An integrated database to measure economic wellbeing

Friderike Oehler, Irene Rioboo & Florian Pallaro
European Commission, Eurostat

7th European User Conference for EU-Microdata
March 2021
Outline

1. Background
2. Joint income-consumption-wealth distribution
3. ICW indicators
4. Final remarks
1. Background
Background

Statistics on ICW – Why?
2016 Conference of DGINS
- Reinforced role of social indicators
- Need to develop a harmonised statistical framework for ICW

Eurostat ICW project
- Joint micro data set of household income, consumption and wealth data
- Micro–macro data comparison and reconciliation for households' income and consumption

International cooperation
- Eurostat/OECD EG ICW (2017-2020)
- OECD/Eurostat EG DNA (2017-2020)
- ECB EG LMM (2015-2019)
- ECB EG DFA (since 2019)
2. Joint ICW distribution
Relevance of ICW statistics

- Show characteristics hidden in a unidimensional analysis
- Find out more about households' economic behaviour
- Measure households’ economic wellbeing
- Support policy analysis on poverty and inequality
Methodological alternatives

Different methods:

• Integrated survey
• Record linking
• Statistical matching (SM)
• Modular approach

Integration of specific variables from several independent data sources (referring to the same target population), using information shared between them as a link.
Sources

**Income**
EU-SILC - Statistics on Income and Living Conditions (NSIs, Eurostat)

**Consumption**
HBS – Household Budget Survey (NSIs, Eurostat)

**Wealth**
HFCS - Household Finance and Consumption Survey (HFCN, ECB)
Main steps in statistical matching

- **Target variables**
  - Total disposal income, EU-SILC
  - Total consumption expenditure, HBS
  - Net wealth, HFCS

- **Common variables**
  - Common variables with the highest explanatory power of targets

- **Matching variables**
  - Stratification by matching variables
  - Random hot-deck method

- **Matching I-C**
  - Recalibration of EU-SILC weights to fit a number of consumption margins from HBS

- **Weights recalibration**
  - Stratification by household type, food consumption quintile, tenure status
  - Random hot-deck method

- **Matching I-C-W**

**Conceptual and statistical consistency**

Scrutinised and improved in 2020
Methodological limitations

Conditional Independence Assumption (CIA)

The relationship between Y and Z is completely explained by X:

Income and consumption are independent once conditioning on the matching variables.

A very strong assumption that:
1. does not seem plausible
2. cannot be tested on the fused data set SILC-HBS
Overcoming limitations

1. **Make the CIA a more justifiable** and plausible assumption by including in the set of matching variables a proxy of one of the target variables

<table>
<thead>
<tr>
<th>SILC-HBS</th>
<th>• Income ventiles as proxy of total disposable income</th>
</tr>
</thead>
<tbody>
<tr>
<td>SILC-HBS-HFCS</td>
<td>• Gross income as proxy of total disposable income</td>
</tr>
</tbody>
</table>

2. **Test the CIA** using auxiliary information: Over-indebtedness, Consumption and Wealth (OCW) testing module for EU-SILC 2017

   | Consumption and wealth | BE, CZ, FI (sub-sampling), IS, IT, LV (sub-sampling), LT, NL, AT (partially and sub-sampling), PT (sub-sampling), SE, UK |
Can we assume independence between income and consumption while controlled by the ordered income class?

... Yes, we can

And between income and wealth while controlled by gross income?

... Yes, we can

<table>
<thead>
<tr>
<th>country</th>
<th>correlation</th>
<th>partial correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td>0.34</td>
<td>0.04</td>
</tr>
<tr>
<td>BE</td>
<td>0.21</td>
<td>0.05</td>
</tr>
<tr>
<td>FI</td>
<td>0.49</td>
<td>0.12</td>
</tr>
<tr>
<td>LT</td>
<td>0.45</td>
<td>0.14</td>
</tr>
<tr>
<td>LV</td>
<td>0.64</td>
<td>0.15</td>
</tr>
<tr>
<td>NL</td>
<td>0.33</td>
<td>0.06</td>
</tr>
<tr>
<td>PT</td>
<td>0.32</td>
<td>-0.02</td>
</tr>
<tr>
<td>SE</td>
<td>0.21</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Therefore, we can assume that CIA holds
An example of I-C matching for Latvia

Matching variables
Household type, age, income ventiles

Explanatory power

<table>
<thead>
<tr>
<th></th>
<th>R2 income</th>
<th>R2 consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.884</td>
<td>0.533</td>
</tr>
</tbody>
</table>

Fréchet bounds

<table>
<thead>
<tr>
<th>Deciles income-consumption</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>01-01</td>
<td>0.03667</td>
<td>0.03719</td>
</tr>
<tr>
<td>10-10</td>
<td>0.05416</td>
<td>0.05417</td>
</tr>
</tbody>
</table>

Original vs matched distributions

Hellinger distance = 0.0007
3. ICW indicators
Topics

Economic resources
- Share of households and economic resources by icw quantiles
- Mean and median economic resources by icw quantiles

Poverty
- Persons with low level of expenditure by risk of income poverty, material deprivation and work intensity
- Persons at two-fold risk of poverty
- Households at risk of asset-based vulnerability
- Proportion of dissaving households

Saving rates
- Median saving rates
- Gini coefficient
- Median consumption by income decile
- Proportion of consumption decile by income decile
- Aggregate propensity to consume

Taxation
- Distribution of direct and indirect taxes paid by households as a percentage of their gross income
- VAT rate by COICOP consumption purpose

Household characteristics
- Structure of household population by: household type, age, educational attainment level, activity status
## Economic resources

**Eurobase, experimental statistics:**
- **icw_res_01:** Share of households and economic resources by income, consumption and wealth quantiles (%)
- **icw_res_02:** Mean and median economic resources of households by income, consumption and wealth quantiles (Euro)

<table>
<thead>
<tr>
<th>Income quantile</th>
<th>Consumption quantile</th>
<th>Wealth quantile</th>
<th>Measure</th>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total population</td>
<td>Total population</td>
<td>Total population</td>
<td>mean</td>
<td>disposable income</td>
</tr>
<tr>
<td>Income Q1</td>
<td>Consumption Q1</td>
<td>Wealth Q1</td>
<td>median</td>
<td>consumption expenditure</td>
</tr>
<tr>
<td>Income Q2</td>
<td>Consumption Q2</td>
<td>Wealth Q2</td>
<td>share</td>
<td>net wealth</td>
</tr>
<tr>
<td>Income Q3</td>
<td>Consumption Q3</td>
<td>Wealth Q3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income Q4</td>
<td>Consumption Q4</td>
<td>Wealth Q4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income Q5</td>
<td>Consumption Q5</td>
<td>Wealth Q5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income Top 10%</td>
<td>Consumption Top 10%</td>
<td>Wealth Top 10%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income Bottom 10%</td>
<td>Consumption Bottom 10%</td>
<td>Wealth Bottom 10%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Full cross-over of all income, consumption and wealth quantiles, plus the top and bottom deciles**
Sticky floors and sticky ceilings

Share of households in the top 20% and bottom 20% of the income, consumption and wealth distribution
Share of resources at IW bottom quintile

Share of households belonging to the 20% most vulnerable in terms of income, expenditure and net wealth, and the share of resources held by these households (%)

4% of French households fall into the bottom 20% of all three of the expenditure, income and wealth distributions. On average, their income is €19 000 per year, with average net wealth of €3 000. They spend €11 000.
Share of households belonging to the 20% most vulnerable in terms of both income and expenditure and the share of disposable income, consumption expenditure and net wealth held by these households, "around 2015"
8 – 25% of women over 75 years are at risk of both income and consumption poverty (4 – 12% of men).

Eurobase, experimental statistics:
icw_pov_10: Persons at two-fold risk of poverty by age and sex
icw_pov_11: Persons at two-fold risk of poverty by household type
icw_pov_12: Persons at two-fold risk of poverty by activity status

- income & consumption poverty
- income & liquid financial wealth poverty
- consumption & liquid financial wealth poverty
Persons at risk of income & liquid financial wealth poverty

20 – 45 % of women over 75 years are at risk of both income and liquid financial wealth poverty.
Persons at risk of income & consumption poverty

5 – 14 % of persons below 35 years are at risk of both income and consumption poverty.
Aggregate propensity to consume by income quintile

Average consumption expenditure exceeds income
The saving rates puzzle
The saving rates puzzle

Median saving rate by age class

- Romania
- Croatia
- Greece
- Bulgaria
- Estonia
- Ireland
- France

< 35 years  35 - 44  45 - 54  55 - 64  65 - 74  75 +
Impact of taxes

Share of disposable income paid in direct and indirect taxes by the median household in the first income quintile

Share of disposable income paid in direct and indirect taxes by the median household in the fifth income quintile
Experimental statistics use **new data sources and methods** in an effort to better respond to our users’ needs.

For example, for the first time Eurostat is estimating **price changes in the food supply chain**, from farm to consumer. Another example is the use of Wikipedia as a new source to produce statistics on the **visits to UNESCO World Heritage Sites**. This is to measure not only the popularity of the sites but also the public’s ‘cultural consumption’.

As these statistics have not reached full maturity in terms of harmonisation, coverage or methodology, they are always marked with a clearly visible **logo** and accompanied by detailed methodological notes.

We are interested in receiving your **feedback** on our experimental statistics in order to further improve their robustness. Join the discussions launched on the **European Statistics User Forum** to share your views and ideas!

All topics covered in this page are summarised in a table under ‘**Published statistics**’. They are also available via the ‘topics’ section below.
4. Final remarks
Recent steps

- The statistical matching was improved
- The quality of the matched dataset was assessed
- Revised statistics were published
- New 2D & 3D indicators were released
- Metadata were published
- All Statistics Explained articles were updated
Thank you