SM2 as a tool, under continued development, for meeting these objective in the international, comparative context. Starting from data on household and personal income given in different forms (including some missing data), and on the basis of the prevailing fiscal system in a country, the model estimates full information on income by component, with breakdown of gross amounts into taxes, social insurance contributions, social transfers, and net and disposable income. Therefore it can be applied to diverse data sets to generate variables (such as the EU-SILC Target Variables) in a standard form. Furthermore, it is designed to be flexible to deal with an annual flux of data in different forms across and within countries and also with periodic changes in the national tax systems, which a longitudinal data source such as EU-SILC must deal with.

Thus an outstanding and unique feature of the SM2 system is that its core consists of a standardised set of routines which can handle a great diversity of input data forms and national tax systems. Country-specific routines are required to convert the input data into standardised forms, and also to specify parameters of the national tax system in an appropriately standardised form. These, then, form inputs to the central core of the system designed to generate the required standardised outputs. The system has been developed to maintain a clear distinction between the common and the country-specific parts, and even more importantly, to maximise the part which can be standardised. This feature makes the system an appropriate and convenient tool for multi-country application.

Given the specific context and objectives of its development, hitherto SM2 is fully "data based" and does not incorporate simulation of benefits or any other income components. It is taken as given that information on all income components has been collected, compiled or imputed in some form, and that the objective is to convert it, under a specified national tax system applicable at the time, to the standard form (specifically that required by EU-SILC). It incorporates generally the same or similar level of detail as other major micro-simulation models - a little less detailed on some points but also more complete on some others, apart from so far being data-based rather than simulation-based
as concerns benefits and similar transfers. Detailed applications have already been developed for France, Italy and Spain using European Community Household Panel (ECHP) data as the input, pending more current data becoming available from EU-SILC.

This paper is the first report at a scientific conference of features and applications of the Siena Micro-Simulation Model. It will focus on the standardised core of the system and explain how it can handle specific features of diverse fiscal systems and forms in which income of households and persons has been recorded.

The model SM2 is described in the remaining sections by introducing complexities step-by-step. Section 2 introduces various terms, and describes the basic relationships between them by considering the model in the simplified situation of a person receiving income from a single source and taxed separately as a single-person tax unit. Section 3 gives a fuller description of the micro-simulation model in the more realistic situation involving more than one income components and multi-person tax units. A number of illustrations (from France, Italy and Spain) are provided. Section 4 deals with issues arising from diverse forms of the input data input. The form (net, gross, etc.) in which the information has been collected may vary from one individual to another in the same survey. Finally, Section 5 introduces the additional complexity resulting from differences in how particular components of income are treated in the tax regime. A outstanding feature of SM2 is that these special features of the tax system can be captured within the general structure of the model simply by appropriately defining special types of 'deductions' and 'tax credits' for the component concerned.

2 Terminology: forms of income and their relationship

Gross income (GG, G) of an individual, household or other tax unit is the total income from all sources received during a reference period, before any deductions for tax or social insurance contribution. Certain irregular or lump-sum receipts may be excluded. In any case, our model does not assume any particular definition. In the model it is useful to distinguish between gross income ("GG") including employer's social insurance contributions (SS), and the gross ("G") including only other SI contributions. This is because SI contributions (including the employers' contributions) are generally a function of G, rather than of GG.

Social insurance contributions (S). Normally, these contributions apply only to income from work and include (i) employer's contribution on behalf of persons in employment; (ii) employee's contributions; (iii) self-employed person's contributions. (iv) In some cases these contributions may also apply to certain other components. (v) In most national systems, the following structure for the social insurance contributions applies: the contributions are component-specific, determined independently of other components of income; and the amount of gross income paid out as social insurance contributions is itself not subject to income tax.

However, the system can be much more elaborate. For instance in France practically all income components are subject to SI contributions, including employment income, self employment income, pensions, invalidity benefits, unemployment benefits, capital income and other incomes each type of income with its own distinct schedule. The only exceptions are certain types (but not all) of capital income and certain types of benefits. Furthermore, a component may be subject to SI contributions of different types. For instance self-employment is subject to seven types of SI contributions schedules: for health insurance, widowhood, old age, complementary pension, unemployment insurance, CSG (Contribution Sociale Généralisée), and CRDS (Contribution au Remboursement de...
la Dette Sociale). For employment income, there is in addition the employer's contribution. Another unique feature is that a certain types of SI contributions are subject to tax: generally CSG in part, and CRDS fully. Furthermore, contributions for a number of components are a function not of the income for that component alone (as is generally the case in other counties), but of the combined income from a set of a number of components. Normally this is in the form of an upper limit applied to the combined income for the set. This applies to the set of incomes from employment, self-employment, pensions and unemployment benefits. Nevertheless, these complications merely make the algorithm specifying the various functional relationships more elaborate, but there is no problem in handling them within the common structure of the model.

**Gross taxable income** (H) is gross income less social insurance contributions: \( H = G - S \).

**Deductions** (D) refers to part of gross taxable income which is exempt from tax.

**Net taxable income** (Y) is obtained by subtracting from gross taxable income the part which is tax exempt (i.e., "deductions"): \( Y = H - D \).

**Tax due** (W). Initial tax due is computed as a function of net taxable income, \( W = W(Y) \). This is determined by the prevailing income tax schedule. For the main part, this normally involves pooling of income across components and individuals in the tax unit to which the common tax schedule applies.

**Tax credits** (C). The tax liability is normally reduced by tax credits.

**Tax paid** (X). Deduction of these tax credits from the tax due gives the final tax to be paid: \( X = W - C \).

**Total net income** (N) is total gross taxable income less tax paid: \( N = H - X \).

In certain systems, income as initially received is subject to retention at source of tax and/or social insurance contributions; for instance, for income subject to both these retentions we have: \( XTS = G-S(G)-T(H) = H-T(H) \). Unlike the amount of 'final tax due' W or X which is determined through complex rules involving pooled income over individuals and components, the relationship \( T(H) \) determining tax retention at source is often component-specific and much simpler.\(^1\) The same applies to any income components which are taxed separately from pooled income as defined above (such as at a flat or some other component-specific rate).

Gross income, SI contributions, gross taxable income, tax and social insurance withheld at source, income received after retentions at source, and some deductions and tax credits are defined at the level of individual income components. Other quantities may involve aggregation over components (and also over individuals in a tax unit). It is useful to first look at the basic relationships between different quantities for the simple case of a single person tax unit, receiving income only from one source. Table 1 summarises these.

---

\(^1\) However, sometimes the amount retained may be determined by individual arrangements rather than on the basis of fixed rules of the fiscal system, in which case the relationship \( T(H) \) has to be determined at the micro-level.
Table 1. Basic relationship among forms of income (one person, single source)

<table>
<thead>
<tr>
<th>form</th>
<th>relationship</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Gross income</td>
<td>G</td>
<td>GG = G + SS(G)</td>
</tr>
<tr>
<td>2 Social insurance contributions</td>
<td>S = S(G)</td>
<td></td>
</tr>
<tr>
<td>3 Gross taxable income</td>
<td>H = G - S</td>
<td></td>
</tr>
<tr>
<td>4 tax and SI contributions at source</td>
<td>XS = H</td>
<td>XST = H - T(H)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>XT = H + S(G) - T(H)</td>
</tr>
<tr>
<td>5 Deductions</td>
<td>D = D(H)</td>
<td></td>
</tr>
<tr>
<td>6 Net taxable income</td>
<td>Y = H - D</td>
<td></td>
</tr>
<tr>
<td>7 Tax due</td>
<td>W = W(Y)</td>
<td></td>
</tr>
<tr>
<td>8 Tax credits</td>
<td>C = C(Y)</td>
<td></td>
</tr>
<tr>
<td>9 Tax paid</td>
<td>X = W - C</td>
<td></td>
</tr>
<tr>
<td>10 Net income</td>
<td>N = H - X</td>
<td></td>
</tr>
</tbody>
</table>

SS: employer's SI contribution. Retentions at source: XS=SI only; XST=tax and SI; XT=tax only

3 Gross-to-Net conversion algorithm

Table 2 shows the basic relationship between gross and net forms of income in the more realistic situation when more than one income components and possibly more than one individuals in the tax unit are involved. Even in this case, the relationships between gross taxable income for a particular component, $H_i$, and quantities like gross income $G_i$ and income after retentions at source $XST_i$ are generally simple, dependent only on the income component concerned (i) and determined independently of other components and other persons in the tax unit. The same applies to the relationship between $H_i$ and net $N_i$ for components which are taxed separately at a flat rate or a rate determined only by the level of income from that component, and of course also for tax exempt components. Sometimes, dependence of the relationship on other sources of income may also be involved, but mostly these are simply in the form of upper limits which may apply to certain quantities pooled over more than one component.

Generally, all or most taxable income is pooled together over components and over persons in the tax units for the purpose of determining the amount of tax due. The relationship between $H_i$ and $N_i$ for components in the pool is, by contrast, more complex. Going from known $H_i$ to $N_i$ is simpler since the relationships (the tax rules) are a function of the former. These relationships are specified in more detail in Table 3. Going from given $N_i$ to $H_i$ required iterative solutions, and are described in the next section.
Because of its more direct relationship with all other quantities, we take $H_i$ as the 'base' form in discussion of the model.

**Social insurance contributions**

The social insurance contributions $S_i$, if applicable to the component, are generally a function of the gross amount $G_i$, but in the case of employment income excluding the employer's contribution. However, some more complex situations can be allowed for in the model while retaining its basic structure. Specifically, it can allow for the dependence of $S_i$ for any particular component $i$ on any set of income components, i.e., a functional relationship of the form $S_i = S_i(G_I)$, where subscript $I$ refers to any set of income components (normally including the particular $i$, of course). In the French system for instance, the pooled contributions for a number of components may be subject to a common maximum limit. The functional relationship $S_i(G_i)$ is specific to the component and the country. This is specified (and "called" as a subroutine in the application programs) separately from the common structure represented in Table 3.

**Deductions**

(Net) taxable income (row 7) is obtained by subtracting from gross taxable income the part which is tax exempt ("deductions"). These deductions are a certain function of gross taxable income. These may be of two types: (i) specific deductions which apply to the particular income components $D_i$ (row 4); and (ii) common deductions which apply to the (remaining taxable) income from all sources together (row 6). In case (i), in most situations the functional relationship $D_i(H_i)$ is specific to the component $i$, i.e., $D_i$ depends on the gross taxable income $H_i$ for the component concerned. As a generalisation, the model can allow for the dependence of $D_i$ for any particular component $i$ on any set of income components, i.e., a functional relationship of the form $D_i = D_i(H_I)$, - or even more
generally as $D_i = D(H_i, G_i)$ - where subscript $I$ refers to any set of income components (normally including the particular $i$, of course). In case (ii), a functional relationship of the form $D_0(H)$ is in terms of total gross taxable income i.e. all components together. Both types of functions are of course country-specific. Again, these relationships can be specified separately from the common structure represented in the table.

**Aggregation**

After the removal of component-specific deductions, it is necessary to pool the income over individuals in the same tax unit and across components which are treated together for taxation purposes. Certain income components may be excluded from this common "pool" and taxed separately; this type of situation is accommodated in the present model (see Section 5).

**Tax credits**

Initial tax due is computed as a function of total taxable income (row 8). This is determined by the countries "basic" income tax schedule, normally applied to pooled income from different sources. This tax liability is normally reduced by tax credits. Tax credits are mostly based on characteristics of the unit (single parent, pensioner, etc.) or are given in compensation for particular expenses (medical, educational, etc.), i.e., are not specific to a particular income source. We refer to these as "common tax credits" (row 9); these are normally expressed as a function of the total taxable income. The result is a more precise expression of "total tax due" (row 10). In addition to the common tax credits, there may also be component-specific tax credits (row 11). Generally, these are based on net taxable income for the component concerned. However, the functional relationship may be more complex: involving other components of income and/or income in other forms (gross, gross taxable, etc.).

**Tax paid and net income**

Deduction of these tax credits from the tax due (as defined in row 10), gives the final tax to be paid (row 12): i.e., total tax to be actually paid is tax due, less all (common as well as component-specific) tax credits. Total net income is total gross taxable income less tax paid (row 13).

The above two quantities, tax paid and net income (rows 12-13) refer at this stage to total income, and not to income by individual components.

**Tax rate**

This refers to the effective tax rate which applies to pooled components. Tax rate in Table 3 has been defined in two forms. (i) The first (row 15) is a descriptive measure. It is the ratio of the total amount of tax to be paid, to the total gross taxable income (row 12/row 3). Hence it is indicative of the overall tax burden. (ii) The second (row 16) provides a more analytical measure in the following sense. It is the ratio of the total amount of tax due before taking into account any component-specific tax credits (row 10), to the total taxable income after removing component-specific deduction (row 5). By removing all known component-specific aspects, that is component-specific deductions and tax credits, it can be seen as the common rate which applies to all taxable income, from whatever source, which has been pooled and subject to a common tax schedule.

---

2 Strictly, this may be referred to as 'disposable income'. Sometimes the term 'net income' as used for gross income less social insurance deductions and tax due, while 'disposable' also takes into account inter-household and some other transfers.
### Table 3. Gross-to-Net conversion algorithm

<table>
<thead>
<tr>
<th>Income measure</th>
<th>total</th>
<th>by component&lt;sup&gt;(1)&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 GROSS(2)</td>
<td></td>
<td>( G = \sum G_i ) ( \leftarrow ) ( G_i )</td>
</tr>
<tr>
<td>2 Social Insurance contribution</td>
<td></td>
<td>( S_i = S_i(G_i) )</td>
</tr>
<tr>
<td>3 GROSS TAXABLE</td>
<td></td>
<td>( H = \sum H_i ) ( \leftarrow ) ( H_i = G_i - S_i )</td>
</tr>
<tr>
<td>4 Component-specific deductions</td>
<td></td>
<td>( D_i = D_i(H_i) )</td>
</tr>
</tbody>
</table>

**Aggregation over components and individuals in tax unit**

| 5 TAXABLE INCOME |       |  \( Y = \sum Y_i \) \( \leftarrow \) \( Y_i = H_i - D_i \) |
| 6 Common deductions |       |  \( D_0 = D_0(H) \) |
| 7 Taxable income(0) |       |  \( Y_0 = Y - D_0 \) |
| 8 Tax due(0) |       |  \( W_0 = W_0(Y_0) \) |
| 9 Common tax credits |       |  \( C_0 = C_0(Y_0) \) |
| 10 TAX DUE |       |  \( W = W_0 - C_0 \) |
| 11 Component-specific tax credits |       |  \( C = \sum C_i \) \( \leftarrow \) \( C_i = C_i(Y_i) \) |
| 12 TAX PAID |       |  \( X = W - C \) |
| 13 TOTAL NET |       |  \( N = H - X \) |
| 14 Tax rate(0) |       |  \( R_0 = X/H \) |
| 15 TAX RATE = TAX DUE/TAXABLE INCOME |       |  \( R = W/Y \) |

**Disaggregation – personal income by component**

| 16 Proportionate tax by component |       |  \( X_i = R*Y_i - C_i \) |
| 17 NET BY COMPONENT |       |  \( N_i = H_i - X_i \) |

<sup>(1)</sup>The functional relationships in this column may be somewhat more complex or varied.

<sup>(2)</sup>Gross including employers' social insurance contribution (SS) is: \( GG = G + SS(G_1) \)

Parameter R has two functions. Firstly, it provides a means for the disaggregation of total tax and net income into component when required (see below). Secondly, it is the parameter of the iteration in going from net to gross, as described in the next section. Its role is even more important in the presence of missing data where modelling has to be used in conjunction with imputation. We explore these issues in a separate paper.

**Disaggregation of tax and net income by component**

This common 'tax rate' can be seen as a rate applying to each component individually, and not merely some average rate applicable only at the level of total income. This permits the decomposition of tax paid by income components (row 16), and consequently the decomposition of total net income into components (row 17). This decomposition is essential for the construction of variables such as net income before and after social transfers. For research and policy purposes, such decomposition is usually required in
much less detail than the breakdown of gross income. In any case, this sort of breakdown does not affect the performance of the rest of the system in the model in any way.

Country-specific schedules

The last two columns of Table 3 define the various income measures in terms of measures defined in the preceding rows; those in the first column concern total income, and in the second concern income components. The table involves six country-specific relationships or tax schedules:

- three concerning total income  
  \[ D_0 = D_0(H), \quad W_0 = W_0(Y_0), \quad C_0 = C_0(Y_0); \]

- another three specific to each component (i)  
  \[ S_i = S_i(G_i), \quad D_i = D_i(H_i), \quad C_i = C_i(Y_i). \]

The functional dependence can be somewhat more complex than indicated above, as explained earlier. In addition, there may be parameters determining retentions at source, taxation of parts of SI contributions, etc. Finally, it should be mentioned that the application of various formulae and relationships requires certain constraints to be met, such as to ensure that all quantities which, to be meaningful, must be non-negative are in fact so. It is not useful to list here such (and many other) programming details.

Some illustrations

The structure in Table 3 is very general and provides a common framework accommodating a wide variety of specific situations. We have found this to be the case at least for the three countries (Spain, Italy and France) for which the fiscal systems by individual income component have been examined in some detail. Here are some examples of how different types of situations can be handled within this common structure.

Spain

Tax-and-deduction information on the major income components fits into the common model in a straightforward way. (i) Most of the income is subject to tax and social insurance retentions at source. (ii) Income from self-employment is subject to specified schedule of SI contributions \( S_i(G_i); \ H_i = G_i - S_i \), with no component-specific deductions or tax credits \( C_i = D_i = 0 \). Real capital income is taxed in entirety, and not subject to SI or component-specific deductions or tax credits \( H_i = G_i; \ C_i = D_i = 0 \). Employment income is subject to specified schedules of employer's and employee's SI contributions, and to component-specific deductions and tax credits \( H_i = G_i - (S_{i1} + S_{i2}); \ C_i = 0; D_i > 0 \). Pensions, financial capital income and property income are not subject to SI contributions or tax credits, but each is allowed component-specific deductions \( H_i = G_i; \ C_i = 0; D_i > 0 \). (iii) The pooled taxable income \( H_i \) from all sources is subject to common deductions \( D_0 \), mostly on the basis of certain expenses made by the tax unit. Then it is subject to a common tax schedule to determine the amount of tax due. Certain common tax credits \( C_0 \) are deducted from the result to obtain the final tax to be paid. (iv) Family benefits are exempt from all retentions, and consequently, they do not contribute to the total taxable base.

Italy

Similarly in the case of Italy, the tax-and-deduction information on the major income components fits into the common model. (i) Income from self-employment is subject to specified schedule of SI contributions, with no component-specific deductions \( D_i = 0 \), but with two types of tax credits \( C_{i1} > 0 \) for "IRAP"; \( C_{i2} > 0 \) for other. Employment income is treated as in the case of Spain, except that no component-specific deductions have been specified \( C_i = 0; D_i > 0 \). (ii) Real capital income is taxed in entirety and treated exactly as in the case of Spain. For property income, the only difference is that in addition to
deductions, component-specific tax credits are also allowed \((C_i>0; D_i>0)\). Pensions are not subject to SI contributions and are allowed tax credits, but no component-specific deductions \((H_i=G_i-S_i; C_i>0; D_i=0)\). (iii) Financial capital income is subject to tax at a specified rate depending on its size, but independent of any income from other sources. Consequently, it does not contribute to the total taxable base. This can be handled in the model structure as described in Section 5. (iv) Family benefits are treated in exactly the same way as in Spain. These are exempt from all retentions, and hence do not contribute to the total taxable base.

**France**

(i) In France, income is generally subject to social insurance contributions at source, but is gross of income tax. This simplifies model implementation. (ii) On the other hand, the system of SI contributions is complex in a number of ways as noted earlier, complicating the programming task. This does not present any substantive problems, however. (iii) A certain type of capital income in France is subject to tax at a specified rate depending on its size, but independent of any income from other sources. Consequently, it does not contribute to the total taxable base, which can be handled in the model structure in the same way as "financial capital income" in the case of Italy. (iv) A special consideration is that certain types of SI contributions are subject to tax: generally CSG in part, and CRDS fully. This feature is easily handled in the common structure of the model (see Section 5).

### 4 The core iterative procedure

**Data forms**

The form in which data on income by component are available may vary from one country (tax regime) to another, and also among individuals and households within the same country. There are two dimensions of variation:

**A.** Whether or not a particular component is subject to social insurance contributions and to income tax. Income tax may apply in various forms. (i) Some components may be pooled together, across components and also across individuals in some appropriately defined tax unit. (ii) Some may be subject to tax separately, each at a certain flat rate. (iii) Some components in the 'pool' may be tax exempt up to a certain flat rate but taxed beyond that if a higher rate applies. (iv) Some may be subject to double taxation, perhaps representing some combination of the other forms.\(^3\) (v) And of course, many types of incomes, in particular social transfers, may be tax exempt. Mostly, the form applicable to each type of income is determined by the national tax regime, normally uniform for all respondents in a country. Hence this information can be compiled at the aggregate level and need not be collected at the micro level. There can be exceptions, however, for persons in special circumstances. There can also be other complications, such as more than one components, otherwise treated separately, being subject to common ceilings. In some systems, individuals have a choice among the various options.

**B.** The form in which the information has been collected. This may generally vary from one individual to another in the same survey, though a uniform reporting form may prevail for some components. In any case, the information on the form in which the data are available is required at the micro-level. The amount may for instances be reported as gross, or net of social insurance contributions and/or tax; and in the case of

---

\(^3\) A unique example is provided by the French system, where a part of social insurance contributions (themselves akin to a 'tax') is subject to income tax.
tax retentions, whether they are "retentions at source" according to some rules or individual arrangements or as the "final retentions" of the tax actually due, in the sense explained below. Table 4 lists the various reporting forms.

**Table 4. Forms of reporting of an income component**

<table>
<thead>
<tr>
<th>Income component (i) subject to tax and social insurance contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>B. Form (X) in which data on the income component have been collected:</strong></td>
</tr>
<tr>
<td>$G_i$ gross income (before tax and SI contributions, if applicable)</td>
</tr>
<tr>
<td>$H_i$ gross taxable (before tax, but after SI contributions, if any)</td>
</tr>
<tr>
<td>$N_i$ net income (after deducting 'final' tax and SI contributions, i.e., as the final amount actually received)</td>
</tr>
<tr>
<td><strong>Income received after retentions at source:</strong></td>
</tr>
<tr>
<td>$XT_i$ taxed at source (but no SI contribution); tax at source $T_i$</td>
</tr>
<tr>
<td>$XS_i$ SI contributions (but not tax) at source; SI contributions at source $S_i$</td>
</tr>
<tr>
<td>$XTS_i$ both tax and SI contributions at source, tax and SI at source $T_i + S_i$</td>
</tr>
</tbody>
</table>

In this section, we describe the standardised 'core' of the SM2 system, taking account of complexities B, but assuming for the moment that form (i) of A applies to the total income, i.e., the information may be reported in diverse forms, but all income components over individuals in the tax unit are pooled together and subject to a common tax schedule. A remarkable feature of the system is that by appropriately defining certain 'deductions' and tax credits, much of the complexity A can also be incorporated into the standardised procedures; this will be explained in Section 5.

**Income net of tax**

As noted above, in the case of tax retentions, an important distinction is to be made between:

(i) "retentions at source" (withholding taxes), and

(ii) the "final retentions" as appropriate for the income source concerned.

This is a very important distinction. It is essential to know what is meant when a component is reported as "net of tax". Does the information on retentions refer to withholding taxes, to final taxes, or even to some mixture? In some systems the withholding tax is quite different in size as well as structure to the final tax liability, and the taxpayers may even be able to choose their withholding rate of tax.\(^4\)

**Tax retention at source**

Among the two, this may be the more common form in which net income is reported. We take "retention at source" to mean that *the amount of tax has been assessed depending only on the income received from the particular source concerned, not taking into account income received from any other sources or the individual's (the tax unit's) personal characteristics.*

\(^4\) By contrast, there can also be systems (such as in Spain) which aim at fine-tuning deductions at source to match closely the final amounts due.
Indeed, in many situations, this retentions may be according to relatively simple and standard rules, which may be express, say, as

\[ T_i = (H_i - XST_i) = T_i(H_i). \]

where tax retention at source (T), being the difference between gross taxable income (H) and the amount received after social insurance and tax retention at source (XST), is some known function of gross taxable income for only the component concerned. Provided that these rules are standard and known, XST; is directly convertible to H; without reference to other components of income for the unit. By comparison, the relationship with H of the "final net" N; is more complex, as it depends on the unit's total income from all sources.

The real difficulty however arises when the rules for retention at source are not standard, are not applied uniformly, or are even non-existent in the sense that the taxpayers can choose or negotiate their withholding tax rates. In such situations, the construction of the gross taxable amount from the reported amount after withholding tax will require separate information on the amount withheld (or the withholding rule applied) in the particular case.

**Final tax retention**

By contrast to the above, the final tax retention is meant to reflect the tax actually due after taking into account the total income situation and characteristics of the tax unit. Consequently, the rules involved in this case tend to be more complex and involve the nature of the unit (individual person, household, or some other tax unit), the unit's particular circumstances and its income from all sources simultaneously.

On the other hand, those rules are supposed to be applied (except for tax evasion and similar factors not considered here) in a standard way, not subject to variations according to individual arrangements as may apply to some retentions at source.

In practice, there may often be some ambiguity as to what a figure reported as "net" by a survey respondent actually represents. For instance, employers often adjust the employee's "tax code" on the basis of tax returns for previous years, such that the amount withheld at source actually approximates the amount of "final tax" which the employee would have to pay on this income in accordance with the prevailing tax rules. In the presence of such ambiguity, it is perhaps safer to interpret the amount reported in the sense of "net after paying the final tax due". Then the term "tax retention at source" would be reserved for situations where the "retention at source" rules or arrangements have been applied more clearly.

**Social insurance contributions**

In contrast to tax retentions, social insurance contributions are essentially component-specific, i.e. determined only or mainly in relation to the income component concerned, so that the above distinction between "retention at source" and "final retention" is generally not relevant. They are usually collected at source in any case.5

---

5 The situation can be more complicated. However, as noted earlier in relation to the French system, generally such complications merely make the algorithm specifying the various functional relationships more elaborate, but there is no problem in handling them within the common structure of the model.
Conversion routines

Table 5 shows the procedure for converting the reported amount with any combinations of the above dimensions of variation into a standard form. For all forms other than 'final net' $N_i$, it is convenient to take 'gross taxable income' $H_i$ as the standard.

$$[G_i, H_i, XS_i, XTS_i, XT_i] \Rightarrow H_i.$$

This conversion involves the component and country-specific functional relationships or schedules, namely

$$S_i = S_i(G_i),$$  social insurance contributions, and

$$T_i = T_i(H_i),$$  tax retention at source.

As noted, tax retentions at source may be according to fixed schedules, or according to arrangements determined at the individual (micro) level.

In a majority of the cases, $H_i$ can be determined directly from the collected amount, for instance from gross amount ($G_i$) reported for an income component $i$ subject to social insurance contributions:

$$H_i = G_i - S_i(G_i).$$

In other cases, an iterative procedure may be required. However, generally the iteration is very simple and converges quickly. This is because by and large component-specific schedules apply to each component separately. There are no other parameters to be estimated. The need for numerical iteration arises simply from the fact that the unknown quantity to be determined ($H_i$) appears in an implicit equation.

Table 5. Calculation of $H_i$ according to the form in which the component is specified.

**Set H**

<table>
<thead>
<tr>
<th>given value $P_i =</th>
<th>XS_i</th>
<th>H_i = XS_i</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_i$</td>
<td>$H_i = G_i - S_i(G_i)$</td>
<td></td>
</tr>
<tr>
<td>$XT_i$</td>
<td>$H_i = G_i - S_i(G_i)$ where $G_i = XT_i + T_i(H_i)$</td>
<td></td>
</tr>
<tr>
<td>$XTS_i$</td>
<td>$H_i = XTS_i + T_i(H_i)$</td>
<td></td>
</tr>
</tbody>
</table>

**Simple iteration, generally separately for each component**

**Set N**

<table>
<thead>
<tr>
<th>given value $P_i =</th>
<th>N_i</th>
<th>H_i = Y_i + D_i(H_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_i$</td>
<td>$Y_i = [H_i - N_i + C_i(Y_i)] / R$</td>
<td></td>
</tr>
</tbody>
</table>

**Double iteration**

(i) with assumed $R$, for each component in turn

(ii) for determining $R$, common to all pooled components

The second panel of the table shows the relationship between $H_i$ and the reported amount in the form 'final net' $N_i$. Going from $N_i$ to $H_i$ in fact involves a double iterative loop. The inner loop of iteration is applied with an assumed value of the parameter "tax rate" ($R$, as defined in Table 3). Once this has been done for every income component in the group (including over all individuals in the same tax unit), an outer iterative loop obtains a convergent value of this parameter which is common to all those components.

The $N_i$ to $H_i$ conversion process is therefore considerably more complex. Furthermore, this complexity is substantially increased in the presence of missing data, where the
modelling and imputation procedures will have to be applied interactively. (We discuss this issue in a separate paper.)

**Iterative procedure**

Table 6 demonstrates the common structure of the iterative procedure. As noted at the bottom of the table, the income components may be divided into two sets, say 'N' and 'H', depending on whether the amount reported is 'final net' ($N_i$), or is in some other form ($G_i, XSi, XTi, XTSi, Hi$) more directly convertible to the 'gross taxable' form $H_i$.

The procedure may be applied as follows. The required $H_i$ quantities for set H are computed (only once) using Table 5, and form an input into the iterative cycle for parameter $R$ required for set N. The parameter is best estimated by using information on all income components from both the sets.

**Table 6**
Common structure of the iterative model

| Reported amount | $\rightarrow$ | Gross taxable | $\rightarrow$ | Net and gross |
|-----------------|---------------|---------------|---------------|
| All data in or convertible to the "H" form: |
| $X_i$ $\rightarrow$ Table 5 $\rightarrow$ $H_i$ $\rightarrow$ $\rightarrow$ Table 3 $\rightarrow$ $\rightarrow$ $G_i, N_i$ |
| Data a mixture of "H" and "N" forms: |
| Set H $X_i$ $\rightarrow$ Table 5 $\rightarrow$ $H_i$ $\rightarrow$ $\rightarrow$ Table 3 $\rightarrow$ $\rightarrow$ $N_i$ $\rightarrow$ $G_i$ $\rightarrow$ $R$ |
| Set N $X_i$ $\rightarrow$ Table 5 $\rightarrow$ $H_i\uparrow R$ $\rightarrow$ $\rightarrow$ $\rightarrow$ $\rightarrow$ $R$ |

Set of variables N: set of income components which are subject to income tax (irrespective of whether the component is also subject to social insurance contributions), and for which the 'final net' amount ($X_i=N_i$) has been specified in the data collected.

Set of variables H: all other income component (not subject to tax, or for which the data has been collected in a form other than the 'final net' amount)

For application in the multi-country comparative context, which motivates the development of SM2, it is very desirable (and possible) to standardise to a high degree the
list of modelling variables (income components) across countries, removing the effect of differences in the list of collection variables. The individual data records with collection variables can be aggregated to the required level of tax units and modelling variables, the latter standardised across countries to the extent possible. For each income component for each unit, the data file should contain information on the amount received and the form in which it has been reported. The various tax and deduction schedules are also country-specific. Beyond these variations, however, the procedures described above have a common structure, applicable to different tax systems.

5 Special deductions and tax credits: a device to accommodate diversity

We return to complexities "A" in the diversity of forms of taxation to which individual components of income may be subject in different systems. As noted, a remarkable feature of SM2 is that by appropriately defining certain 'deductions' and tax credits, much of these complexities can be incorporated into the standardised procedures described in the previous section without altering them in any way.

Specification of special deductions and tax credits

Deductions refer to the part of gross taxable income which is tax exempt (see Section 2). These deductions are a certain function of gross taxable income, and the difference between the two quantities gives the net taxable income, i.e. income subject to tax. These deductions may be component-specific or common deductions which apply to taxable income as a whole. Initial tax due is computed as a function of total net taxable income. This tax liability is normally reduced by tax credits. Again, these may be component-specific or common credits which apply to the initial tax due as a whole. In addition to these 'normal' deductions and tax credits, we can define 'special' component-specific deductions and tax credits to accommodate variations in the form in which the component is taxed without altering any other aspect of model specification. Consider for instance the common situation with one component tax exempt, and the remaining components pooled together and subject to a common tax regime.

By simply specifying special deduction for the tax exempt component as \( D_i = H_i \), its gross taxable amount, we automatically retain its tax-exempt nature and it is no more necessary to separate it from rest of the pool. It makes no contribution to the total net taxable income, and its original gross taxable income appears automatically as a part of the final net income. Similarly, if a component is taxed at a flat rate (say \( f \)) separately from the pool, we can simply specify its special deduction as \( D_i = H_i \) and its special tax credit as a negative quantity \( C_i = -f \times H_i \). It makes no contribution to the tax liability of the pool, but the final tax liability is automatically increased by the appropriate amount. Again, no other treatment separate from the pool is required for this component. Table 7 lists a number of such possibilities. These cover all the situations we have encountered.
Table 7: Examples of special deductions and tax credits

<table>
<thead>
<tr>
<th>Form of taxation of component i</th>
<th>Special deduction</th>
<th>Special tax credit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Tax exempt</td>
<td>$D_i=H_i$</td>
<td>-</td>
</tr>
<tr>
<td>2 Taxed at flat rate $f_i$</td>
<td>$D_i=H_i$</td>
<td>$C_i=-f_i*H_i$</td>
</tr>
<tr>
<td>3 Tax-exempt at flat rate $f_i$</td>
<td>-</td>
<td>$C_i=+f_i*H_i$</td>
</tr>
<tr>
<td>4 Deductions for expenses</td>
<td>+common deductions</td>
<td>-</td>
</tr>
<tr>
<td>5 Tax credit for expenses</td>
<td>-</td>
<td>+common tax credits</td>
</tr>
<tr>
<td>6 Special tax not related to income</td>
<td>-</td>
<td>-common tax credits</td>
</tr>
<tr>
<td>7 Double taxation at flat rate $f_i$</td>
<td>-</td>
<td>$C_i=-f_i*H_i$</td>
</tr>
<tr>
<td>8 Part $\Delta S_i$ of social insurance contributions subject to tax</td>
<td>$-\Delta S_i$</td>
<td>-</td>
</tr>
</tbody>
</table>

Different forms may apply to cases like 3, 4 and 7: for instance the tax rate being a function of the amount of income involved for the component concerned.

Examples

**Taxable SI contributions.** The last case (8) is an important one, as it handles a special and complicating factor in the treatment of social insurance contributions in, for instance, France. Certain types of SI contributions are subject to tax: generally CSG in part, and CRDS fully (or sometimes not at all, depending on the income component). By specifying the taxable part of SI contributions as negative deductions from (i.e. in effect as additions to) gross taxable income defined in the same way as in other case ($H_i=G_i-S_i$), the net taxable income (the amount actually subject to tax) is augmented by the taxable part $\Delta S_i$, $Y_i=H_i+\Delta S_i$. No further special treatment of this feature of the system is required in the model.

**Family benefits.** Family benefits are commonly exempt from all retentions. Consequently, they do not contribute to the total taxable base. As noted, this can be handled in the model structure simply by specifying component-specific deduction to equal the amount of income received, making their contribution to the net taxable income zero. Hence the relationships are ($G_i=H_i=N_i=X_i$; $C_i=0$; $D_i=H_i$; hence $Y_i=0$), where $X_i$ is the amount of income received specified in the survey, and $Y_i$ is (net) taxable income for the component. As another example, in Spain some sickness invalidity benefits are tax exempt, e.g. when the individual is completely disabled, or receiving sickness benefit due to HIV. These can be treated in the same way.

**Double taxation.** In Italy, self employment income is liable to income taxation and to an additional tax, IRAP (*Imposta sul reddito delle attività produttive*).

**Flat rate or other separate taxation.** Financial capital income in Italy is subject to tax at a specified rate depending on its size, but independent of any income from other sources. Consequently, it does not contribute to the total taxable base. This can be handled in the model structure by making component-specific deduction equal to the amount of income received, and introducing a negative tax credit, equal in size to the tax to be paid on financial capital. Hence the relationships are ($D_i=H_i$; $C_i=-T_i(H_i)$), where $T_i$ is the amount of tax due as a function of the amount of capital income received. Similarly in Spain, self
employment income, capital income and income from renting out buildings or lands are taxed at source at a fixed (flat) rate, depending on the component concerned: \( (D_i=H_i; C_i=-f_i^H_i) \).

A certain type of capital income in France is subject to tax at a specified rate depending on its size, but independent of any income from other sources. Consequently, it can be handled in the model structure in the same way as "financial capital income" in the case of Italy described above.

**Tax credits.** Examples are encountered of two types of tax credits. One type of tax credits relate to rebates on certain types of expenditures or outgoings, such as mortgage interest payments in certain situations. These are not components of income, but are handled as a part of the common tax credits \((C_0)\) as appropriate, the magnitude of which would normally be a function of the amount of expenditure/outgoing involved.

The same applies to certain types of expenditures which can be deducted from taxable income before the application of the tax schedule, for example certain medical expenses. Here, these can be handled as a part of the common deductions \((D_0)\).

Certain income components may be allowed a tax credit at some specified rate, but otherwise are included in the determination of the unit's total income to which the overall tax schedule is to be applied. This is the complement of the situation with regard to financial capital income described above in the case of Italy and France. It can be handled in the model by including the component concerned fully in the pool of the total income for the computation of tax \((D_i=0)\); and subsequently allowing component-specific tax credit for an appropriate amount \((C_i>0)\), normally a function of the amount of taxable income for the component concerned.

**The system in Spain**

We conclude by providing a detailed illustration of how a system with very special feature has been handled within the general structure of SM2.

**Dual system**

In Spain, the tax rate for retention at source is determined on the basis separate income from each component. There is also a system (not necessarily obligator for all persons) in which the final tax rate (IRPF) is determined on the basis of pooled income from all taxable components. Hence the characteristic feature of the Spanish system is that two systems of tax computation are applied in parallel: a system of tax withholding at source; and a system based on tax returns giving the 'final' tax liability. The former system aims at making it possible for a large part of the taxpayers to fulfil their tax obligations through the payment of withholdings (tax withheld at source), which the system aims to closely adjust to the right amount of their liability under the later (tax-return) system. In this way, the number of tax payers who need to make a tax return can be significantly reduced. Essentially, the system operates as follows.

- Income recipients with a total income above a certain level (depending on various personal characteristics) are subject to tax *withholding at source*.
- A subset of those subject to withholding at source are not required to make a *tax return*, and the taxation at source can be seen as the actual or net taxation. These people do not have to present the tax return, limiting their tax burden exclusively to the monthly withholding. However, if that is convenient or advantageous to them, these
persons may nevertheless choose to make a tax return, and hence have their final tax liability adjusted to that return.

- Persons who are obliged to make a tax return have their final tax liability adjusted to that return in any case. The tax return gives the 'final' tax liability. Some of the taxpayers may not actually have to pay any income tax because of the various deductions which may apply.

- Persons who are obliged to or who, while not obliged, nevertheless choose to make a tax return, have the option of choosing between individual and joint taxation, provided they meet the required conditions to qualify for joint taxation.

**Main differences**

The two systems have a similar structure, and in fact aim to replicate each other in the final outcome. The main differences are the following.

1. The withholding system takes individual taxpayer as the unit; the tax return may be on an individual or, where applicable, a joint basis.

2. The common income base in the tax-return system is the total taxable income of the individual, or where applicable, of the joint unit. For the withholding system, each component is considered separately.

3. The tax-return system allows for certain tax credits (tax refunds); these do not apply in the tax-at-source regime.

The social insurance and tax schedules are the same or very similar in the two systems, apart from the effect of the choice of tax unit (individual versus joint) in the tax return system. The effect (2) and (3) are compensatory, helping in the aim of making the two systems replicate each other in the final outcome.

**Options**

In summary, the following are the options available to Spanish taxpayers:

<table>
<thead>
<tr>
<th>Family situation</th>
<th>Whether obliged to make tax return</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single, i.e. not in a fiscal family unit</td>
<td>Obliged</td>
<td>Must make individual return (no option)</td>
</tr>
<tr>
<td></td>
<td>Not obliged</td>
<td>1 Taxed at source (no tax return)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Makes tax return (individual)</td>
</tr>
<tr>
<td>In fiscal family unit</td>
<td>Obliged</td>
<td>1 Makes tax return (individual)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Makes tax return (joint)</td>
</tr>
<tr>
<td></td>
<td>Not obliged</td>
<td>1 Taxed at source (no tax return)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 Makes tax return (individual)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 Makes tax return (joint)</td>
</tr>
</tbody>
</table>

What do individuals and households choose when more than one options are available? The ECHP data which we have used in our work, for instance, provide no information on whether a particular respondent has elected to make a tax return, and to be taxed individually or jointly. We have used the following strategy in dealing with this issue. In the absence of information on what a person or household actually did, it seems
reasonable to expect that individuals and households will choose the arrangement least disadvantageous to them. Where a choice exists between making an individual or a joint return, the model can determine which of those two results in a lower tax liability. Similarly, where a choice exists between accepting taxation at source or making a tax return (individual or joint, whichever is advantageous), the model can determine which of those two results in a lower tax liability.

The specification of the two systems in SM2 is very similar. The same model structure applies to the two very different systems just through specifying, apart from the possibly different tax units, the deduction and tax credit schedules appropriately in each case.

References


Eurostat (2002), Commission Regulation concerning EU-Silc as regards the list of primary target variables, SILC107.


7. B. Contini, *Wage structures in Europe and in the USA: are they rigid, are they flexible?*, Gennaio 1999.


Le richieste di working papers vanno indirizzate alla Sig.ra Patrizia Piacentini, Dipartimento di Scienze Statistiche, via C. Battisti 241-243, 35121 Padova, oppure all’indirizzo e-mail patrizia.piacentini@unipd.it.