“Flash estimates” of income distribution indicators for the European Union: methods, assessment and future prospects

Aura Leulescu, Chrysa Leventi, Cristina Lopez-Vilaplana, Olga Rastrigina and Holly Sutherland

Presented by Sofiya Stoyanova, Office for National Statistics UK

34th IARIW General Conference
Dresden, Germany, August, 2016
Outline

• Introduction
• Methodological framework
• Quality framework
• Conclusions
• Comments & questions
Introduction

- Importance of the European Union Statistics on Income and Living Conditions (EU-SILC)
- Income data available with significant lag due to their complexity
- Timeliness of indicators is crucial:
  - keeping track of the effectiveness of policies
  - Evaluating the impact of macroeconomic conditions on poverty and the income distribution
- Current strategy for providing more timely income estimates is based on two pillars:
  - Flash estimates on income distribution and poverty
  - Final EU-SILC microdata
Methodological framework

1. Adjustment for changes in population characteristics

2. Reproducing the evolution of market income components

3. Accounting for changes in taxes and benefits using a microsimulation model
Methodological framework (1): Changes in population characteristic

Data source:
- Labour Force Survey (LFS)

(1) Labour market transitions

Two types of transitions modelled:
- from non-employment into employment
- from employment into short/long-term unemployment

Logit models to estimate probability of being employed:
- explanatory variables include: age, marital status, education level, etc.
- model estimated separately for men and women

Unemployment benefits are simulated according to country rules
(2) Reweighting

Allows controlling for a wider set of population characteristics

Three alternative methods tested:

- Calibration at household level based on marginal distributions of a set of variables from LFS
- Calibration at household level based on changes in shares for same set of variables
- Calibration at individual level based on specific socio-demographic groups

Reweighting not suitable in times of rapid economic changes
Methodological framework (2): Updating income sources

Analysis makes use of EUROMOD (microsimulation model based on EU-SILC data)

**EUROMOD uprating factors**

- Based on admin or survey data
- Country-specific uprating factors derived for each income source

**Model-based factors for socio-demographic groups**

- Introduces differential growth rates of income via a model-based approach
- EU-SILC time series for 2009-2012 used to compute average growth rates for pre-defined socio-demographic categories
- Use decision trees and logistic regression models to choose the categories
- Estimate current growth rates of the categories using dynamic factor modelling approach
Methodological framework(3): Tax-benefit simulation

EUROMOD used for simulation
   Income elements simulated: universal and targeted cash benefits, social insurance contributions, direct taxes.

Incorporate tax evasion and benefit non take-up wherever possible

Adjustment made to account for differences in EUROMOD and EU-SILC household income estimates
   - assume discrepancy between the two is stable over time
Quality framework

- Consistency of trends in auxiliary data sources

- Retrospective assessment based on:
  - Intermediate checks for all production stages
  - Ability of the model to reproduce past estimates for main economic indicators

- Quality measures for flash estimates:
  - Incorporate the role of uncertainty
  - Integrate information from different methods and their historical performance to produce a measure of quality of flash estimates
## Assessment of results for flash estimates

### Average consistency by indicator and methods (2012/11 and 2013/12)

<table>
<thead>
<tr>
<th>Indicator /Method</th>
<th>AT</th>
<th>CZ</th>
<th>FI</th>
<th>FR</th>
<th>IT</th>
<th>LU</th>
<th>LV</th>
<th>PL</th>
<th>PT</th>
</tr>
</thead>
<tbody>
<tr>
<td>AROP Labour transitions</td>
<td>95%</td>
<td>83%</td>
<td>89%</td>
<td>92%</td>
<td>98%</td>
<td>71%</td>
<td>94%</td>
<td>99%</td>
<td>99%</td>
</tr>
<tr>
<td>AROP CAL_H_A</td>
<td>97%</td>
<td>86%</td>
<td>90%</td>
<td>93%</td>
<td>97%</td>
<td>79%</td>
<td>96%</td>
<td>97%</td>
<td>97%</td>
</tr>
<tr>
<td>AROP CAL_H</td>
<td>98%</td>
<td>91%</td>
<td>90%</td>
<td>91%</td>
<td>96%</td>
<td>81%</td>
<td>97%</td>
<td>97%</td>
<td>96%</td>
</tr>
<tr>
<td>AROP CAL_I</td>
<td>97%</td>
<td>88%</td>
<td>90%</td>
<td>93%</td>
<td>97%</td>
<td>79%</td>
<td>95%</td>
<td>96%</td>
<td>96%</td>
</tr>
<tr>
<td>D10 Labour transitions</td>
<td>96%</td>
<td>97%</td>
<td>99%</td>
<td>99%</td>
<td>92%</td>
<td>96%</td>
<td>96%</td>
<td>99%</td>
<td>97%</td>
</tr>
<tr>
<td>D10 CAL_H_A</td>
<td>96%</td>
<td>97%</td>
<td>98%</td>
<td>98%</td>
<td>96%</td>
<td>95%</td>
<td>94%</td>
<td>97%</td>
<td>97%</td>
</tr>
<tr>
<td>D10 CAL_H</td>
<td>96%</td>
<td>98%</td>
<td>98%</td>
<td>97%</td>
<td>94%</td>
<td>95%</td>
<td>94%</td>
<td>97%</td>
<td>97%</td>
</tr>
<tr>
<td>D10 CAL_I</td>
<td>96%</td>
<td>98%</td>
<td>98%</td>
<td>98%</td>
<td>96%</td>
<td>91%</td>
<td>96%</td>
<td>96%</td>
<td>91%</td>
</tr>
<tr>
<td>MEAN Labour transitions</td>
<td>94%</td>
<td>98%</td>
<td>98%</td>
<td>97%</td>
<td>99%</td>
<td>99%</td>
<td>96%</td>
<td>99%</td>
<td>98%</td>
</tr>
<tr>
<td>MEAN CAL_H_A</td>
<td>95%</td>
<td>99%</td>
<td>99%</td>
<td>90%</td>
<td>98%</td>
<td>99%</td>
<td>96%</td>
<td>98%</td>
<td>99%</td>
</tr>
<tr>
<td>MEAN CAL_H</td>
<td>95%</td>
<td>99%</td>
<td>99%</td>
<td>89%</td>
<td>97%</td>
<td>98%</td>
<td>94%</td>
<td>98%</td>
<td>99%</td>
</tr>
<tr>
<td>MEAN CAL_I</td>
<td>95%</td>
<td>99%</td>
<td>99%</td>
<td>89%</td>
<td>98%</td>
<td>99%</td>
<td>94%</td>
<td>99%</td>
<td>99%</td>
</tr>
<tr>
<td>MEDIAN Labour transitions</td>
<td>96%</td>
<td>98%</td>
<td>98%</td>
<td>96%</td>
<td>99%</td>
<td>99%</td>
<td>95%</td>
<td>99%</td>
<td>97%</td>
</tr>
<tr>
<td>MEDIAN CAL_H_A</td>
<td>97%</td>
<td>99%</td>
<td>99%</td>
<td>97%</td>
<td>98%</td>
<td>98%</td>
<td>96%</td>
<td>98%</td>
<td>98%</td>
</tr>
<tr>
<td>MEDIAN CAL_H</td>
<td>96%</td>
<td>99%</td>
<td>98%</td>
<td>97%</td>
<td>97%</td>
<td>98%</td>
<td>96%</td>
<td>98%</td>
<td>98%</td>
</tr>
<tr>
<td>MEDIAN CAL_I</td>
<td>97%</td>
<td>99%</td>
<td>98%</td>
<td>98%</td>
<td>98%</td>
<td>99%</td>
<td>94%</td>
<td>99%</td>
<td>99%</td>
</tr>
<tr>
<td>QSR Labour transitions</td>
<td>91%</td>
<td>93%</td>
<td>96%</td>
<td>97%</td>
<td>87%</td>
<td>93%</td>
<td>93%</td>
<td>97%</td>
<td>96%</td>
</tr>
<tr>
<td>QSR CAL_H_A</td>
<td>95%</td>
<td>95%</td>
<td>97%</td>
<td>72%</td>
<td>99%</td>
<td>92%</td>
<td>91%</td>
<td>97%</td>
<td>96%</td>
</tr>
<tr>
<td>QSR CAL_H</td>
<td>95%</td>
<td>95%</td>
<td>97%</td>
<td>97%</td>
<td>95%</td>
<td>93%</td>
<td>93%</td>
<td>97%</td>
<td>96%</td>
</tr>
<tr>
<td>QSR CAL_I</td>
<td>95%</td>
<td>96%</td>
<td>97%</td>
<td>72%</td>
<td>98%</td>
<td>93%</td>
<td>93%</td>
<td>93%</td>
<td>96%</td>
</tr>
</tbody>
</table>

Note: The methods considered are: (1) Labour transitions; (2) Calibration at household level: Cal_H; (3) Calibration at household level adjusted: Cal_H_A; (4) Calibration at individual level: Cal_I. Calibration at individual level is not available for Poland. The indicators are: AROP (at-risk-of-poverty rate); D10 (decile 10), mean and median equivalentised household disposable income; QSR (income quintile share ratio).
Conclusions

- No single method shows better performance for all indicators and all years
- Further work will focus on the development of uncertainty measures
- Alternative estimates from EU Member States will help inform the decision for a set of flash estimates at a EU level
• An alternative strategy may be to estimate year-on-year change and apply to the last observed value
• Different methods may be appropriate for different countries
• Issues with data availability in different countries
• How realistic is the assumption that the discrepancy in income estimates between EU-SILC and EUROMOD is constant over time?