“Flash Estimates” of Income Distribution Indicators for the European Union: Methods, Assessment and Future Prospects

Aura Leulescu (Eurostat), Chrysa Leventi (ISER, University of Essex, United Kingdom), Cristina Lopez-Vilaplana (Eurostat), Olga Rastrigina (ISER, University of Essex, United Kingdom), and Holly Sutherland (ISER, University of Essex, United Kingdom)

Paper prepared for the 34th IARIW General Conference

Dresden, Germany, August 21-27, 2016

Session 4A: Household Accounts

Time: Tuesday, August 23, 2016 [Afternoon]

Paper prepared for the IARIW 34th General Conference

Dresden, Germany, August 21-27, 2016

"Flash estimates" of income distribution indicators for the European Union: methods, assessment and future prospects[1]

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

** Eurostat*

*** ISER, University of Essex*

August 2016

Abstract

The indicators on poverty and income inequality based on the European Union Statistics on Income and Living Conditions (EU-SILC) are an important part of the toolkit for the European Semester. However, the indicators on income of year N are only available in the autumn of year N+2, which comes too late for the EU's policy agenda. Timeliness of income distribution statistics is critical for monitoring the effectiveness of policies. Therefore, an improvement in this area represents an important priority at the EU level. This paper presents the framework for the development of flash estimates on income distribution in EU countries based on microsimulation techniques. The EU tax-benefit microsimulation model EUROMOD is used for this purpose in combination with timelier macro-level statistics on changes in demographics, employment characteristics and income. A central element of the exercise is the quality framework which would allow building a common platform together with the EU Member States for comparing and assessing results. This paper aims to make an outline of this framework and to illustrate the performance of several alternative methods of flash estimates production based on their historical performance. We also outline our plans

[1] The ISER work on nowcasting was financed by the European Commission, Directorate General for Employment, Social Affairs and Inclusion through the Social Situation Monitor. An earlier version of this work is published as part of Social Situation Monitor Research Note 1/2015. The version of EUROMOD used in this paper is G3.0+. We acknowledge the contribution of all past and current members of the EUROMOD consortium. The process of extending and updating the version of EUROMOD used in this paper was financially supported by the Directorate General for Employment, Social Affairs and Inclusion of the European Commission [Progress grant no. VS/2011/0445]. For Czech Republic, Latvia, Portugal, and Finland we make use of microdata from the EU Statistics on Incomes and Living Conditions (EU-SILC) made available by Eurostat (59/2013-EU-SILC-LFS); for Luxembourg and Poland the EU-SILC together with national variables provided by respective national statistical offices; for France, Italy, and Austria the national EU-SILC PDB data made available by respective national statistical offices. The usual disclaimers apply. Email address for correspondence: orastr@essex.ac.uk

for further methodological developments and discuss issues of quality assessment and communication that emerge in the context of this work.

Keywords: Flash Estimates, Nowcasting, European Union, Income distribution, Microsimulation.

1. Introduction

Policy makers in the European Union (EU) are facing an increasing demand for monitoring changes in social conditions at national and EU level, especially during periods of rapid economic changes. The European Union Statistics on Income and Living Conditions (EU-SILC) is an instrument aimed at collecting comparable multidimensional microdata on income, poverty, social exclusion and living conditions for EU countries. EU-SILC indicators on poverty and income inequality are a key part of the toolkit for the European Semester, the yearly cycle of economic policy coordination among EU member states. The timeliness of these indicators is crucial for keeping track of the effectiveness of policies and the impact of macroeconomic conditions on poverty and income distribution. However, partly due to the complexity of the data collection process, income data for year N is only available in the autumn of year N+2, which comes too late for the policy agenda.

The current strategy at the level of the European Statistical System for providing more timely data on income is based on two pillars. The first pillar refers to the production of flash estimates on income distribution and poverty. It states that flash estimates on income of year N should be available in time to prepare for the European Semester in the autumn of year N+1. The second pillar concerns the final EU-SILC microdata, and it states that the microdata on income of year N should be available during the European Semester (i.e. in the end of year N+1 or early N+2).

This paper aims to provide an overview of the process put in place for the development of flash estimates for income distribution at EU level. The approach presented here relies on microsimulation modelling and is being developed by Eurostat in collaboration with the Institute for Social and Economic Research (ISER) at the University of Essex. Nowcasting refers to the estimation of current income distribution using microsimulation modelling, statistical techniques or other appropriate methods based on household income microdata from a previous period in combination with the most up-to-date macro-level statistics. In this methodological frame several nowcasting methods are presented and assessed.

Our analysis makes use of EUROMOD, the microsimulation model based on EU-SILC data which estimates in a comparable way the effects of taxes and benefits on the income distribution in each of the EU Member States. For the purposes of the nowcasting exercise standard EUROMOD policy simulation routines are enhanced with additional adjustments to the input data to take into account changes in the population structure. Therefore, the methodological framework is composed of three main stages: 1) adjustments for changes in population characteristics (e.g. demographic structure of the population, labour market characteristics, etc.); 2) reproducing the evolution of market income components (such as income from employment, self-employment, property, etc.); and 3) accounting for changes in taxes and benefits using EUROMOD.

In the context of development of flash estimates the quality framework becomes a central part of the production process in order to select the methods and auxiliary sources that perform best and are able to reproduce EU-SILC indicators for previous years. The quality framework is composed of two parts: (1) the quality assurance, which focuses on analysing inconsistencies in the input data and includes several intermediate quality checks along the process; (2) the quality assessment, which focuses on the historical performance of different methods. The quality framework is an essential element for both the scientific and political acceptance of the flash estimates. For this it is important that the flash estimates are disseminated together with methodological guidelines concerning their development, interpretation and appropriate use.

The quality framework is illustrated based on the flash estimates for 2012 and 2013 for nine EU countries (Austria, Czech Republic, Finland, France, Italy, Latvia, Luxembourg, Portugal, and Poland) produced using microdata on income distributions as of 2011 (based on SILC 2012) as a starting point. Different target indicators, such as income quintiles, at-risk-of-poverty rate, deciles and quantile share ratio, are examined.

The structure of the paper is the following: in Section 2 the nowcasting methodology is explained. Section 3 presents the quality framework. Section 4 concludes by summarising the most important findings and the next steps for the development of the flash estimates on income distribution.

2. Methodology

The nowcasting methodology presented in this paper is based on microsimulation techniques used in combination with the latest macro-level statistics. It aims at developing a generic approach that can be applied to all EU countries in a straightforward, flexible and transparent way. By doing so, it ensures the comparability and consistency of results both across countries and through time.

Microsimulation models have been widely used for assessing the distributional impact of current and future tax-benefit policy reforms, as well as the impact of the evolution of market incomes, changes in the labour market and in the demographic structure of the population.[2] Using microsimulation techniques based on representative household data enables changes in the distribution of market income to be distinguished and the effects of the tax-benefit system to be identified taking into account the complex ways in which these factors interact with each other (Peichl, 2008; Immervoll et al., 2006). Combined macro-micro modelling has also been used for analysing the impact of macroeconomic policies and shocks on poverty and income distribution.[3]

In this paper in order to produce flash estimates for income indicators, the microsimulation approach is used to update the structure of a micro dataset to account for changes to the main components of income variables over time. This is based on the following stages: 1) adjustment for changes to the demographic structure of the population and for changes to the presence of income sources determined by labour market characteristics; 2) uprating the level of market income components; and 3) changes in taxes and benefits due to policy reforms

[2] Some examples include Brewer et al. (2013) for the UK, Keane et al. (2013) for Ireland, Brandolini et al. (2013) for Italy, Matsaganis & Leventi (2014) for Greece and Narayan & Sánchez-Páramo (2012) for Bangladesh, Mexico, Philippines and Poland.

[3] A detailed review is provided in Bourguignon et al. (2008) and Essama-Nssah (2005). See also Figari et al. (2015) for a discussion.

(O'Donoghue and Loughrey, 2014). The remaining of this Section explains each of these stages in detail.

2.1 Changes in population characteristics

There are two main approaches to take into account changes in population characteristics: static and dynamic. The static approach is based on reweighting (or calibration). It consists of the derivation of a new vector of sample weights that brings the marginal distributions from the base year for a set of main socio-demographic variables (e.g. age, labour, gender) to the level of the target year. In the dynamic process individual trajectories are modelled and individuals in the sample undergo transitions. The paper tests both approaches. The main auxiliary source of information used to obtain the population characteristics in the target year is the Labour Force Survey (LFS) statistics. LFS macro-level statistics for year N are usually available in April N+1. This allows the production of flash estimates for year N based on the updated structure for labour and demographics in time for the European Semester.

2.1.1 Modelling labour market transitions

The dynamic approach to take into account changes in population characteristics is based on modelling net employment transitions. It accounts for changes in labour market characteristics, while other population characteristics (such as demographics) are kept constant. A detailed discussion of this approach can be found in Navicke et al. (2013) and Rastrigina et al. (2016). Changes in employment are modelled by explicitly simulating transitions between labour market states (Figari et al., 2011; Fernandez Salgado et al., 2013; Avram et al., 2011). Two types of transitions are modelled: (i) from non-employment into employment and (ii) from employment into short-term/long-term unemployment (or inactivity). Observations are selected for transitions based on their conditional probabilities of being employed rather than being unemployed or inactive. A logit model is used for estimating these probabilities for working age (16-64) individuals in the EUROMOD input data. In order to account for gender differences in the labour market situation, the model is estimated separately for men and women. Students, working-age individuals with permanent disability or in retirement and mothers with children aged below 2 are excluded from the estimation, unless they report employment income in the underlying data. Explanatory variables include age, marital status, education level, country of birth, employment status of partner, unemployment spells of other household members, household size, number of children and their age, home ownership, region of residence and urban (or rural) location. The specification of the logit model used and the estimated coefficients are reported in Rastrigina et al. (2016).

The weighted total number of observations that are selected to go through transitions corresponds to the relative net yearly change in employment rates by age group and gender (a total of 6 strata) as shown in the LFS statistics. Changes from short-term to long-term unemployment are modelled based on a similar selection procedure but using LFS figures on long-term unemployment (with unemployment duration more than one year) as an external source of information. This transition is critical due to its implications for eligibility and receipt of unemployment benefits. Transitions to and from inactivity are modelled implicitly through restricting eligibility for unemployment benefits, according to the country-specific rules.

Labour market characteristics and sources of income are adjusted for those observations that are subject to transitions. In particular, employment and self-employment income is set to zero for individuals moving out of employment. For individuals moving into employment, earnings are set equal to the mean among those already employed within the same stratum. Unemployment benefits are simulated for those moving out of employment in case they are eligible for such benefits according to the country rules. If the rules require assessment of earnings and number of months in work for several years preceding unemployment, we assume that these remain unchanged throughout the assessment period and are equal to the values observed in the income reference period. For those moving into long-term unemployment the eligibility is adjusted assuming that the duration of unemployment spell is more than one year. In some countries the long-term unemployed are not eligible for any unemployment benefits (e.g. Latvia); in other countries they are not eligible for unemployment insurance but still qualify for unemployment assistance (e.g. Portugal); in countries with fairly long duration of unemployment insurance (e.g. Finland) we assume that the long-term unemployed continue to receive unemployment insurance.

2.1.2 Reweighting

The static approach to account for changes in population characteristics is based on reweighting and consists of the derivation of a new vector of sample weights that brings the marginal distributions from the base year for a set of main socio-demographic variables to the level of the target year. This approach allows controlling for a wider range of population characteristics including retirement and demographic changes. In theory, this can be done also through modelling individual transitions but the process is much more complex. However, there are limitations related to the use of reweighting in the context of nowcasting: given that it only adjusts the structure of the population to some marginal distributions it may perform worse in times of rapid economic changes. For example, reweighting cannot capture if individuals entering a particular state have characteristics completely different from the characteristics of the people observed in that state in the base year. Such changes were often observed during recent unemployment shocks.

The variables that are more likely to impact the income distribution over time should be both highly related to income and volatile: thus the main relevant variables are related to labour market information and changes in household composition. Other relevant controlled characteristics include regional information. The reweighting can be done at household or individual level. Both options are explored.

The variables at household level are based on demographic and employment characteristics of individuals within households: number of members in the household by age group, gender and activity status; household size and number of dependent children; region and degree of urbanisation. The target distributions of the relevant variables are obtained from the LFS. However, the initial distributions observed in EU-SILC and LFS are not always consistent.

This implies that reweighting based on LFS margins can introduce a bias. For example, in Portugal there is a 10% difference in the number of self-employed in 2011. Nevertheless in both sources there was a decrease of 7% between 2011 and 2012. In such cases we adjusted the margins in LFS by adding up the percentage change from LFS to the SILC base year. Hence, the adjustment reflects the change to a more recent structure while systematic source inconsistencies are ruled out.

At individual level, the reweighting is done based on joint distributions of individual level variables. We start off with the premise that the growth in income of a given individual over a given period of time is the combined outcome of a set of characteristics which the individual shares with other individuals (such as activity status, level of education, gender, age, etc.) and elements which are purely idiosyncratic to that individual. According to this assumption, the income growth rate of any individual can be decomposed into two distinct components. The first component depends on the socio-demographic category an individual belongs to and is identical for each individual within that category. At this stage only this is taken into account in the model framework. The second component is specific to each individual and can be included as a random error term. Based on this assumption individuals are classified into groups in order to take into account the within and between variance of income changes. The specific socio-demographics groups are built based on gender, geographical regions, degree of urbanisation, activity status and age group. The margins for the socio-demographics groups are updated according to the LFS data.

Overall, three alternative reweighting methods were tested in this paper:

a. CAL_H: calibration at household level based on the marginal distributions (*levels*) from LFS for the following variables: household size, number of people by age group and sex, number of people part/full-time employed/self-employed/retired, region and degree of urbanisation.
b. CAL_H_ADJ: calibration at household level based on the *changes* in the shares for the same variables.
c. CAL_I: calibration at individual level based on specific socio-demographic groups formed on the basis of: age, gender, labour status, region and degree of urbanisation. Thus the calibration is based on joint distributions.

The principle of calibration is that it minimizes the distance between the vectors of weights so that the other distributions are not distorted. Further work should investigate not only the effect on the income distribution but also possible side effects of calibration:

- distortions of characteristics not controlled for;
- implications of individual-level calibration (i.e. the new weight is not the same for all individuals in the household) for household-level analysis.

2.2 Updating non-simulated income sources

After adjusting the input data for changes in the population characteristics, the next step is to update non-simulated income beyond the income data reference period. This approach applies uprating coefficients to market incomes[4] and non-simulated social benefits (or taxes). The coefficients are based on more timely data sources from the target year which reflect indexation rules or the change in the average income per recipient. Two approaches are tested in the paper.

2.2.1 EUROMOD uprating factors

EUROMOD contains uprating factors based on available administrative or survey statistics. Country-specific updating factors are derived for each income source, reflecting statutory rules (such as indexation rules) or the change in the average amount per recipient between the income data reference period and the target year. The latter is preferred for the nowcasting exercise, especially for pensions. The evolution of average pensions can capture important changes in the population of pensioners (e.g. inflow of newly retired pensioners with higher average pensions). In order to capture differential growth rates in employment income, updating factors are disaggregated by economic activity and/or by economic sector if such information is available.

2.2.2 Model-based factors for socio-demographic groups

An alternative way to update income also uses EUROMOD uprating factors as a base but introduces differential growth rates in the income distribution for some important income components (such as employment and self-employment income) via a model-based approach. The modelling approach starts from the premise that the growth rates of income of individuals belonging to a given socio-demographic category follow the same probability distribution. A socio-demographic category is defined as one specific outcome of all possible combinations of a collection of categorical variables. The estimation algorithm is divided into two steps: (i) the classification step and (ii) the modelling step.

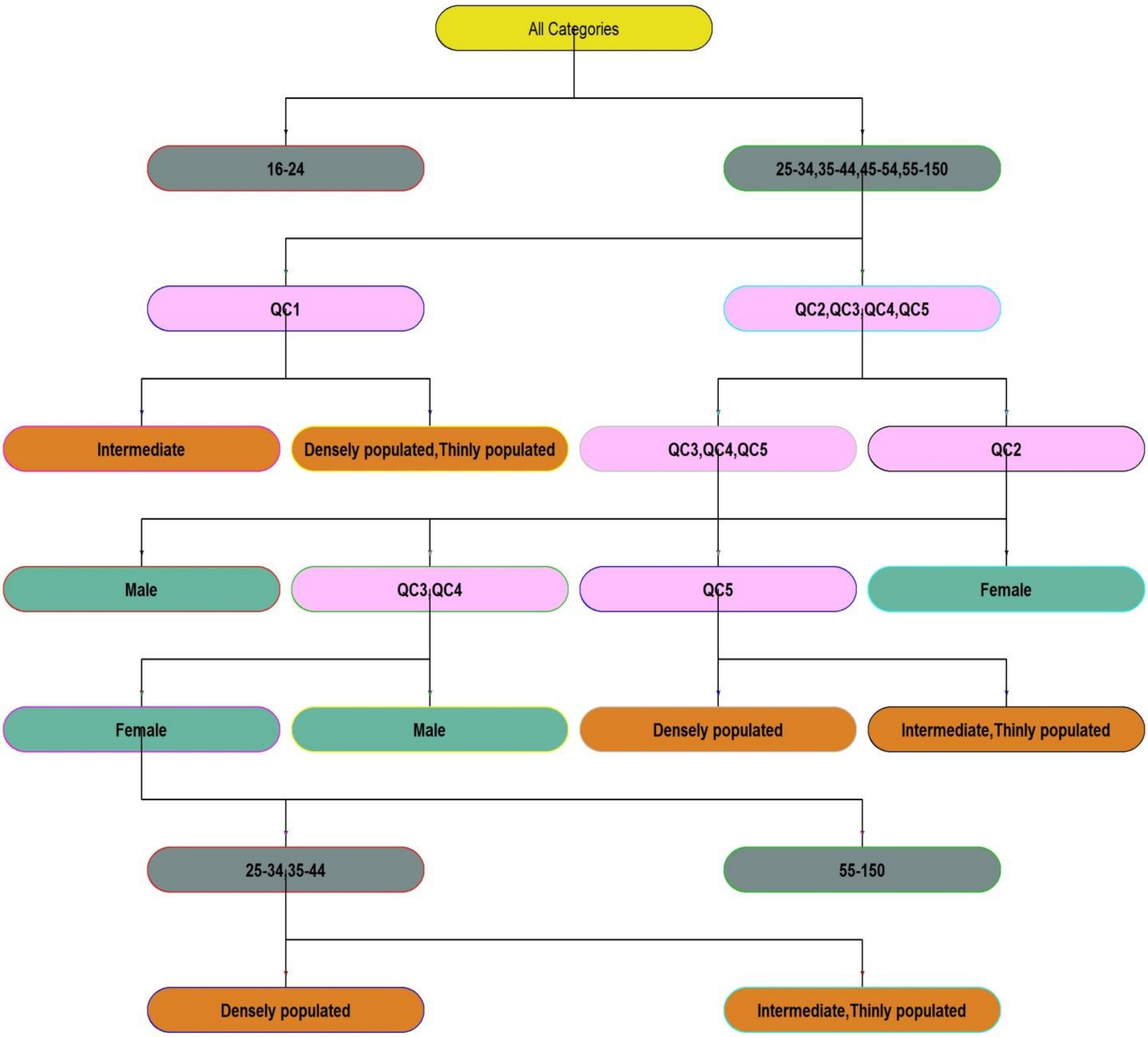
The classification step aims at detecting those socio-demographic categories whose time series of growth rates display similar patterns and to regroup these categories into larger classes. We use EU-SILC time series for 2009-2012 to compute the average growth rates of each of the socio-demographic categories. For wages we use the following variables: gender; geographical regions; degree of urbanisation (densely populated, intermediate, and thinly populated area); type of employment (full-time, part-time, other); age groups (16-24, 25-34, 35-44, 45-54, 55+); wage quantile group in the base year (QC1, QC2, QC3, QC4, QC5). At each node of the decision tree there are several possibilities to build the subclasses using different categorical variables. In order to choose between these potential subclasses we need to define an optimality criterion.

To obtain the optimality criterion we estimate different logistic regression models and compute the associate Wald significance test for the regression parameters. The explanatory variables of the different models are identical and are given by the growth rates from 2009 to

[4] Market incomes are wages and salaries, self-employment income, property income, income from capital, etc.

2011. However, the models differ from each other with respect to dependent variables. Each model uses a different categorical variable as a dependent variable. Finally, we choose the categorical variable associated with the lowest p-value for the Wald test of significance indicating the maximum discriminatory power for this variable. Figure 1 illustrates the shape of the decision tree for Austria using SILC data from 2009 to 2012.

Figure 1: Decision tree for building socio-demographic groups in Austria (SILC 2009-2012)



The modelling step consists of estimating the underlying statistical properties of the growth time series of the classes using a dynamic factor modelling approach. In order to produce estimates of income growth rates, the factor model includes explanatory variables which are more timely than the EU-SILC data, e.g. non-financial quarterly sector accounts such as gross domestic product, gross saving rate, compensation of employees and social contributions and benefits.[5] The estimates are obtained through a combined use of the expectation-maximization algorithm and the Kalman filter/smoother (see Shumway and Stoffer, 1982 and Banbura et al., 2013).

2.3 Simulating changes in tax-benefit policies

After updating market income and other non-simulated income sources, we simulate tax-benefit policies for each year from the base year up to the target year.

We use EUROMOD to simulate changes in the income distribution within the period of analysis. Income elements simulated by the model include universal and targeted cash benefits, social insurance contributions and personal direct taxes. Income elements that cannot be simulated mostly concern benefits for which entitlement is based on previous contribution history (e.g. pensions) or unobserved characteristics (e.g. disability benefits). These are read from the data and updated according to statutory rules (such as indexation rules) or changes in their average levels over time (see Section 2.2). Both contributory and non-contributory unemployment benefits are simulated in the model; but e.g. severance payments are not. Detailed information on EUROMOD and its applications can be found in Sutherland & Figari (2013).

All simulations are carried out on the basis of the tax-benefit rules in place on the 30th June of the given policy year. The exception to this rule is Portugal (in 2012), where policy changes after the 30th of June were taken into account to better match the annual income observed in the EU-SILC data. In order to enhance the credibility of estimates, an effort has been made to address issues such as tax evasion (e.g. in Italy) and benefit non-take-up (e.g. in France and Portugal). However, such adjustments are not possible to implement in all countries due to data limitations.[6] For Italy self-employment income has been calibrated in order to take into account tax evasion behaviour. For France random non-take-up corrections are simulated for the main social assistance benefit. In Portugal non-take-up adjustments were implemented for the social solidarity supplement for the elderly.

The last methodological step involves an attempt to account for differences between EUROMOD and EU-SILC estimates of household income in the data reference year (here 2011). The main reasons for these discrepancies are related to the precision of simulations when information in the EU-SILC data is limited, issues of benefit non-take-up and tax evasion, under-reporting of income components, and small differences in income concepts and definitions.[7]

In order to account for these differences, an alignment factor is calculated for each household. The factor is equal to the absolute difference between the value of equivalised household

[5] Source: Eurostat, non-financial transactions, code “nasq_10_nf_tr”.

[6] Detailed information on the scope of simulations, updating factors, non-take-up and tax evasion adjustments is provided in the EUROMOD Country Reports (see: https://www.iser.essex.ac.uk/euromod/resources-for-euromod-users/country-reports).

[7] For more detailed information on these issues see Figari et al. (2012) and Jara and Leventi (2014).

disposable income in EU-SILC 2012 and the EUROMOD estimate for the same period and income concept. For consistency reasons, the same household specific factor is applied to all later policy years. This is based on the assumption that the discrepancy between EUROMOD and EU-SILC estimates remains stable over time.

3. Quality framework

In the context of the development of flash estimates the quality framework has an essential role as a tool for designing the production process. Therefore, the quality framework doesn't focus only on the final results but includes the following steps:

- Consistency analysis of auxiliary data sources;
- A retrospective assessment based on:
  - intermediate checks for the three stages of production of flash estimates;
  - the ability of the model to reproduce past estimates for main income indicators taking into account the uncertainty related to the sampling variance;
- Quality measures for flash estimates for the target year that takes into account the past performance of the model, the uncertainty in the estimation process related to sampling and model variance.

In this section we illustrate the quality framework using the example of the flash estimates for 2012-2013 for the nine EU countries (Austria, Czech Republic, Finland, France, Italy, Latvia, Luxembourg, Portugal, and Poland) produced based on SILC 2012 data (2011 incomes). The conclusions drawn from this assessment will form the basis for the production of flash estimates for 2014 and 2015 and their respective assessment.

3.1 Consistency analysis of auxiliary data sources

An important element of the nowcasting exercise is the use of auxiliary information in order to estimate changes in income distribution based on the evolution of related indicators such as income components from National Accounts, labour market changes etc. Therefore, this estimation relies on the assumption that the trends observed in EU-SILC data and in auxiliary data sources are similar. A retrospective assessment of consistency was performed for the auxiliary information used in the estimation process: LFS and National Accounts. In addition, the consistency of income components simulated with EUROMOD and observed in EU-SILC was tested for the base year.

3.1.1 Labour Force Survey data

In the first stage the socio-demographic structure of the input data is updated in line with the labour and demographic information from the LFS data for the target year. In order to assess the similarity of two probabilistic distributions (V,V') of the relevant characteristics in EU-SILC and LFS, we use the Hellinger distance (HD):

$$HD(V,V') = \sqrt{\frac{1}{2}\cdot\sum_{i=1}^{K}\left(\sqrt{p(V=i)} - \sqrt{p(V'=i)}\right)^2} = \sqrt{\frac{1}{2}\cdot\sum_{i=1}^{K}\left(\sqrt{\frac{n_{Oi}}{N_O}} - \sqrt{\frac{n_{Pi}}{N_P}}\right)^2}$$

where:

K is the total number of cells in the contingency table;

n_{Oi} is the frequency of cell i in the original data O;

n_{Pi} is the frequency of cell i in the recipient data P;

N is the total size of the specific sources.

A HD value of 0% indicates a perfect similarity between the two probabilistic distributions, whereas a HD value of 100% indicates a total discrepancy.[8] In general, distributions of demographic variables are well-aligned in EU-SILC and LFS, whereas higher discrepancies are observed for labour variables. Table 1 provides a summary for the main demographic and labour variables by country.

Table 1: Average HD (%) for main variables used in the estimation by country: EU-SILC vs. LFS (2011-2014)

Variables	AT	CZ	FI	FR	IT	LU	LV	PT	PL
Main demographic variables (average)	1.2	0.9	1.0	1.0	1.0	1.4	1.0	1.1	1.3
Number of employed (15-74)	1.2	1.4	2.3	0.7	1.1	3.1	0.6	2.0	1.0
Number of retired	0.7	1.0	2.0	2.6	0.7	1.0	1.1	3.6	2.1
Number of self-employed (15-74)	0.8	0.8	0.4	1.7	0.7	1.4	2.1	2.6	1.1
Number of unemployed (15-74)	1.6	3.3	0.8	1.7	3.7	5.3	1.5	1.5	1.1

We can observe that in general the HD values are rather low for the main demographic variables with HD ranging from 0.9 to 1.2. However, the degree of urbanisation for Italy, Finland and Latvia and region for France had a high HD so they were not used further on in the estimation. The HD is higher for labour variables at household level especially for the number of unemployed in the Czech Republic, Italy and Luxembourg and the number retired in Portugal. The inconsistencies are even higher for the self-declared labour status at individual level (the average around 5%). Therefore, for calibration, an adjustment is done so that the update takes into account only the changes in LFS from the base to the target year. In this way the source inconsistencies which are systematic (i.e. constant across time) are ruled out.

A similar approach is taken for modelling employment transitions: the total number of simulated labour market transitions in EUROMOD input data and their direction are determined by *relative* changes in employment rates as shown in the LFS. However, as noted in Rastrigina et al. (2015), employment changes are not always the same in LFS and EU-SILC. The evolution of employment rates in LFS and EU-SILC for 2011-2013 follows

[8] Usually, in statistical analysis a Hellinger distance smaller than 5% is considered acceptable.

different trends in Luxembourg and for specific categories in the Czech Republic, Latvia and France. There are several reasons for these discrepancies, such as differences in definitions, imputations, survey methodology, as well as operational differences that may affect the nature of non-response and sampling errors. A detailed discussion on these issues can be found in Rastrigina et al. (2015).

3.1.2 National Accounts data

For uprating income components, we test two alternative options: (1) the uprating factors included in EUROMOD and (2) the classification method based on modelling time series using macro indicators and National Accounts data on income components. In the latter case only income from employment and self-employment are uprated based on the classification method, for the other income components uprating factors from Euromod are applied. Data from National Accounts and EU-SILC were checked in terms of levels and trends for the period 2009-2014. The intermediate checks showed that for property and self-employment income, the two sources are very different. Moreover, these components vary significantly between the quarterly National Accounts data available for the flash estimates and later revisions. Therefore, only income from employment, social benefits and taxes from the National accounts are used for time series modelling. Further work is ongoing on the reconciliation of macro-micro data on income.

3.1.3 EUROMOD data

EUROMOD input data are based on cross-sectional EU-SILC data. However, these data undergo some transformations: e.g. missing values are imputed, children born after the income reference period are dropped from the data, employment characteristics observed in the data collection period are adjusted to match the income observed in the income reference period; in some countries gross incomes are recalculated from net incomes using different (arguably more precise) algorithms than in the original EU-SILC (e.g. in Italy). In the case of Luxembourg the discrepancies between the nowcasted estimates and the Eurostat indicators in the base year are caused by the fact that households with at least one international civil servant have been excluded from the EUROMOD input data (645 households), as they have a specific tax-benefit system which is different from the national one.[9]

Further discrepancies between EU-SILC and EUROMOD output data for the base year may arise due to precision of simulations when information in the EU-SILC data is limited, issues of benefit non take-up or tax evasion, under-reporting of income components in the EU-SILC data, as well as small differences in income concepts and definitions.

3.2 Intermediate quality checks

In the retrospective quality assessment, intermediate quality checks were implemented throughout the process. This allows for a more precise identification of the problematic

[9] This limitation is going to be addressed in the next update of the Luxembourg model based on SILC 2015 data.

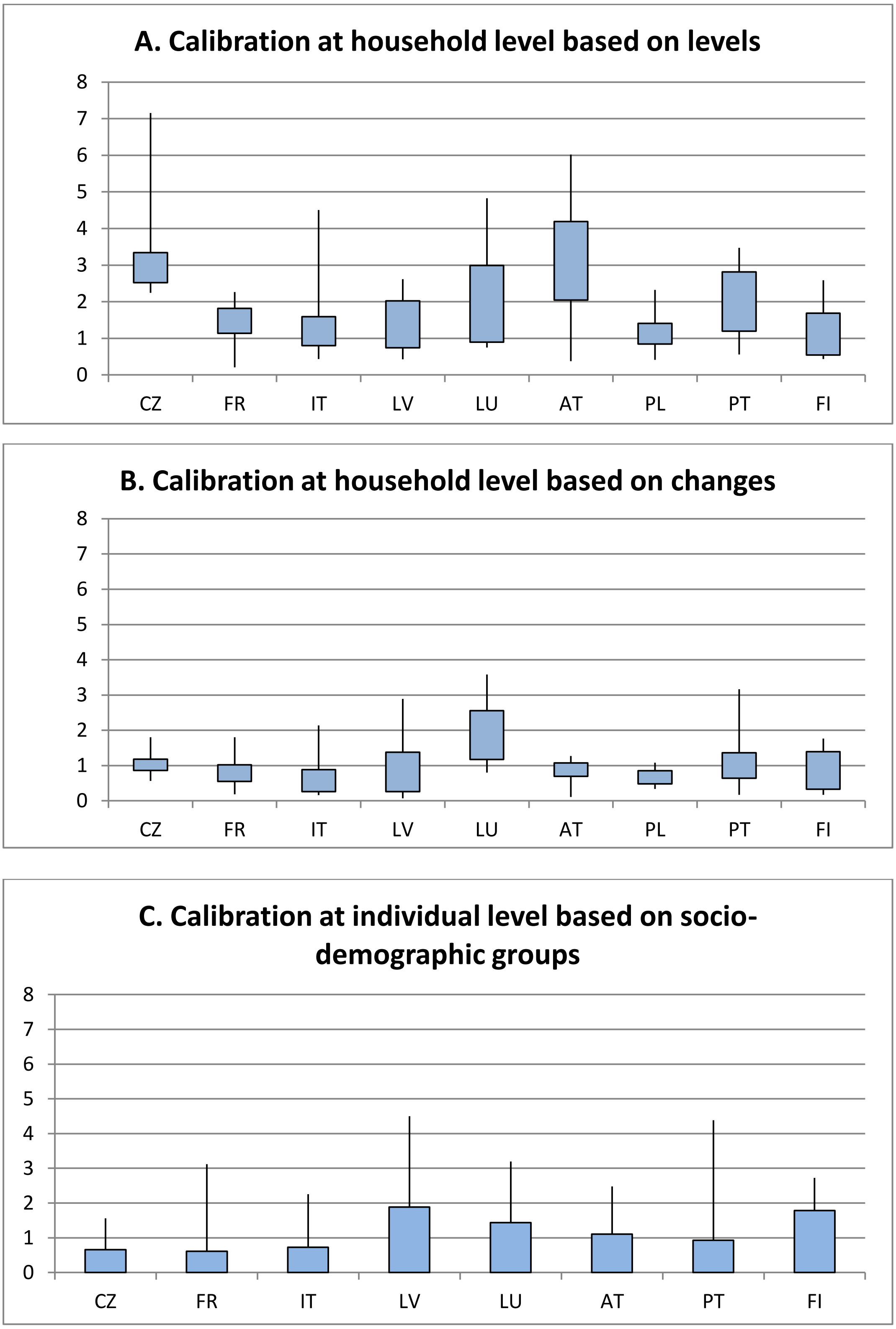
issues: e.g. identify discrepancies for a particular income component or a particular population group.

3.2.1 Demographic and labour characteristics

For categorical demographic or labour market variables we measure the similarity of distributions using the Hellinger distance (HD). We compare variables from the base year (2011) after calibration to the target year (2012 or 2013) with figures from the actual EU-SILC for the corresponding year. This is a measure of performance after the first stage of input data transformations aimed at assessing how close we are to EU-SILC in terms of labour and demographic characteristics. The results are in line with the previous section: the variables with a higher HD are the ones related to the labour information.

Figure 2 shows the box plots of the HD values for the three calibration methods for the following labour variables: number of employed (full-time and part-time), self-employed (full-time and part-time), retired and unemployed by age groups. It can be noticed that in general the calibration at the household level with adjustment (Panel B) improves the consistency between the 'updated' labour variables and EU-SILC data from the target year. For several countries the values are really close for different variables while in others we have a larger spread because of discrepancies for particular categories: e.g. unemployed in Luxembourg, retired in Austria. For the individual level reweighting there are larger discrepancies in Finland and Latvia.

Figure 2: Similarity of distributions for main labour variables: calibrated SILC and actual SILC for 2012 and 2013 income years (HD in %)



Note: Calibration at individual level is not available for Poland.

3.2.2 Key income indicators

For the retrospective assessment of the key income indicators the original EU-SILC indicators and the estimated ones are compared based on two main metrics:

(1) Accuracy - measuring the accuracy of point estimates based on the mean absolute percentage error (MAPE):

$$accuracy = 1 - MAPE = 1 - average_{i=1}^{n}\left(abs\left(\frac{EST_i}{REF_i} - 1\right)\right)$$

(2) Consistency - measuring the extent to which the year-on-year rates of changes are similar across the time series:

$$consistency = 1 - average_{i=1}^{n}\left(abs\left(\frac{EST_i}{EST_{i-1}} - \frac{REF_i}{REF_{i-1}}\right)\right)$$

Both measures involve subtraction from 1 in order to have higher values indicating better performance. The two performance metrics allow conducting an extensive comparative analysis across countries, years and methods and to summarise a large amount of information.

In Table 2 the average accuracy for each income component is calculated by comparing the EU-SILC 2012 (income 2011) with the EUROMOD output file based on the same data. If there is no difference between them, the accuracy values are equal to 1.

Table 2. Average accuracy for income components in EUROMOD by country in the base year (2011)

Country	AT	CZ	FI	FR	IT	LU	LV	PL	PT	**Average income**
Employment	1	1	1	1	0.97	0.98	1	1	1	0.99
Private pension	1	1	1	1	0.87	1	1	1	1	0.99
Inter-household transfers received	1	1	1	1	0.88	0.99	1	1	1	0.99
Inter-household transfers paid	1	1	1	1	0.97	0.99	1	1	0.87	0.98
Self-employment	1	1	1	1	0.8	0.99	1	1	1	0.98
Property	1	1	0.97	0.79	0.99	0.97	0.98	1	1	0.97
Social benefits	0.99	0.93	0.95	0.99	0.64	0.96	0.96	0.94	0.97	0.93
Public pension	0.85	0.96	0.86	0.84	0.86	0.72	0.98	0.9	0.87	0.87
Taxes	0.89	0.71	0.89	0.67	0.97	0.94	0.58	0.86	0.83	0.82
Unemployment benefits	0.83	0.79	0.76	0.38	0.92	0.94	0.96	0.71	0.82	0.79
Average country	0.96	0.94	0.94	0.87	0.89	0.95	0.95	0.94	0.94	0.93

Note: Average for indicators: D10, D30, MEDIAN, D70, D90.
Source: EUROMOD 3.0+

Overall, the non-simulated income components[10] are very similar to their original SILC values. However, there are some country-specific differences due to imputations and other data transformations: e.g. in Italy for private pensions, self-employment income and inter-household transfers received, in France for property income and in Portugal for inter-household transfers paid. Larger differences are observed for the simulated income components with accuracy being below 0.90 for most of the countries for unemployment benefits and taxes.

Table 3 shows the consistency indicators for the same income components comparing the evolution of the main indicators based on three alternative nowcasting methods and the actual evolution in EU-SILC.

Table 3. Average consistency of methods by indicator and income component (2011-2013)

	2011-2012					2012-2013					**Average income**
	D10	D30	D50	D70	D90	D10	D30	D50	D70	D90	
Employment	0.94	0.96	0.98	0.97	0.97	0.91	0.97	0.98	0.98	0.97	0.96
Inter-household transfers paid	0.76	0.81	0.82	0.83	0.86	0.81	0.87	0.94	0.94	0.9	0.85
Inter-household transfers received	0.78	0.77	0.77	0.66	0.8	0.77	0.91	0.93	0.88	0.96	0.82
Property	0.52	0.61	0.75	0.79	0.66	0.79	0.77	0.82	0.85	0.78	0.73
Private pension	0.61	0.38	0.67	0.67	0.61	0.55	0.26	0.78	0.75	0.62	0.59
Public pension	0.95	0.97	0.97	0.97	0.97	0.95	0.99	0.98	0.97	0.98	0.97
Social benefits	0.84	0.86	0.96	0.96	0.93	0.94	0.94	0.99	0.98	0.99	0.94
Self-employment	0.86	0.84	0.92	0.84	0.87	0.8	0.79	0.95	0.11	0.83	0.78
Taxes	0.9	0.93	0.96	0.97	0.95	0.91	0.93	0.97	0.96	0.97	0.95
Unemployment benefits	0.89	0.89	0.93	0.88	0.9	0.82	0.92	0.91	0.92	0.9	0.9
Average indicator	0.8	0.8	0.87	0.85	0.85	0.83	0.83	0.93	0.84	0.89	0.85

Note: Average for 9 countries for the following methods: (1) calibration and detailed model-based uprating factors for employment and self-employment income; (2) calibration and EUROMOD uprating factors; (3) labour transitions and EUROMOD uprating factors.

The consistency is quite high for the median for employment income (0.98) and public pensions (0.97), but relatively lower for self-employment (0.92) and, particularly, for property income (0.75). The latter is highly volatile and difficult to predict. The consistency is generally worse for smaller-scale income components such as inter-households transfers paid/received. The changes in the left tail of distribution (D10 and D30) are also more challenging to estimate.

[10] The non-simulated components are employment and self-employment income, inter-household transfers received/paid, property income, private pension and most public pensions.

3.2.3 Income distribution

Finally, a different approach can be used to compare the nowcasted income distribution (of total household disposable income or of separate income components) with the corresponding EU-SILC income distribution. In order to assess the performance of a given nowcasting procedure we first need to define an appropriate measure of similarity between the empirical distribution function provided by the nowcasting procedure and the distribution function derived from the corresponding EU-SILC dataset. Moreover, we need to establish a decision rule indicating whether a given value of this measure reflects a strong or a weak similarity.

To do this, we can use some of the concepts of two-sample distribution testing. One of the most widely used distribution test is the Kolmogorov–Smirnov test (KS test). The similarity measure used by the KS test takes the form of the maximum distance between the empirical cumulative distribution functions (ECDF) of the flash estimate and the one derived from EU-SILC. Hence, if we denote as $\hat{F}(x)$ the ECDF of the flash estimate and as $\tilde{F}(x)$ the ecdf of the income component in EU-SILC we have that the test statistics of the KS test is given by:

$$D(x) = \sup_x |\hat{F}(x) - \tilde{F}(x)|.$$

The decision rule based on the KS test provides an evaluation of the overall fit of the nowcasted probability distribution with respect to the EU-SILC distribution. In some cases, however, the overall goodness-of-fit might not be sufficient in order to draw conclusions on the quality of the flash estimates of the income indicators. It is possible for the p-value of the KS test to be below the chosen threshold and hence for the null hypothesis of a common population distribution to be rejected even when the at-risk-of-poverty (AROP) rate and the S80/S20 are well-approximated (note that these two indicators do not depend directly on the central location of the respective probability distribution). Hence, if the nowcasted distribution of the disposable income only differs from the SILC distribution by a central shift, the AROP and the S80/S20 of the two distributions are identical. However, the decision rule based on the KS test will indicate a low degree of similarity. One possible remedy in order to make the KS-based decision rule shift-insensitive would be to standardize both distributions by their respective medians.

In order to assess the goodness of the income distribution estimation, we have standardised the results of the tests to the initial result of the base year (see Table 4). After controlling for the initial similarity of the income distribution, none of the income components performs better on average than the initial distributions. The similarity to the target SILC distribution is worse on average after 2 years than after 1 year. This highlights the importance of using the latest SILC income distribution in the production of flash estimates. A further more detailed analysis needs to be done by method and income component to assess the impact of our methods on the estimation of the SILC income distribution.

Table 4. Proportion of results for which the null hypothesis that the two distributions are the same is accepted (with α = 0.01)

	2011 (BASE)	2012	2013	**Average**
Employment	100	86	59	72
Inter-household transfers paid	100	69	60	65
Inter-household transfers received	100	82	107	95
Property	100	15	27	21
Private pension	100	42	63	52
Public pension	100	76	94	85
Social benefits	100	53	94	73
Self-employment	100	94	85	90
Taxes	100	83	82	83
Unemployment benefits	100	101	78	90

Note: Average for 9 countries for the following methods: (1) calibration and detailed model-based uprating factors for employment and self-employment income; (2) calibration and EUROMOD uprating factors.

3.3 Assessment of results for flash estimates 2012-2013

As part of the assessment procedure, we do a retrospective analysis of flash estimates for 2012 and 2013 income years. In particular we look at:

(a) the two performance metrics for comparing point estimates and year-on-year changes for the key income indicators;

(b) additional quality measures for different flash estimates under uncertainty.

3.3.1 Performance metrics

Table 5 provides a quick overview on the performance of different methods by country for several indicators: the first decile (D10), median, mean, quintile share ratio (QSR) and the at-risk-of-poverty rate. This can give a first basis for comparison of different methods and their selection for the production of flash estimates in the future. We can notice that the results are indicator and country dependent. We highlight mainly differences across methods for updating the demographic structure. The different methods of uprating income components seem to have a less significant impact on the final indicators. The method based on labour market transitions works better in several countries and for more difficult indicators such as the quintile share ratio (QSR) and AROP. However, there is a certain complementarity as for example calibration methods work better in Italy and the Czech Republic. In general, adjusting for source inconsistencies in LFS improves accuracy but has little effect on consistency. Therefore, it is probably optimal to use a set of methods for the production of flash estimates rather than relying on a single method. Further work will need to be done in order to define the ‘conditions of use’ of these methods which will be country-dependent but might depend also on other factors, such as changes in the economic conditions. We know from the literature that reweighting might work better in times of stability while explicitly modelling labour market transitions is better suited for turbulent years.

Figure 3 (see Appendix) provides the results for the at-risk-of-poverty indicator. In the upper part of each graph we show the levels of flash estimates in 2012 and 2013 for the different nowcasting methods. In the lower part we report the nowcasted changes between 2012 and 2013 which are compared with the actual changes in EU-SILC. Consistency scores are generally higher than accuracy measures so we can assume that the model bias is systematic and focus on the estimation of year-on-year changes rather than levels.

While AROP is an extremely relevant indicator (because it is one of the Europe 2020 target indicators), it is also one of the most difficult to estimate as it can cumulate different dynamics across the distribution. Positional indicators such as the at-risk-of-poverty threshold and the deciles cut-off points are in general better predicted than the at-risk-of-poverty rate or the quintile share ratio. In general, the adjustment for the changes in population characteristics seems to have a larger impact than the uprating factors. We can also notice that there is no single method that shows better performance for all indicators and all years. This raises questions with respect to the method selection procedure which would need to be country specific, taking into account also the consistency of input sources. An important issue is the estimation of very small changes that often are not statistically significant. The next section discusses the analysis of the historical performance under uncertainty and proposes a method for taking into account the sampling error in the assessment framework.

Table 5: Average consistency by indicator and by method (2012/11 and 2013/12)

Indicator /Method	AT	CZ	FI	FR	IT	LU	LV	PL	PT
AROP Labour transitions	95%	83%	89%	92%	98%	71%	94%	99%	99%
AROP CAL_H_A	97%	86%	90%	93%	97%	79%	96%	97%	97%
AROP CAL_H	98%	91%	90%	91%	96%	81%	97%	97%	96%
AROP CAL_I	97%	88%	90%	93%	97%	79%	95%		96%
D10 Labour transitions	96%	97%	99%	99%	92%	96%	96%	99%	97%
D10 CAL_H_A	96%	97%	98%	98%	96%	95%	94%	97%	97%
D10 CAL_H	96%	98%	98%	97%	94%	95%	94%	97%	97%
D10 CAL_I	96%	98%	98%	98%	96%	96%	91%		96%
MEAN Labour transitions	94%	98%	98%	97%	99%	99%	96%	99%	98%
MEAN CAL_H_A	95%	99%	99%	90%	98%	99%	96%	98%	99%
MEAN CAL_H	95%	99%	99%	89%	97%	98%	96%	98%	99%
MEAN CAL_I	95%	99%	99%	89%	98%	99%	94%		99%
MEDIAN Labour transitions	96%	98%	98%	96%	99%	99%	95%	99%	97%
MEDIAN CAL_H_A	97%	99%	99%	97%	98%	98%	96%	98%	98%
MEDIAN CAL_H	96%	99%	98%	97%	97%	98%	96%	98%	98%
MEDIAN CAL_I	97%	99%	98%	98%	98%	99%	94%		99%
QSR Labour transitions	91%	93%	96%	97%	87%	93%	93%	97%	96%
QSR CAL_H_A	95%	95%	97%	72%	99%	92%	91%	97%	96%
QSR CAL_H	95%	97%	97%	69%	95%	93%	93%	97%	96%
QSR CAL_I	95%	96%	97%	72%	98%	93%	93%		96%

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 equivalised household disposable income; QSR (income quintile share ratio).

3.3.2 Performance measures under uncertainty

The objective of statistical estimation of the income distribution is to find an estimate of the population income probability distribution function using samples of income data drawn from the population. The estimate of the distribution function allows deriving estimates of various parameters of interest, for example, income indicators measuring income inequality within a population such as the at-risk-of-poverty rate, the quintile share ratio or the Gini coefficient.

An important element of uncertainty in flash estimates is related to the sampling variance coming from EU-SILC data. For assessing if the change observed in EU-SILC between two periods is significant Eurostat relies on the multivariate regression approach developed by Berger and Priam (2016). This approach takes into account the fact that the comparison is based on two waves of cross-sectional data that are partially overlapping. The analysis of the AROP changes between 2012 and 2013 shows that only 14 out of the 28 EU countries have a significant change during this time period according to Eurostat calculations.

The current work focuses on the development of performance measures, complementary to accuracy and consistency, which aim at establishing a link between flash estimation and concepts from the probability theory. A performance analysis which aims at accounting for the random nature of the underlying data generating process would not aim at computing the distance between the reference value and the flash estimate but would take into account the uncertainty of the estimates being compared.

Taking into account these considerations of uncertainty we test if our results from different methods can give an accurate indication of the change in the indicators on a magnitude-direction scale with 6 classes (referred to as "MD6"):

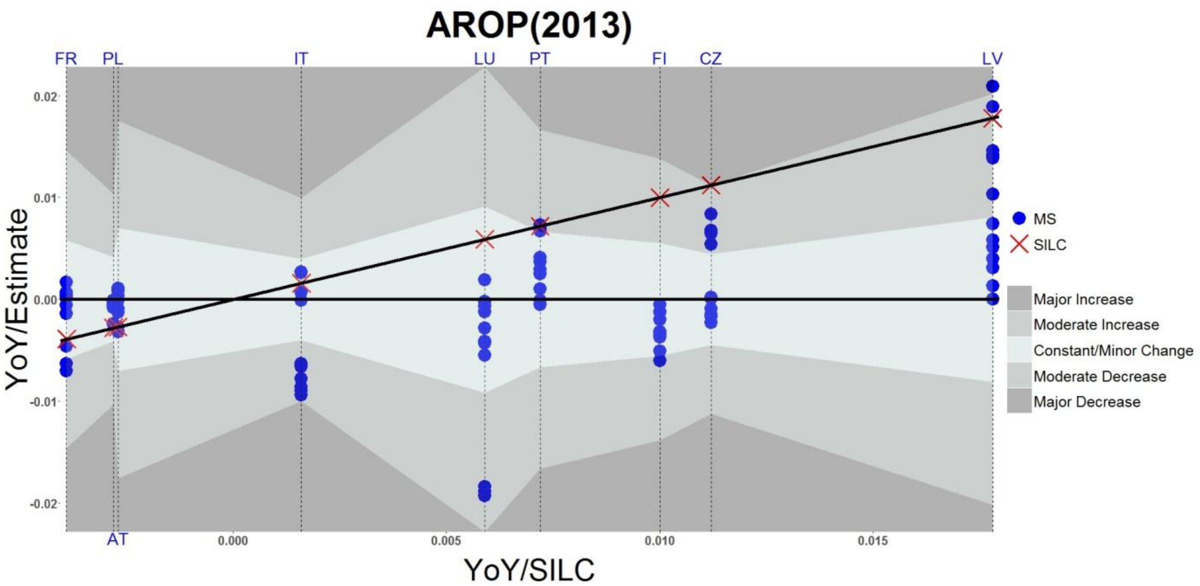
(1) major increase [+++]
(2) moderate increase [+]
(3) (quasi) stable / minor changes [O]
(4) moderate decrease [-]
(5) major decrease [---]
(6) *no conclusion* (when we get contradictory signals from different methods with good past performance).

Figure 4 illustrates the performance of the different flash estimation approaches based on the MD6 scale for the AROP indicator by country. The x-axis shows the year-on-year (YoY) change of the SILC indicators observed between 2012 and 2013. The y-axis indicates the YoY change of the flash estimate during the same period. The 45-degree line contains all those points for which the YoY change of the flash estimate is identical to the YoY change in the original SILC. Hence, the closer a given flash estimate is located to the 45-degree line the higher is the consistency of that estimate. The shaded areas indicate the five classes of the MD6 rule indicated above. Each of the points refers to a given flash estimate.

We use the MD6 scale where the classes are derived from the average sampling error (ASE) for each indicator for the period 2008-2013; the thresholds that define the classes are multiples of ASE:

- major increase: > 5 ASE
- moderate increase: between 2 and 5 ASE
- (quasi) stable / minor changes [O]: within ±2 ASE
- moderate decrease: between -2 and -5 ASE
- major decrease: < - 5 ASE

Figure 4: Performance of AROP flash estimates on a magnitude-direction scale with 6 classes (2012-2013)



The results differ by country but in general for Austria, France, Italy, Luxembourg, and Portugal the YoY of the flash estimates tend to cluster in the same class of the MD6 rule or in one of the adjacent classes. On the contrary, in case of Finland all flash estimates for AROP are in different classes. A further selection of the methods will be done by country based on their historical performance. The estimates might also be based on several methods that had a good historical performance. The convergence of the methods towards a certain estimate for the year-on-year change can be used as an additional criterion of quality for the flash estimate in the target year.

3.4 Uncertainty measurement and statistical significance of changes

Even though the results of the estimation process are expressed as point estimates for a set of indicators related to the income distribution, the primary interest for users and policy makers often lies in the changes or trends from one time period to another: e.g. they would like to know if there were changes in the at-risk-of-poverty rate and if these changes were statistically significant.

The summary of the actual EU-SILC changes in the last years shows that it is extremely important to take into account the sampling variance. Therefore, current work focuses on the inclusion of uncertainty in the retrospective assessment so that the different methods are "scored" not just on the basis of their ability to reproduce point estimates but also to take into account the sampling variance.

The last step in the quality framework is to integrate the information on the results from different methods and their historical performance in a measure of quality of flash estimates. These should be based on three elements: 1) the historical performance of the model taking into account the sampling error; 2) the model variance; and 3) the extent to which the different methods with a good historical performance converge towards similar estimates for the target year. The latter point refers to the results obtained through different methods. The larger the distance between the values of the year-on-year changes of the different flash estimates the less informative the set of flash estimates. The decision whether a set of flash estimates is informative or not should therefore take into account the results obtained at the target year and the historical performance of the different estimation procedures. Indeed, if for instance all the flash estimates have values close to each other but on the other hand these flash estimates have very low historical performance, the informative power of this type of scenario might still be very low.

4. Conclusions

The first stage of quality assessment has focused mainly on comparing different methodologies and auxiliary sources of data based on their historical performance. Some general conclusions can be drawn from the performance analysis:

- There is no single method that shows better performance for all indicators and all years. There is a strong country effect, related to the auxiliary sources used in the estimation process and the times series in EU-SILC. The performance can depend also on other exogenous factors such as economic conditions; and some methods might be more suitable for capturing the trend in stable conditions whereas others in times of crisis. This raises questions with respect to the method selection procedure. There are two potential ways to address this issue: 1) quantify the loss of accuracy taking into account in a systematic way all indicators together with model uncertainty/stability and choose the best method considering all parameters; 2) consider several methods together and produce an estimate based on their convergence or based on a combination of several methods (e.g. an average weighted according to performance).

- Positional indicators such as the at-risk-of-poverty threshold and the decile cut-off points are in general better predicted than the at-risk-of-poverty rate or the quintile share ratio. This should be taken into account when deciding on the targeted indicators for flash estimates.

- Consistency scores are generally higher than accuracy measures; this indicates that a more efficient strategy of producing flash estimates would be to estimate the year-on-year change and apply it to the last observed value. The lower accuracy scores for high-consistency models might be indicative of a model bias; the development of procedures for adjusting for such model biases is a high-priority next step.

- The consistency analysis should be considered in relation to the actual changes in EU-SILC data. For example, from 2012 to 2013 only 14 EU countries had significant changes in the AROP indicator according to Eurostat calculations. A further refinement of the analysis should consider separately the consistency analysis for countries with significant changes and without such changes. The method should include a more clear assessment of the ability of the model to capture the direction and magnitude of changes when the year-on-year changes are statistically significant.

- It is important to notice that consistency – the ability to correctly estimate year-on-year changes – decreases as the magnitude of changes increases. In other words, our approaches do not work particularly well in estimating larger jumps.

Further work will focus on developing uncertainty measurement and include it in the assessment framework. This raises not only a question of quality but also of communication of the results to the public. In some cases the time series in SILC are rather stable so flash estimates, even if small changes are registered, can give only a general message of stability. In other cases the direction and magnitude of change might be significant and a more precise and clear quantification of such change can be presented.

The quality framework presented in this paper will be used to assess the flash estimates produced using microsimulation techniques and alternative estimates coming from parallel streams of work at Eurostat: a macro-approach for estimating the income distribution indicators. In addition several Member States have their own national models which are used to produce similar types of estimates. All these alternative estimates will enter a common platform based on a single quality framework that would support the selection of a set of flash estimates at EU level. The ultimate goal is to include the flash estimates on income distribution as a part of the European Semester toolkit conditional on the flash estimates passing the assessment framework and consultation with the main users.

References

Anderson, T. W. and Darling D. A. (1954) A test of goodness-of-fit, Journal of the American Statistical Association, vol. 49, no. 268, pp. 765-769.

Avram, S., Sutherland, H., Tasseva, I. and Tumino, A. (2011) Income protection and poverty risk for the unemployed in Europe, Research Note 1/2011 of the European Observatory on the Social Situation and Demography, European Commission.

Banbura, M., Giannone, D., Modugno, M. and Rechlin, L. (2013) Now-casting and the real-time data flow, ECB Working Paper Series, no. 1564, pp. 1-53.

Berger, Y.G., and Priam, R. (2016) A simple variance estimator of change for rotating repeated surveys: an application to the European Union Statistics on Income and Living Conditions household surveys, Journal of the Royal Statistical Society, Series A Statistics in Society, vol. 179, issue , January 2016, pp. 251–272

Bourguignon, F., Bussolo, M. and Da Silva, L.P. (2008) The impact of macro-economic policies on poverty and income distribution Macro-Micro Evaluation Techniques and Tools, The World Bank and Palgrave-Macmillan, New York.

Brandolini, A., D'Amuri, F. and Faiella, I. (2013) "Country case study – Italy", Chapter 5 in Jenkins et al, The Great Recession and the Distribution of Household Income, Oxford: Oxford University Press.

Brewer, M., Browne, J., Hood, A., Joyce, R., Sibieta, L. (2013) The Short- and Medium-Term Impacts of the Recession on the UK Income Distribution, Fiscal Studies, 34(2): 179–201.

Essama-Nssah, B. (2005) The Poverty and Distributional Impact of Macroeconomic Shocks and Policies: A Review of Modeling Approaches, World Bank Policy Research Working Paper 3682, Washington, DC.

Fernandez Salgado M., Figari, F., Sutherland, H. and Tumino, A. (2014) Welfare compensation for unemployment in the Great Recession, The Review of Income and Wealth, 60(S1): 177–204.

Figari, F., Iacovou, M., Skew, A. and Sutherland, H. (2012) Approximations to the truth: comparing survey and microsimulation approaches to measuring income for social indicators, Social Indicators Research, 105(3): 387-407.

Figari, F., A. Paulus and H. Sutherland (2015) "Microsimulation and Policy Analysis" in A.B. Atkinson and F. Bourguignon (Eds.) Handbook of Income Distribution, Vol 2B. Amsterdam: Elsevier, pp. 2141-2221. ISBN 978-0-444-59430-3

Figari, F., Salvatori, A. and Sutherland, H. (2011) Economic downturn and stress testing European welfare systems, Research in Labor Economics, 32: 257-286.

Immervoll, H., Levy, H., Lietz, C., Mantovani, D. and Sutherland, H. (2006) The sensitivity of poverty rates in the European Union to macro-level changes, Cambridge Journal of Economics, 30: 181-199.

Jara, X.H. and Leventi, C. (2014) Baseline results from the EU27 EUROMOD (2009-2013), EUROMOD Working Paper EM18/14, Colchester: ISER, University of Essex.

Kalman, R. E. (1960) A new approach to linear filtering and prediction problems, Transactions of the ASME-Journal of Basic Engineering, no. 8, pp. 35-45.

Keane C., Callan, T., Savage, M., Walsh, J.R. and Timoney, K. (2013) Identifying Policy

Impacts in the Crisis: Microsimulation Evidence on Tax and Welfare, Journal of the Statistical and Social Inquiry Society of Ireland, XLII: 1-14.

MacDonald, J. B. and Xu, Y. J. (1995) A generalization of the beta distribution with applications, Journal of Econometrics, vol. 66, no. 1-2, pp. 133-152.

Massey, J. (1951) The Kolmogorov-Smirnov Test for Goodness of Fit, Journal of the American Statistical Association, vol. 46, no. 253, pp. 68-78.

Matsaganis, M. and Leventi C. (2014) Poverty and inequality during the Great Recession in Greece, Political Studies Review, 12: 209 -223.

Narayan, A. and Sánchez-Páramo C. (2012) Knowing, when you don't know: using microsimulation models to assess the poverty and distributional impacts of macroeconomic shocks, The World Bank, Washington DC.

Navicke, J., Rastrigina, O. and Sutherland, H. (2013) Nowcasting Indicators of Poverty Risk in the European Union: A Microsimulation Approach, Social Indicators Research, 119(1): 101-119.

O'Donoghue, C., Loughrey, J. (2014) Nowcasting in Microsimulation Models: A Methodological Survey, Journal of Artificial Societies and Social Simulation 17 (4) 12.

Peichl, A. (2008) The Benefits of Linking CGE and Microsimulation Models: Evidence from a Flat Tax Analysis, IZA Discussion Paper No. 3715.

Rastrigina, O., Leventi, C., Vujackov S. and Sutherland, H. (2016) Nowcasting: estimating developments in median household income and risk of poverty in 2014 and 2015, Research Note 1/2015, Social Situation Monitor, European Commission.

Rastrigina, O., Leventi, C. and Sutherland, H. (2015) Nowcasting: estimating developments in the risk of poverty and income distribution in 2013 and 2014, Research Note 1/2014, Social Situation Monitor, European Commission.

Shumway, R. H. and Stoffer, D. S. (1982) An approach to time series smoothing and forecasting using the EM algorithm, Journal of Time Series Analysis, vol. 3, no. 4, pp. 253-264.

Sutherland, H. and Figari, F. (2013) EUROMOD: the European Union tax-benefit microsimulation model, International Journal of Microsimulation, 6(1): 4-26.

Appendix

Figure 3: Flash AROP estimates: 2012-2013 levels and year-on-year change

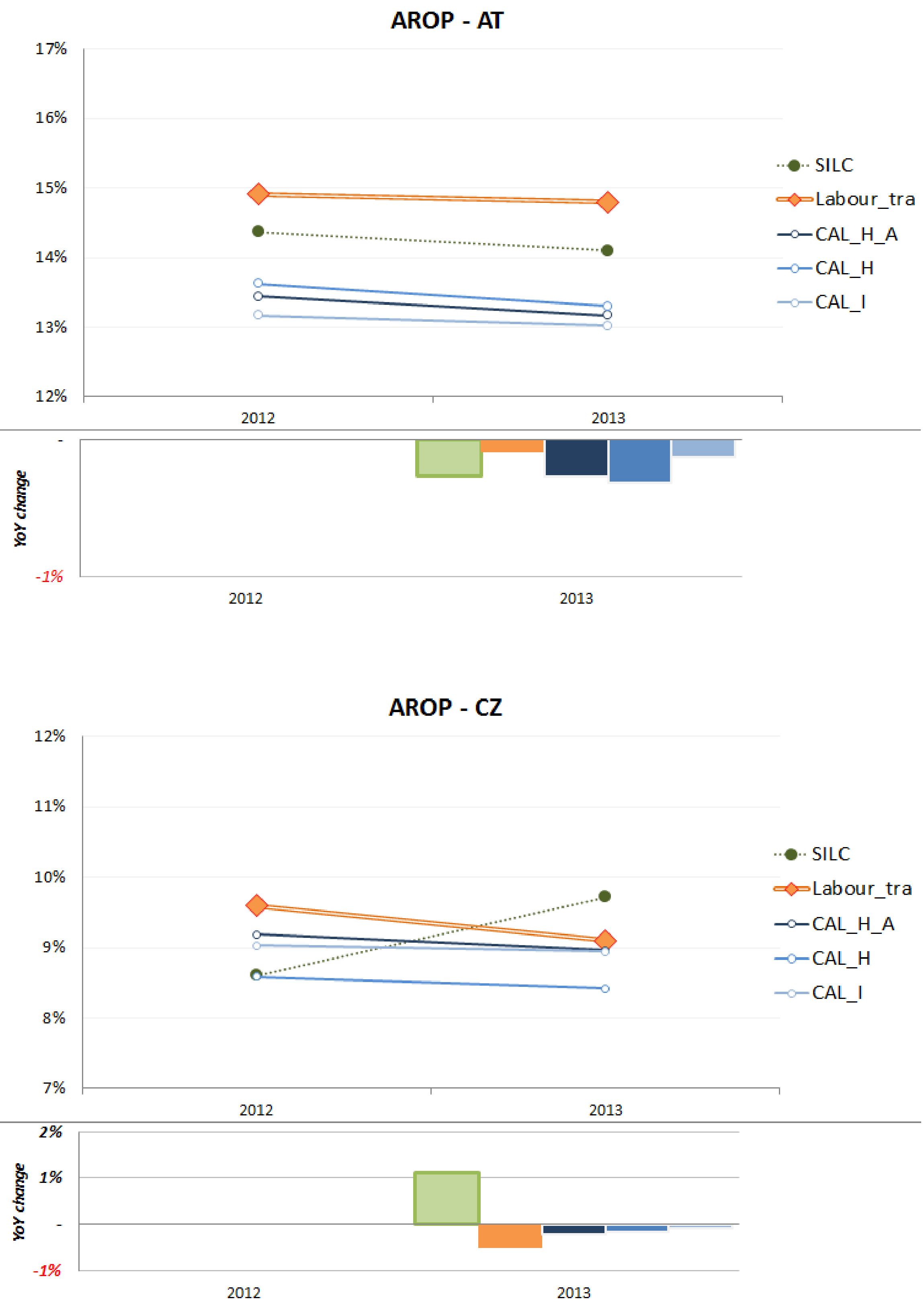
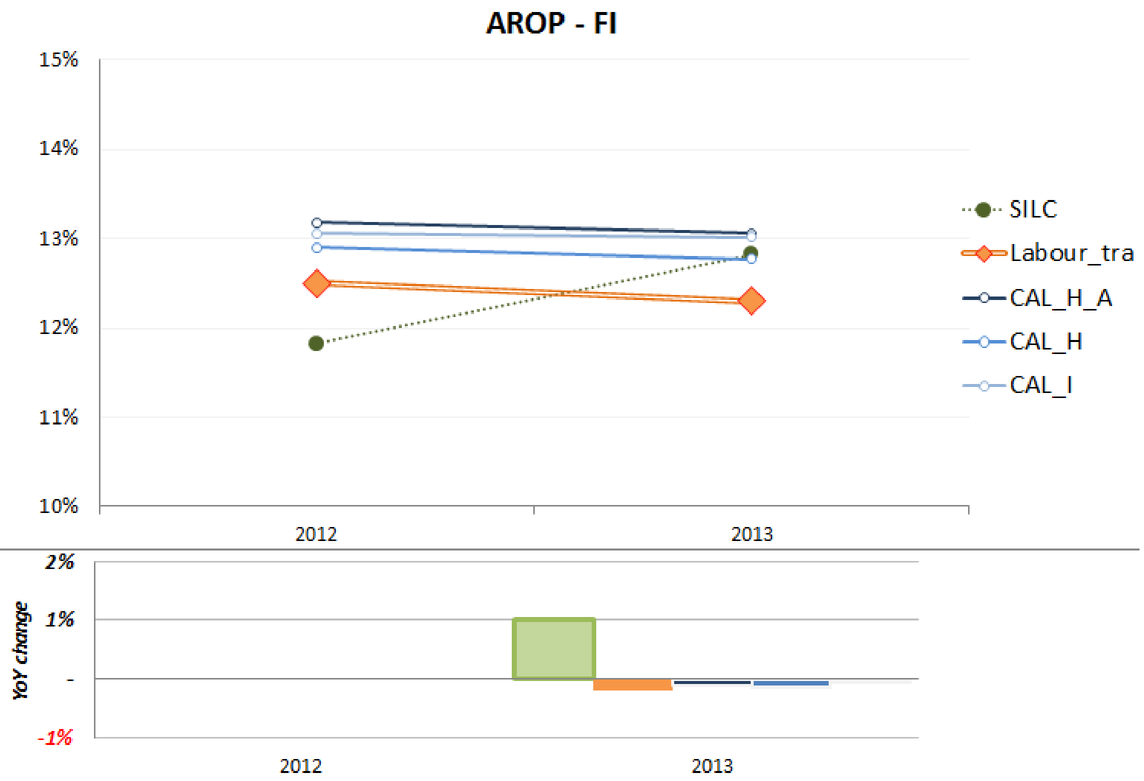


Figure 3 (contd): Flash AROP estimates: 2012-2013 levels and year-on-year change



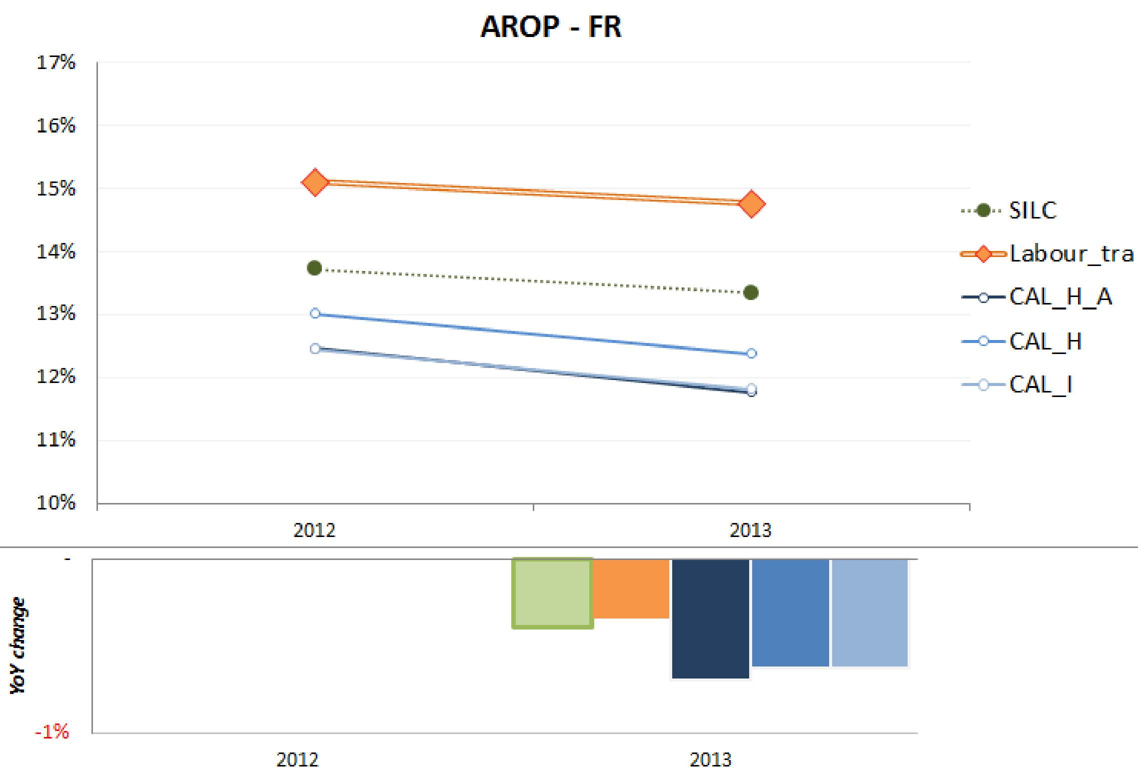
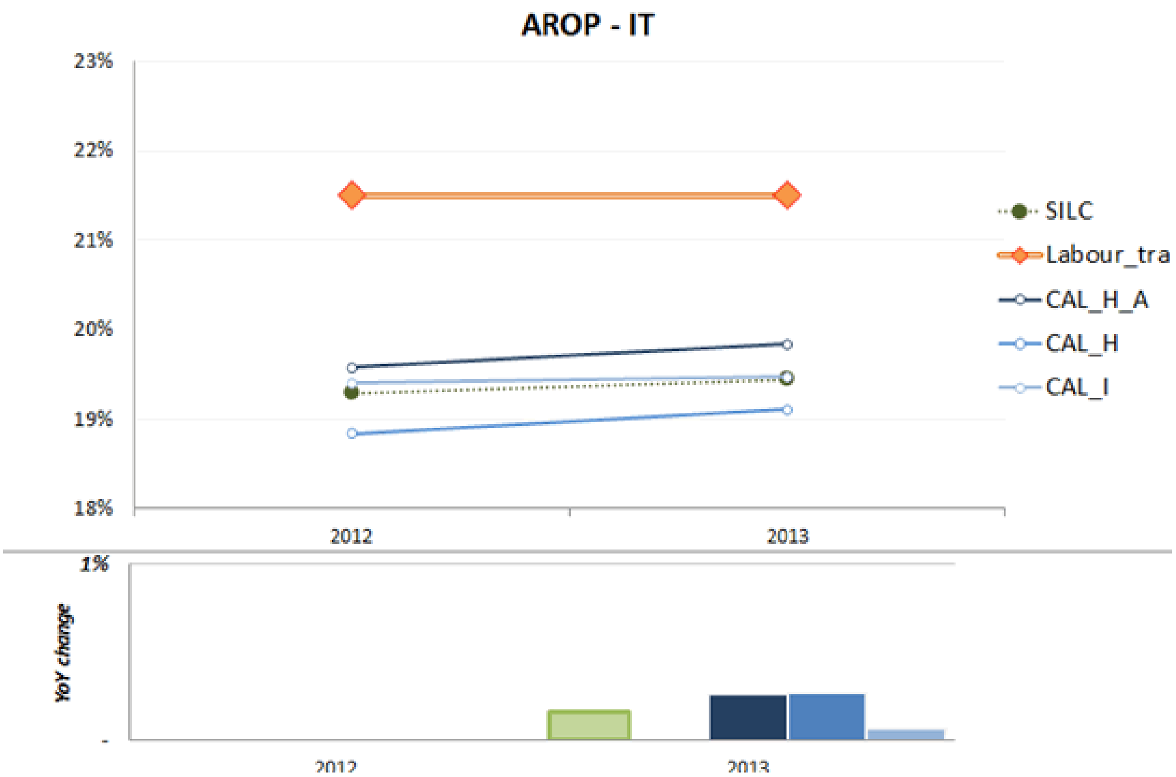


Figure 3 (contd): Flash AROP estimates: 2012-2013 levels and year-on-year change



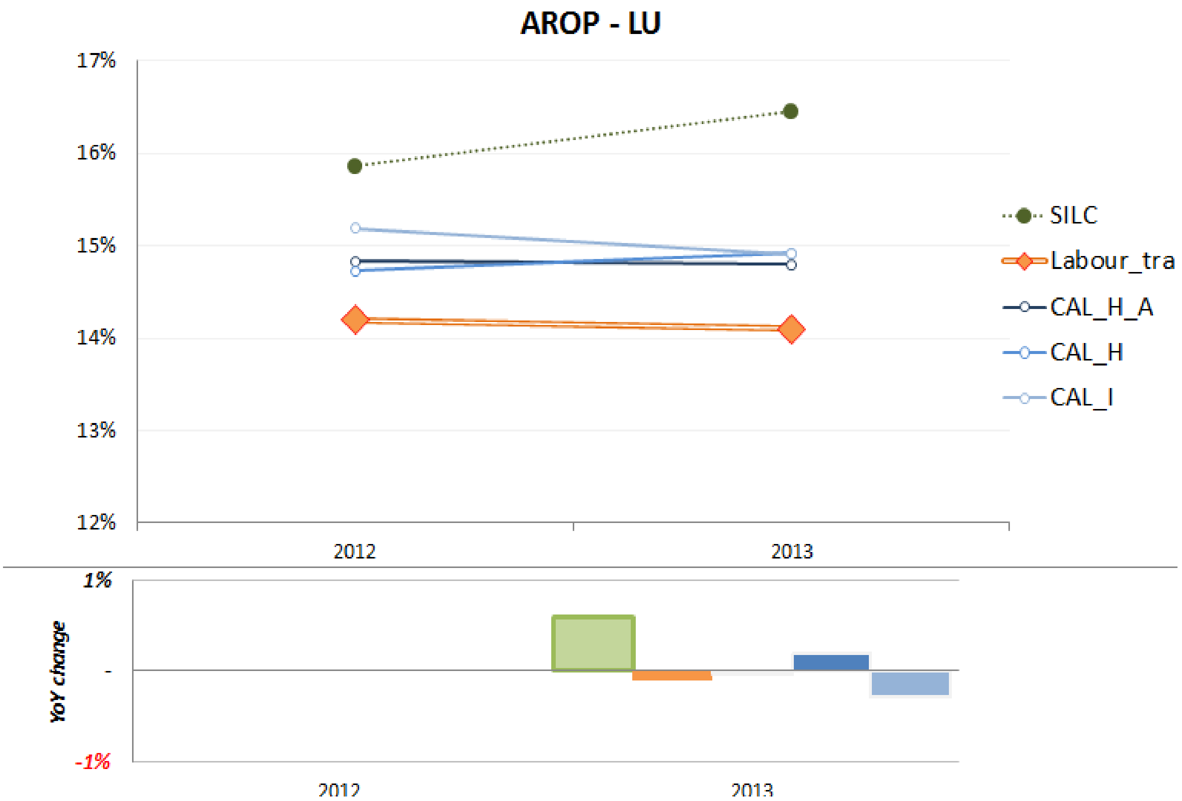
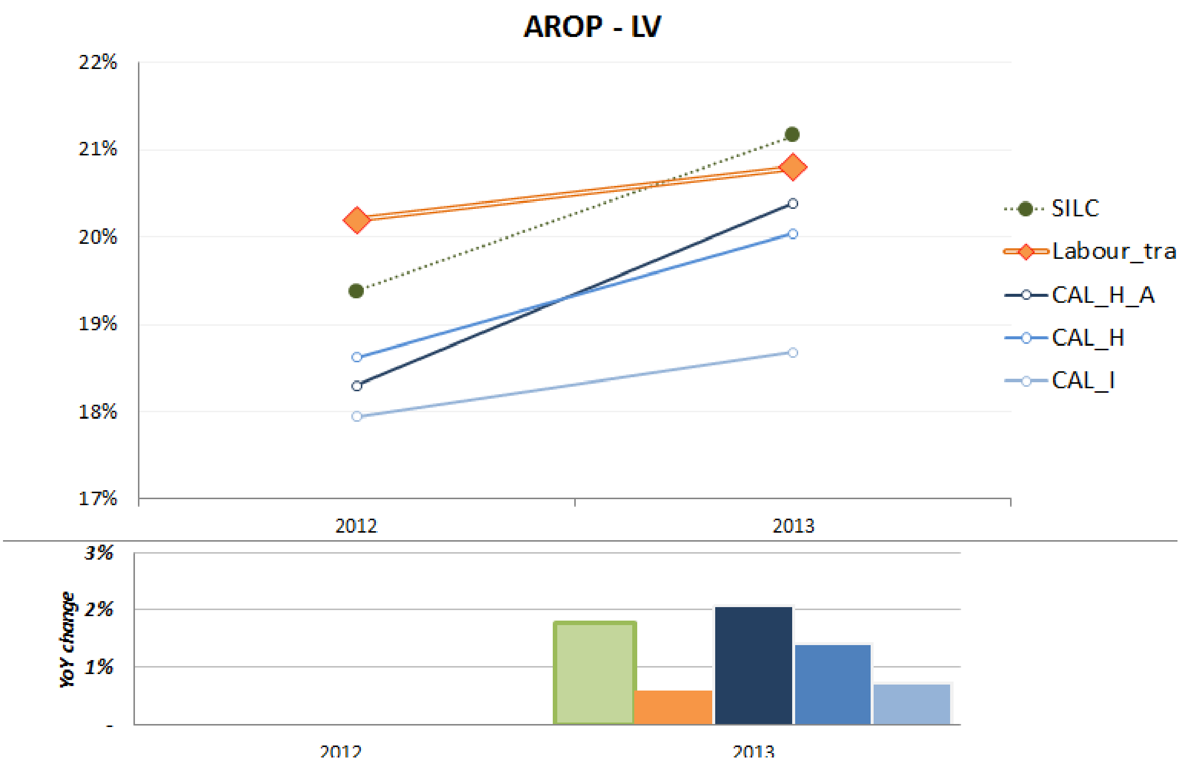


Figure 3 (contd): Flash AROP estimates: 2012-2013 levels and year-on-year change



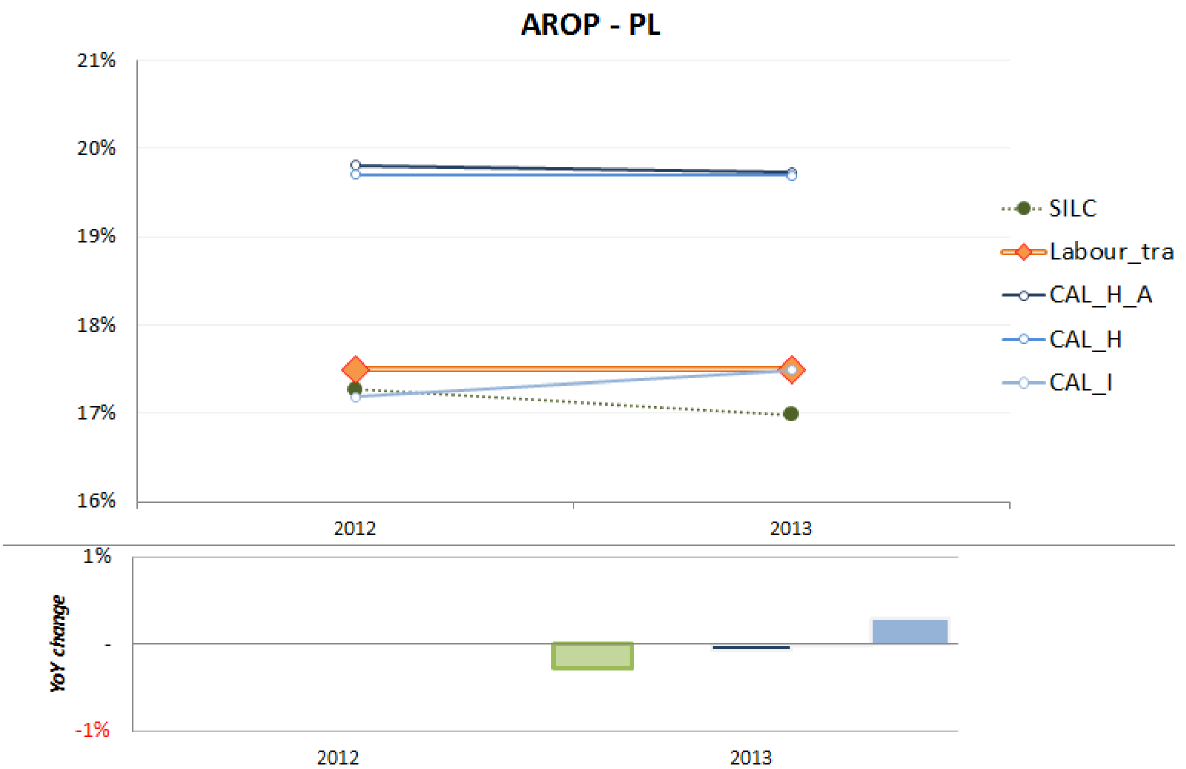
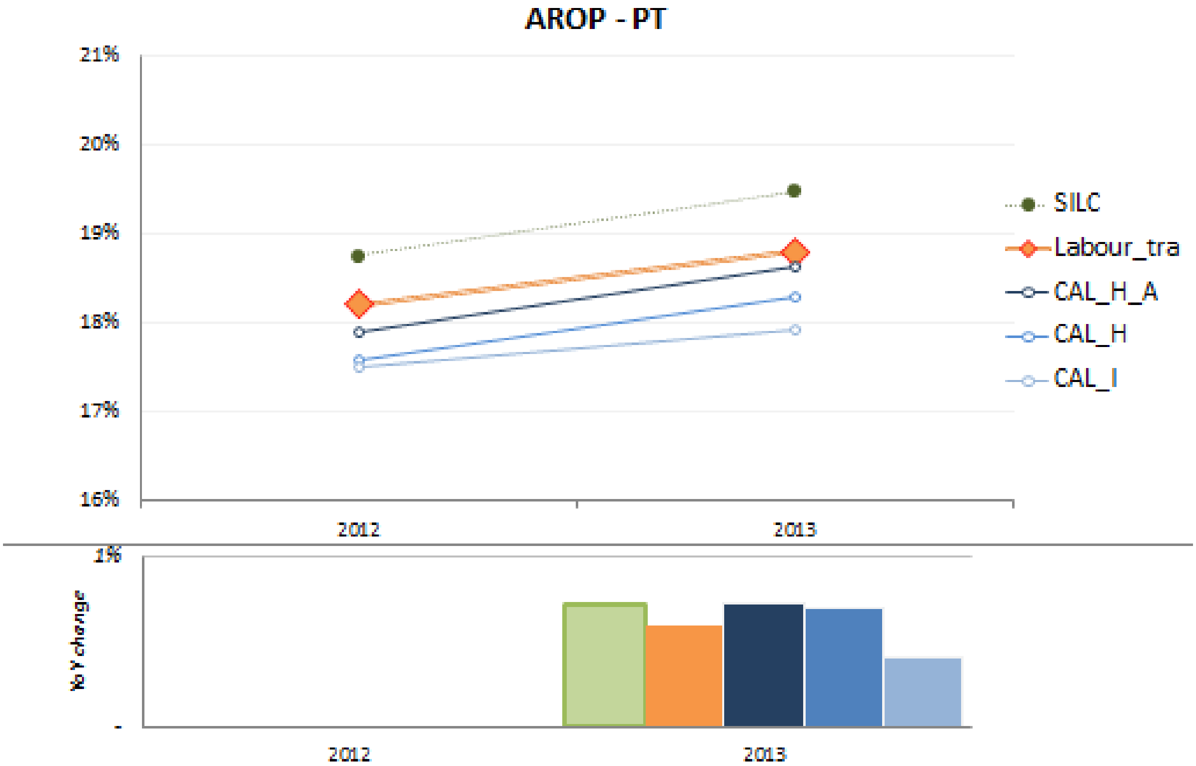


Figure 3 (contd): Flash AROP estimates: 2012-2013 levels and year-on-year change



Note: The methods considered are: (1) Labour transitions: Labour_tra; (2) Calibration at household level adjusted: Cal_H_A; (3) Calibration at household level: Cal_H; (4) Calibration at individual level