Measuring Income, Consumption and Wealth at the micro-level

Workshop on fiscal policy and microsimulation
Outline

1. Background
2. Objectives
3. Strategy
4. Results
5. Conclusions
1. Background

- Communication by the EC "GDP and beyond – Measuring progress in a changing world" (2009);

- Sen-Stiglitz-Fitoussi report on the Measurement of Economic Performance and Social progress (2009);

- HLEG on Measuring Progress, Well-Being and Sustainable Development (2011);

- Eurostat "GDP and beyond" action plan.
2. Objectives

Better joint distributions at micro level: ideally a joint micro-data set, from which meso-indicators could be calculated;

Better consistency between micro-data and aggregates, when concepts and data are compatible;

Parallel, but strong links
Addressing data consistency is essential

Note: The x-axis shows the aggregate saving rates as measured in the surveys. The y-axis shows the aggregate saving rates as obtained in national accounts.
3. Strategy

The current options are the following ones:

- Survey integration
- Multi-source approach
- Statistical matching and modelling
- Modular approach

Every option has to be assessed; the most likely approach consists of a combination of the different possibilities.
An overview of available data

HBS
- Consumption
- Income

SILC
- Income

HFCS
- Wealth
- (gross) Income
- Consumption

Harmonized at EU level
Aggregates, no distributional element
Multi-source

Mind the gap
Statistical matching

• Relying on cumulated experience by Eurostat on the topic

• Test various methods:
  • One purely based on hot-deck: take observations in HBS with the same $X$ and allocate to EU-SILC
    Multi-dimensional, non-parametric, stochastic
  • A mixed approach: estimate the conditional mean of consumption according to $X$, then allocate observations in HBS to EU-SILC using the cumulative distribution function
    Unidimensional, semi-parametric, randomness goes through the estimation of the equation
Choice of covariates $X$

- Following D'Orazio et al. (2006), comparability between EU-SILC and HBS is key:
  - Ex-post harmonization process
  - Selection of candidate variables based on Hellinger distance
- Automatic selection procedure implemented for method 1; comprehensive set of covariates for method 2.
Assessment of the uncertainty

Monte-Carlo approach: replicate hot-deck 1,000 times

- Facing multiply-imputed data → impact on the estimation method
- Computationally demanding; does not relax CIA

"Frechet bounds" approach

- Relaxing CIA
- Relying on "granular" data
Estimation of the VAT

*Use of data from EC's DG for Taxation and Customs Union (freely available)*

**Challenge:** mapping COICOP classification and the ad-hoc classification applied in the data

**Issue:** adding an additional layer of assumptions to an already 'heroic' dataset
Data from DG Taxud

| Category                        | BE | BG | CZ | DK | DE | EE | FI | EL | ES | FR | HR | IT | CY | LV | LT | LU | HU | MT | NL | AT | PL | PT | RO | SI | SK | FI | SE | UK |
|--------------------------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1. Foodstuffs                  | 6  | 20 | 13 | 20 | 10 | 15 | 7  | 20 | 6  | 20 | 15 | 13 | 19 | 20 | 10 | 5  | 4  | 21 | 5  | 5  | 21 | 21 | 13 | 21 | 5  | 5  | 21 | 21 | 5  | 21 |
| 2. Water supplies              | 6  | 20 | 13 | 20 | 10 | 15 | 7  | 20 | 6  | 20 | 15 | 13 | 19 | 20 | 10 | 5  | 4  | 21 | 5  | 5  | 21 | 21 | 13 | 21 | 5  | 5  | 21 | 21 | 5  | 21 |
| 3. Pharmaceutical products    | 6  | 20 | 13 | 20 | 10 | 15 | 7  | 20 | 6  | 20 | 15 | 13 | 19 | 20 | 10 | 5  | 4  | 21 | 5  | 5  | 21 | 21 | 13 | 21 | 5  | 5  | 21 | 21 | 5  | 21 |
| 4. Medical equipment for       | 6  | 20 | 13 | 20 | 10 | 15 | 7  | 20 | 6  | 20 | 15 | 13 | 19 | 20 | 10 | 5  | 4  | 21 | 5  | 5  | 21 | 21 | 13 | 21 | 5  | 5  | 21 | 21 | 5  | 21 |
| disabled persons               | 6  | 20 | 13 | 20 | 10 | 15 | 7  | 20 | 6  | 20 | 15 | 13 | 19 | 20 | 10 | 5  | 4  | 21 | 5  | 5  | 21 | 21 | 13 | 21 | 5  | 5  | 21 | 21 | 5  | 21 |
|DVDs, etc. or similar           | 6  | 20 | 13 | 20 | 10 | 15 | 7  | 20 | 6  | 20 | 15 | 13 | 19 | 20 | 10 | 5  | 4  | 21 | 5  | 5  | 21 | 21 | 13 | 21 | 5  | 5  | 21 | 21 | 5  | 21 |
How we proceed

1. Perform the mapping
2. For each COICOP 5-digit category, apply the corresponding VAT rate as for the reference year
3. Compute the VAT supposedly paid by the households out of consumer prices (at the household level)
4. Aggregate the VAT across all COICOP categories (and deal with data quality issues...)
Aim of the exercise

The fused dataset contains information on VAT and direct taxes at the household level.

The computation of VAT can be performed on HBS data itself; but then we only have the VAT as a share of expenditures.

Having SILC information makes it possible to analyze VAT with respect to gross/disposable income.
4. Results

Data that cannot qualify for 'official statistics' yet, hence in the experimental section

Multi-level communication:

- 3 'simple' articles explaining the data and the method
- 1 more elaborate note + 1 working paper

Data available on Eurobase

The distribution of consumption is well reproduced...
...the distribution of savings may be a bit distorted
What do National Accounts say about VAT?

Source: Eurostat (online data codes: hbs_exp_t111, hbs_str_t211, icw_tax_10, gov_f10a_taxag and nasa_10_nf_br), 2010.
VAT (as a share of expenditure) along income distribution

Source: Eurostat (online data codes: hbs_exp_f133, hbs_str_f223 and icw_tax_10)
VAT (as a share of gross income) along income distribution

Source: Eurostat (online data code: icw_tax_02)
VAT (as a share of expenditure) along the life cycle

Source: Eurostat (online data codes: hbs_exp_f135, hbs_str_f225 and icw_tax_10)
VAT (as a share of gross income) along the life cycle

Source: Eurostat (online data code: icw_tax_01)
Taking direct taxation into account

Can be done thanks to SILC data

Yet, reported by the households: may suffer from underreporting (affecting much more other components of income than taxes)

Aim: making the cross-country comparison more convincing
What do National Accounts say?

Note: The median rate is computed from the ratio of VAT out of gross income at the household level. The aggregate ratio is derived from the sum of disposable income (B60) and taxes (D6) paid by households, as well as social contributions (D613). Taxation is the sum of income taxes received by the government (D51A), social contributions (D613), other taxes (D59A) and VAT (D211).
Source: Eurostat (online data codes: icw_tax_07, gov_10a_taxag and masa_10_ni_hr)
VAT and direct taxation along income distribution

Source: Eurostat (online data code icw_tax_09)
VAT and direct taxation along the life cycle

Source: Eurostat (online data code: icw_tax_07)
5. Conclusions

- Collection of feedback coming from users
- Publication of SILC-HBS 2015 results by mid-2018
- Analysis of the OCW module
- Continuation of the micro-macro linkage exercise
5. Conclusions (continued)

- Data quality issues to be borne in mind (ideally to be dealt with);
- General progressivity of the tax system in the EU seems to hold; nevertheless it also depends on the focus on the distribution;
- The inclusion of excises, as well as improving the VAT estimation, is key;
- Micro-macro linkage is essential; so is the full measurement of the ICW joint distribution,