The Euro Area Inflation Flash Estimate Procedure

Pedro Martins Ferreira

European Commission, Eurostat, Luxembourg

pedro-jorge.martins-ferreira@ec.europa.eu

Luca Gramaglia

European Commission, Eurostat, Luxembourg

luca.gramaglia@ec.europa.eu

Abstract

The flash estimates of the euro area Harmonised Indices of Consumer Prices (HICP) are statistical indicators that give an early insight into monthly inflation in the euro area. From a methodological point of view, the flash estimates are special forecasts. Literature on forecasting inflation covers very diverging methods, from purely statistical to more econometrical ones. Most of these methods, however, focus on long-term forecast horizons (one to two years) and often use complex statistical or econometric analysis, which is both time and resource consuming. The procedure of the euro area HICP flash estimates differs from these more traditional forecasting designs in two particular aspects, namely that only one-step ahead forecasts are computed and that the compilation is run in a tight timeframe. These elements were decisive factors when choosing the forecasting procedures and setting up the system. In the production of the flash estimates, Eurostat is using preliminary price index data from the euro area national statistical offices supplemented with an econometric one-step ahead forecasting model for those countries not able to send the data. The missing data estimation model combines timely energy price information with HICP back series and the preliminary data of those euro area countries available for the reference month. The preliminary data and the estimates of missing countries are aggregated to the euro area level to produce the nowcasts for the all-items inflation as well as for four main components 'food', 'non-energy industrial goods', 'energy' and 'services'. This paper describes the methodology developed by Eurostat for producing a consistent set of Harmonised Indices of Consumer Prices flash estimates for the euro area.

Keywords: flash estimate, inflation, HICP, forecast, nowcast

1. Introduction

The literature on forecasting inflation is vast and covers several methods, from purely statistical to more econometrical approaches. Since most of the methods focus on long-term forecasting (one to two years), they rely on heavy statistical and/or econometric analysis, which need time and resources to attain maximum forecasting accuracy. To produce the euro area inflation flash estimates, timeliness is crucial and this narrows the choice of forecasting method. Time is such an important constraint that the estimation procedure has to be automated in order to meet the narrow time window available (few hours) for producing the estimates, which in turn, may lead to the use of sub-optimal forecasting techniques. However, only a 1-step ahead forecast needs to be made to produce the flash estimates. More importantly, actual data on energy prices and preliminary data from a fair share of countries, which prove to be significantly better estimates than model-based ones, can be used to improve the accuracy of the forecasts without compromising timeliness. This paper presents a new methodological framework for producing a consistent set of flash estimates for the euro area inflation and its main components, as developed at Eurostat. The procedure combines preliminary data, energy prices and forecasting techniques to produce flash estimates for the inflation main components 'food', 'non-energy industrial goods', 'energy' and 'services', as well as for the headline inflation (all-items). The flash estimate procedure is in place since September 2012 and proves to be a sucessful tool in terms of timeliness and accuracy.

In the following, a terminological distinction is made between the terms "flash estimate", considered as the overall product released as result of the estimation procedure; "nowcast", as the pure result of the estimation procedure and "early estimate", intended as the input (flash estimate) provided by countries.

Section 2 defines the terms of the problem; section 3 and 4 describe, respectively, the availability of input data and the theoretical framework. Section 5 illustrates the obtained results which are complemented in section 6 by summary conclusions. Eventually, section 7 outlines the future work in this area.

2. Problem statement

Eurostat has been publishing the euro area Harmonised Index of Consumer Prices (HICP) flash estimate (all-items) every month since October 2001. Although the flash estimate is one of the most followed European economic indicators, its relevance has been somewhat limited to the users as only the HICP all-items was released and no flash indication on its main components were made available (Pasanen and Ferreira [11]).

The HICP all-items flash estimate was based on euro area countries flash estimates (national HICP all-items) complemented by nowcasts for the missing countries. Technically speaking, the inflation flash estimate was a univariate nowcast of the HICP all-items.

The need to produce the nowcasts of the main components of inflation moved the target of the estimation procedure from one single indicator (HICP all-items – univariate approach) to a consistent set of indicators (multivariate approach). Estimating a consistent set of nowcasts instead of a single one is challenging. Inflation main components, i.e., 'food', 'non-energy industrial goods' (igoodsxe), 'energy' and 'services' (serv) are, in principle, more difficult to forecast. In addition, the outcome of the nowcasting procedure has to be internally consistent, i.e., the aggregation of the nowcasts for the main components has to equal the nowcast of the all-items index. Internal consistency is mainly important for the clarity of the published data, even if it results in the loss of some accuracy on the single nowcasts. These new challenges stressed the fact that the methodology in place in Eurostat ("existing" methodology) was not adequate to calculate a fully consistent set of nowcasts and, therefore, a "new" nowcasting system had to be developed. The new system combines early information from the euro area Member States with 1-step ahead forecasts and timely price data on specific energy products, as did the existing all-items nowcasting system.

Nevertheless, it is extended to the main components and ensures the overall coherence of the nowcasts of the set of indicators (HICP all-items and components).

The research work has been guided by two questions: "how useful is preliminary data to nowcast inflation and what is the best way to use it" and "what is the best way to forecast missing data and combine it with preliminary data to produce a consistent set of nowcasts". To answer the first question, preliminary data were analysed in detail to test their predictive power. The second question is more complex: an important issue in forecasting aggregated data is the decision whether to forecast the aggregate series (direct forecast) or to aggregate the forecasts of its sub-components (indirect forecast). Forecasters have given a lot of attention to this question.

Nicholai Benalal [10] conclude that the direct forecast approach provides decidedly better results than the indirect forecast approach for timespans of 12 and 18 months but for shorter horizons, the results are mixed. These results are in line with Hubrich [9], who shows that for many of these methods, when applied to HICP, "forecasting aggregate euro area year-on-year inflation directly results in higher forecast accuracy for the medium-term forecast horizons of 12 months".

In contrast, Bermingham and D'Agostino [1] show that the direct approach often has the least satisfactory performance. They endorse clearly the indirect approach and explain their contradictory conclusions, as compared to the earlier studies, with the short span of data used in those previous works.

It is generally accepted, however, that the choice between the direct and indirect approaches is essentially an empirical issue. Such an empirical exercise was carried out at Eurostat, where forecasts were made for a set of countries that did not provide preliminary data on time for the inflation flash estimate and not for the euro area as a whole. The results did not clearly favour one approach over the other. For 1-step ahead forecasts, the two approaches were rather similar and their performance depended on the main component under analysis, the time period and the number of countries to be forecasted. However, there was a slight tendency for the direct approach to perform better and thus, the direct approach is, in general, preferred. Moreover, with the direct approach there are fewer models to take into account, which fits better in a production environment. The flash estimates are produced under a strict time constraint, and the forecasts are just meant to complement the preliminary data, only a small part of which is missing.

The complexity of the forecasting model presents another potential problem. Pursuing forecasting accuracy at all cost, without taking into account the resources needed to do that, is not the wisest thing to do. This is even more relevant if the forecasts have to be integrated in a production system that is usually under pressure to respect the pre-announced release calendar. This is the case of the new nowcasting system: there is a very short period between receiving the data and calculating a consistent set of nowcasts. Due to the time constraints to produce the flash estimates, every model used has to be 'simple' enough and not 'resource intensive'. It needs to produce the results quickly and, preferably, to be fully automated.

Fortunately, only 1-step ahead forecasts need to be made, and Buelens [2] refers to a number of authors who have shown that for short forecast horizons simple models have often performed as well as more complex approaches.

As a consequence, the new nowcasting system needs to combine preliminary data with forecasted data, to be fully automated (which implies the existence of an automatic model selection procedure) to be applied in general to aggregated data and to make use of simple forecasting methods. In addition, the system needs to be implemented in 'modules' for a proper maintenance and for future improvements. For example, if the specific part of the procedure is not 'hard coded' but an independent and autonomous module, it allows the forecaster to research better alternative approaches and if one is found, to replace the old module in the running system.

3. Available data

Three different types of data are available for the nowcast: HICP historical data, energy prices and HICP preliminary data. HICP historical data is the collection of all monthly indices, per country and per 4-digit Classification of Individual Consumption According to Purpose (COICOP), from January 1996 until period $t-1$. The 4-digit COICOP are the building blocks for any type of aggregate including the HICP all-items and the four main components.

HICP preliminary data is a set of indices for the reference period t. These data are transmitted by the Member States at the end of the month and they are early HICP estimates. These data can be full 4-digit COICOP data sets, aggregated indices for the main components, or even one single figure for the all-items.

As for energy prices, each Monday, Member States communicate to the European Commission consumer prices of petroleum products with and without duties and taxes. Prices are available for 'euro-super 95', 'automotive gas oil' and 'heating gas oil' from January 1995 onwards. Prices are collected weekly but there are occasional missing values in the series. Furthermore, the last week of the month is usually not available on time to produce the flash estimates. In order to have complete time series, missing values are either estimated by interpolation (for past values) or, for the last week figures, forecasted to complete the month. When the series are complete, monthly values are calculated as the average of each week of the month. Monthly series are then available until period t.

4. Theoretical framework

4.1. Assumptions

The theoretical framework presented in this paper stands on the assumption that preliminary data sent by the Member States are significantly more accurate than any model-based forecast. Although this assumption seems reasonable *a priori*, it needs to be tested. This implies that the flash estimate procedure is to be more of a 'missing data estimation' procedure than a 'forecasting' procedure. Another assumption is that the accuracy of the all-items forecast is generally higher than the accuracy of the main components. Thus, to produce the most accurate headline inflation nowcast and a consistent set of flash estimates the nowcasts of the main components are calibrated at the end of the procedure in order to aggregate to the nowcast of the all-item. Changing the forecasts so that they meet the consistency constraint could potentially impact their accuracy. However, if the nowcasts of the main components are unbiased and sufficiently accurate, the impact of the final calibration should be residual.

Pursuant to our analysis on the choice between direct and indirect forecast, the former is preferred. As a result, if for example 'food', an aggregation of fifteen 4-digit level COICOP indices, needs to be forecasted for five countries, the approach would be to aggregate the $15 \times 5 = 75$ indices first and make the forecast for that aggregate as a whole.

4.2. Class of models

At the time the flash estimates have to be calculated, there is additional information available, as was mentioned in section 3, that can substantially improve the accuracy of the forecasts. The availability of additional information leads us to search for a multivariate model that captures the relationship between the set of countries that needs to be forecasted, Y_t and all present data available that might be useful as a predictor, Xt. Chatfield [3] presents a variety of multivariate time series models and the respective forecasting methods.

The choice of a forecasting model is restricted to single-equation multivariate models, in particular, transfer function models. This class of models is very general and it includes auto-distributed lagged (ADL) models, ARIMA, regression with ARIMA errors or even linear regression as special cases. This choice is a compromise between having a simple model, which tends to be easier to automate and fits better into a heavily time constrained statistical production environment, and a more accurate model whose complexity would be too resource intensive and / or too time consuming. Since more complex models are not necessarily more accurate for the short term forecasts and since only 1-step ahead forecasts are needed, our choice is to have relatively simple forecasting models.

The general model is given by

$$Y_t = \delta^{-1}(B)\, w(B)\, X_t + n_t \tag{4.1}$$

where B is the lag operator, $\delta^{-1}(B)\, w(B)\, X_t$ is the linear system part of the model and n_t is the noise. The noise component is allowed to have a seasonal ARIMA structure $(ARIMA\,(p,d,q)\,(P,D,Q)\,[s])$:

$$\Phi(B)_s\, \Phi(B^s)\, \Delta_s^D \Delta^d n_t = \theta(B)_s\, \Theta(B^s)\, a_t \tag{4.2}$$

where a_t denotes a white noise process. Combining both equations, we have the chosen class of models

$$Y_t = \delta^{-1}(B)\, w(B)\, X_t + \frac{\phi_n^{-1}(B)\, \theta_n(B)}{\Phi(B)_s\, \Phi(B^s)\, \Delta_s^D \Delta^d} a_t \tag{4.3}$$

Depending on $\delta^{-1}(B)\, w(B)$, $\phi_n^{-1}(B)\, \theta_n(B)$ and $\Phi(B)_s\, \Phi(B^s)\, \Delta_s^D \Delta^d$, the model could be seen as a linear regression, a pure seasonal ARIMA model, a linear regression with seasonal ARIMA errors or an ADL model. For the sake of simplicity we decided for the time being to restrict the pool of models to the ones where $\delta^{-1}(B)\, w(B)$ is of order 0 in B.

To forecast a function transfer, one needs to forecast the linear system part and the noise part of the model independently and then add them up at the end. Since only 1-step ahead forecasts are needed and since there will always be an extra observation of X_t available (either preliminary data from the countries, energy prices or both), there is no need to forecast X_t in order to produce a forecast for Y_t.

4.3. Automatic model selection procedure

4.3.1. Searching for the best model

To incorporate any structural break in the model as soon as possible, an automatic model selection procedure was built. Each month, it checks regularly for the 'best' model to be used in the next flash estimate production. In the context of the flash estimate calculation, 'best model' is defined as **the model that produces unbiased estimates and has the lowest 1-step ahead out-of-sample mean square error (MSE)**.

The MSE is defined as

$$MSE = \sum_{t=N-m+2}^{N} \frac{e_t^2}{m} \tag{4.4}$$

with $e_t = y_t - \hat{y}_t$, i.e., the difference between the y_t and the forecast $\hat{y}_t$. The MSE implies an underlying quadratic loss function which tend to be dominated by outliers (the final sum will basically be the result of particularly large e_t values) but in this context this is a desirable feature. From a statistical production point of view, it is preferable to have a certain amount of nowcasting error each month that should be published than to have slightly smaller nowcasting errors that from time to time result in flash estimates that are totally off the final estimate.

The overall forecast bias (if any) will be measured by the mean error (ME):

$$ME = \sum_{t=N-m+2}^{N} \frac{e_t}{m} \tag{4.5}$$

The model selection procedure developed is a computer-intensive procedure that scans different parameter combinations from the transfer function class of models described above and performs a pseudo-out-of-sample forecast over the last 24 months for each parameter combination. At the end of the scan, both the ME and MSE are recorded and sorted by their MSE. A simple t-test is used to check if ME is significantly different from zero, i.e. to check for biases in the forecasts. If a model has an ME significantly different from zero, it is considered 'biased' and is discarded. The best model is the first model from the sorted list that is not rejected by the t-test, denoted from now on as 'candidate model'.

4.3.2. Model stability

In order to maintain a certain model stability over time and not change the forecasting models every month, the candidate model's pseudo-out-of-sample forecasting errors are compared with the pseudo-out-of-sample forecasts from the model used in the last flash estimate production (denoted as 'reference model'). This assessment is made using the Diebold-Mariano statistic (DM) whose 'use and abuse' is not free from criticism. As stated by Diebold [6], the DM statistic compares forecasts rather than models, and it is quite common to misuse this statistic when choosing forecasting models. Costantini and Kunst [4] point out that the use of DM statistic for model selection tends to favour the null model, usually a simple benchmark model. That is, a more sophisticated procedure is recommended only if it is 'significantly' better than the benchmark and not simply if it has better accuracy statistics. This creates a kind of bias towards choosing simpler but not necessarily more accurate models.

However, while we are aware of the main criticisms on the use of the DM statistic, the flash estimates procedure is a specific case regarding model selection. The models used for the production of the first set of flash estimates were not 'simple benchmark' models, but the result of a thorough analysis on the choice of models to be used for the regular production of flash estimates. The intention is to use the models defined during the research phase as much as possible and change them only if there is strong evidence that an alternative set of parameters would have produced significantly better 1-step ahead out-of-sample forecasts. It is important to check a candidate model so as to be able to incorporate any potential structural break as soon as possible. The use of the automatic model selection procedure does not imply that a deeper and non-automatic model analysis could not be carried out in the future. In fact, such an analysis should be performed with some frequency (e.g., annually).

The asymptotic test proposed by Diebold and Mariano [7] involves testing the null hypothesis of equal forecast accuracy. Two competing forecasts are made and have forecast errors e_{1t} and e_{2t} $(t = 1, ..., n)$. If $g(e_t)$ is a loss function, then the loss differential series d_t can be constructed. The desired null can be written has $H_0 : E\left(d_t\right) = 0$.

The DM statistic, corrected for its finite sample oversizing is given by

$$DM = \frac{\bar{d}}{\sqrt{v\hat{a}r_m\left(\bar{d}\right)}} \tag{4.6}$$

where

$$\bar{d} = n^{-1}\sum_{t=1}^{n} d_t$$
$$v\hat{a}r\left(\bar{d}\right) = \left[n + 1 - 2h + n^{-1}h\left(h-1\right)\right]^{-1}\left[\hat{\gamma}_0 + 2\sum_{k=1}^{h-1}\hat{\gamma}_k\right]$$
$$\hat{\gamma}_k = n^{-1}\sum_{t=k+1}^{n}\left(d_t - \bar{d}\right)\left(d_{t-k} - \bar{d}\right)$$

Since the flash estimates only use 1-step ahead forecasts, the variance of the loss differential sample mean can be simplified to

$$v\hat{a}r\left(\bar{d}\right) = \left(n-1\right)^{-1}\hat{\gamma}_0 \tag{4.7}$$

and the corrected DM statistics follows a t_{n-1} distribution.

To test if a candidate model should replace the model that was used in the last production of the flash estimates, the following algorithm was defined:

Algorithm: find the best model to be used in the next production round of the flash estimate.

1. For each parameter model combination, perform a pseudo-out-of-sample forecast over the last 24 months ($n = 24$). Compute and record ME and MSE;
2. Sort the results by MSE in ascending order;
3. Test if the ME of the model of the ordered list is significantly different from zero using the t-test. If the ME is significantly different from zero, discard that model and test the next model from the sorted list. Repeat while t-test is rejected;
4. Perform the pseudo-out-of-sample for the last 24 months using both the reference model and the candidate model and save the two vectors of forecasting errors;
5. Compare the forecast errors vectors using the DM statistic. If the MSE of the forecast errors of the candidate model is significantly lower than that of the reference model, replace the reference model by the candidate model.

The following notation is a convenient way to present the model specifications:

$$\{X_t\}\,(p, d, q)\,(P, D, Q)\,[\text{period}]$$

where X_t is the set of independent variables, $(p, d, q)\,(P, D, Q)$ are the parameters of an ARIMA model and $[\text{period}] = \{6, 12\}$ is the period considered for the seasonal part of the ARIMA model.

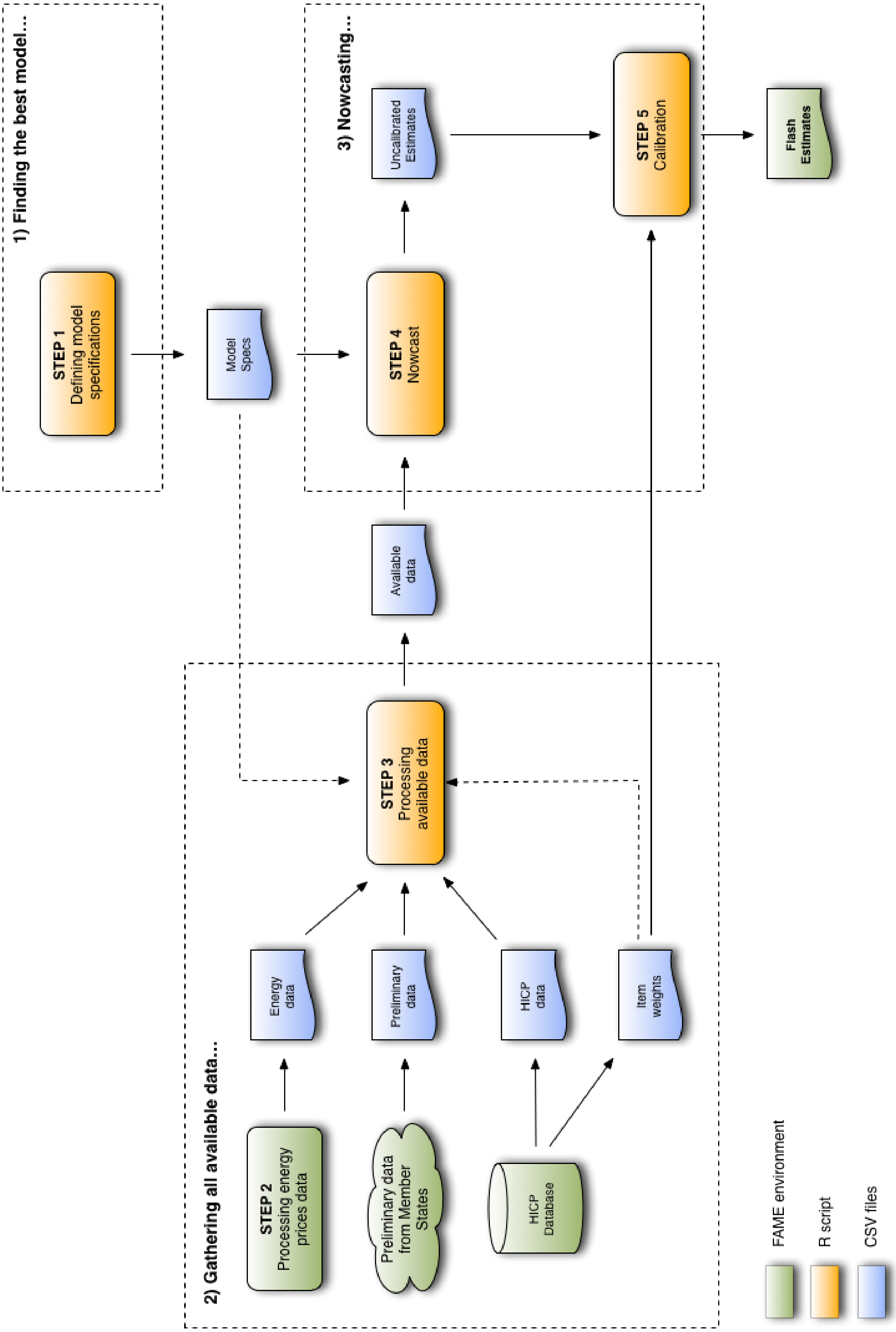


Figure 4.1: Flash estimate procedure. Five main steps are needed to produce the flash estimates; they can be grouped in three types of procedures: automated model selection, data processing and nowcasting.

4.4. *Flash Estimate procedure*

Figure 4.4 show graphically the main steps of the flash estimate (FE) procedure:

- Step 1 is run before each production round of the FE. In this step the automatic model selection procedure described in 4.3 sets the model specifications that are used in the next production round;
- Step 2 forecasts prices for petrol, diesel and heat energy for the last week of the reference month. These data are taken from the Weekly Oil Bulletin of the Directorate-General for Energy (DG ENER). Since the last week of the reference month is usually not available at the time of the flash estimates production, it needs to be estimated. Monthly averages are then computed;
- Step 3 gathers all data necessary for the computation. Euro area Member States are allocated into two groups: group A for those who have provided preliminary data for the reference month, and group N for those who did not provide any preliminary data. For each aggregate that is going to be nowcasted, a single country-aggregate price index is computed for each one of the two groups: $A_t, t = 1, ..., T$ and $N_t, t = 1, ..., T-1$. A_t has one extra observation which is the preliminary data received from the Member States;
- Step 4 produces the 1-step ahead forecast using the data processed in Step 3 and using the model specifications found in Step 1. The missing data forecast, $\hat{N}_t$ is afterwards aggregated with the preliminary data, A_t, giving the (non-calibrated) nowcast for that COICOP aggregate;
- Step 5 ensures the balancing of the estimated nowcasts of the four main components 'food', 'non-energy industrial goods', 'energy' and 'services', i.e. it ensures that the nowcasts of the main components aggregate correctly to the all-items nowcast, in order to produce a consistent set of flash estimates.

It was always assumed, since the beginning of this project, that the methodology would not be definitive but rather the 'best-so-far' methodology. However, it was also evident that after the first flash estimates were made public, Eurostat would continuously produce them each month according to the agreed release calendars. Therefore, the nowcasting system in production was built in a modular way, which allows Eurostat to make changes into a clone of any module, test those changes and replace the existing module by the improved one in case it significantly improves the quality, without stopping the production of flash estimates.

5. Results

This section shows the results of nowcasting euro area inflation for the all-items and main components using the new procedure.

Aggregate	period	coverage	MAPE	max	RMSE	Rel. RMSE
all-items	**P1**	**95%**	**0.051**	**0.117**	**0.034**	**0.073**
	P2	**98%**	**0.053**	**0.089**	**0.028**	**0.060**
food	P2	85%	0.053	0.189	0.066	0.216
igoodsxe	P2	85%	0.110	0.305	0.106	0.285
energy	P2	85%	0.136	0.358	0.148	0.086
serv	P2	85%	0.100	0.207	0.070	0.149

Table 1: Preliminary data accuracy. P1: from Jan-2011 to Feb-2012; P2: from Mar-2012 to Jan-2013; MAPE: weighted mean absolute percentage error; max: maximum absolute difference in percentage error; RMSE: root mean square error; Rel. RMSE: root mean square error relative to a benchmark model.

5.1. Preliminary data accuracy

As mentioned before, the basic assumption of this nowcasting procedure is that preliminary data sent by Member States are better estimates than any model-based forecast. Although this assumption seems reasonable, it is nevertheless important to check if the assumption holds. To check this assumption, preliminary data sent by Member States per special aggregate are compared with the respective values published in the HICP release (usually, two weeks after the flash estimates are published). The comparison time period presented in this paper is from January 2011 to January 2013.

Table 1 shows the accuracy of preliminary data for the all-items and main components measured by the weighted Mean Absolute Percentage Deviation (MAPE) for indices, the maximum absolute difference in percentage points for indices and the weighted root mean square error (RMSE) for the annual rates[1] and the relative RMSE using as a benchmark a simple random walk with drift[2]. The comparison can be split into two distinct time periods: from January 2011 to February 2012 (P1), where less preliminary data was available, and March 2012 to January 2013 (P2), where a more complete set of preliminary data was available. The accuracy of the main components is analysed for period P2 only since the amount of preliminary data available before period P2 is insufficient.

Looking at Table 1 we can conclude that **RMSE for the all-items are significantly smaller**, both for period P1 and P2, **than 1-step ahead forecasts described in the literature**. It is usual to see in the literature RMSE for the all-items around 0.2 - 0.3 for 1-step ahead forecasts (see for example Hubrich [9], Duarte and Rua [8]). This means that preliminary data sent by Member States has a 7 to 10 times lower RMSE, which corroborates our assumption that preliminary data is a better forecast than any model-based forecast. Unsurprisingly, the biggest gains in accuracy for using preliminary data instead of a model-based forecast are expected to be for the all-items and energy, where the relative RMSE is

[1]Weighted MAPE is the weighted average of each country's absolute percentage error using as weights the country's expenditure in percentage of the total expenditure of the aggregate. Weighted RMSE is computed by comparing the annual rates for the aggregation of countries who provided preliminary data with the annual rate of the respective final HICP data for that aggregate. As an example, if only 6 countries had provided data for a specific month, then the comparison is made between the preliminary annual rate for that 6-country-aggregate with the 'final' annual rate for that 6-country-aggregate.

[2]The strong seasonal pattern for non-energy industrial goods was taken into account: while in general the benchmark model was $y_t = y_{t-1} + d + \epsilon_t$, for non-energy industrial goods the benchmark was $y_t = y_{t-12} + d + \epsilon_t$.

less than 10% of the RMSE of the benchmark model.

It is also possible to see that the RMSE for the main components is higher than for the all-items. This supports another assumption made before which is that the accuracy for the all-items preliminary data is higher than for the main components and, therefore, the former can be used as a benchmark to calibrate the aggregation of the four main components.

It was also stated that, in principle, the final calibration procedure to ensure a consistent set of nowcasts was expected to have a minor impact on the non-calibrated estimates.

The final calibration procedure used is a simple **proportional allocation**, where the difference between the nowcast of the all-items and the aggregation of the main components nowcasts are proportionally distributed over the main components. The calibration factor is computed by:

$$\phi = \frac{I_t^{\text{All items}}}{\text{aggregation}\left(I_t^{\text{Main component}}\right)} \tag{5.1}$$

The calibrated and final nowcasts are calculated by multiplying each non-calibrated nowcasts with the same calibration factor ϕ. As such, the calibration factor can be interpreted as the percentage change that each non-calibrated nowcasts have to have in order to make them consistent with the all-items. For example, $\phi = 1.01$ means that each non-calibrated nowcasts have to increase 1% in their value in order to make the set of nowcasts consistent.

Figure 5.1 shows the density plot of the final calibration for the period March 2012 to February 2013, where we can see that in general the calibration factor is close to 1. The average calibration factor for this period is 0.9998 which means that, in average, each main component nowcast had to be lowered 0.02% in average in order to make them consistent with the all-items nowcast. This result corroborates the assumption that the final calibration that makes the set of nowcasts consistent doesn't have a significant impact on the accuracy of the nowcasts.

5.2. Nowcasting energy

Among all main components, 'energy' is the one that needs more attention, since it is the most volatile component and, consequently, the most difficult to forecast. Figure 5.2 plots the Δ^{12} log of the euro area HICP indices for the all-items and main components, where it is easy to see that the variance for energy is much higher than for any other index.

The main component 'energy' is composed of two sub-aggregates:

- **elgas**: accounts for around 48% of the 'energy' aggregate and is the aggregation of 'electricity', 'gas', 'solid fuels' and 'heat energy' (COICOP 0451, 0452, 0454 and 0455);

- **fuels**: accounts for around 52% of the 'energy' aggregate and is the aggregation of 'liquid fuels' and 'fuels and lubricants for personal transport equipment (COICOP 0453 and 0722).

Energy prices collected by DG ENER are very useful to forecast the sub-aggregate 'fuels'. In particular, the price of 'heating gas oil' is strongly correlated with the HICP 'liquid fuels' (COICOP 0453) while the HICP 'fuels and lubricants for personal transport equipment' is strongly correlated with a linear combination of the energy prices for 'euro-super 95',

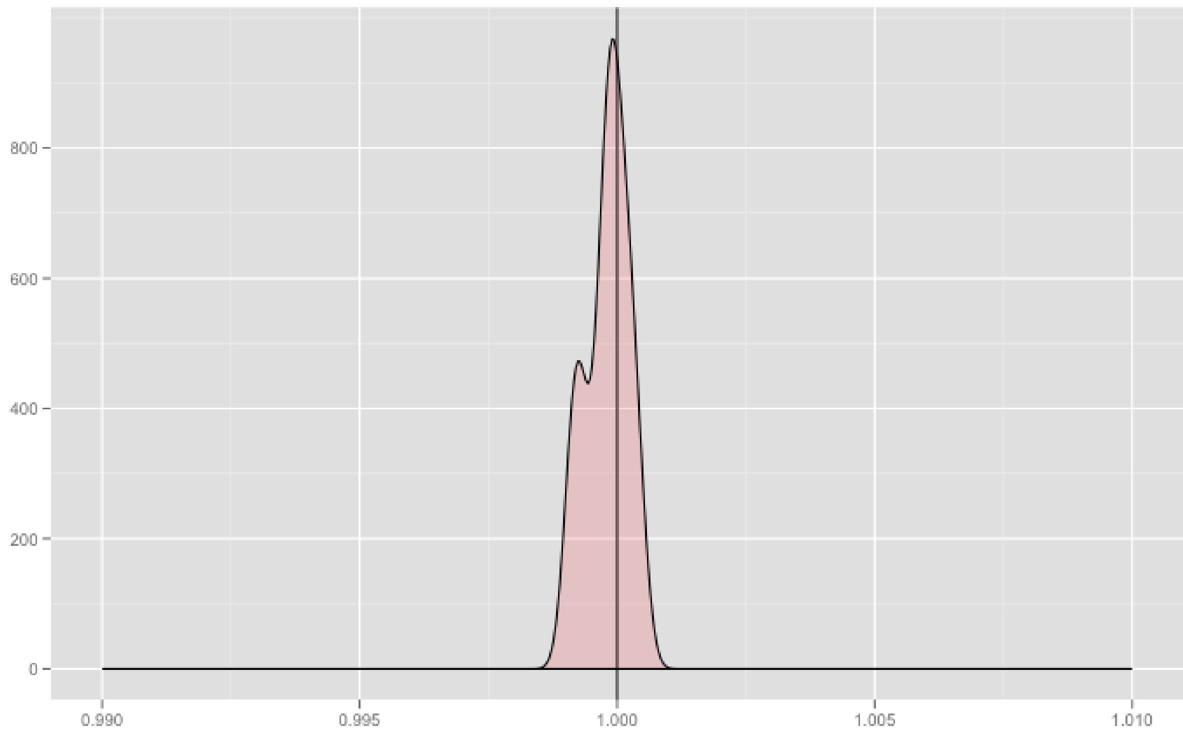


Figure 5.1: Calibration factor density plot. The calibration factor ϕ is the scalar that needs to be multiplied with the non-calibrated main components nowcasts to make them aggregate to the all-items nowcast. The closer ϕ is to 1, the less the main components nowcasts need to be changed to meet the consistency constraint.

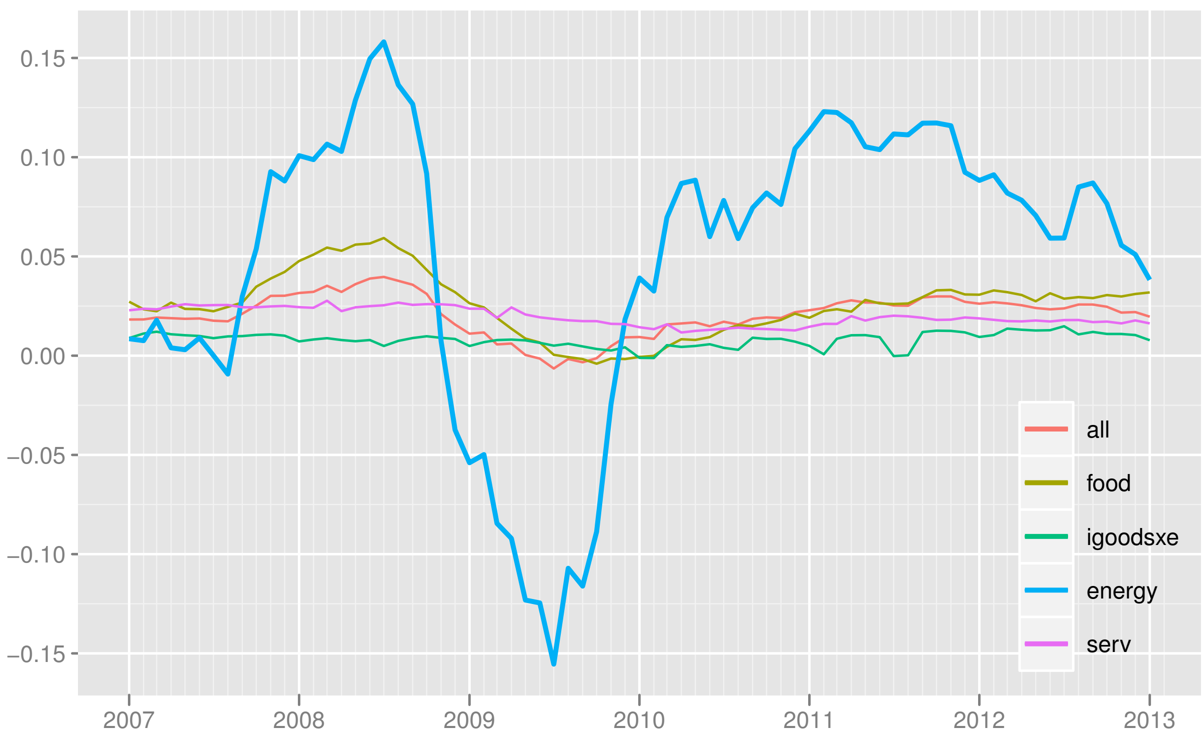
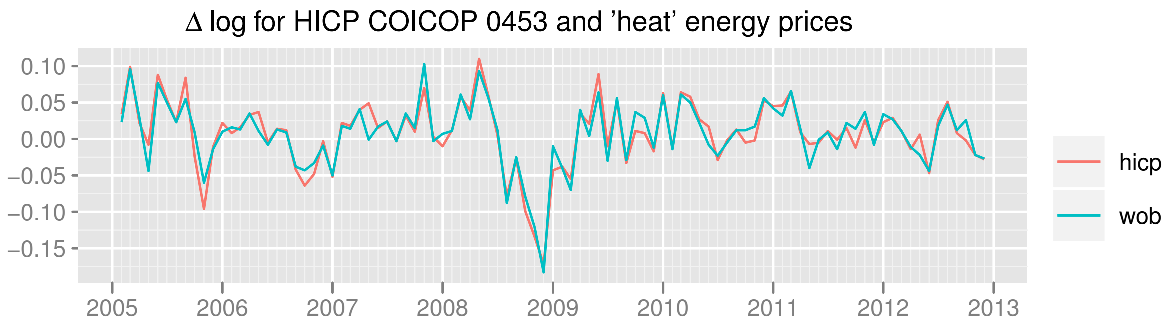


Figure 5.2: Δ^{12} log for the euro area HICP for the all-items and main components.



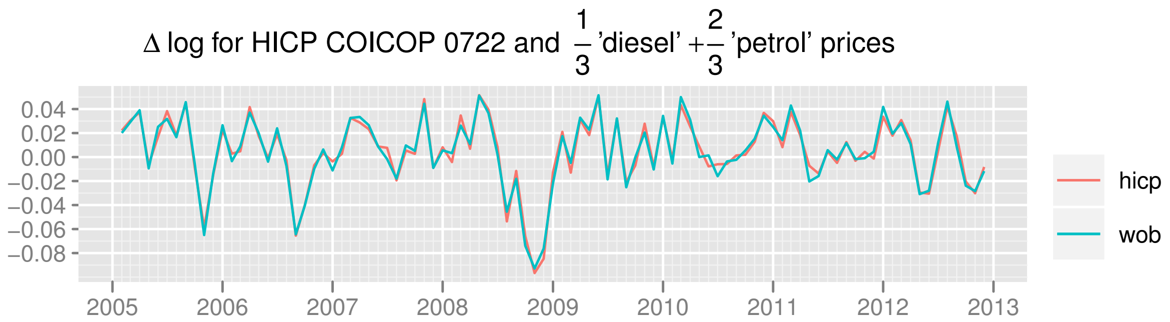


Figure 5.3: HICP versus energy prices. Since COICOP 0722 is essentially a mix of 'diesel' and 'petrol', a linear combination of these two energy products was calculated as $C_t = \alpha \cdot \text{diesel} + (1-\alpha) \cdot \text{petrol}$, where α was chosen in order to maximize the correlation between $\Delta \log HICP$ and $\Delta \log C_t$. The value of α that maximizes the correlation is around $\alpha \approx 0.333$.

'automotive gas oil'. Figure 5.3, which plots the Δ log for the HICP and DG ENER data, shows how strongly correlated these two data sets are.

This strong correlation led us to choose a different approach to forecast energy. While the direct approach is always preferred, 'energy' is forecasted indirectly, i.e., 'elgas' and 'fuels' are forecasted separately and then aggregated into 'energy'.

5.3. Nowcast accuracy

The nowcast accuracy is evaluated for the time period March 2012 to February 2013. Before starting the evaluation, a pseudo-out-of-sample forecast for the expected N_t (that is, for the aggregate of countries that usually don't provide preliminary data on time of the nowcasts calculation) was conducted in order to check which model specifications produced the lowest MSE for a period of two years. This test was based on data covering a period of 24 months, from March 2010 to February 2012. The resulting models that are used to assess the nowcasts accuracy are the following:

- all-items: $\{P_t\}\,(0,1,0)\,(0,0,0)$; $RMSE = 0.182$
- food: $\{D_t, P_t, H_t\}\,(2,1,2)\,(1,0,1)$; $RMSE = 0.431$
- igoodsxe: $\{A_t, D_t, P_t, H_t\}\,(1,1,2)\,(0,1,0)$; $RMSE = 0.559$
- elgas: $\{A_t\}\,(0,1,1)\,(0,0,0)$; $RMSE = 1.057$

- fuels: $\{A_t, P_t, H_t\}\,(1,1,0)\,(1,0,1)$; $RMSE = 1.115$
- serv: $\{A_t, D_t, H_t\}\,(2,1,1)\,(0,0,0)$; $RMSE = 0.192$

Table 2 shows the nowcasts accuracy for the analysis period. Table 3 is divided in two distinct periods: 1) between October 2012 to January 2013, where more preliminary data for the main components were available and where a calculation error, unfortunately only identified after the September release, was corrected; 2) from March 2012 to September 2012.

Some conclusions can be taken from Table 2 and 3. First, the accuracy has improved from October 2012 onwards, even when September 2012 is not taken into account. Second, the headline inflation (all-items) is the most accurate. Third, the maximum and minimum errors stay within an acceptable range. The accuracy recorded from October 2012 to February 2013, in spite of this time period being still too short, are nevertheless encouraging. Looking to results of the complete analysis period, and taking into consideration the fact that Eurostat has now more preliminary data available, led us to conclude that the nowcasts are of good quality.

6. Conclusions

A new methodology to produce a consistent set of nowcasts for the euro area inflation of 'all-items', 'food', 'non-energy industrial goods', 'energy' and 'services' was developed at Eurostat. Since September 2012, Eurostat has published at the end each reference month or beginning of the following month the five nowcasts for the indices from which the annual rates and monthly rates are derived. The calculation procedure now in place was designed specifically to fit Eurostat's production environment and time and data availability constraints. The nowcasting system needed to be fast, as automated as possible, accurate and flexible enough to allow methodological adaptations when needed. We believe that the procedure presented in this paper fulfils these requirements. The basic assumption that the flash estimates calculation stands on is that preliminary data sent by Member States is the best available information to nowcast euro area inflation. It was shown that this assumption holds and, as a result, combining preliminary data with forecasting data is a valid approach to nowcasting inflation for the euro area. Euro area inflation flash estimates are considered to have an acceptable quality, both in terms of timeliness and accuracy. **From October 2012 onwards[3] Eurostat produced timely and accurate flash estimates. The maximum recorded error was 0.2 percentage points for 'energy', the most volatile component, and no error[4] for the headline inflation (all-item) was recorded so far**.

[3]After the first flash estimate breakdown release in September 2012, a calculation error was discovered that lead the nowcasts for non-energy industrial goods to be underestimated by 0.4 percentage points and services to be overestimated by 0.3 percentage points. The error was promptly corrected and therefore October 2012 starts s period from where the accuracy of the presented methodology can be better assessed.

[4]In this context, 'error' is the difference between the flash estimate and the final annual rate, each rounded to 1 decimal, which is the annual rate precision published by Eurostat. Performances of the flash estimates are compared to the first regular release of HICP for the euro area.

		Mar-12	Apr-12	May-12	Jun-12	Jul-12	Aug-12	Sep-12
all-items	**fe**	**2.64**	**2.58**	**2.45**	**2.38**	**2.44**	**2.62**	**2.66**
	hicp	2.67	2.57	2.43	2.36	2.41	2.61	2.61
	diff.	*-0.03*	*0.01*	*0.02*	*0.02*	*0.03*	*0.01*	*0.05*
food	**fe**	**3.26**	**3.01**	**2.82**	**3.40**	**2.98**	**2.99**	**2.94**
	hicp	3.25	3.10	2.78	3.20	2.91	2.99	2.94
	diff.	*0.01*	*-0.01*	*0.04*	*0.21*	*0.07*	*0.00*	*0.00*
igoodsxe	**fe**	**1.31**	**1.37**	**1.37**	**1.38**	**1.36**	**1.04**	**0.83**
	hicp	1.37	1.32	1.28	1.29	1.50	1.08	1.20
	diff.	*-0.07*	*0.06*	*0.09*	*0.09*	*-0.14*	*-0.04*	*0.20*
energy	**fe**	**8.32**	**7.75**	**7.46**	**6.26**	**5.64**	**8.82**	**9.29**
	hicp	8.54	8.14	7.32	6.10	6.11	8.87	9.09
	diff.	*-0.22*	*-0.39*	*0.14*	*0.16*	*-0.47*	*-0.04*	*0.20*
serv	**fe**	**1.73**	**1.61**	**1.83**	**1.68**	**1.78**	**1.80**	**2.03**
	hicp	1.76	1.74	1.79	1.74	1.81	1.81	1.70
	diff.	*-0.03*	*-0.13*	*0.04*	*-0.05*	*-0.03*	*-0.01*	*0.33*

		Oct-12	Nov-12	Dec-12	Jan-13	Feb-13
all-items	**fe**	**2.51**	**2.25**	**2.23**	**2.02**	**1.84**
	hicp	2.49	2.19	2.22	1.98	1.84
	diff.	*0.02*	*0.05*	*0.01*	*0.04*	*0.00*
food	**fe**	**3.18**	**3.04**	**3.15**	**3.23**	**2.73**
	hicp	3.10	3.03	3.16	3.24	2.72
	diff.	*0.08*	*0.01*	*-0.01*	*-0.01*	*0.01*
igoodsxe	**fe**	**1.14**	**1.15**	**1.09**	**0.83**	**0.82**
	hicp	1.10	1.10	1.05	0.78	0.78
	diff.	*0.04*	*0.05*	*0.05*	*0.05*	*0.04*
energy	**fe**	**7.75**	**5.80**	**5.25**	**3.93**	**3.99**
	hicp	7.96	5.72	5.23	3.90	3.92
	diff.	*-0.21*	*0.08*	*0.01*	*0.04*	*0.07*
serv	**fe**	**1.78**	**1.70**	**1.78**	**1.75**	**1.59**
	hicp	1.73	1.64	1.79	1.64	1.54
	diff.	*0.04*	*0.06*	*-0.01*	*0.11*	*0.04*

Table 2: Nowcast results. fe: flash estimate annual rate; hicp: final annual rate; diff.: difference in percentage points between the flash estimate and the final figure. igoodsxe: 'non-energy industrial goods'. Unrounded differences.

	Mar-2012 to Sep-2012			Oct-2012 to Feb-2013		
	RMSE	**min**	**max**	**RMSE**	**min**	**max**
all-items	**0.027**	**-0.03**	**0.05**	**0.030**	**0.00**	**0.05**
food	0.092	-0.10	0.21	0.035	-0.01	0.08
igoodsxe	0.161	-0.37	0.09	0.044	0.04	0.05
energy	0.269	-0.47	0.20	0.107	-0.21	0.08
serv	0.137	-0.13	0.33	0.062	-0.01	0.11

Table 3: RMSE, minimum and maximum errors of the flash estimates. RMSE: root mean square error; min (max): minimum (maximum) difference between the flash estimate and the final figure; igoodsxe: 'non-energy industrial goods'.

7. Future plans

Eurostat has planned further improvements regarding the euro area inflation flash estimates. The first one, while not directly visible to external users, would be internally useful for model performance analysis. A reporting tool focussing on accuracy, both for the preliminary data sent by Member States and for model base forecasts, would give a deeper insight into the overall accuracy of the model. This in turn may give some ideas on how to reduce estimation errors and where improvements or changes should be made. This will also allow us to analyse the possibility of extending the pool of models to include more general transfer function models.

Moreover, a thorough analysis on preliminary data sent by the Member States will be carried out. When this project began, not many preliminary data were available, particularly for special aggregates. As a consequence, it was not possible to make a proper quality assessment, e.g., to identify possible systematic biases in the preliminary data. The amount of preliminary data is increasing each month and now a minimum amount of data is starting to be available for this type of analysis. The development of methodology to correct potential biases or inconsistencies on preliminary data is also foreseen. It would lead to an overall improvement in the accuracy of the euro area flash estimates.

Finally, we plan to extend the flash estimates to two extra figures: 'all-items excluding energy' and 'all-items excluding food and energy'. We will test if nowcasting the all-items excluding energy directly leads to a more accurate estimate than making a simple aggregation of 'food', 'non-energy industrial goods' and 'services'. If the indirect approach does not perform significantly worse than the direct approach, then it would be trivial to produce estimates for two extra components. Otherwise, the methodology would have to be adapted in order to make the existing main components estimates consistent with the two additional special aggregates.

Acknowledgements

This paper presents the methodology behind the euro area HICP flash estimates, which is only one part of the flash estimates production process. The authors would like to express their gratitude to Jarko Pasanen and Roberto Barcellan who gave valuable input through comments, ideas, suggestions, guidance or simply through their pertinent questions that made us think more deeply on the methodology being developed. The other part of the production process which involves, among other things, the publication material, would not have been possible without the participation of Svetoslava Pavlova and Sílvia Santos.

Disclaimer

The content of this paper does not reflect the official opinion of the Eurostat. Responsibility for the information and views expressed therein lies entirely with the authors.

References

[1] Bermingham, C., D'Agostino, A., Ausgust 2011. Understanding and forecasting aggregate and disaggregate price dynamics. ECB Working Paper Series (1365).

[2] Buelens, C., 2012. Inflation forecasting and the crisis: assessing the impact on the performance of different forecasting models and methods. European Economy. Economic Papers (451).

[3] Chatfield, C., Dec. 2000. Time-Series Forecasting. CRC, Chapman Hall.

[4] Costantini, M., Kunst, R. M., Nov. 2011. On the usefulness of the diebold-mariano test in the selection of prediction models. Economics Series 276, Institute for Advanced Studies.
URL `http://ideas.repec.org/p/ihs/ihsesp/276.html`

[5] DG-ENER, 2009. Oil bulletin data collection methodology. Tech. rep., Market Observatory for Energy - European Commission.

[6] Diebold, F. X., September 2012. Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of diebold-mariano tests. Working Paper 18391, National Bureau of Economic Research.
URL `http://www.nber.org/papers/w18391`

[7] Diebold, F. X., Mariano, R. S., July 1995. Comparing predictive accuracy. Journal of Business & Economic Statistics 13 (3), 253–63.
URL `http://ideas.repec.org/a/bes/jnlbes/v13y1995i3p253-63.html`

[8] Duarte, C., Rua, A., 2007. Forecasting inflation through a bottom-up approach: How bottom is bottom? Economic Modelling 24 (6), 941 – 953.
URL `http://www.sciencedirect.com/science/article/pii/S0264999307000351`

[9] Hubrich, K., 2005. Forecasting euro area inflation: Does aggregating forecasts by hicp component improve forecast accuracy? International Journal of Forecasting 21 (1), 119 – 136.
URL `http://www.sciencedirect.com/science/article/pii/S0169207004000536`

[10] Nicholai Benalal, Juan Luis Diaz del Hoyo, B. l. M. R. F. S., July 2004. To aggregate or not to aggregate? euro area inflation forecasting. ECB Working Paper Series (374).

[11] Pasanen, J., Ferreira, P. M., May 2012. Euro area hicp flash estimate breakdown. Meeting of the Group of Experts on Consumer Price Indices, UNECE - United Nations Economic Commision for Europe.
URL `http://www.unece.org/stats/documents/2012.05.cpi.html`