

Conference and Methodological
Workshop on
"Comparative EU-Statistics and
Living Conditions:
Issues and Challenges"

**Gross-net
conversion issues**

Gianni Betti and Vijay Verma

Università degli Studi di Siena



1

AIMS of the PAPER

- Develop a tool for constructing EU-SILC income target variables
- Present a Micro Simulation Model with these characteristics
- Test the Model with ECHP for France, Italy and Spain
- Discuss Micro Simulation and Imputation in conjunction

2

**SM2: Siena Micro
Simulation Model**

Multi-country and flexible tool portable across countries
Flexible to deal with an annual flux of data in different forms

Objective:

- To convert information on income (under specified national tax system) to a standard form
- **It comprises:**
 - *Standardised set of routines*
(to handle diversity of input data forms and national tax systems)
 - *Country-specific routines*
(to convert the input data into standardised forms to specify parameters of the national tax system in standardised form)
- SM2 is fully "data based"
(does not incorporate simulation of benefits or other income components)

3

**Terminology: forms of
income and their
relationship**

- Gross income (GG, G): GG includes also employer's social insurance contributions (SS); G includes other SI contributions.
- Social insurance contributions (S): includes (i) employer's contribution on behalf of persons in employment; (ii) employee's contributions; (iii) self-employed person's contributions. (iv) other income
- Gross taxable income (H): gross income less social insurance contributions:
 $H = G - S$.
- Deductions (D): gross taxable income exempt from tax
- Net taxable income (Y) gross taxable income less deduction:
 $Y = H - D$
- Tax due (W): function of net taxable income, $W = W(Y)$
- Tax credits (C): The tax liability is normally reduced by tax credits
- Tax paid (X): Tax due less tax credits: $X = W - C$
- Total net income (N): total gross taxable income less tax paid: $N = H - X$
- Income subject to retention at source of tax and/or social insurance contributions:
 $XTS = G - S(G) - T(H) = H - T(H)$

4

Basic Relationship among Forms of Income (one person, single source)

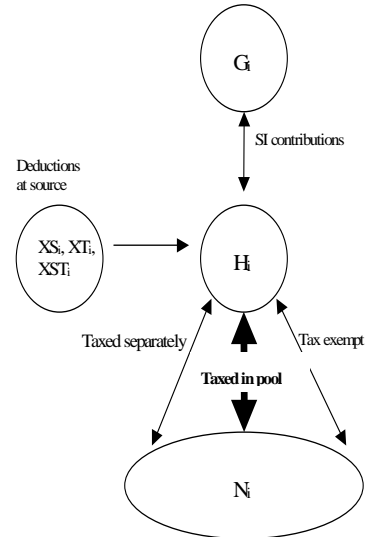
Table 1 - Basic relationship among forms of income (one person, single source)

| | form | relationship | comment |
|----|--------------------------------|------------------------------------|--|
| 1 | Gross income | G | $GG = G + SS(G)$ |
| 2 | Social insurance contributions | $S = S(G)$ | |
| 3 | Gross taxable income | $H = G - S$ | |
| 4 | ↓ | tax and other deductions at source | $XS = H$ $XST = H - T(H)$ $XT = H + S(G) - T(H)$ |
| 5 | Deductions | $D = D(H)$ | |
| 6 | Net taxable income | $Y = H - D$ | |
| 7 | Tax due | $W = W(Y)$ | |
| 8 | Tax credits | $C = C(Y)$ | |
| 9 | Tax paid | $X = W - C$ | |
| 10 | Net income | $N = H - X$ | |

SS: employer's social insurance contribution. Retentions at source: XS=social insurance only, XST=tax and social insurance, XT=tax only.

5

Basic Relationship Between Net and Gross Amounts



6

Gross-to-net algorithm

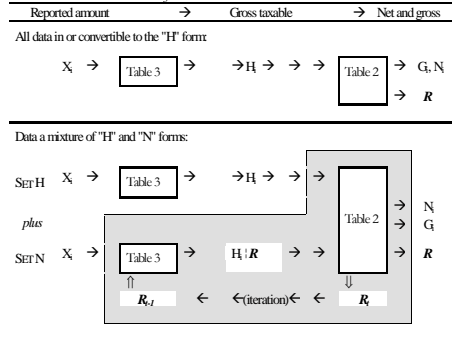
Table 2 - Gross-to-Net conversion algorithm

| | Income measure | total | by component ⁽¹⁾ |
|---|------------------------------------|------------------|-----------------------------|
| 1 | GROSS ⁽²⁾ | $G = \sum G_i$ | G_i |
| 2 | Social Insurance contribution | | $S_i = S_i(G_i)$ |
| 3 | GROSS TAXABLE | $H = \sum H_i$ | $H_i = G_i - S_i$ |
| 4 | Component-specific deductions | | $D_i = D_i(H_i)$ |
| Aggregation over components and individuals in tax unit | | | |
| 5 | TAXABLE INCOME | $Y = \sum Y_i$ | $Y_i = H_i - D_i$ |
| 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 | $C_i = \sum C_i$ | $C_i = C_i(Y_i)$ |
| 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 | | $N_i = H_i - X_i$ |

7

The core iterative procedure

Table 4 - Common structure of the iterative model



8

Tax Rate: R

R refers to the effective tax rate which applies to pooled components.

- An analytical measure.
Ratio of:
total amount of tax due before taking into account any component-specific tax credits
to the total taxable income after removing component-specific deductions.

It can be seen as the *common rate* which applies to all taxable income which has been pooled and subject to a common tax schedule (W/Y).

- It permits the *decomposition* of tax paid by income components, and the decomposition of total net income into components.
- It is a *parameter in the iterative procedure* of the model
- It is a critically useful in the application of *imputation and modelling in conjunction*

9

Handling diverse situations in SM2

The concept of 'special' deductions and tax credits

Table 5 - Examples of special deductions and tax credits

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

10

Applications

| country | Data source and reference year | Fiscal system reference year | Main form of input data from the survey (by component) | |
|---------|--|---------------------------------------|---|-----|
| France | ECHP1998 | 1998 | Gross of tax but net of social insurance deductions | XS |
| Italy | ECHP1998 Supplementary: ISTAT household budget survey Bank of Italy income survey | 2003 | Final net income | N |
| Spain | ECHP1999 | 1999 (applied) 2003 (also studied) | Net of taxation and social insurance deducted at source | XTS |

11

Application to France: comparison with INSEE

Table 6 - Application to France: comparison with INSEE

| | Composition of gross income (INSEE) | | comparison: | |
|---|-------------------------------------|--------------|-----------------------|--------|
| | mean amount | distribution | SM2 | INSEE |
| [1] net-household tax | 67,858 | 67.2 | | |
| [1a] net | | 66.4 | 86.6 | 85.7 |
| [1b] household tax | | 0.8 | 1.1 | 1.1 |
| [2] personal income tax | 4,086 | 4.0 | 5.3 | 5.9 |
| [3] CSG-CRDS, SI on capital income | 5,431 | 5.4 | 7.0 | 7.3 |
| subtotal [1]+[2]+[3] (revenu avant impôts) | 77,375 | 76.7 | 100.0 | 100.0 |
| [4] other personal SI contributions | 6,180 | 6.1 | Mean/household (Euro) | |
| [5] employer's SI contribution | 17,297 | 17.2 | 30,200 | 29,856 |
| total social insurance contributions | 29,008 | 28.7 | | |
| total gross | 100,952 | 100.0 | | |

Source: INSEE: INSEE PREMIERE, n. 916, August 2003, 'Des ménages modestes aux ménages aisés: des sources de revenus différentes'. SM2: France ECHP Wave 6.

12

EU-SILC target variables in France

Table 7 – France EU-SILC target variables: distribution of income by component

| | mean amount | | ratio net/gross | % distribution | |
|--|---------------|---------------|--------------------|----------------|--------------|
| | gross (1) | net (2) | | gross (4) | net (5) |
| income from work | 74,269 | 44,121 | 59.4 | 75.0 | 66.3 |
| PY010 employee cash or near cash income | 41,228 | 38,951 | 94.5 | 41.6 | 58.5 |
| employee's SI contribution | 17,059 | | | 17.2 | |
| employee's SI contribution | 9,549 | | | 9.6 | |
| PY050 cash benefits or losses from self-employment | 5,576 | 5,170 | 92.7 | 5.6 | 7.8 |
| Self-employed SI contribution | 857 | | | 0.9 | |
| property income | 2,664 | 2,217 | 83.2 | 2.7 | 3.3 |
| HY030 interest, dividends, profit from capital | 1,704 | 1,531 | 89.9 | 1.7 | 2.3 |
| capital income SI contribution | 185 | | | 0.2 | |
| HY040 income from rental of a property or land | 775 | 685 | 88.5 | 0.8 | 1.0 |
| taxable benefits | 20,383 | 18,526 | 90.9 | 20.6 | 27.8 |
| PY090 unemployment benefits | 1,693 | 1,620 | 95.7 | 1.7 | 2.4 |
| unemployed SI contribution | 95 | | | 0.1 | |
| PY100 old-age benefits | 14,170 | 13,278 | 93.7 | 14.3 | 20.0 |
| PY110 survivor benefits | 897 | 822 | 91.6 | 0.9 | 1.2 |
| pension SI contribution | 681 | | | 0.7 | |
| PY120 sickness benefits | 669 | 669 | 100.0 | 0.7 | 1.0 |
| PY130 disability benefits | 394 | 372 | 94.3 | 0.4 | 0.6 |
| disability SI contribution | 18 | | | 0.0 | |
| HY050 family related allowances | 1,741 | 1,741 | 100.0 | 1.8 | 2.6 |
| PY150 other personal benefits | 25 | 25 | 100.0 | 0.0 | 0.0 |
| tax-exempt social transfers | 1,674 | 1,674 | 100.0 | 1.7 | 2.5 |
| PY140 education-related allowances | 152 | 152 | 100.0 | 0.2 | 0.2 |
| HY060 social assistance | 152 | 152 | 100.0 | 0.2 | 0.3 |
| HY070 housing allowances | 720 | 720 | 100.0 | 0.7 | 1.1 |
| HY080 regular inter-household cash transfer received | 621 | 621 | 100.0 | 0.6 | 0.9 |
| total | 98,990 | 66,539 | 67.2 | 100.0 | 100.0 |

French France 1998, per capita

13

Application to Italy: comparison with ISTAT

Table 8 – Application to Italy: comparison with ISTAT

| | SM2 (data: ECHP 1998) (tax system: 2003) | ISTAT | Error (% points) | |
|---------------------------------------|--|-------|---------------------|-------|
| Gross including SI | 10,241 | 100.0 | 100.0 | |
| SI contributions | | | | |
| - Employers' contribution | 1,361 | 13.3 | 11.1 | 2.2 |
| - Employees' contribution | 416 | 4.1 | 3.2 | 0.9 |
| - Self-employment contribution | 202 | 2.0 | 1.6 | 0.4 |
| gross taxable | 8,261 | 80.7 | 84.1 | (3.4) |
| Personal income tax and financial tax | 1,044 | 10.2 | 11.6 | (1.4) |
| net income | 7,217 | 70.5 | 72.5 | (2.0) |

Euro 2003, per capita.

Sources: ISTAT: National Account (1998), SM2: Italy ECHP Wave 6 Our Model.

14

Social Insurance contributions in Italy

Table 9 – Distribution of total gross income excluding employers' SI contributions

| | SM2 (data: ECHP 1998) (tax system: 2003) | ISTAT | Error (% points) | |
|--|--|-------|---------------------|-------|
| Gross excluding employers' SI contribution | 8,880 | 100.0 | 100.0 | |
| - Employees' contribution | 416 | 4.7 | 3.6 | 1.1 |
| - Self-employment contribution | 202 | 2.3 | 1.8 | 0.5 |
| gross taxable | 8,261 | 93.0 | 94.6 | (1.6) |
| Personal income tax and financial tax | 1,044 | 11.8 | 13.0 | (1.3) |
| net income | 7,217 | 81.3 | 81.6 | (0.3) |

Sources: ISTAT: National Account (1998), SM2: Italy ECHP Wave 6 Our Model.

15

EU-SILC target variables in Italy

Table 10 – Italy EU-SILC target variables: distribution of income by component

| | mean amount | | ratio net/gross | % distribution | |
|--|---------------|--------------|--------------------|----------------|--------------|
| | gross (1) | net (2) | | gross (4) | net (5) |
| income from work | 7,436 | 4,691 | 63.1 | 72.6 | 65.0 |
| PY010 employee cash or near cash income | 4,319 | 3,774 | 87.4 | 42.2 | 52.3 |
| employee's SI contribution | 1,361 | | | 13.3 | |
| employee's SI contribution | 416 | | | 4.1 | |
| PY050 cash benefits or losses from self-employment | 1,137 | 917 | 80.7 | 11.1 | 12.7 |
| Self-employed SI contribution | 202 | | | 2.0 | |
| property income | 275 | 228 | 83.1 | 2.7 | 3.2 |
| HY090 interest, dividends, profit from capital investments in unincorporated business | 161 | 134 | 83.7 | 1.6 | 1.9 |
| HY040 income from rental of a property or land | 114 | 94 | 82.4 | 1.1 | 1.3 |
| taxable benefits | 2,454 | 2,221 | 90.5 | 24.0 | 30.8 |
| PY090 unemployment benefits | 74 | 65 | 88.2 | 0.7 | 0.9 |
| PY100 old-age benefits | 1,866 | 1,685 | 90.3 | 18.2 | 23.4 |
| PY110 survivor benefits | 316 | 291 | 92.0 | 3.1 | 4.0 |
| PY120 sickness benefits | 3 | 2 | 90.2 | 0.0 | 0.0 |
| PY130 disability benefits | 160 | 142 | 89.0 | 1.6 | 2.0 |
| HY050 family related allowances | 34 | 34 | 100.0 | 0.3 | 0.5 |
| PY150 other personal benefits | 1 | 1 | 100.0 | 0.0 | 0.0 |
| tax-exempt social transfers | 76 | 76 | 100.0 | 0.9 | 1.2 |
| PY140 education-related allowances | 5 | 5 | 100.0 | 0.0 | 0.1 |
| HY060 social assistance | 6 | 6 | 100.0 | 0.1 | 0.1 |
| HY070 housing allowances | 2 | 2 | 100.0 | 0.0 | 0.0 |
| HY080 regular inter-household cash transfer received | 64 | 64 | 100.0 | 0.6 | 0.9 |
| total | 10,241 | 7,217 | 70.5 | 100.0 | 100.0 |

Euro 2003, per capita.

16

Application to Spain: comparison with OECD

Table 11 – Application to Spain: comparison with OECD

| | SM2 (ECHP1999) | | OECD (for 1997) | | |
|----------------------------------|----------------|--------------|------------------------------------|-------|-------|
| | mean amount* | distribution | Assumed ratio private income / GDP | | |
| | | | (1) | (2) | (3) |
| | | | 0.67 | 0.69 | 0.71 |
| Gross including social insurance | 1,377,492 | 400.0 | 100.0 | 100.0 | 100.0 |
| Employers' contribution | 157,072 | 11.4 | 12.4 | 12.0 | 11.7 |
| Employees' contribution | 47,287 | 3.4 | 2.8 | 2.7 | 2.7 |
| Self-employment contribution | 30,450 | 2.2 | 2.4 | 2.3 | 2.3 |
| gross taxable | 1,142,683 | 83.0 | 82.4 | 83.0 | 83.3 |
| Personal income tax | 145,647 | 10.6 | 11.0 | 10.7 | 10.4 |
| net income | 997,040 | 72.4 | 71.4 | 72.2 | 73.0 |

*Pesetas 1999 per capita. Sources: OECD: *Revenue Statistics*, Paris (1999). (1)-(3) exemplify sets of figures which can be derived from the published results, depending on the assumed ratio of private income to GDP per capita. SM2: Spain ECHP Wave 7.

17

EU-SILC target variables in Spain

Table 12 – Spain EU-SILC target variables: distribution of income by component

| | mean amount | | ratio net/gross (3) | % distribution | |
|---|------------------|----------------|---------------------|----------------|--------------|
| | gross (1) | net (2) | | gross (4) | net (5) |
| income from work | 1,096,818 | 737,057 | 67.2 | 79.6 | 73.9 |
| PY010 employee cash or near cash income | 719,829 | 615,794 | 85.5 | 52.3 | 61.8 |
| employee's SI contribution | 157,072 | | | 11.4 | |
| employee's SI contribution | 47,287 | | | 3.4 | |
| PY050 cash benefits or losses from self-employer | 142,179 | 121,263 | 85.3 | 10.3 | 12.2 |
| Self-employed SI contribution | 30,450 | | | 2.2 | |
| property income | 36,948 | 31,915 | 81.9 | 2.8 | 3.2 |
| HY030 interest, dividends, profit from capital invest | 26,659 | 21,653 | 81.2 | 1.9 | 2.2 |
| HY040 income from rental of a property or land | 12,289 | 10,262 | 83.5 | 0.9 | 1.0 |
| taxable benefits | 236,660 | 223,003 | 94.2 | 17.2 | 22.4 |
| PY100 unemployment benefits | 14,011 | 13,567 | 96.8 | 1.0 | 1.4 |
| PY100 old-age benefits | 154,061 | 143,645 | 93.2 | 11.2 | 14.4 |
| PY110 survivor benefits | 36,406 | 35,395 | 97.2 | 2.6 | 3.5 |
| PY120 sickness benefits | 6,174 | 6,120 | 99.1 | 0.4 | 0.6 |
| PY130 disability benefits | 21,773 | 20,248 | 93.0 | 1.6 | 2.0 |
| HY050 family related allowances | 1,632 | 1,527 | 93.3 | 0.1 | 0.2 |
| PY150 other personal benefits | 2,634 | 2,510 | 95.3 | 0.2 | 0.3 |
| tax-exempt social transfers | 5,066 | 5,066 | 100.0 | 0.4 | 0.6 |
| PY140 education related allowances | 12 | 12 | 100.0 | 0.0 | 0.0 |
| HY060 social assistance | 115 | 115 | 100.0 | 0.0 | 0.0 |
| HY070 housing allowances | 348 | 348 | 100.0 | 0.0 | 0.0 |
| HY080 regular inter-household cash transfer receiv | 4,591 | 4,591 | 100.0 | 0.3 | 0.5 |
| total | 1,377,492 | 997,040 | 72.4 | 100.0 | 100.0 |

[1]: Pesetas 1999 per capita

18

Microsimulation and imputation in conjunction

- Microsimulation requires no missing information in income data
- Imputation requires income variables in the same form
- This section describes the two actions in conjunction

19

Step (0). Initial data conversion and modelling

- 0.1 Variables (i) are ordered according increasing proportion of missing values among applicable cases.
- 0.2 Units in Y are divided into two subsets:
 - A: complete information set (units with information available on all variables);
 - B: set with missing values on one or more variables.
- 0.3 The starting point of this process is provided by the 'complete information' set of units (set A), for which there are no missing values on any variable and hence modelling can be carried out without involving imputation, and the data for set A reduced to the following form, giving amount both in gross taxable (H_i) and in net (N_i) forms for each income component (i), and also the unit's 'tax rate' R: $[0, H_i, \text{and}, N_i, \text{and}, R]_{i,u}$
- 0.4 Using the model conversion routine, the available information for units in set B is reduced to the form: $[0, N_i, H_i, X]_{i,u}$

20

Step 1. Conversion to a uniform form

- (1) Determine for each variable the **predominating reporting form** (Y_i) in set B:
- (2) Assign the **average** of R_j values for units in set A to every unit in set B, so as to convert all reported values (N_i or H_i) its "predominating form" Y_i
- **Set A** \rightarrow [0, Y_i] from Step(0)
- **Set B** \rightarrow [0, Y_i , ($Y_i;R_0$), X].
- (3) Impose the same "structure" (variable ordering, Y form, R values) on the lagged variables Y_i for converting each variable to the relevant predominating form.

21

Step 2. Imputation of regressors and lagged variables

- Now we have , covering all current (Y), lagged ($Y^{(t-1)}$) and other auxiliary variables (U), with cases in Set B converted to standard form using R_0 values.
- (1) We apply the normal **imputation** procedure to remove missing values in Set B. Variables ordered according increasing proportion of missing values.
- (2) We apply the **micro-simulation** model to produce improved estimates of R_j for all units in set B. All previously imputed values are rejected: the remaining original values in both Y and $Y^{(t-1)}$ are transformed to the predominating form Y_i using the improved R_j values.
- (3) The **imputation** procedure is repeated on the resulting data set. These are taken to be the final results for the lagged and regressor variable sets.

22

Step 2. Imputation of regressors and lagged variables

- (4) For the main data set Y, the results are used in **micro-simulation** to re-estimate R_j values; imputed values of y_i themselves are rejected. The remaining original values are transformed to the predominating form Y_i using these improved R_j values.
- The data form is now the same as that in Step (1):
 - all available values converted to the predominating form Y_i for variable i
 - conditional on current values of parameter R_j specific to each case j, *the cell values being of the form* \rightarrow .

23

Step 3. Imputation of target variables Y

- **The first ("triangular") imputation**
- (1) The first ("triangular") cycle of **imputation** is performed, without distinguishing between the two forms of known values Y_i , giving cell values in the form
 - - prime (') indicates imputed values.
- (2) The resulting complete set is used to **simulate** improved values of R_j for each unit. These R_j values are used to re-estimate the conditional values ($Y_i;R_j$).
- **Imputation cycle for the full set**
- (1) **Imputation cycle** performed iteratively on variables in order of increasing proportion of missing values - using as auxiliaries *all other variables previously imputed* and the full U and $Y^{(t-1)}$ sets.
- (2) The resulting complete set used to **simulate** improved values of R_j for each unit. These R_j values used to re-estimate the conditional values ($Y_i;R_j$).
- The sequence is repeated **iteratively** a number of times

....

24