Estimation process of Finnish EU-SILC

This presentation shows some special characteristics of the sampling design, rotation and calibration in Finnish EU-SILC.

1) Cross-sectional SILC data based on two waves
2) Selection of a master sample
3) Allocation for strata
4) Calibration process

Pauli Ollila  Income and Living Conditions / SILC Team, Statistics Finland
Cross-sectional SILC data based on two waves

- Eurostat proposes an integrated design, where the cross-sectional and longitudinal statistics are produced from essentially the same set of observations.

Two waves of the IDS are utilised in the cross-sectional part of EU-SILC. In the future third and fourth waves are surveyed separately from the IDS, not included in the cross-sectional part of EU-SILC.

- parallel results from the IDS & SILC
- ”wasting” two rotations
Selection of a master sample / Allocation for strata

+ The dwelling unit preparations for several surveys at the same time
+ Master sample calibration used for creating household distributions
- More complexity to the sampling design
+ Allocation emphasising well-earning, farmers and entrepreneurs ensures the sample size of these vital groups
+ Improves the quality of “income-based” indicators (gini, share ratio)
- Allocation not good for poverty indicators
- Proportional allocation outperforms the others

Table by Guillaume Osier, Eurostat

The Finnish Income Distribution Survey describes the level, composition, and distribution of income among households and the economically active population.

Example: collection year 2004

<table>
<thead>
<tr>
<th></th>
<th>Variance phase 2 (current allocation)</th>
<th>Variance phase 2 (simple random sampling)</th>
<th>Variance phase 2 (proportional allocation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poverty threshold</td>
<td>1217</td>
<td>1088</td>
<td>907</td>
</tr>
<tr>
<td>Poverty rate</td>
<td>0.095</td>
<td>0.069</td>
<td>0.067</td>
</tr>
<tr>
<td>Income quantile share ratio</td>
<td>7.2×10^-4</td>
<td>11.4×10^-4</td>
<td>6.3×10^-4</td>
</tr>
<tr>
<td>Gini coefficient</td>
<td>0.023</td>
<td>0.034</td>
<td>0.024</td>
</tr>
<tr>
<td>Relative median gap</td>
<td>0.31</td>
<td>0.23</td>
<td>0.22</td>
</tr>
<tr>
<td>Mean equivalized income</td>
<td>1312</td>
<td>2467</td>
<td>1288</td>
</tr>
</tbody>
</table>
Calibration process

- Income variables from registers, e.g. Cash or near cash employee income, Pensions, Unemployment benefits, Income from self-employment, Capital income, Income from agriculture, Income from property and forestry, Capital gains, Mortgage interests

### Weighting Procedure of the IDS

- Population register
- DU marginal distributions for DU size, province and type of municipality
- Income information from registers
- Master sample with DU inclusion probabilities
- Calibration weights for master sample
- Calibration for wave 1
- Calibration for wave 2
- Accepted households of the IDS with two phase inclusion probabilities
- Final cross-sectional weights

### Standard Errors

<table>
<thead>
<tr>
<th></th>
<th>Poulpe</th>
<th>Bootstrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>At-risk-... after social tr.</td>
<td>0.38</td>
<td>0.35</td>
</tr>
<tr>
<td>At-risk-... before s.tr. exc. old</td>
<td>0.40</td>
<td>0.42</td>
</tr>
<tr>
<td>At-risk-... before social tr.</td>
<td>0.36</td>
<td>0.40</td>
</tr>
<tr>
<td>Inequality of inc.dist. S80/S20</td>
<td>0.034</td>
<td>0.088</td>
</tr>
<tr>
<td>Relative median at-risk-of-p.</td>
<td>0.68</td>
<td>0.70</td>
</tr>
<tr>
<td>Dispersion around the risk...</td>
<td>0.13</td>
<td>0.14</td>
</tr>
<tr>
<td>Gini coefficient</td>
<td>0.20</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Poulpe calculations (Eurostat), calibration included

Bootstrap (StatFi), calibration process not in replications

- Integrative calibration improving significantly the estimation of e.g. S80/S20 and gini coefficient

- The results with very narrow confidence intervals may cause disbelief at the estimation methods