Nowcasting real economic activity in the euro area: Assessing the impact of qualitative surveys

Working Paper Research

by Raïsa Basselier, David de Antonio Liedo and Geert Langenus

December 2017 No 331

NationalBank OF BELGIUM

Eurosystem

Editor
Jan Smets, Governor of the National Bank of Belgium

Statement of purpose:

The purpose of these working papers is to promote the circulation of research results (Research Series) and analytical studies (Documents Series) made within the National Bank of Belgium or presented by external economists in seminars, conferences and conventions organised by the Bank. The aim is therefore to provide a platform for discussion. The opinions expressed are strictly those of the authors and do not necessarily reflect the views of the National Bank of Belgium.

Orders

For orders and information on subscriptions and reductions: National Bank of Belgium,
Documentation - Publications service, boulevard de Berlaimont 14, 1000 Brussels

Tel +32 2 221 20 33 - Fax +32 2 21 30 42

The Working Papers are available on the website of the Bank: http://www.nbb.be

ISSN: 1375-680X (print)
ISSN: 1784-2476 (online)

Abstract

This paper analyses the contribution of survey data, in particular various sentiment indicators, to nowcasts of quarterly euro area GDP. It uses a genuine real-time dataset that is constructed from original press releases in order to transform the actual dataflow into an interpretable flow of news. The latter is defined as the difference between the released values and the prediction of a mixed-frequency dynamic factor model.

Our purpose is twofold. First, we aim to quantify the specific value added for nowcasting GDP from a set of heterogeneous data releases including not only sentiment indicators constructed by Eurostat, Markit, the National Bank of Belgium, IFO, ZEW, GfK or Sentix, but also hard data regarding industrial production or retail sales in the aggregate euro area and individually in some of the largest euro area countries. Second, our quantitative analysis is used to draw up an overall ranking of the indicators, on the basis of their average contribution to updates of the nowcast.

Among the survey indicators, we find the strongest impact for the Markit Manufacturing PMI and the Business Climate Indicator in the euro area, and the IFO Business Climate and IFO Expectations in Germany. The widely monitored consumer confidence indicators, on the other hand, typically do not lead to significant revisions of the nowcast. In addition, even if euro area industrial production is a relevant predictor, hard data generally contribute less to the nowcasts: they may be more closely correlated with GDP but their relatively late availability implies that they can to a large extent be anticipated by nowcasting on the basis of survey data and, hence, their 'news' component is smaller. Finally, we also show that, in line with the previous literature, the NBB's own business confidence indicator appears to be useful for predicting euro area GDP. The prevalence of survey data remains also under a counterfactual scenario in which hard data are released without any delay. This finding confirms that, in addition to being available in a more timely manner, survey data also contain relevant information that does not seem to be captured by hard data.

JEL classification: C32, C55, C53, C87

Key words: JDemetra+Nowcasting, surveys, news, dynamic factor models, press releases, real-time data, Bloomberg, Forex Factory, Kalman gain .

Authors:

Raïsa Basselier, Economics and Research Department, NBB – e-mail: raisa.basselier@nbb.be.

David de Antonio Liedo, R&D Statistics, NBB – e-mail: david.deantonioliedo@nbb.be.

Geert Langenus, Economics and Research Department, NBB – e-mail: geert.langenus@nbb.be.

TABLE OF CONTENTS

1 Introduction

In the field of economics, the term *nowcasting* generally refers to methods for monitoring the current state of the economy and developments in the short term, and it has become increasingly popular since the work by Evans (2005) and Giannone, Reichlin and Small (2008). Real-time estimates of economic growth are particularly relevant for both policy makers and market participants, as official national accounts data only come with a substantial delay. For instance, Eurostat releases the gross domestic product (GDP) 'flash' figure for the aggregate euro area only with an approximate delay of 45 days[1]. Also for individual euro area countries, flash estimates for GDP growth are published with a lag of at least one month. However, there is a wide range of higher-frequency variables containing either 'hard' or 'soft' information, that is being released at an earlier stage and journalists, analysts and financial market participants can already form their expectations on the basis of this dataflow. In this context, market participants and observers will continuously react to data releases in function of whether those are above or below their expectations, also depending on the perceived quality of the underlying data. The approach presented in this paper to formalise this behaviour will enable conclusions to be drawn about the relevance of every data release.

As stressed by Bańbura, Giannone and Reichlin (2011), the practice of nowcasting goes beyond the simple production of an early estimate and also requires an assessment of the impact of new data on forecasting updates over the time horizon. This paper aims to do so by converting the heterogeneous dataflow that enters the forecaster's information set into a newsflow that can be interpreted and, most importantly, quantified. The news is defined as the difference between the released values and the predictions of a mixed-frequencies dynamic factor model (DFM). Models of this sort are successful at capturing the business cycle comovements in terms of few underlying factors and have been applied for many countries[2]. The analysis of contributions will be obtained in function of the Kalman

[1]As of 29 April 2016, Eurostat also publishes a *preliminary* flash GDP with a delay of about 30 days. A second, more final, GDP flash estimate will continue to be published about 45 days after the end of the reference quarter. In this paper, we will not take on board the preliminary flash publication, as the time span covered is too limited.

[2]Please refer to Bańbura, Giannone, Modugno and Reichlin (2013) for an overview. More recent GDP forecasting examples based on DFM include de Antonio Liedo (2015) for Belgium, D'Agostino, McQuinn and O'Brien (2013) for Ireland, Bragoli, Metelli and Modugno (2015) for Brazil, Franta, Havrlant and Rusnák(2016) for the Czech Republic, Modugno, Soybilgen and Yazgan (2016) for Turkey, Bragoli and Modugno (2015) for Canada and Bragoli (2017) for Japan. Also, the EUROMIND model developed by

filter gain, which translates news into forecasting updates. From a methodological point of view, however, our proposal is not restricted to DFMs, but it simply requires a fully specified dynamic system that can be written in state-space form. Hence, linking news and forecasting updates is impossible in the context of partial models such as bridge equations, MIDAS regressions and univariate models in general, which remain widely used in nowcasting applications.

The incremental contributions of the news resulting from our empirical results then allow us to create a ranking for the multiple press releases. Such a ranking, which is determined by both the model parameters and the schedule for data releases, could provide a powerful tool for analysts and observers that are keeping track of the newsflow on a regular basis. The ranking would allow them to filter the huge amount of incoming information and mainly focus on the most important indicators. From a more formal point of view, we will investigate whether such a ranking could be used as a variable selection criterion.

Modelling news and calculating the impact on forecasting revisions requires dealing with two key characteristics of real-time data. First, the dataset typically has a ragged edge, as potential predictors are released with different lags. Hence, any assessment of the contribution of a given predictor to GDP nowcasts will have to take into account the real-time data availability, as determined by the actual data release schedule. Second, data series are often revised: the current value may deviate from the first release, while the latter is used for nowcasts in real time. Hence, those first-release data should be taken into account when evaluating the importance of a given predictor. While addressing the first issue has become standard practice in nowcasting applications – e.g. in Giannone *et al.* (2008), Angelini, Camba-Mendez, Giannone, Rünstler and Reichlin (2011) and Gayer, Girardi and Reuter (2015)–, the presence of data revisions is often ignored and current values are used for the analysis. This is usually referred to as a *pseudo real-time* approach. To the best of our knowledge, we are the first to exploit information from a genuine real-time dataset of time series covering 15 years, constructed on the basis of original press releases. The use of first-release data is in our view necessary to quantify the precise impact of the various indicators in an actual nowcasting context.

Our investigation of the role of qualitative surveys data when forecasting macroeco-

Frale, Marcellino, Mazzi, and Proietti (2011) models the individual components of euro GDP separately.

nomic variables provides specific weights to all indicators included in the analysis. Thus, our work is connected to the results by Abberger (2007), Claveria, Pons and Ramos (2007), Giannone, Reichlin and Simonelli (2009), Lui, Mitchell and Weale (2011), Martinsen, Ravazzolo and Wulfsberg (2014), Piette and Langenus (2014), de Antonio Liedo (2015) and Gayer *et al.* (2015), to name a few. However, we are the first ones to quantify a *direct* measure of surveys' informative content. Gayer et al. (2015), for example, provide an *indirect* assessment of the usefulness of a whole block of survey data by quantifying the forecasting accuracy losses resulting from a model without this block. In turn, we use a unique model to determine the exact contribution of each individual predictor in the forecasts.

Going beyond the precise quantification of each indicator's contribution on the forecast, this paper also explores whether the incremental impacts defining our ranking could be used as a variable selection criterion, as Rünstler (2016) proposes in his alternative analysis of contributions. Interestingly, we find that the dynamic factor model containing only the highest-ranked indicators produces less accurate forecasts than the benchmark model. Hence, we argue that relying on a limited set of indicators may prove to be somewhat less beneficial in the real-time environment, as opposed to the conclusion reached by Rünstler (2016), who exploits the bridging with factors framework popularised by Giannone *et al.* (2008).

Finally, we show that our workhorse dynamic factor model based on both hard and soft data produces euro area growth nowcasts that improve over time in function of the news distilled by the model. Those forecasting updates are not significantly different from the ones produced by the market. This implies that the parametric assumptions incorporated in our news-reading machinery are not at odds with reality, and our results can indeed be used as a measure of the relative importance of the various indicators.

The paper is structured as follows. Section 2 describes in detail the real-time dataflow that is relevant for monitoring the business cycle in the euro area, and defines the *standard impact* concept that will be used to rank the different indicators. Section 3 presents the model, i.e. a dynamic factor model that is flexible enough to account for a substantial proportion of the dynamic interactions between all indicators. It is shown that the nowcasts provided by this model perform well in terms of forecasting accuracy, also relative to well-known benchmarks in the field. Section 4 discusses the standard impact on the euro area GDP nowcasts that result from this model. Those results allow us to construct our ranking for the data releases. Section 4 also shows the outcome of different counterfactual scenarios. These counterfactual exercises make it possible to disentangle the part of the impact that comes from the timeliness of the indicator, from the part that is driven by its quality. The fifth section investigates whether our ranking can serve as a tool for selecting variables, as suggested by Rünstler's (2016) analysis. The last section concludes. The results presented in this paper can be easily reproduced and extended by installing a nowcasting plug-in into the JDemetra+ software, which was developed at the National Bank of Belgium[3].

[3] JDemetra+ is free and open-source software written in Java. Download it here: `https://github.com/jdemetra/jdemetra-app/releases`
The Nowcasting plugin can also be downloaded here: `https://github.com/nbbrd/jdemetra-nowcasting/releases`. After downloading it, go to the Tools option in JDemetra+ and select *plugins* and proceed with the installation. The software is portable and platform independent so it could even be launched using any operating system from a USB key.

2 Dataset and Analysis of News

Giannone *et al.* (2008) introduced the real-time dataflow as an essential element in nowcasting at a time when the literature on real-time analysis of business cycles payed little attention to the fact that data sets are unbalanced at the end of the sample. Bańbura *et al.* (2011, 2013) emphasize that nowcasting has been taken to a whole new level by defining a mapping from surprises in new data releases onto forecasting revisions. Against this backdrop, we first describe the release calendar for the relevant predictors of euro area GDP. We then formalize the concept of news, without yet making any assumption on the exact data-generating process.

2.1 Dataset and the Real-Time Dataflow

All data used in this paper, including the flash GDP and the hard indicators, were taken from original press releases from the statistical agencies. We only take on board the new figures that are provided by each publication and disregard any revisions to earlier figures. Hence, we use a genuine real-time dataset, including both soft data (survey variables) and hard data (industrial production, factory orders, retail sales and consumer spending) for a selection of euro area countries, as shown in Table A.4. The table contains an overview of the variables' definition, their frequency, the publication lag and the start of the sample.

The data selection method is very simple. The dataset contains 34 monthly and quarterly series that meet one condition: they have been regularly distributed through the economic calendar of either Bloomberg or Forex Factory, and hence, can be considered as indicators that are relevant for market observers.

There are two additional considerations in the construction of this dataset. First, all series are taken exactly as they were distributed through the Bloomberg platform. Hard indicators such as industrial production, for example, were distributed by Bloomberg in terms of both year-on-year growth rates and seasonally adjusted month-on-month rates. Due to the presence of data revisions, both transformations provide us with complementary information[4]. In turn, survey data have been distributed in terms of seasonally-adjusted balances, without further amendments.

[4]In case revisions have occurred to past data, the year-on-year growth rate is likely to give some information on this as well.

Second, the time span has been extended back in time using the official sources when it was deemed appropriate. We argue that such an extension of our sample does not violate the conditions for a genuine real-time experiment, especially in the case of surveys, because those series have not been subject to revisions, i.e. the historical data that were used to extend the sample fully coincide with the official first releases. As shown in Table A.4 in Annex III, we have extended six survey indicators back to 2000, or 2003 in the case of investor confidence for the euro area. We have also extended year-on-year German industrial production, which was only available on the Bloomberg platform from the year 2013, by combining some of the real-time information on the first releases of the month-on-month rates with the latest available vintage of Destatis historical rates, to obtain a realistic estimate of the year-on-year growth rates that would have been realised at that time.

For the GDP variables, we follow Kishor and Koenig (2012) and distinguish the first release from subsequent ones, but, unlike them, we do not make any assumption regarding the nature of data revisions. Thus, our real-time approach differs from earlier papers that exploit conventional panels of real-time data *à la* Croushore and Stark (2002). Their real-time panel for a given variable consists of increasingly large time-series in each column, which is associated with the date on which that series has been made available. In this case, the last column available – which is the one used by a forecaster today – would mix the most recent figures with past values that have already been subject to revisions. This is the reason we deviate from this approach, which has been followed by Diron (2008) and Camacho and Pérez-Quirós (2010) for euro area data[5].

[Insert Table A.4 here]

The dataset generally exhibits a ragged-edge pattern. The 34 variables are subject to different publication lags, as specified in the last column of Table A.4. Moreover, two data releases distributed around the same time may refer to an entirely different reference period, especially as our dataset includes both soft and hard data, with the latter typically

[5]Interestingly, Camacho and Pérez-Quirós (2010) do acknowledge that there is a problem with the use of vintages. For only one of the variables, GDP, they build a time series composed of flash estimates, and a different time series containing revised values. Such refinement, however, is not considered for the remaining time series. De Antonio Liedo (2015), who places more emphasis on the Belgian economy than on the euro area, follows the same approach, but also fails to reconstruct the monthly releases of hard data.

subject to longer publication lags. To make it easier to visualise the characteristics of the dataset used, a simple description of the dataflow is displayed in Figure 1, taking the August-November 2015 period as a reference. The horizontal axis represents the sequence of variables that are released. The vertical axis represents the dates at which each one of those variables is released (see triangles) and at the same time shows the period referred to using bars. Figure 1 is a more simple representation of Figure A.2, which fully demonstrates the complexity of the problem. Although the data are released in a continuous manner, we simplify the analysis by proposing a regular updating scheme with a two-weekly frequency. As noted by Angelini *et al.* (2011), European data releases tend to be clustered at the end or the middle of the month anyway. The vertical lines in the graphs represent dates where GDP nowcasts will be updated. The impact of the multiple variables will be calculated by breaking down the news in different *blocks*.

- The first block of news, which is *read* on 1 August 2015, consists of the Gfk survey for the month of August and the remaining surveys for July, i.e. by the EC, Markit, NBB, IFO and ZEW. It also contains some hard indicators (retail sales and industrial production) for Italy for the month of May. Each piece of news will contribute to updating the growth forecasts for the euro area.

- Two weeks later, the second block of data (2.0) will be read. This block contains mainly hard indicators that still refer to the month of June: industrial production, factory orders and retail sales for Germany, consumer spending and industrial production for France, as well as industrial production for the aggregate euro area. Only one survey variable is included: the Sentix release regarding uncertainty for the month of August. Apart from those monthly indicators, the release of Belgian flash GDP for the second quarter, which is typically published by the National Accounts Institute one month after the end of each quarter, also comes under the second block of news. Whether such national releases can help to forecast the euro area aggregate is an empirical question for which the method described in this paper provides a straightforward answer.

- Separately from the second block, we have the official flash estimate of real GDP growth rates for the euro area and Germany (block 2.2), which are often read in terms of how close they are to the forecast expected by market participants. Such a forecast is also represented in the graph (block 2.1), and we will assume it is known right before the official flash release.

- Blocks $3-4$ and $5-6$ represent the sequential arrivals of additional soft and hard data, while blocks $7-8.0-8.1-8.2$ are a repetition of the previous two points.

[Insert Figure 1 here]

2.2 Defining the Newsflow and its Impact on Euro Area Real GDP Growth

This subsection clarifies the concept of news, and how the real-time dataflow is translated into a newsflow.

The *news* associated with a given release is represented by the discrepancy between the published figure and the expected value. Thus, this concept depends on expectations, which in our case will be model-consistent. The words *news*, *innovations* or *forecast errors* will be used interchangeably (see Durbin and Koopman, 2001). Once the concept of news is clarified, we will show how the *Kalman gain* induces the model to update the forecast path for GDP or any other variable of interest after a given piece of news becomes available. The impact of the news that gradually enters the forecaster's information set is given not only by its quality, but also by its timeliness, which crucially depends on the release calendar. Our particular schedule for data releases and the publication lags associated to each indicator are represented in Figure A.2.

Let's consider a generic recursive representation of the factor f_t driving the observables y_t:

$$\mathrm{y_t} = \Lambda \mathrm{f_t} + \mathrm{e_t} \tag{1}$$

$$\mathrm{f_{t+1}} = A\mathrm{f_t} + \eta_t \tag{2}$$

with normally distributed measurement errors and shocks to the factors: $e_t \sim N(0, R_t)$, $\eta_t \sim N(0, Q_t)$.

Defining the information sets

The concept of news can be formalised by specifying information sets that enter the model. In order to simplify the notation and without loss of generality, we will assume in this exposition that there are only two news components and there are no revisions.

$\mathcal{F}^{old}$ contains all time series available right before the publication of the news. For the sake of simplicity, all observations are considered to be available until time t.

$\mathcal{F}^{new}$ includes the previous information set, $\mathcal{F}^{old}$, plus new data corresponding to a given macroeconomic release. Again for the sake of simplicity, one can assume that the release extends by one month, $(t+1)$, two of the indicators referring to sales $(\mathrm{y}^{\mathrm{s}}_{\mathrm{t+1}})$ and manufacturing $(\mathrm{y}^{\mathrm{m}}_{\mathrm{t+1}})$.

The forecast for the whole vector of variables y_{t+h} is formulated in our framework in terms of model-consistent conditional expectations:

$$E_\theta[y_{t+h}|\mathcal{F}^{new}] = E_\theta[y_{t+h}|\mathcal{F}^{old}, \{V_{t+1}\}] \tag{3}$$

where the expression on the right-hand side decomposes the new conditioning information set in two orthogonal parts. In this particular example, $V_{t+1} \equiv \mathcal{F}^{new} - \mathcal{F}^{old} = [\mathrm{v}^{\mathrm{m}}_{\mathrm{t+1}}\ \ \mathrm{v}^{\mathrm{s}}_{\mathrm{t+1}}]'$ incorporates two innovations or pieces of news, which are defined as the difference between the released figures and the model's forecasts:

$$\mathrm{v}^{\mathrm{m}}_{\mathrm{t+1}} = \mathrm{y}^{\mathrm{m}}_{\mathrm{t+1}} - E_\theta[\mathrm{y}^{\mathrm{m}}_{\mathrm{t+1}}|\mathcal{F}^{old}]$$

$$\mathrm{v}^{\mathrm{s}}_{\mathrm{t+1}} = \mathrm{y}^{\mathrm{s}}_{\mathrm{t+1}} - E_\theta[\mathrm{y}^{\mathrm{s}}_{\mathrm{t+1}}|\mathcal{F}^{old}]$$

Thus, one could state that, even if the released figures have declined with respect to the recent past, the model could interpret them as good news as long as they are above the values that the model was expecting.

The Kalman filter gain

This news is then exploited by the Kalman filter gain in order to update GDP forecasts together with the remaining variables. If we could observe $\mathrm{f}_{\mathrm{t+h}}$, obtaining the forecast would be straightforward: $E_\theta[y_{t+h}|\mathcal{F}^{new}] = \Lambda A^{h-1} f_{t+1}$. However, the factor $\mathrm{f}_{\mathrm{t+1}}$ can not be observed because only two data releases for $t+1$ are available and they are assumed to contain measurement errors. Thus, the conditional expectation in expression 3 needs to be developed further:

$$\begin{aligned} E_\theta[y_{t+h}|\left\{\mathcal{F}^{old}, V_{t+1}\right\}] &= \Lambda A^{h-1} E_\theta[\mathrm{f}_{\mathrm{t+1}}|\left\{\mathcal{F}^{old}, V_{t+1}\right\}] \\ &= \underbrace{\Lambda A^{h-1} E_\theta[\mathrm{f}_{\mathrm{t+1}}|\mathcal{F}^{old}]}_{old\,forecast} \\ &+ \Lambda A^{h-1} \underbrace{E_\theta[\mathrm{f}_{\mathrm{t+1}} V'_{t+1}] E_\theta[V_{t+1}V'_{t+1}]^{-1}}_{Gain\,(quality,\,timeliness)} \underbrace{V_{t+1}}_{news} \end{aligned} \tag{4}$$

The product of expectations shown in the expression 4 above defines how the *news* induces an update[6] of the state of the economy, which is represented by $\mathrm{f}_{\mathrm{t+1}}$.

In our general framework where we have news for alternative variables concerning different reference dates, notation may get more complex[7] but the two main computations required to calculate the gain can be easily understood within the context of our stylized example.

Gain 1: News Covariance. Regarding the first element of the gain, the covariance of the news,

$$E_\theta[V_{t+1}V'_{t+1}] = \begin{bmatrix} E_\theta[\mathrm{v}^{\mathrm{m}}_{\mathrm{t+1}}\mathrm{v}^{\mathrm{m}}_{\mathrm{t+1}}] & E_\theta[\mathrm{v}^{\mathrm{m}}_{\mathrm{t+1}}\mathrm{v}^{\mathrm{s}}_{\mathrm{t+1}}] \\ E_\theta[\mathrm{v}^{\mathrm{m}}_{\mathrm{t+1}}\mathrm{v}^{\mathrm{s}}_{\mathrm{t+1}}] & E_\theta[\mathrm{v}^{\mathrm{s}}_{\mathrm{t+1}}\mathrm{v}^{\mathrm{s}}_{\mathrm{t+1}}] \end{bmatrix},$$

can be shown to have a simple form where the off-diagonal elements are written as:

$$E_\theta[\mathrm{v}^{\mathrm{m}}_{\mathrm{t+1}}\mathrm{v}^{\mathrm{s}}_{\mathrm{t+1}}] = \Lambda_m \underbrace{E_\theta\left[\left(\mathrm{f}_{\mathrm{t+1}} - E_\theta[\mathrm{f}_{\mathrm{t+1}}|\mathcal{F}^{old}]\right)\left(\mathrm{f}_{\mathrm{t+1}} - E_\theta[\mathrm{f}_{\mathrm{t+1}}|\mathcal{F}^{old}]\right)'\right]}_{Precision\,Matrix} \Lambda'_s + E_\theta\left[\mathrm{e}^{\mathrm{m}}_{\mathrm{t+1}}\mathrm{e}^{\mathrm{s}}_{\mathrm{t+1}}\right],$$

with $E_\theta\left[\mathrm{e}^{\mathrm{m}}_{\mathrm{t}}\mathrm{e}^{\mathrm{s}}_{\mathrm{t}}\right] = 0$ because the measurement errors of our model are idiosyncratic. The diagonal elements will contain the covariance of the measurement errors, which is non-zero. Λ_m and Λ_s contain the row of Λ that corresponds to the two indicators for which we have news, i.e. y^m_{t+1} and y^s_{t+1}, respectively. The precision matrix above reflects the fact that factors are unobserved and it turns out to be an output of the Kalman smoother iterations. However, if one of the news items were to refer to a more distant period of

[6]This update takes the same form of a simple OLS regression of the factors on the innovations. Note that the size of the news vector V_{t+1} may be large in practical applications when many variables are released at the same time or many observations for the same variable are made available simultaneously. However, the resulting gain remains a function of a reduced number of model parameters θ.

[7]Banbura and Modugno (2010) propose a general notation in their section 2.3. and use it to expand expression 4.

time, e.g. if y^s_{t-h} is revised, the relevant precision matrix would contain $\mathrm{f_{t+1}}$ and $\mathrm{f_{t-h}}$:

$$E_\theta[\mathrm{v^m_{t+1} v^s_{t-h}}] = \Lambda_m \underbrace{E_\theta\left[\left(\mathrm{f_{t+1}} - E_\theta[\mathrm{f_{t+1}}|\mathcal{F}^{old}]\right)\left(\mathrm{f_{t-h}} - E_\theta[\mathrm{f_{t-h}}|\mathcal{F}^{old}]\right)'\right]}_{Precision\,Matrix} \Lambda_s' + E_\theta\left[\mathrm{e^m_{t+1} e^s_{t-h}}\right],$$

The role of timeliness, which determines patterns of missing observations that enter the gain though the news covariance matrix, is crucial for defining the weights. Thus, it can be easily understood that variables that enter the model's information set first will receive a larger weight than if they had have been part of a larger group of data releases. The reason is that, in the presence of strong collinearity where all variables incorporate the same signal, the first variable entering the information set may be sufficient, rendering the remaining releases irrelevant.

Gain 2: Correlation of the News with the Factors. The precise weight of each one of the innovations when updating the expectations about the state of the economy also depends on the quality of the indicators. By quality, we refer to the correlation of the factor with the innovation. However, this definition of quality may be ambiguous for two reasons. First, in models with more than one factor, the news associated with a certain indicator may be correlated with one factor but not with the others. In this case, it is the aggregate impact that matters. Second, this definition of quality is not invariant to timeliness. Think of an indicator that may contain news that is highly correlated with the factors, i.e. it is an indicator of good quality. If the same indicator happens to be published with an unusual delay right after the release of competing indicators with similar informative content, the properties of the resulting news component may be totally different. Therefore, quality can be unambiguously defined only when the timeliness dimension is fixed.

The second element of the gain shown in expression 4, which we relate to quality, can be expanded as follows[8]:

$$
\begin{aligned}
E_\theta[\mathrm{f}_{\mathrm{t+1}} V'_{t+1}] &= \begin{bmatrix} E_\theta[\mathrm{f}_{\mathrm{t+1}} \mathrm{v}^{\mathrm{m}}_{\mathrm{t+1}}] \\ E_\theta[\mathrm{f}_{\mathrm{t+1}} \mathrm{v}^{\mathrm{s}}_{\mathrm{t+1}}] \end{bmatrix}' \\
&= \begin{bmatrix} E_\theta\left[\left(\mathrm{f}_{\mathrm{t+1}} - E_\theta[\mathrm{f}_{\mathrm{t+1}}|\mathcal{F}^{old}]\right)\left(\mathrm{f}_{\mathrm{t+1}} - E_\theta[\mathrm{f}_{\mathrm{t+1}}|\mathcal{F}^{old}]\right)'\right]\Lambda'_m \\ E_\theta\left[\left(\mathrm{f}_{\mathrm{t+1}} - E_\theta[\mathrm{f}_{\mathrm{t+1}}|\mathcal{F}^{old}]\right)\left(\mathrm{f}_{\mathrm{t+1}} - E_\theta[\mathrm{f}_{\mathrm{t+1}}|\mathcal{F}^{old}]\right)'\right]\Lambda'_s \end{bmatrix}'
\end{aligned} \tag{5}
$$

Defining the standard impact of news. Let's continue with our simple example with only two data releases and one factor. Our goal now is to decompose the gain in terms of a few parameters that can help us to understand the logic underlying the updating mechanism. The last term of expression 4 can be written in terms of parameters $\sigma^2_m = var_\theta(\mathrm{v}^{\mathrm{m}}_{\mathrm{t+1}})$, $\sigma^2_s = var_\theta(\mathrm{v}^{\mathrm{s}}_{\mathrm{t+1}})$, $\sigma^2_{ms} = cov_\theta[\mathrm{v}^{\mathrm{m}}_{\mathrm{t+1}}, \mathrm{v}^{\mathrm{s}}_{\mathrm{t+1}}]$, $\beta_m = cov_\theta[\mathrm{f}_{\mathrm{t+1}}, \mathrm{v}^{\mathrm{m}}_{\mathrm{t+1}}]$, $\beta_s = cov_\theta[\mathrm{f}_{\mathrm{t+1}}, \mathrm{v}^{\mathrm{s}}_{\mathrm{t+1}}]$. This will allow us to illustrate the interaction of quality and timeliness in the definition of the Kalman gain:

$$
\begin{aligned}
E_\theta[y_{t+h}|\left\{\mathcal{F}^{old}, V_{t+1}\right\}] - E_\theta[y_{t+h}|\left\{\mathcal{F}^{old}\right\}] &= \underbrace{\Lambda A^{h-1}\frac{\beta_m\sigma^2_s - \beta_s\sigma^2_{ms}}{\sigma^2_m\sigma^2_s - \sigma^2_{ms}\sigma^2_{ms}}}_{impact\ of\ manufacturing} \mathrm{v}^{\mathrm{m}}_{\mathrm{t+1}} \\
&+ \underbrace{\Lambda A^{h-1}\frac{\beta_s\sigma^2_m - \beta_m\sigma^2_{ms}}{\sigma^2_m\sigma^2_s - \sigma^2_{ms}\sigma^2_{ms}}}_{impact\ of\ sales} \mathrm{v}^{\mathrm{s}}_{\mathrm{t+1}}
\end{aligned} \tag{6}
$$

This exposition shows that the whole set of news, i.e. the vector of innovations V_{t+1}, induces an update of the path for all variables in y_t. The extent to which all the individual pieces of news induce change expectations for GDP growth rates in the euro area depends on the quality of the data, i.e. the correlation of each news component with the factor (β_m, β_s) and on the particularities of the calendar of data releases, which defines the resulting pattern in expression 6. Interestingly, when the correlation of the two news components (σ^2_{ms}) is different from zero, the impact of each news component is reduced by a factor

[8] The formulation we are using for the case where we have only two news components can be easily generalised to handle more realistic situations where we have a larger set of news concerning different reference dates. Expression 5 will simply grow to incorporate the relevant precision matrix multiplied by the corresponding loadings. If the news refers to different periods of time, the only change relates to the time indices in the second parenthesis. As mentioned earlier, the increased complexity does not prevent us from exploiting the Kalman smoother.

equal to σ^2_{ms} times the β of the competing indicator. This logic may lead data publishing agencies to release their data earlier than the competition. A further discussion using the same notation can be found in the technical appendix of de Antonio Liedo (2015). Moving away from the stylised exposition above, Figure 2 displays how two consecutive information sets would look like in a practical application for news 8.0, for example.

[Insert Figure 2 here]

The next section will introduce our assumption about the data generating process, which will help us in distilling model-consistent news factors from the dataflow. Quantifying the precise role of all the pieces of news is the goal of Section 4. By multiplying the impacts defined in the equations above by the *standard deviation of the news* associated with each data release (e.g. σ_m, σ_s), we will obtain a measure that enables us to compare the *average* informative content of the different indicators when the object of interest is real economic growth, as measured by GDP.

3 A Model for Reading News

In this section, we present a common approach to link GDP, which is the quarterly variable that we use as a target, with the unobserved factors, which are specified at a monthly frequency and determine the joint dynamics of the whole system.

3.1 A State-Space Representation

We describe here a joint model including euro area, German and Belgian GDP. Germany was included because it could be considered as one of the main drivers of the euro area business cycle, given its size and the high share of the manufacturing sector in total value added in Germany (ECB, 2011). Belgian GDP was introduced mainly for practical reasons, as the model is intended to be used within the National Bank of Belgium for monthly nowcasting of the country's GDP. Also, given the evidence reported by the Wall Street Journal (1999) and Camacho and Pérez-Quirós (2010) suggesting that Belgian data may be informative about the euro area, one could argue that the Belgian flash GDP could help to anticipate the forthcoming euro area GDP release. In our actual model, we assume that there is more than one factor, but for the sake of simplicity, the following expression links the monthly growth rates of the variables to only one monthly factor:

$$y_t \quad = \quad \bar{y} + \Lambda_y f_{1,t} + \psi_t \qquad\qquad Measurement \;\; equation \tag{7}$$

The error term ψ_t is assumed to be uncorrelated with the factor at all leads and lags. It is also assumed to be independent and identically distributed (iid) and following a normal distribution: $\psi_t \sim N(0, R_\psi)$. The covariance matrix is assumed to be diagonal, which implies that the factor will account for 100% of the comovements implicit in the model. As suggested by Doz, Giannone and Reichlin (2012), this assumption may lead to model misspecification if it turns out that there are cross-correlation patterns in the error term[9]. However, in the presence of weak correlation patterns, they show that a Quasi-ML estimation, i.e. fixing the correlation across measurement errors to zero, would still yield

[9]The nowcasting plugin of *J*Demetra+ contains visual tools such as the schemaball to inspect for correlation patterns. If instead of having a few measurement errors correlated with each other we identify a pattern of cross-correlation that affects most of the measurement errors in the panel, it would be impossible to distinguish whether the correlation in the data comes from the common factors or from the measurement errors. Thus, estimates would not be consistent in this case.

a consistent estimator of the factors.

Because the model has been designed for short-term analysis, it makes sense to represent all these series, including GDP, in terms of monthly growth rates or differences. However, the monthly growth rates of official GDP figures are not published, so equation 7 will need to be modified. Thus, GDP growth rates published by the statistical agencies (i.e. y_t^Q for the euro area, for example) are linked to the quarterly growth rate of the underlying factor, which can be expressed as a moving average of the monthly growth rates:

$$y_t^Q = \bar{y}^Q + \Lambda_y^Q f_{1,t}^Q + \psi_t^Q, \qquad t = 3, 6, 9, \ldots \tag{8}$$

where

$$f_{1,t}^Q = \frac{1}{3} f_{1,t} + \frac{2}{3} f_{1,t-1} + f_{1,t-2} + \frac{2}{3} f_{1,t-3} + \frac{1}{3} f_{1,t-4}$$

As mentioned above, $f_{1,t}$ represents the monthly growth rate of the latent factor. The last expression for $f_{1,t}^Q$ is based on the technical assumption that the quarterly level of the factors can be represented by the geometric mean of the latent monthly levels[10]. This assumption makes it possible to obtain a simple expression for the quarterly growth rate of the factors as a moving average of the latent monthly growth rates. Because we apply the Mariano and Murasawa (2003) approximation to the factors alone, and not to the observables, the error term ψ_t^Q is assumed to be iid normally distributed and uncorrelated with the factor at all leads and lags.

So far, we have described the measurement equation, which defines the link between the unobserved factors and the two types of observable time series: monthly variables and quarterly variables (e.g. GDP). Specifying the joint dynamics of all variables requires a second equation representing the factor as a vector autoregressive (VAR) process with a

[10]The approximation proposed by Mariano and Murasawa (2003) is applied to the factors. Let F_t be the monthly *level* of the economy and let $f_t = lnF_t - lnF_{t-1}$ be its monthly growth rate. Now, define F_t^Q as the geometric mean of the last three levels. This implies that $lnF_t^Q = \frac{1}{3}(lnF_t + lnF_{t-1} + lnF_{t-2})$. The resulting quarterly growth rate of the factors, which we denote as f_t^Q, can be expressed as $lnF_t^Q - lnF_{t-3}^Q$. By substituting both terms by the geometric mean approximation we obtain $f_t^Q = \frac{1}{3}(lnF_t - lnF_{t-3}) + \frac{1}{3}(lnF_{t-1} - lnF_{t-4}) + \frac{1}{3}(lnF_{t-2} - lnF_{t-5})$. Finally, a simple expression for the quarterly growth rate of the factors in terms of their monthly growth rates can be obtained as follows: $f_t^Q = \frac{1}{3}(f_t + f_{t-1} + f_{t-2}) + \frac{1}{3}(f_{t-1} + f_{t-2} + f_{t-3}) + \frac{1}{3}(f_{t-2} + f_{t-3} + f_{t-4})$. Rearranging terms yields the expression $f_t^Q = \frac{1}{3} f_t + \frac{2}{3} f_{t-1} + f_{t-2} + \frac{2}{3} f_{t-3} + \frac{1}{3} f_{t-4}$ presented above.

non-diagonal covariance matrix for the error term. To sum up, the representation given by equations 9 and 10 conforms to the so-called state-space representation of this model:

$$\begin{pmatrix} y_t^Q - \bar{y}^Q \\ y_t - \bar{y} \end{pmatrix} = \begin{pmatrix} \Lambda_y^Q & 2\Lambda_y^Q & 3\Lambda_y^Q & 2\Lambda_y^Q & \Lambda_y^Q \\ \Lambda_y & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} f_{1,t} \\ f_{1,t-1} \\ f_{1,t-2} \\ f_{1,t-3} \\ f_{1,t-4} \end{pmatrix} + \begin{pmatrix} \psi_t^Q \\ \psi_t \end{pmatrix} \tag{9}$$

$$\begin{pmatrix} f_{1,t} \\ f_{1,t-1} \\ f_{1,t-2} \\ f_{1,t-3} \\ f_{1,t-4} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} & 0 \\ I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \end{pmatrix} \begin{pmatrix} f_{1,t-1} \\ f_{1,t-2} \\ f_{1,t-3} \\ f_{1,t-4} \\ f_{1,t-5} \end{pmatrix} + \begin{pmatrix} u_t^f \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{10}$$

This equation represents a VAR of order 4 for the factor, but one could restrict it to a VAR(1) by setting $A_{12} = A_{13} = A_{14} = 0$. The error component is uncorrelated with the measurement error term, in line with the literature on factor models. For simplicity, and in contrast to the model built by Mariano and Murasawa (2003), we do not incorporate autocorrelation in the measurement errors. This helps to keep the size of the state vector as small as possible without restricting the extent to which the factors can account for the business cycle comovements[11].

3.2 Estimation in the Context of Missing Observations

Once the building blocks of the model have been defined, we need to tackle the problem of estimation. As mentioned earlier, the model is estimated under the restriction that the off-diagonal elements of the measurement error covariance matrix are equal to zero. This has the practical implication that the cross-correlation patterns generated by the model will be fully accounted for by the factor(s).

The model is estimated at monthly frequency with maximum-likelihood even in the presence of missing observations. The presence of quarterly data generates missing observations, since they are treated as indicators that are observed every third month of the quarter, i.e. y_t^Q as a missing variable for $t \neq 3, 6, \ldots$. Finally, as in most macroeconomic forecasting applications, the relevant information set is based on indicators that arrive

[11]Series that behave erratically or are not significantly correlated with the factors would yield a poor forecast. Specifying an ARMA process for the measurement component of those series would clearly improve the forecast, but this does not change the fact that the correlation of those series (and hence, their news component) with the factors is small. Thus, the weight associated with those series is likely to remain low.

gradually throughout the quarter and with important delays with respect to the period of time to which they refer, i.e. the real-time dataflow. Thus, in practice, missing values at the end of the sample are unavoidable. For a detailed overview of the estimation method used in this paper, the reader is referred to Bańbura and Modugno (2010). Below, we summarise the most important concepts underlying the approach with special emphasis on the aspects that are particularly relevant in our nowcasting framework:

- **Maximum-likelihood**. In this application, a slightly more complex[12] version of the state-space model represented by equations 9 and 10 is first estimated with the Expectation-Maximization (EM) algorithm. The Kalman (1960) filter and smoothing recursions, however, need to be slightly modified so that only the actual observations can be taken into account in the estimation of the factors and the evaluation of the likelihood. The EM algorithm was derived by Shumway and Stoffer (1982) only for the case where the factor loadings multiplying the factors in the measurement equation are known. Bańbura and Modugno (2010) were the first ones to apply this algorithm to a set-up similar to ours, where the loadings need to be estimated in the context of missing observations. They show that their method is consistent and computationally feasible even when the number of variables is large. The outcome of the EM algorithm is used to initialise a numerical optimisation routine to maximise the likelihood of the model[13].

- **Identification of the factors**. The strongest assumption, which is key for identification, is that the measurement errors in expression 9 are uncorrelated with the factor innovations in the transition equation 10. This allows for a clear-cut separation of the measurement errors and the signal provided by the factors. In the absence of the restrictions we impose in the factor loadings, the model would be identified only up to an invertible linear transformation. That is, applying the following transformation, $g_t^M = Gf_t^M$, the transition equation $g_t^M = GA_1G^{-1}g_{t-1}^M + \ldots + GA_pG^{-1}g_{t-p}^M + Gu_t$ would be observationally equivalent to the one represented

[12] Our actual model uses 4 factors instead of one and some surveys are linked to the cummulative sum of the factors over 12 months, as in Camacho and Pérez-Quirós (2010).

[13] Numerical optimisation of the likelihood, which is feasible for parsimonious models, has the advantage that it does not require the Kalman smoother. Moreover, the multithreading ability of most software packages is able to reduce the execution time by exploiting multiple processors. For example, the current estimation of dynamic factor models in *J*Demetra+ is feasible without the need to apply the EM algorithm.

by equation 10. Nevertheless, Dempster, Laird and Rubien (1977) suggest that the EM algorithm is not affected by this lack of identification. The space generated by the factors, and thereby the projections on such space, are unaffected by the choice of G. This identification issue is well known in factor analysis and does not distort any of the results presented in this paper.

3.3 Nowcasting Accuracy

Before discussing the impact that corresponds to the news associated with each of the indicators, it is useful to make sure that the chosen model is at least able to provide a realistic representation of the news for the whole set of indicators. In a first step, we show that this particular dynamic factor model performs well in terms of forecasting accuracy. It is also demonstrated that our model's nowcasts are improving when moving closer to the publication date of the target variable, as more news is being released. In a second step, we show that our model's nowcasts can compete with some of the well-known benchmarks in the field.

Model Selection

Our analysis of news is based on a state-space representation with four factors. Thus, our model has a larger stochastic dimension than the Euro-Sting model developed by Camacho and Pérez-Quirós (2010) or the model designed by Aruoba, Diebod and Scotti (2009) and subsequent versions[14] for the US, which relies on a single factor and the performance of which crucially depends on the variables that were selected. Figure 3 shows that the model with four factors and three lags delivers the best results in terms of the root mean squared error for the euro area flash GDP. Furthermore, the graph clearly shows the stepwise-looking pattern of the decreasing root mean squared error (RMSE) for real flash GDP growth in the euro area over the period 2006Q1-2015Q1 as more information becomes available. The error is based on forecasts obtained by re-estimating the model once a year and updating the forecasts with every data release.

[Insert Figure 3 here]

As one can see more clearly in Figure 4, the RMSE associated with updates that take

[14] The original model is maintained, modified and updated by the Federal Reserve Bank of Philadelphia.

place with every data release (solid line) decreases gradually until the arrival of the last piece of news corresponding to each block, which is represented by the circle at the end of each news block. The statistical significance of the reduction in RMSE is highlighted with asterisks. This would imply that the strategy of exploiting the newsflow in real time is clearly better than waiting for the last data release of each block before updating the model, in line with the conclusion reached by Bragoli *et al.* (2015). Visually, this is most clear for news blocks 4 and 6 in Figure 4, containing releases for industrial production, sales, factory orders, consumer spending and the Sentix index. Within news block 6, the specific variable that causes the RMSE to plummet at horizon 5 is the factory orders in Germany, concerning the second month of the reference quarter. However, such an important reduction in the RMSE does not turn out to be statistically significant, so it could be driven by just a few forecasts that turned out to be particularly accurate. As a matter of fact, factory orders in Germany is not among the best ranked indicators according to the estimated weights that will be reported in Section 4, so it is not reasonable to expect systematic improvements. Thus, the good luck hypothesis is actually compatible with the updating behaviour of the model.

On aggregate, all the news represented in the figure has ended up reducing the RMSE by 0.34, up to 0.22. The graph also reveals further accuracy gains over the evaluation sample when Bloomberg forecasts, which represent the only quantitative survey used in this paper, are processed by the model (see the small dip in the RMSE at horizon 44). The particularities of the evaluation sample, which include the Great Recession and the sovereign debt crisis, may not be representative of a typical business cycle with long expansion periods with stable growth rates, so the 0.22 figure may be considered as an upper bound of the forecasting uncertainty that can be expected from the model.

Besides a visual inspection of our results, it is worth investigating more formally how the forecasting accuracy improves when more news becomes available. The classical tests described in Annex I are used to determine which news blocks significantly improve the RMSE of our nowcasts. Over a long sample, we expect the forecasting accuracy of the model to improve after reading each block of news. Using the results of the Diebold-Mariano test included in Table A.2 in Annex I, it is possible to distinguish which news blocks provide the most significant updates in terms of forecasting accuracy. Even though every update brings a decrease in RMSE, only some of them actually bring statistically significant gains. The outcome of the Diebold-Mariano test may be considered jointly with the results of two encompassing tests, also included in Table A.2. In the first case, the null hypothesis states that the updated forecast encompasses all the relevant information from the old forecast. In the outmost right hand column of the table, the null hypothesis works the other way around. As expected, in most of the cases, one is able to reject the hypothesis that the old forecast encompasses the updated one. This confirms that the updated forecast adds significant information that is not present in the previous forecast.

[Insert Figure 4 here]

Performance Relative to Benchmarks

As a first indication, Figure 5 shows the target series (euro area flash GDP), compared to the dynamic factor model's (DFM4) nowcast at a horizon of 30 days after the end of the reference quarter (i.e. about two weeks before the actual - more final - flash is published by Eurostat). The nowcast provided by Bloomberg, which is usually available right before the flash publication, is also plotted in this graph. Visual inspection suggest that the DFM4 is performing relatively well: it seems to capture the most important movements of the euro area flash GDP, some two weeks before the official flash estimate is actually released. The performance of the DFM4 also seems comparable with that of the Bloomberg forecast.

[Insert Figure 5 here]

Table (A.3) in Annex I shows the results of a more formal investigation of the forecasting accuracy of our DFM4, relative to some of the well-known benchmarks in the field. A first comparison was made between our nowcasts and those coming from Now-Casting.com, an economic forecasting business that publishes their nowcasts online. Our model appears to register significantly worse forecasts than the Now-Casting.com benchmark at the middle of the quarter (-45 days), but this could be considered rather normal as their platform makes use of a much larger dataset. One day prior to the flash release (+44 days), our model significantly improves over the Now-Casting.com benchmark, but according to the encompassing test, they both contain valuable information that should not be neglected.

Other important benchmarks are the PMI-implied growth rates for a given quarter[15] or the Bloomberg forecast that is available one day before the flash release. The second part of Table (A.3) suggests that Bloomberg and the PMI rule are comparable to our model in terms of accuracy. The first encompassing test suggests that both benchmarks contain useful information that is not present in our DFM nowcast. But at the same time, the second encompassing test indicates that the DFM contains useful information that is not captured by the PMI-implied forecast, nor by the Bloomberg benchmark. This means that both models under comparison (i.e. our DFM4 relative to the benchmark) contain complementary information and the forecasts could be improved by combining them.

[15]Please refer to the publication by Williamson (2015) of Markit Economic Research entitled 'Using PMI survey data to predict official eurozone GDP growth rates' to find out more about these PMI-implied growth rates.

4 Empirical Results and Counterfactual Exercises

4.1 Benchmark case

This section quantifies the impact of all data releases with respect to updating the forecasts for real GDP growth rates in the euro area on the basis of the methodology explained in Section 2 and the model presented in Section 3. We will rank all the predictors according to their expected impact on updating real GDP growth forecasts for the euro area. The expected impact will be given by a parametric dynamic factor model representing all monthly and quarterly variables.

We have defined the *standard impact* of each predictor as the product of the impact coefficients defined as in equation 6 and the standard deviation of the respective innovations, i.e. the root mean squared error (RMSE) associated with the release of each series. The flow of information within the quarter has been clarified by Figures 1 and A.2 in Section 2.

The resulting standard impacts on the prediction for the current and the next quarter GDP of all data releases are displayed in Figure A.3. The standard impacts are a function of the real-time dataflow. Thus, the function is *constant over time* as long as the order of the data releases remains unchanged. The graph shows that some indicators consistently have a substantial impact on the revision of real GDP growth expectations in the euro area in the current quarter, or even the next. Impacts also generally have the expected sign: positive (negative) news about the predictor leads to an upward (downward) revision of the GDP growth estimate. This for example is the case for the Markit PMI releases for the euro area, the German IFO surveys, the Business Climate Indicator for the euro area and also industrial production. Other indicators, such as the retail sales indices or the headline consumer confidence indicators always show low impacts[16]. In this connection, Piette and Langenus (2014) have already warned that the headline consumer confidence indicator is not very useful for nowcasting. However, they find that certain sub-components of the headline consumer confidence indicator (in particular the unemployment expectations) are more relevant for the early estimates of GDP growth.

[16]Oddly, the impact of the GfK consumer confidence even displays the wrong sign: a positive surprise in consumer confidence actually leads to a downward revision of the GDP nowcast. We believe that this result should rather be interpreted as a *zero* impact.

The graph can be read in a more detailed way. For example, inside the first block of news, we find the Markit PMI release that is *read* by our model on the first of August (but the information it holds still refers to the month of July). This publication, which comes out in the second month of the current quarter, is expected to induce a revision of real GDP growth expectations in the euro area close to 0.06 in absolute value for that quarter, but also 0.02 for the next quarter already. Markit's euro area PMI release could thus clearly be considered as a timely indicator. The graph also reveals that GDP growth expectations for Q3 will be affected by the August, September and October euro area PMI releases as well, although the impact is gradually decreasing. Interestingly, Figure A.3 also demonstrates that the industrial production release in Germany and the euro area have a non-trivial standard impact even if they are published with a time lag of more than one month. This result may question the empirical findings of Camacho and Pérez-Quirós (2010) or de Antonio Liedo (2015), where the impact of hard data is found to lean too much in favour of surveys. Our results rely on a modelling approach based on a relatively large number of factors, which aim to give the model a fair chance to fit a very heterogeneous dataset that combines surveys coming from multiple sources together with the hard data expressed in terms of both monthly and yearly growth rates. The graph illustrates that the industrial production release in the block of news 4, which corresponds to the first month of Q3, i.e. July, still has a significant impact on the euro area GDP for that quarter. The standard impact of 0.07 for industrial production in the euro area turns out to be comparable to the sum of the impacts of the other hard indicators combined. Interestingly, industrial production in other countries (France, Italy) seems to be a much less relevant predictor of euro area GDP (and sometimes even has the wrong sign).

[Insert Figure A.3 here]

This graph also reveals that the Belgian flash GDP release (within the block of news 8.0) has a very small impact on aggregate euro area growth referring to the same quarter, even though it is published two weeks earlier. The estimated impact on the euro area GDP nowcast is very small, which suggests that it does not incorporate much added value beyond all indicators that have been previously released.

Despite the small impact of the Belgian flash GDP publication, the results do prove that certain national (survey) indicators can nonetheless be useful for the short-term prediction of the euro area aggregate business activity. The Business Confidence Indicator published by the National Bank of Belgium, for example, turns out to be among the releases with the largest impact, together with the IFO surveys for Germany.

As Figure A.3 may be somewhat difficult to read or interpret, Figure 6 provides a more straight-forward ranking of the releases, based on their *cumulative impact over the entire half-year*. This figure confirms the dominance of soft data, but industrial production in the euro area and Germany occupy the fifth and twelfth place, respectively, in the overall ranking. Two things should be kept in mind when interpreting our results. First, by analysing the data releases by blocks, we are neglecting the fact that the Markit euro area PMI and the NBB Business Survey, for example, are published somewhat earlier than the competing indicators. Had we taken that aspect into account, the standard impact associated with both indicators would have been even larger. Second, this ranking is constructed using standard impacts that depend on how much each piece of news is weighted by the model for the prediction of our target variable, euro area GDP (see Figure 2). The weights associated with the news are a function of the model parameters, and more specifically, of their correlation with the target variable. This implies that if the target variable was, say, German GDP instead of euro area GDP, the resulting ranking would be different (see Annex II).

[Insert Figure 6 here]

The standard impacts of the news, summarised in Figure 6 and represented in detail in Figure A.3, capture a measure of the relevance of each piece of news for updating the growth expectations. They are calculated with the full sample to make sure the numbers we report are as precise as possible. Before the complete newsflow is incorporated, the uncertainty surrounding the GDP estimates is given by the sum of the squared impacts. We focus on the standard impacts instead of the squared impacts because it is easier to interpret in terms of how much each piece of news is expected to revise the forecasts. Figure 7 represents the real process of updating the GDP growth expectations for the euro area in our specific example for 2015Q3. Every data release holds a positive or negative surprise relative to the model's estimates, which leads to forecasting revisions given by the product of the Kalman gain and the news (see equation 4). In this specific quarter, the impacts turn out to be negative on sum and, hence, the GDP nowcast is revised downwards as time progresses. Ultimately, right before the flash is released, we end up with a nowcast that is slightly above 0.2. One day later, the official flash estimate for euro area GDP growth is released and turns out to be 0.3.

[Insert Figure 7 here]

4.2 Counterfactuals

In addition to the benchmark exercise presented above, we will investigate how the results (referred to as Analysis A) regarding the ranking would change, under a number of counterfactual hypotheses. More specifically, we will study whether the large impact we have found for surveys remains under the assumption that hard data are published without delay (Analysis B). Furthermore, we will re-calculate the results under the hypothesis that the hard indicators are published without revisions. That is, instead of the original data releases, which contain preliminary data that will be subject to significant revisions in the case of hard data, we use the series that have already been revised (Analysis C). Finally, we investigate the *combined* impact of having fully revised hard data that are published without any delay on fully revised GDP (analysis D).

[Insert Table 1 here]

4.2.1 Timeliness

Timeliness is a unique characteristic of soft data that probably contributes to their large impact, as discussed in the benchmark case. However, in this counterfactual exercise, we analyse whether (and by how much) those impacts are reduced when we assume that hard data are published already at the end of the reference month, i.e. the publication delay equals zero. Figure 8 compares the benchmark ranking from Section 4.1 to the results obtained in this counterfactual exercise. The first thing that strikes the eye is that one of the variables, industrial production in the euro area, has now a much bigger impact. Industrial production in Germany also gains some importance, and its standard impact is now close to that of the manufacturing PMI indicator for the euro area. Industrial production in Italy or France continue to have smaller impacts, in spite of their improved timeliness in this counterfactual exercise. Surprisingly, the sum of the impacts of all soft indicators in Figure 8 is still larger than the aggregate impact of all hard indicators that are represented in the ranking. This implies that timeliness is not the only characteristic of soft data that accounts for the large impact of survey data. This result has also been discussed by de Antonio Liedo (2015) in the context of Belgian data using a different model, although the role of hard indicators is practically negligible according to his findings. This finding confirms the conclusion by Piette and Langenus (2014), who show that even when a broader set of hard indicators is available, survey data still contain relevant information that is not captured by the usual set of hard data. Gayer *et al.* (2015) argue more specifically that survey data have other characteristics besides their timeliness that can possibly improve the nowcasting accuracy: they are often forward-looking and also tend to have a broader sectoral scope.

[Insert Figure 8 here]

4.2.2 Revisions

An important innovation of this paper is the real-time dataflow obtained directly from press releases. Earlier work on nowcasting usually deals with *pseudo real-time* datasets that make use of the most recent (i.e. already revised) data that are treated as if they were real-time (i.e. the dataset also has a ragged edge, because the publication delay is respected). This counterfactual exercise C mimics such a pseudo real-time dataset: data for the hard indicators and the target variable GDP were replaced by the most recent data found via Thomson Reuters Datastream and are therefore likely to have been *revised* compared to the initial vintage. It may be important to note that only the month-on-month growth rates were replaced by their revised counterpart, while the year-on-year growth rates were kept unchanged (i.e. first vintage). Otherwise, this would imply feeding the same information twice into the model and this could cause the dynamic factor model to attribute an abnormally large weight to certain hard indicators.[17]

Figure 9 combines the benchmark ranking with the standard impacts of the current

[17] This was less of a risk in the benchmark case, because the annual growth rate refers to a level of industrial production of one year ago, which already incorporates some revision. Hence, while the monthly growth rate only provides information on the most recent observation, the annual growth rate already gives an indication of past revisions.

counterfactual analysis C. When comparing these two cases, it is necessary to take into account that the impacts for the benchmark case were calculated with regard to euro area flash GDP, while the impacts for this counterfactual analysis C are calculated with regard to euro area *revised GDP*, which is the only relevant target variable to use in this counterfactual exercise. Certain indicators, such as the Sentix investor confidence and the NBB business confidence barometer for Belgium, gained some importance. Their simulated impact in the counterfactual analysis is wider than in the benchmark case: this would imply that they are closer to the final euro area GDP than to the flash estimate. Although, overall, the ranking in the counterfactual exercise does not change too much relative to the benchmark, we do warn that analysing the release impacts based on a pseudo real-time dataset may be somewhat misleading. First, the resulting rankings do not fully coincide, but, more importantly, from a logical point of view, it may not be very relevant to discuss the influence of news on the real-time target (flash GDP), if the revised series of GDP are used for the analysis instead.

[Insert Figure 9 here]

4.2.3 Timeliness and Revisions

The dataset in this last counterfactual scenario represents the most ‘extreme’ scenario: all hard data are revised, while they are also assumed to be published without any delay. The resulting ranking is based on the standard impact of the news on the model’s estimate for *revised* euro area GDP. Figure 10 confirms that industrial production in the euro area is the big winner in this scenario, as already concluded in scenario B. However, as seen before, the survey indicators also maintain their significant impact for the prediction of revised GDP. This is an important result, because it may appear to be somewhat contradictory with the way that ‘final’ (i.e. revised) GDP is supposed to be assembled. According to the producers of the national accounts, only hard data are taken into account when the revised GDP series are being made. Hence, survey data should prove to be irrelevant for the revised euro area GDP series. We see two possible reasons to explain why this is not the case. First, for some reason, statisticians may prefer to keep their final GDP estimate as close as possible to the flash GDP estimate. As the flash estimate relies mainly on survey variables, the impact of surveys may simply be propagated onto the final GDP as well. Second, as hard data may not be fully exhaustive or contain certain irregularities, statisticians may decide to apply some *judgment* to the mechanical estimates. In this case, our results would suggest that their expert judgment is influenced by the information from soft data.

[Insert Figure 10 here]

5 Variable selection based on expected impacts

Based on the ranking of the predictors established in Section 4, one may be tempted to use this methodology in order to identify a reduced set of variables for the estimation of a small dynamic factor model. This idea could be considered as a refinement of the method proposed by Rünstler (2016). Rünstler's analysis is based on earlier work by Bańbura and Rünstler (2011) and exploits the weights of the different predictor variables in the factors.[18]

In Rünstler's empirical set-up, those weights are given by a measure of the historical correlations of the *revised* predictor variables with the factors extracted from them. This means that the results could possibly be distorted by data revisions that have taken place years after the actual data releases. Second, from a theoretical point of view, selecting variables in function of those weights can also be misleading when the correlation across predictor variables is neglected.[19]

Regardless of the underlying methodology, the idea is to select the indicators with the highest weight in the forecasts. In our set-up, we select the indicators that entail the largest updates in the model's GDP nowcast in the benchmark scenario. Thus, we ensure that timeliness does not bias our analysis of contributions to the detriment of quality by considering the correlation patterns across all news as an essential part of our modelling framework. A second difference from Rünstler's approach is that we perform the analysis in a genuine real-time environment, whereas he uses a pseudo-real-time dataset.

Although the idea of variable selection is appealing, and it has worked in Rünstler's simulations, it does not seem to work in our real-time case. As shown in Figure 11, selecting the highest-ranked indicators deteriorates the forecasting performance of the model. The RMSE function corresponding to our workhorse model with four factors deteriorates remarkably when the same model is re-estimated using only the highest-ranked variables. In particular, the reduction in the RMSE that takes place as a consequence of the news defined in block 6, is now smaller, and subsequent updates do not improve the forecasts

[18]See Harvey and Koopman (2003) for details on the calculation of observation weights.

[19]Camacho and Pérez-Quirós (2010), for example, exploited the same idea and found, within the context of their 'Eurosting' model, that the euro area GDP forecast obtained on 24 January 2007 for the first quarter of that year was fully driven by the NBB Business Survey, simply because it was the only indicator available for January. This can be misleading because that figure does not necessarily change the forecast and it could have been anticipated to some extent on the basis of other indicators that were available.

anymore.

The phenomenon reported above can be easily explained. Our dynamic factor model was originally estimated via maximum likelihood using the full set of variables. Thus, the factors of the model, which is only an approximate representation of the data, are determined in such a way that they help to match the dynamics of all series. Those that turn out to have a smaller weight for GDP, which we have discarded along the lines of Rünstler (2016), may still help to forecast the series that actually have a big impact for GDP. Given the complex interactions between all series, which we aim to capture with a 4-factor model, using the standard impacts on GDP as a criterion to reduce the set of indicators may not be a good idea.

[Insert Figure 11 here]

6 Conclusion

This paper provides a formal way of quantifying the incremental value of alternative business cycle indicators that are often monitored for nowcasting GDP growth in the euro area. The objective is to rank all those indicators according to their importance for the growth forecasts in order to facilitate the complex task of interpreting such a heterogeneous and asynchronous flow of data releases. To do so, we propose a state-space representation for a dynamic factor model with a sufficiently large number of factors in order to account for the joint dynamics of all the series, which are constructed using the first available data vintage as reflected in the original press releases. Such a representation enables us to define the news or unexpected component of each data release in terms of the Kalman filter innovations. Their precise impact on the GDP growth estimate is determined by both the timeliness and quality of the news, which is captured by the Kalman gain. Thus, GDP forecasts are updated after interpreting the news – or the 'surprise' component– of each data release. The model-based news is conceptually equivalent to the forecast errors made by analysts monitoring the data releases.

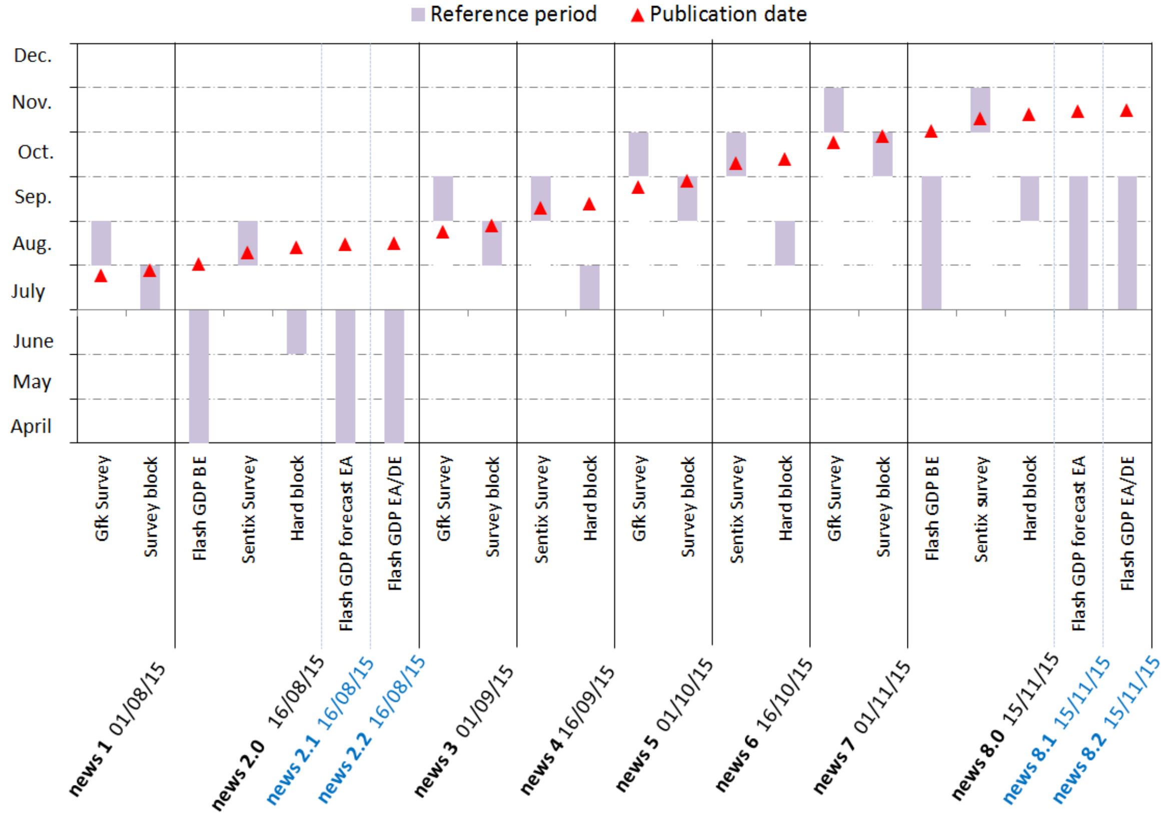
There are various aspects of our methodology that represent a novel and powerful approach to think about the real-time impact of predictor variables on a given target. First, we believe that time series econometrics that rely on the revised history of a given indicator (i.e. the pseudo real-time analysis) cannot answer the same question due to the large size of data revisions in certain series such as industrial production or sales. Our empirical results are based on time series constructed from real-time press releases in order to correct for artificial correlation patterns that may be present in historical time series in surveys and, especially, in hard data. This implies that in contrast to standard evaluations based on the historical time series available at a given point in time, our results will not be artificially distorted by revisions, such as seasonal adjustments, redefinitions or reweighing, which are typically incorporated with the benefit of hindsight.

Second, the formulation of the research question in this paper is very specific and rather unique. We aim to determine the *incremental* information content of a given press release, rather than of a block of releases as is the case in, for example, Gayer *et al.* (2015). While earlier literature has indeed already confirmed the importance of a given block of certain indicators (e.g. surveys) in order to reduce forecast errors, this information does not provide us with any clue regarding the impact of each one of the elements in the block. After all, a few particular surveys may be determining the performance of the whole block.

More specifically, we find that, in the process of updating nowcasts for euro area GDP growth, the strongest impact corresponds to the Markit Manufacturing PMI and the Business Climate Indicator for the euro area, followed by the IFO Business Climate and IFO Expectations for Germany. Interestingly, the NBB's own business confidence indicator for Belgium is following closely those survey variables and obtains seventh place in the overall ranking – a position that improves when targeting 'final' euro area GDP. More generally, it is quite remarkable that none of the consumer confidence indicators feature in the overall ranking. These findings have not been presented before. When it comes to hard data, euro area industrial production occupies fifth place in that ranking and is actually the only hard variable that makes it into the top ten. In the counterfactual scenario where hard data for a given month is released exactly at the end of that month, industrial production in the euro area and Germany would rank first and third, respectively, while the Manufacturing PMI for the euro area would still have the second largest impact. This suggests that, in addition to being available in a more timely manner, survey data also contain relevant information that is not captured by hard data. Having an overview of the impact of the different data releases helps to make the models more transparent for the user, but as shown in Section 5, it is not necessarily a good idea to simply rely on our ranking as a data selection tool when producing estimates with formal nowcasting models.

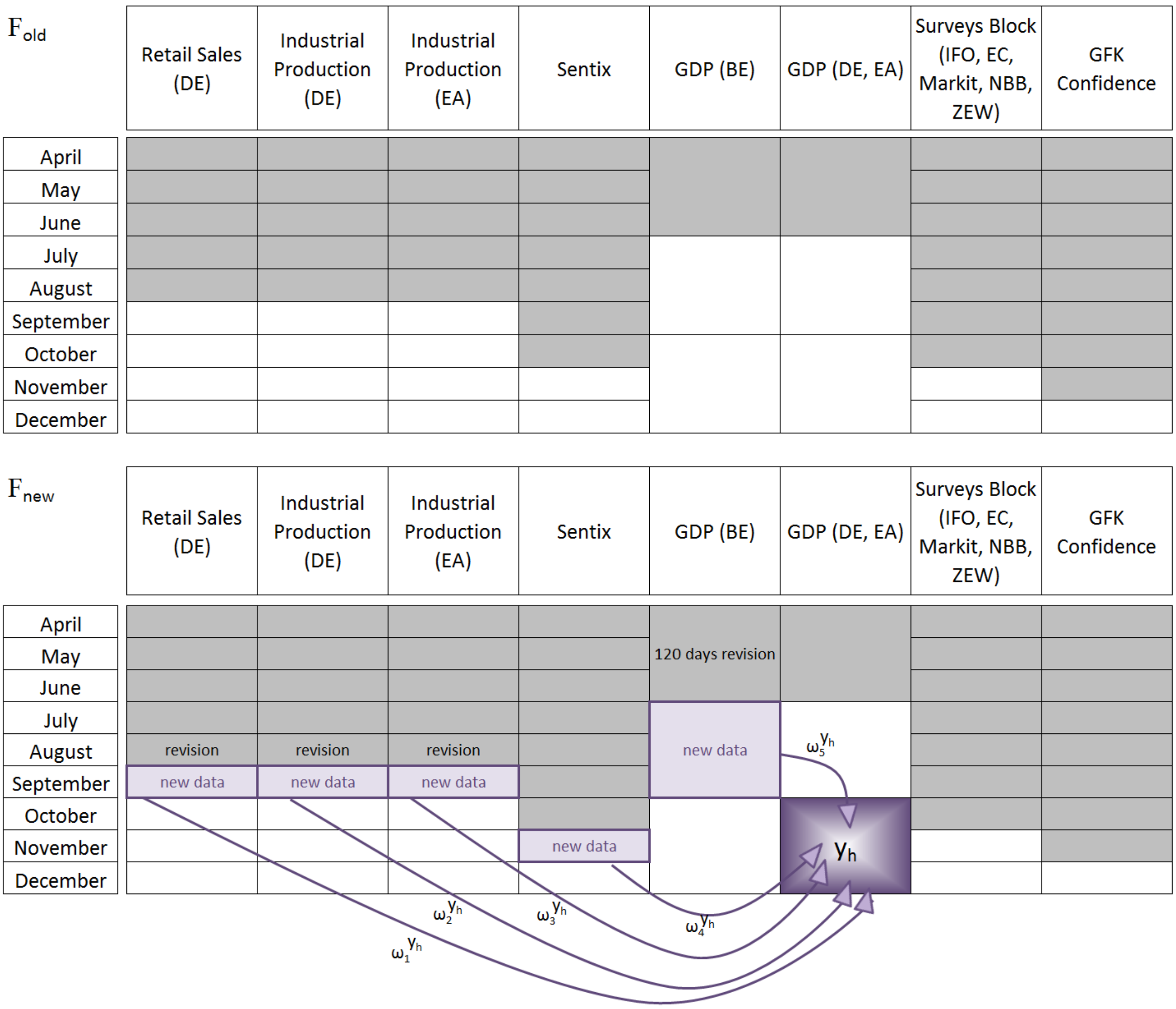
Figures and Tables

Figure 1: Representation of the Dataflow



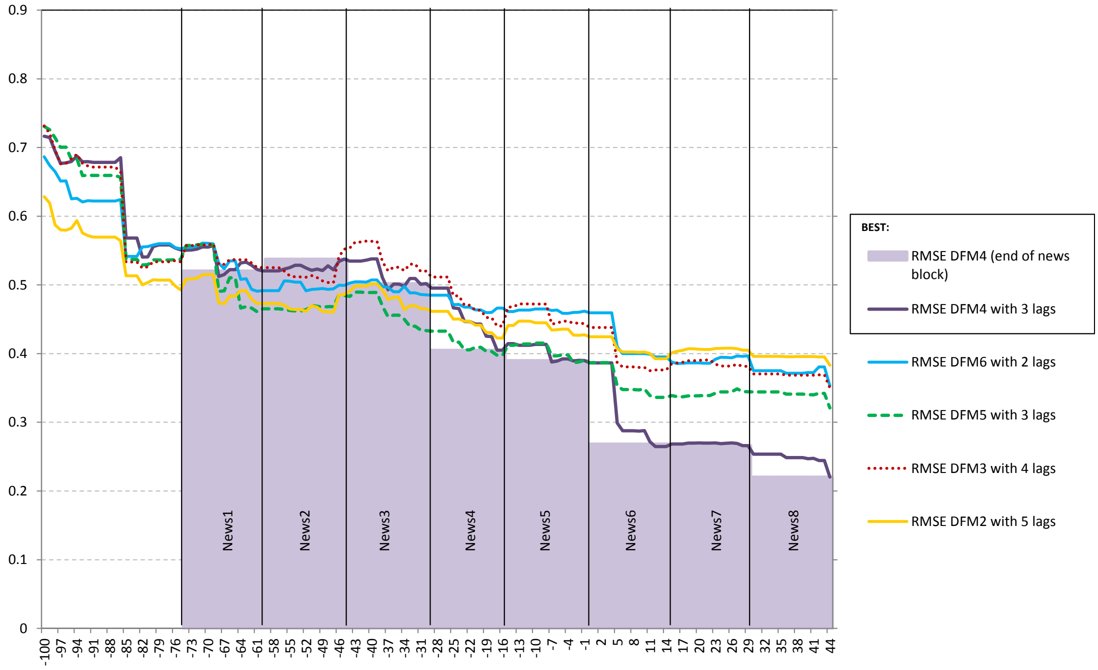
This figure represents the arrival of macroeconomic news that market participants are faced with. Although the data are actually released in a continuous manner, we aim to simplify the analysis by proposing a regular updating scheme that takes place only twice per month: around the middle of each month, i.e. after most hard indicators are published, and two weeks later, i.e. when most surveys have been released. Notice that the flash GDP for the euro area in this graph still refers to the flash with a publication delay of 45 days, and does not yet incorporate the recently introduced *preliminary flash release* with a delay of only 30 days. In this paper, we define the news as the forecast errors obtained for all variables when the forecasts are updated twice a month. We will also emphasise the news content of the GDP releases for the euro area and Germany, which take place 45 days after each quarter ends.

Figure 2: *News 8.0* Block's weights for updating euro area GDP for next quarter



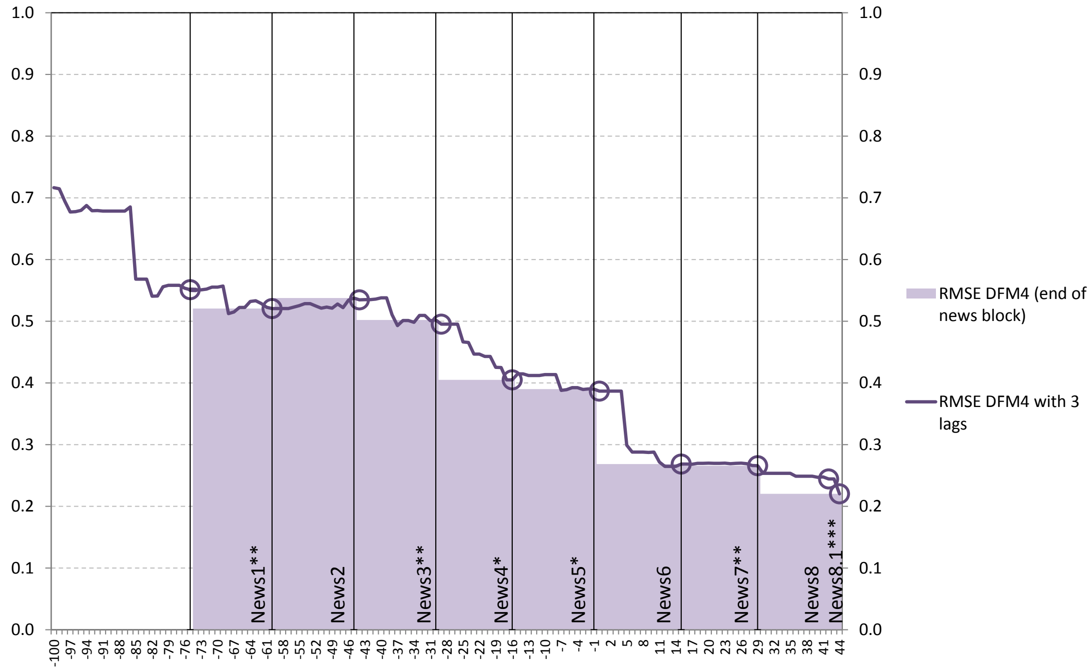
This figure represents two consecutive information sets and the weights associated to the news represented in Figure 1 as news 8.0. Note that the weight's subindex corresponds to each one of the pieces of news, while the upper index refers to the target variable, which is next quarter's GDP growth (y_h). Thus, the weight of those elements of news at updating current quarter's GDP would be represented with the superindex y_{h-1}. Although press releases may contain revisions for certain (hard) indicators, these will not be taken on board in our empirical application as we are only interested in the first vintages (i.e. new data).

Figure 3: RMSE over 2006Q1-2015Q1 for different models and as a function of the expanding information set



This graph shows that the DFM4 with 3 lags performs best in terms of forecasting accuracy, relative to other models tested that include more or less factors and lags.

Figure 4: Updates causing a significant change in RMSE over 2006Q1-2015Q1 for the selected DFM4



This graph shows which blocks of news updates cause the RMSE of the DFM4 to significantly change. *, ** and *** are used to indicate significance at the 5%, 10% and 20% level using the fixed-smoothing (FS) asymptotics, as proposed by Coroneo and Iacone (2015). More detailed results can be found in Table A.2 in Annex I.

Figure 5: Euro area flash GDP and the model nowcasts compared to the Bloomberg forecast

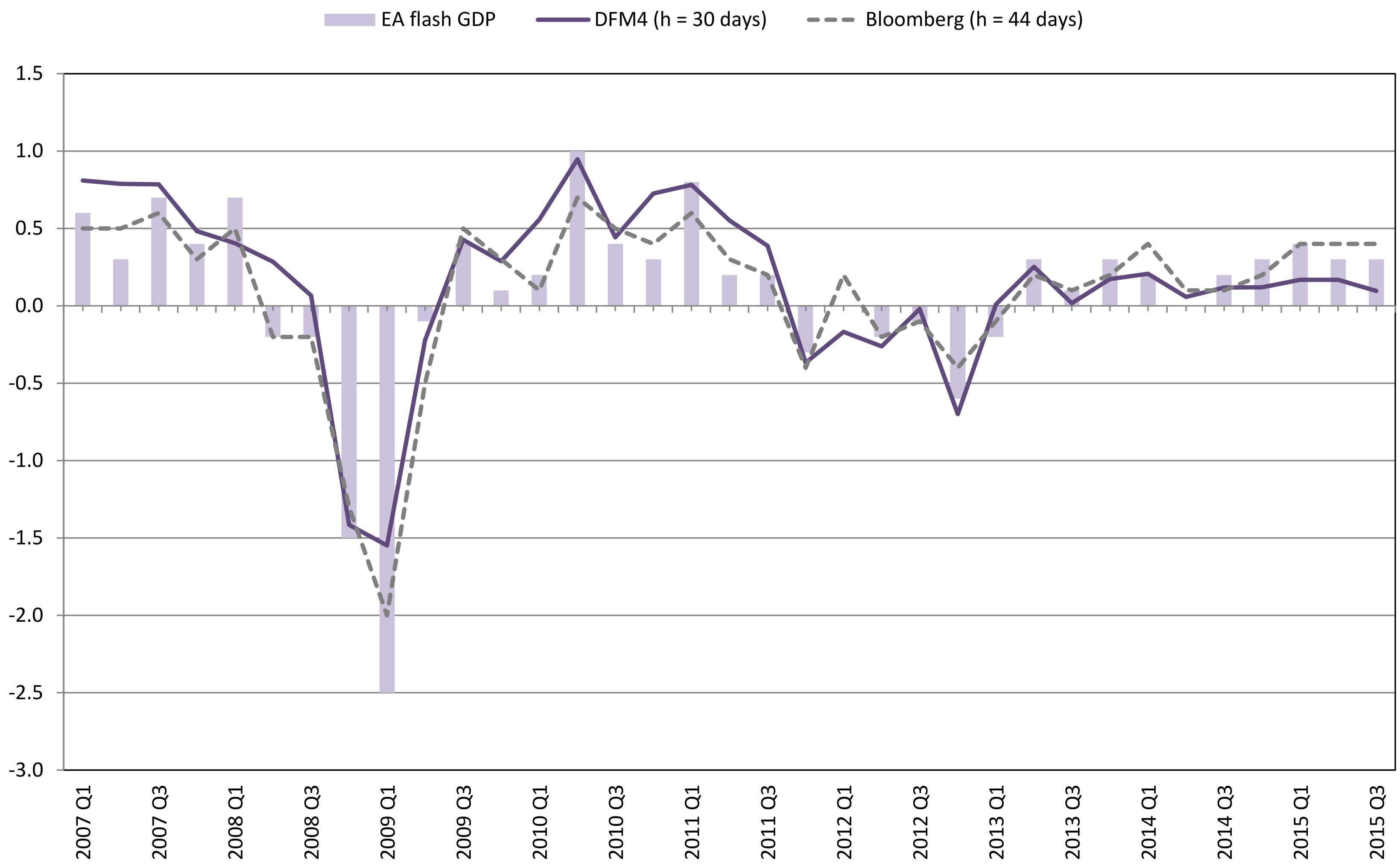
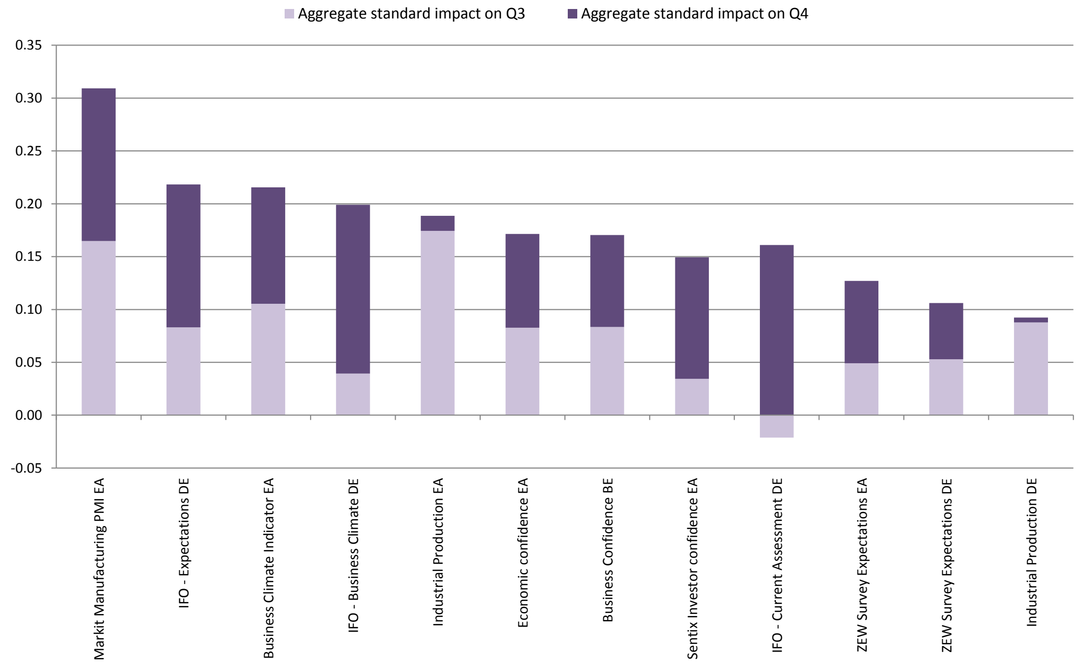


Figure 6: Ranking According to the Standard Impacts for euro area GDP Flash

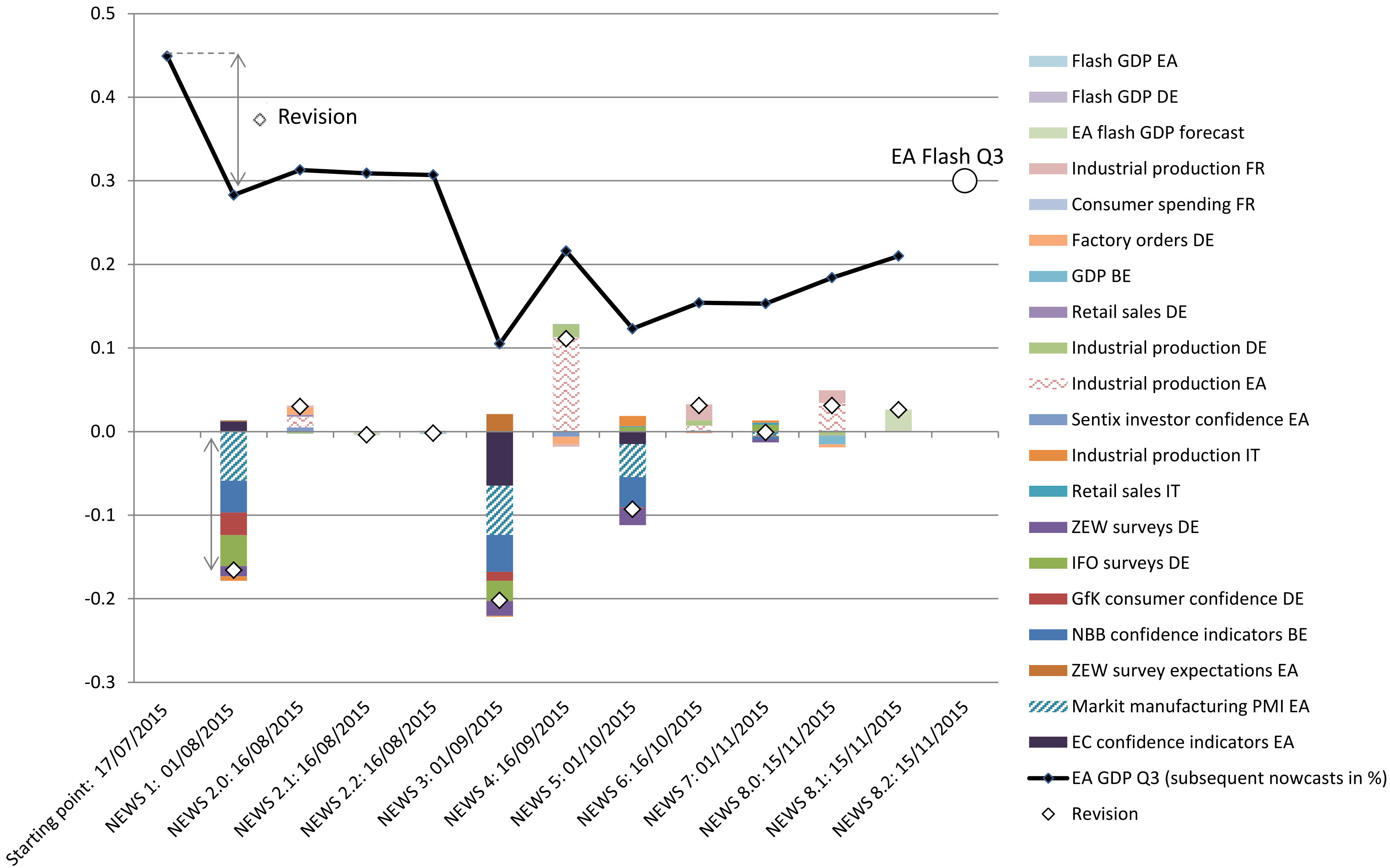


This graph provides some sort of a 'horizontal summation' of the impacts that occur in Figure A.3, i.e. for every indicator, the sum is made of the impact of the news that is being read between the first of August and mid-November. Only the twelve highest-ranked indicators are shown. The distribution of the colours in each bar (light vs dark) also gives a first indication about the timeliness of the indicator. The fact that news about the hard data mainly impacts the expectation for Q3 is a reflection of the release calendar (Figure A.2), as hard data that can be read by the model between August and mid-November actually still refer to the months between May and September.

Table 1: Robustness Analysis

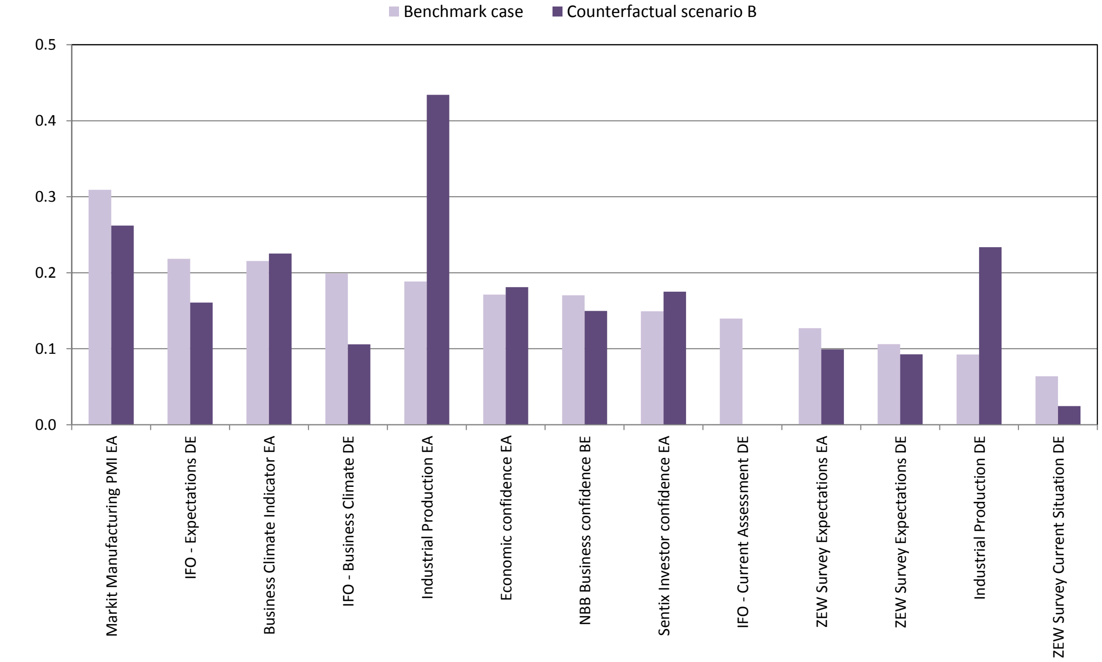
Analysis	Objective	Details	Results
A	Impacts on euro area GDP flash (see subsection 4.1)	Real-time dataflow	Figures A.3 and 6
B	Counterfactual impact on euro area GDP flash	Hard data published without any delay	Figure 8
C	Counterfactual impact of revised hard data on revised euro area GDP	Hard data are revised, but published according to the actual real-time calendar	Figure 9
D	Counterfactual impact of revised hard data on revised euro area GDP	Hard data are fully revised *and* published without any delay	Figure 10

Figure 7: Real-Time Updates for 2015Q3



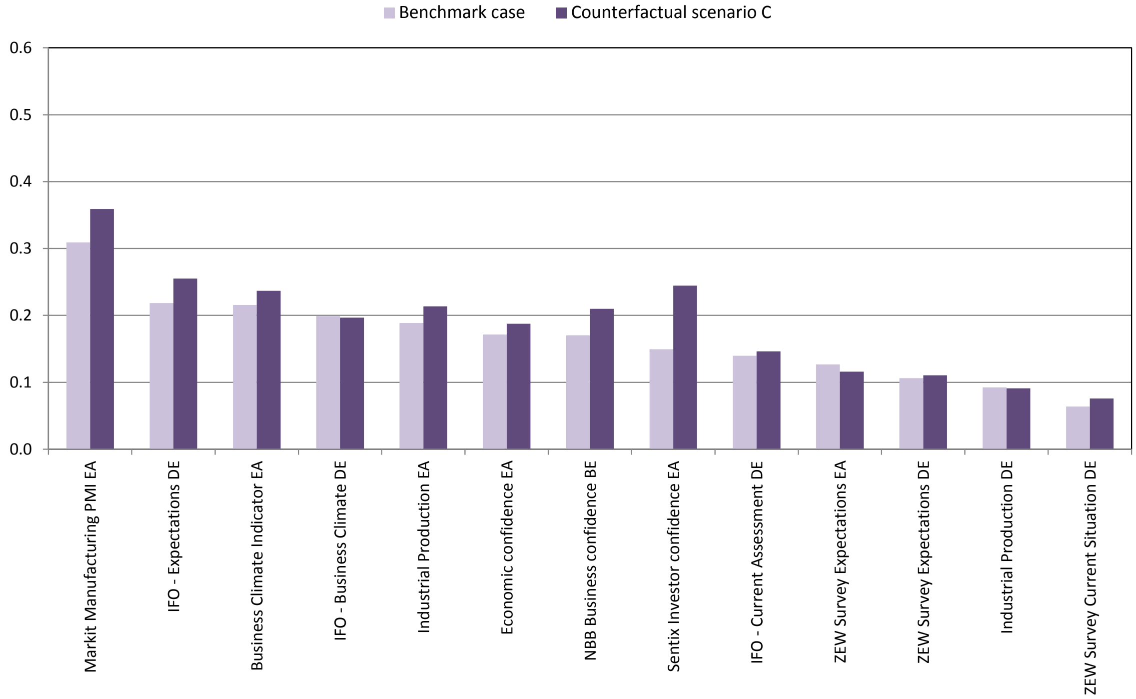
This figure represents the updating process of the GDP nowcast every time news arrives. In order to facilitate the reading of the figure, certain variables were taken together (e.g. IFO surveys DE is compiled of three separate IFO indicators).

Figure 8: Ranking According to the Counterfactual B (timeliness) Standard Impacts for euro area GDP



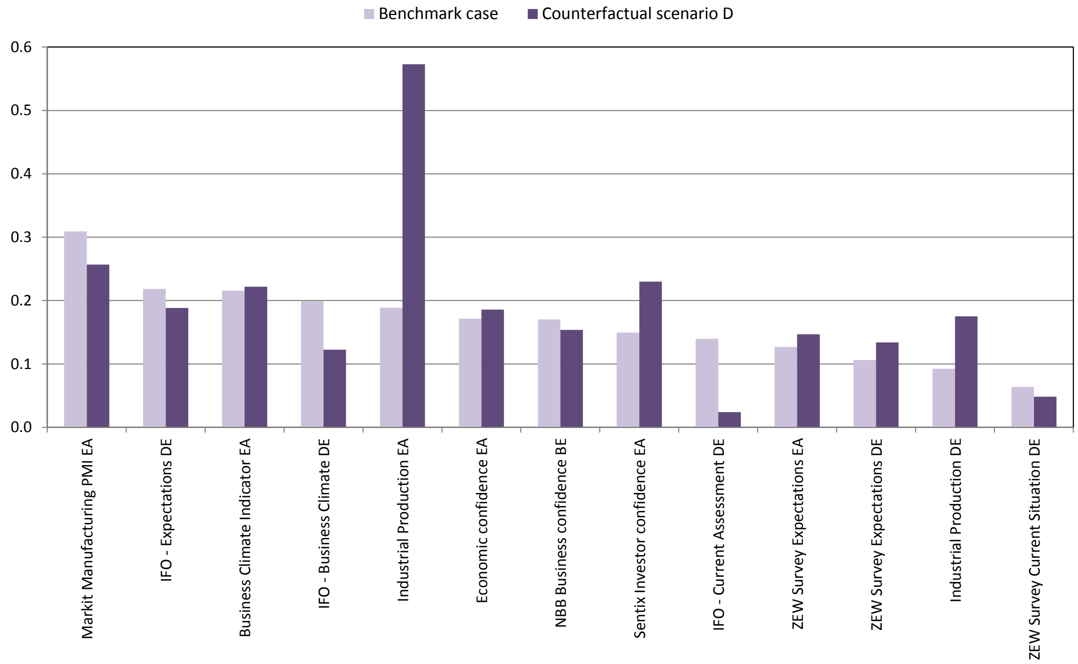
This figure shows the results of the benchmark case (light purple) and those of robustness analysis B (dark purple). Note that each bar represents the aggregated standard impact over an entire semester (i.e. the impact on both Q3 and Q4).

Figure 9: Ranking According to the Counterfactual C (revisions) Standard Impacts for euro area revised GDP



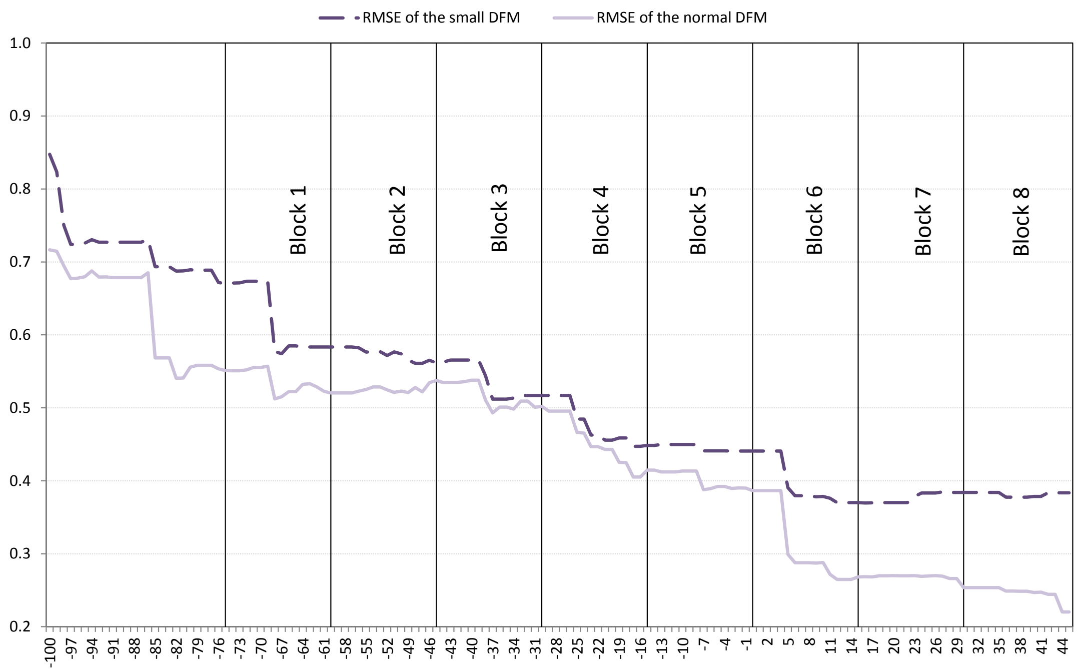
This figure shows the results of robustness analysis C (cf. scenarios described in Table 1), compared to the benchmark results. Note that each bar represents the aggregated standard impact over an entire semester (i.e. the impact on both Q3 and Q4).

Figure 10: Ranking according to the Counterfactual D (timeliness + revisions) Standard Impacts for *revised* euro area GDP



This figure shows the results of robustness analysis D (cf. scenarios described in Table 1), compared to the benchmark results. Note that each bar represents the aggregated standard impact over an entire semester (i.e. the impact on both Q3 and Q4).

Figure 11: RMSE functions of the normal model and the small model



References

[1] Abberger, K. (2007). "Qualitative business surveys and the assessment of employment – a case study for Germany". *International Journal of Forecasting*, **23**, 249-258.

[2] Angelini, E., G. Camba-Mendez, D. Giannone, G. Rünstler and L. Reichlin (2011). "Short-Term Forecasts of Euro Area GDP Growth". *The Econometrics Journal*, **1**, C25-C44.

[3] Aruoba, S.B., F.X. Diebold and C. Scotti (2009). "Real-Time Measurement of Business Conditions". *Journal of Business and Economic Statistics*, American Statistical Association, **27(4)**, 417-427.

[4] Bańbura, M. and M. Modugno (2010). "Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data". *ECB Working Paper Series*, **1189**

[5] Bańbura, M., D. Giannone and L. Reichlin (2011). "Nowcasting". *Oxford Handbook on economic Forecasting*, in Michael P. Clements and David F. Hendy.

[6] Bańbura, M., D. Giannone, M. Modugno and L. Reichlin (2013). "Nowcasting and the Real-Time Data Flow". *Handbook of Economic Forecasting*, in G. Elliot and A. Timmermann, **2**, Elsevier.

[7] Banbura, M. and G. Rünstler (2011). "A look into the factor model black box: publication lags and the role of hard and soft data in forecasting GDP". *International Journal of Forecasting*, **27(2)**, 333-346.

[8] Bragoli, D., L. Metelli and M. Modugno (2015). "The Importance of Updating: Evidence from a Brazilian Nowcasting Model". *OECD Journal: Journal of Business Cycle Measurement and Analysis*, **1**, 5-22.

[9] Bragoli, D. and M. Modugno (2017). "A Now-Casting Model for Canada: Do U.S. Variables Matter?". *International Journal of Forecasting*, **33**, 786–800.

[10] Bragoli, D. (2017). "Now-casting the Japanese economy". *International Journal of Forecasting*, **33**, 390-402.

[11] Camacho, M. and G. Pérez-Quirós (2010). "Introducing the Euro-Sting". *Journal of Applied Econometrics*, **25**, 663-694.

[12] Claveria, O., Pons, E. and R. Ramos (2007). "Business and consumer expectations and macroeconomic forecasts". *International Journal of Forecasting*, **23**, 47-69.

[13] Coroneo, L. and F. Iacone (2016). "Comparing predictive accuracy in small samples using fixed-smoothing asymptotics". *Discussion Papers University of York*, **15/15**.

[14] Croushore, D. and T. Stark (2002). "Forecasting with a Real-Time Data Set for Macroeconomists". *Journal of Macroeconomics*, **24**, 507-531.

[15] D'Agostino, K. McQuinn and D. O'Brien (2013). "Nowcasting Irish GDP". *OECD Journal: Journal of Business Cycle Measurement and Analysis*, **7**, 21–31.

[16] de Antonio Liedo, D. (2015). "Nowcasting Belgium". *Eurostat review on National Accounts and Macroeconomic Indicators* , **75**, 7-48

[17] Dempster, A.P., N.M. Laird and D.B. Rubin (1977). "Maximum Likelihood from Incomplete Data via the EM Algorithm". *Journal of the Royal Statistical Society. Series B (Methodological)*, **39:1**, 1-38.

[18] Diebold, F.X. and R.S. Mariano (1995). "Comparing Predictive Accuracy". *Journal of Business and Economic Statistics*, **13**, 253-265.

[19] Diron, M. (2008). "Short-Term Forecasts of Euro Area Real GDP Growth: An Assessment of Real-Time Performance Based on Vintage Data". *Journal of Forecasting*, **27**, 371-390.

[20] Doz, C., D. Giannone and L. Reichlin (2012). "A Quasi Maximum Likelihood Approach for Large Approximate Dynamic Factor Models". *Review of Economics and Statistics*, **94**, 1014-1024.

[21] Durbin, J and S.J. Koopman (2001). "Time Series Analysis by State Space Methods ". Oxford University Press.

[22] European Central Bank (2011). "The measurement and prediction of the euro area business cycle ". *Monthly Bulletin*, May 2011.

[23] Evans, M.D.D. (2005). “Where are we now? Real-Time Estimates of the Macroeconomy”. *International Journal of Central Banking*, **1**.

[24] Frale, C., M. Marcellino, G.L. Mazzi, and T. Proietti (2011). “ EUROMIND: a monthly indicator of the euro area economic conditions”. *Journal of the Royal Statistical Society: Series A*, **174**, 439-470.

[25] Franta, D., D. Havrlant and M. Runsák (2016). “Forecasting Czech GDP Using Mixed-Frequency Data Models”. *Journal of Business Cycle Research*, **12**, 165-185.

[26] Gayer, A., A. Girardi and A. Reuter (2015). “The role of survey data in nowcasting euro area GDP growth”. *Journal of Forecasting*, **35**, 400-418.

[27] Giannone, D., L. Reichlin and D. Small (2008). “Nowcasting: The Real-Time Informational Content of Macroeconomic Data Releases”. *Journal of Monetary Economics*, **55**, 665-676.

[28] Giannone, D., L. Reichlin and S. Simonelli (2009). “Nowcasting Euro Area Economic Activity in Real Time: The Role of Confidence Indicators”. *National Institute Economic Review*, **210**, 90-97.

[29] Harvey, D., S. Leybourne and P. Newbold (1997). “Testing the equality of prediction mean squared errors”. *International Journal of Forecasting*, **13**, 281-291.

[30] Kalman, R.E. (1960). “A new approach to linear filtering and prediction problems”. *Journal of Basic Engineering*, **82**, 35-45.

[31] Kiefer, N.M. and T.J. Vogelsang (2005). “A new asymptotic theory for heteroskedasticity-autocorrelation robust tests”. *Econometric Theory*, **21(6)**, 1130-1164.

[32] Kishor, N.K. and E.F. Koenig (2012). “VAR Estimation and Forecasting When Data are Subject to Revision”. *Journal of Business and Economic Statistics*, **30**, 182-190.

[33] Lui, S., Mitchell, J. and M. Weale (2011).“Qualitative business surveys: signal or noise?”. *Journal of the Royal Statistical Society: Series A*, **174**, 327-348.

[34] Mariano, R.S. and Y. Murasawa (2003). "A new coincident index of business cycles based on monthly and quarterly series". *Journal of Applied Econometrics*, **18**, 427-443.

[35] Martinsen, K., F. Ravazzolo and F. Wulfsberg (2014). "Forecasting Macroeconomic Variables Using Disaggregate Survey Data". *International Journal of Forecasting*, **30**, 65-77.

[36] Modugno, M., B. Soybilgen and E. Yazgan (2016). "Nowcasting Turkish GDP and News Decomposition". *International Journal of Forecasting*, **32**, 1369-1384.

[37] Newey, W.K. and K.D. West (1987). "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix". *Econometrica*, **55(3)**, 703-708.

[38] Patton, A.J. and A. Timmerman (2012). "Forecast rationality tests based on multi-horizon bounds". *Journal of Business and Economic Statistics*, **30(1)**.

[39] Piette, C. and G. Langenus (2014). "Using BREL to now-cast the Belgian business cycle: the role of survey data". *Economic Review*, National Bank of Belgium, 75-98.

[40] Rünstler, G. (2016). "On the Design of Datasets for Forecasting with Dynamic Factor Models". *ECB Working Paper Series*, **1893**

[41] Shumway, R. and D. Stoffer (1982). "An approach to time series smoothing and forecasting using the EM algorithm". *Journal of Time Series Analysis*, **3**, 253-264.

[42] The Wall Street Journal (1999) "Euroland discovers a surprise indicator: Belgian confidence", 14 July.

[43] Williamson, C. (2015). "Using PMI survey data to predict official eurozone GDP growth rates". *Markit Economic Research*, 17 August 2015.

Annex I - Evaluating Forecasting Accuracy

The prediction errors are defined with a reference i to the information set available at the time the forecast was made:

$$e_{t|i} = y_t - \hat{y}_{t|\mathcal{F}_i} \tag{11}$$

where $\mathcal{F}_i$ need not only include lags of y_t. In practice, the information that will be actually used may be a small subset of $\mathcal{F}_i$.

The properties of these forecast errors can be assessed in isolation or relative to a benchmark, which we will define as $\breve{e}_{t|i}$. The benchmark may be a naive forecast, e.g. random walk, in which case $\breve{y}_{t|\mathcal{F}_i}$ would be equal to $\breve{y}_{t|y_{t-1}} = y_{t-1}$. However, the benchmark could also be a prediction that is regularly published by a forecasting institute or market analysts, i.e. Bloomberg, which is not necessarily model-based. In that case, $\breve{y}_{t|\mathcal{F}_i}$ would be given by methods and a subset of $\mathcal{F}_i$ which is unknown to us.

For model-based forecasts, we use the following notation: $\hat{y}_{t|\mathcal{F}_i} = E_\theta[y_t|\mathcal{F}_i]$ to highlight the fact that they are based on model-consistent expectations given by the parameter vector θ.

In forecasting comparisons involving competing forecasts resulting from the same information set, the subindex i will be removed because it does not play a role. We will first test the following hypotheses involving forecast errors:

$$Unbiasedness: \quad E[e_t] = 0 \tag{12}$$

$$Autocorrelation: \quad E[e_t e_{t-1}] = 0 \tag{13}$$

$$Equality\ in\ squared\ errors: \quad E[e_t^2 - \breve{e}_t^2] = 0 \tag{14}$$

$$Equality\ in\ absolute\ errors: \quad E[|e_t| - |\breve{e}_t|] = 0 \tag{15}$$

$$Forecast\ \hat{y}_t\ encompasses\ \breve{y}_t: \quad E[(e_t - \breve{e}_t)e_t] = 0 \tag{16}$$

$$Forecast\ \breve{y}_t\ encompasses\ \hat{y}_t: \quad E[(\breve{e}_t - e_t)\breve{e}_t] = 0 \tag{17}$$

An overview of the tests can also be found in Table A.1.

[Insert Table A.1 here]

Table A.1: Forecasting Evaluation Tests

Test	Null hypothesis	Statistic	Asym. theory	Finite sample
Bias	$E[e_t] = 0$	$\mathrm{B} = \dfrac{\bar{e}}{\sqrt{\dfrac{2\pi \hat{f}_e(0)}{T}}}$	$N(0,1)$	KV(2005)
Autocorrelation	$E[e_t e_{t-1}] = 0$	$\mathrm{AR} = \dfrac{\bar{\rho}}{\sqrt{\dfrac{2\pi \hat{f}_\rho(0)}{T}}}$	$N(0,1)$	KV(2005)
Diebold-Mariano	$d_t \equiv L_{1,t} - L_{2,t} = 0$	$\mathrm{DM} = \dfrac{\bar{d}}{\sqrt{\dfrac{2\pi \hat{f}_d(0)}{T}}}$	$N(0,1)$	KV(2005)
Encompassing 1	$d^e_{1,t} \equiv E[(e_t - \breve{e}_t)e_t] = 0$	$\mathrm{E}_1 = \dfrac{\bar{d}_1}{\sqrt{\dfrac{2\pi \hat{f}_{d_1}(0)}{T}}}$	$N(0,1)$	KV(2005)
Encompassing 2	$d^e_{2,t} \equiv E[(\breve{e}_t - e_t)\breve{e}_t] = 0$	$\mathrm{E}_2 = \dfrac{\bar{d}_2}{\sqrt{\dfrac{2\pi \hat{f}_{d_2}(0)}{T}}}$	$N(0,1)$	KV(2005)

Diebold-Mariano Test

The test originally proposed by Diebold and Mariano (1995) considers a sample path of loss differentials $\{d_t\}_{t=1}^{T}$. In the case of a squared loss function, we have $d_t = e_t^2 - \breve{e}_t^2$. Under the assumption that the loss differential is a covariance stationary series, the sample average, $\bar{d}$, converges asymptotically to a normal distribution:

$$\sqrt{T}\bar{d} \ \underset{\rightarrow}{d} \ N(\mu, 2\pi f_d(0)) \tag{18}$$

In particular, they proposed to test the null hypothesis that the forecast errors coming from the two forecasts bring about the same loss: $E[e_t^2 - \breve{e}_t^2] = 0$ against the two-sided alternative. Thus, the resulting p-values represent the probability of obtaining the realized forecast error differential or a more extreme one in a new experiment if the null hypothesis was actually true. The test-statistic that will be used to calculate our p-values is computed as follows:

$$DM = \frac{\bar{d}}{\sqrt{\dfrac{2\pi \hat{f}_d(0)}{T}}} \tag{19}$$

where $2\pi\hat{f}_d(0)$ is a consistent estimate of the variance of $\bar{d}$. Consider $2\pi\hat{f}_d(0) = \sum_{\tau=-(T-1)}^{(T-1)} w_\tau \gamma_d(\tau)$, where $\gamma_d(\tau) = \dfrac{1}{T}\sum_{t=|\tau|+1}^{T}(d_t - \bar{d})(d_{t-|\tau|} - \bar{d})$. Under the assumption that $\gamma_d(\tau) = 0$ for $\tau \geq h$, we can use a rectangular lag window estimator by setting $w_\tau = 0$ for $\tau \geq h$. Another option is to use the Heteroscedasticity and Autocorrelation Consistent (HAC) estimator proposed by Newey and West (1987). In this case, the weights could be given by a triangular window, $w_\tau = 1 - \dfrac{\tau}{h}$ for $\tau < h$. In this case, however, the consistency property only remains valid when the truncation lag h or bandwidth is a function of the sample size T.

The idea is to test the statistical significance of the regression of $e_t^2 - \breve{e}_t^2$ on an intercept. In order to determine the statistical significance of the intercept, its associated standard errors need to take into account the autocorrelation patterns of the regression error, which are considered in the denominator of equation (19). *JDemetra+* exploits the same unified framework to conduct all tests listed in Table A.1. But given the small sample sizes that are typical in real-time forecasting applications, which leads to an over-rejection of the null hypothesis, we follow the fixed-smoothing asympotics proposed by Coroneo and Iacone (2015) exploiting the finite sample distributions of Kiefer and Vogelsang (2005). The distribution of the test statistic (19) will depend on kernel (triangular in our case) and the bandwidth chosen, which is set by default equal to $T^{0.5}$, as suggested by Coroneo and Iacone (2015). The results can be very different than those resulting from the traditional asymptotic theory, where the test statistic would have the same distribution under the null independently of the kernel and the bandwidth used.

Tables A.2 and A.3 contain the results of this test together with the encompassing test and two efficiency tests, which are described below.

Encompassing Test

Independently of whether the null hypothesis $E[e_t^2 - \breve{e}_t^2] = 0$ is rejected or not, it is relevant to understand to what extent our model encompasses all the relevant information of the benchmark, and the other way around. Because of the obvious symmetry of both statements, we consider only the first one. If our forecast $y_{t|\mathcal{F}_i}$ encompasses a given benchmark $\breve{y}_{t|\mathcal{F}_i}$, the difference between those benchmark forecasts and ours will not be a relevant factor in explaining our own forecast error. In other words, the regression coefficient λ will not be significantly different from zero in the following regression:

$$\underbrace{y_t - y_{t|\mathcal{F}_i}}_{e_t} = \lambda \underbrace{(\breve{y}_{t|\mathcal{F}_i} - y_{t|\mathcal{F}_i})}_{e_t - \breve{e}_t} + \xi_t \tag{20}$$

$$\Updownarrow$$

$$y_t = \lambda \breve{y}_{t|\mathcal{F}_i} + (1-\lambda) y_{t|\mathcal{F}_i} + \xi_t \tag{21}$$

Following Harvey, Leybourne and Newbold (1997), the statistical significance of the λ coefficient in regression 20 can be used to reject the null hypothesis that our model encompasses the benchmark. In this case of rejection, equation 21 suggests that a combination of the two forecasts would yield a more informative forecast.

By construction, the value of the coefficient of a regression $\breve{e}_t = \alpha(\breve{e}_t - e_t) + \xi_t$ is equal to $1 - \lambda$, but it is not necessarily true that the *rejection* of the null hypothesis in the first case implies the *acceptance* of the symmetric statement.

The test-statistic is computed as follows. When the null hypothesis is that our model encompasses the benchmark, we define the sequence $\{d_t\}_{t=1}^{T}$, where $d_t = e_t(e_t - \breve{e}_t)$, and we compute $E1 = \frac{\bar{d}}{\sqrt{\frac{2\pi \hat{f}_d(0)}{T}}}$, exactly as in equation 19.

Efficiency: Bias Test

In order to assess whether our forecasts are unbiased, we will simply test the statistical significance of the average error. In some cases, the time series of forecast errors $\{e_t\}_{t=1}^{T}$ may be autocorrelated to some extent even when they are based on a model with IID innovations. In such cases, the variance associated to the estimate of the average forecast error may be large. The test statistic has exactly the same form as the previous tests discussed so far.

Efficiency: Autocorrelation Test

We will test here a second necessary condition for our forecasts to be efficient: absence of autocorrelation. In the same spirit as the tests described above, we will assess the statistical significance of the forecast errors' autocorrelation. Thus, our sequence $\{d_t\}_{t=1}^{T}$ will be defined with $d_t = e_t e_{t-1}$.

Testing the rationality of nowcasting updates

Patton and Timmerman (2012) suggest testing whether the mean squared forecast error is actually decreasing when the horizon decreases. This idea could be applied in our set-up by replacing the concept of forecast horizon with the number of days from the moment in which we update the forecasts for GDP until the day it is realized, i.e. the release date.

In our set-up, the size and power of that test would be too much dependent on the number of times we update the model. We can update it every time we have a new data release, or update it every two weeks, for example. However, what is relevant for us is not whether the model produces rational multi-horizon forecasts, which is likely because they are based on a unique model with parameters obtained via maximum likelihood. Instead, we ask what are the forecasting updates that are most likely to yield significant improvements in forecasting accuracy. The results are available in Table A.2.

Application: Evaluating the forecasts of our DFM

The tests described in Table A.1 are applied to two cases. In the first case, the aim is to determine which news blocks lead to a significant improvement of the forecasting accuracy. Results displayed in Table A.2 include bias, autocorrelation, RMSE and the λ coefficient defined above, which is the weight given to a benchmark forecasts that competes with our model's. Statistical significance is highlighted with shades. Grey shaded areas in column FS-DM demonstrate which news blocks have induced a significant change in the RMSE of the model, i.e. the null of equal accuracy between old (O) and updated (U) forecasts is rejected. The outcome of the DM test may be considered jointly with the results of the encompassing tests. For a certain news block to be considered relevant, the corresponding nowcasting update (U) should hold a larger amount of information than the older nowcast (O) based on the previous information set, while the old nowcast does not incorporate any useful information absent in the new update. The last two colums of the table show that this is generally the case, with some exceptions. That is, the null *U encompases O* is not rejected while *O encompasses U* is rejected.

In the second case, displayed in Table A.3, we compare the forecasting accuracy of our dynamic factor model (labeled DFM in the table) with that of relevant benchmarks in the field (e.g. now-casts from the web based service Now-Casting.com, PMI-based forecasts and Bloomberg expectations). Once again, the DM test may be considered together with the encompassing test. Ideally, the nowcasts from our DFM should encompass the information contained in competing forecasts, and not the other way round. Thus, the grey shaded areas in the first column (i.e. DFM encompasses Benchmark) show that the null hypothesis can be always rejected and therefore it is not true that the competitors do not add value. However, the inverse also holds (Benchmark encompasses DFM): the null that the benchmark nowcasts encompass our DFM forecast is also rejected, with only one exception. Hence, the forecasting accuracy of the now-casts could possibly be improved by combining the two information sets together.

[Insert Table A.2 here]

[Insert Table A.3 here]

Table A.2: Statistical significance of each update based on fixed-smoothing (FS) asymptotics

Evaluation period: 2007.Q1 - 2015.Q1, T=25

Real-Time Updates	**FS-Efficiency** bias	**FS-Efficiency** corr	**FS-DM** RMSE	**FS-Encompassing (U)pdate vs (O)ld** U enc O	**FS-Encompassing (U)pdate vs (O)ld** O enc U
ARIMA	-0.27	0.50	-	-	-
DFM -90 (d)ays	-0.22	0.41	0.68	0.60	0.39
DFM -75 d	-0.19	0.47	0.55	-0.60	1.59
DFM -60 d	-0.12	0.55	0.52	0.26	0.51
DFM -45 d	-0.14	0.54	0.54	1.48	-0.54
DFM -30 d	-0.08	0.58	0.50	-0.20	1.07
DFM -15 d	-0.13	0.46	0.41	-0.65	1.59
DFM 0 d (end of quarter)	-0.06	0.45	0.38	-0.13	0.82
DFM +15 d	-0.09	-0.11	0.27	-0.02	1.01
DFM +30 d	-0.07	-0.08	0.26	-0.39	1.23
DFM +42 d	-0.10	-0.06	0.26	0.27	0.66
DFM +44 d	-0.06	-0.18	0.23	-0.17	1.03

Note: The FS-Efficiency multicolumn of his table reports bias and autocorrelation for the forecast errors obtained at different horizons. The FS-DM and FS-Encompassing blocks should be considered simultaneously. They aim to determine for each forecasting update (U) whether there is any added value with respect to the old/last available forecast (O). The null hypothesis of the Diebold-Mariano (DM) test is rejected when the *difference in the squared errors of U and O* is significantly different from zero. For the two encompassing tests, the null hypothesis states that the updated forecast (U) encompasses all the relevant information from the old forecast (O) (*or vice versa*). When the null hypothesis can be rejected, this implies that *U can be improved by combining it with O*. The combination weight associated to O (*or U*) is therefore reported below the "U enc O" test. In order to assess the added value of the updated forecast, the DM null of equal forecast accuracy should be rejected and at the same time the null "U enc O" and "O enc U" should be, respectively, not rejected and rejected. Given the small size of our evaluation sample and the time-series correlation patterns, we determine significance at the 5%, 10% and 20% level using the fixed-smoothing (FS) asymptotics, as proposed by Coroneo and Iacone (2015).

Table A.3: Our DFM compared to competitive benchmarks

Evaluation period: 2011.Q3 - 2015.Q1, T=15

Now-Casting.com (N-C)

	FS-Efficiency		**FS-DM**	**FS-Encompassing DFM vs Benchmark**	
Nowcasts	bias	corr	Rel RMSE	DFM enc Bench	Bench enc DFM
DFM -45 d	0.00	-0.05	1.21		0.34
N-C -45 d	0.02	-0.22		0.66	
DFM 0 d	0.11	0.04	1.31		0.33
N-C 0 d	0.00	-0.13		0.43	
DFM +44 d	0.00	0.14	0.70		0.56
N-C +44 d	-0.08	0.58		0.28	

Bloomberg (BLO) and Markit rule (PMI)

	FS-Efficiency		**FS-DM**	**FS-Encompassing DFM vs Benchmark**	
Nowcasts	bias	corr	Rel RMSE	DFM enc Bench	Bench enc DFM
DFM 0 d	0.11	0.04	0.83		0.68
PMI 0 d	0.08	0.11		0.31	
DFM +44 d	0.00	0.14	1.13		0.37
BLO +44 d	-0.02	-0.09		0.60	

Note: The FS-Efficiency multicolumn of his table reports bias and autocorrelation for the forecast errors obtained at different horizons. The FS-DM and FS-Encompassing blocks should be considered simultaneously. They aim to determine whether forecasts based on the DFM and the corresponding benchmarks are significantly different. The null hypothesis of the Diebold-Mariano (DM) test is rejected when the *difference in the RMSE* is significantly different from zero. In this table, the relative RMSE, defined as the RMSE of the DFM divided by the RMSE of the benchmark, will indicate that the forecast performance of the DFM is better than that of the benchmark when the fraction is smaller than one. For the two encompassing tests, we reject the null hypothesis that the DFM encompasses all the relevant information from the benchmark (*or vice versa*) when *the DFM can be improved by combining it with the benchmark*. The combination weight associated to the benchmark (*or DFM*) is therefore reported below the "DFM enc Bench" test. In order to assess the added value of the DFM, the DM null of equal forecast accuracy should be rejected and at the same time the null "DFM enc Bench" and "Bench enc DFM" should be, respectively, not rejected and rejected. Given the small size of our evaluation sample and the time-series correlation patterns, we determine significance at the 5%, 10% and 20% level using the fixed-smoothing (FS) asymptotics, as proposed by Coroneo and Iacone (2015).

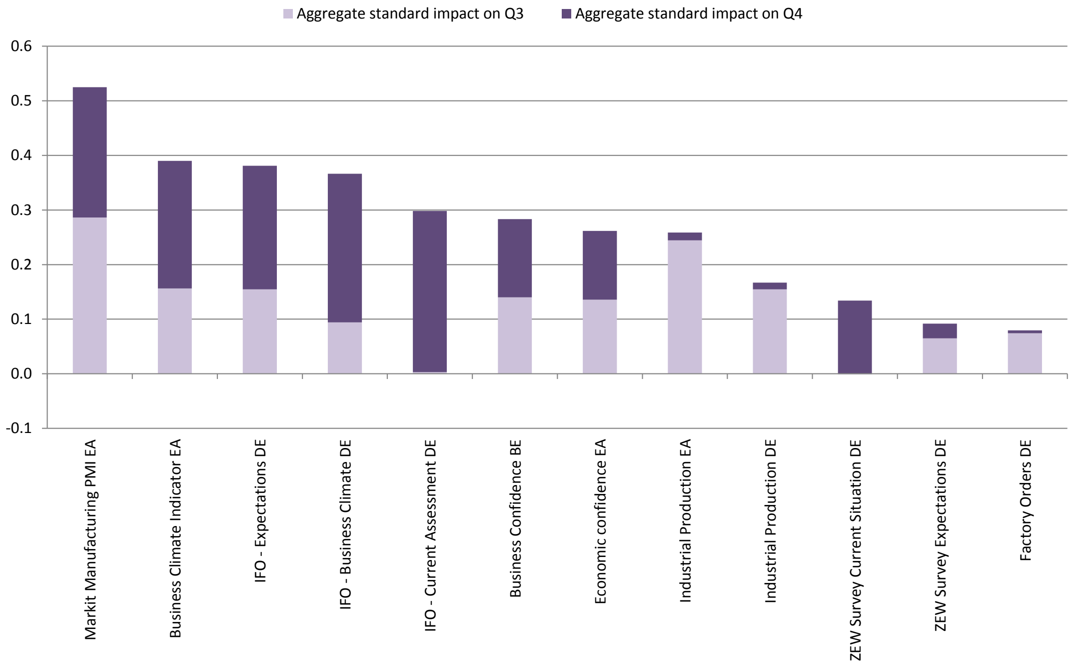
Annex II - Robustness exercise

Standard impacts when the target becomes German GDP

In this section, we re-calculate the standard impacts depicted in Figure A.3 and the resulting ranking in Figure 6 in the case that our target is German flash GDP instead of the euro area flash. The ranking of indicators is shown in Figure A.1. The top four of best-ranked indicators remains unchanged, lead by the Markit PMI. Industrial production in the euro area loses a few spots in the ranking, but remains in the top ten. Industrial production in Germany is now following more closely that of the euro area, in terms of ranking. The NBB Business Confidence has moved to the sixth position after the IFO Business Climate and Expectations for Germany.

[Insert Figure A.1 here]

Figure A.1: Ranking According to the Standard Impacts for German GDP



This figure shows the results of robustness analysis B (cf. scenarios described in Table 1). Only the twelve highest-ranked indicators are shown.

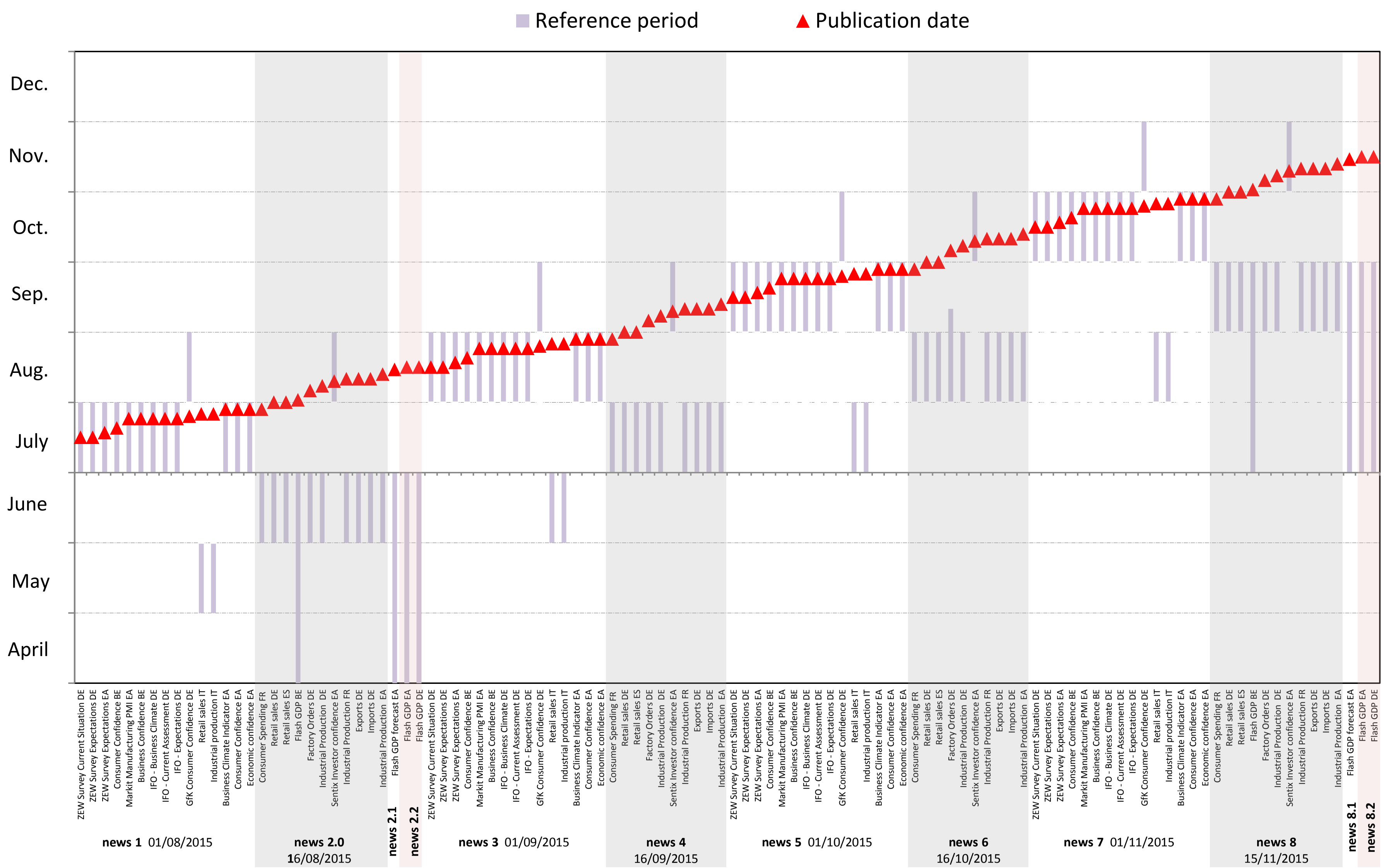
Annex III - Additional Tables and Figures

Table A.4: Dataset

	Indicator	Country or Region	Source	Available on Bloomberg	Available on Forex Factory	Definition	More information	Start of our sample	Sample extension	Frequency	Released
soft data						**For survey variables: Level of an index based on…**					
	Business climate indicator	Euro area	EC	✓	×	surveyed manufacturers, builders, wholesalers, and retailers	Survey of businesses in euro area countries	May 2001		M	around 3 weeks into the current month
	Consumer confidence	Euro area	EC	✓	✓	surveyed consumers	Survey of about 2,300 consumers in euro area countries that asks respondents to rate the relative level of past and future economic conditions, including personal financial situation, employment, inflation, and climate for major purchases	May 2001		M	around 22 days into the current month
	Economic sentiment indicator	Euro area	EC	✓	×	combined surveys	The economic sentiment indicator is composed of the industrial confidence indicator (40%), the service confidence indicator (30%), the consumer confidence indicator (20%), the construction confidence indicator (5%), and the retail trade confidence indicator (5%). Its long term average (1990-2015) equals 100.	May 2001		M	around 3 weeks into the current month
	Manufacturing PMI	Euro area	Markit	✓	✓	surveyed purchasing managers in the manufacturing industry	Survey of about 3000 purchasing managers that asks respondents to rate the relative level of business conditions including employment, production, new orders, prices, supplier deliveries, and inventories	December 2004	January 2000	M	around 3 weeks into the current month
	Investor confidence	Euro area	Sentix	✓	✓	surveyed investors and analysts	Survey of about 2,800 investors and analysts that asks respondents to rate the relative 6-month economic outlook for the Eurozone	January 2007	January 2003	M	on the first or second Monday of the current month
	Economic sentiment	Euro area	ZEW	✓	✓	surveyed German investors and analysts	Survey of up to 350 German institutional investors and analysts that asks respondents to rate the relative 6-month economic outlook for the Eurozone	November 2003	January 2000	M	on the second or third Tuesday of the current month
	Business confidence	Belgium	NBB	✓	✓	surveyed manufacturers, builders, services and trade-related firms	Survey of about 6,000 businesses that asks respondents to rate the relative level of current business conditions and expectations for the next 3 months	April 2001		M	around 3 weeks into the current month
	Consumer confidence	Belgium	NBB	✓	×	surveyed consumers	Survey of about 1850 households on their current appreciation and expectations for the next 12 months on the outlook for the general economy and regarding their own situation	January 2004	January 2000	M	around 19 days into the current month
	Consumer confidence	Germany	GfK	✓	✓	idem	Survey of about 2,000 consumers that asks respondents to rate the relative level of past and future economic conditions, including personal financial situation, climate for major purchases, and overall economic situation	January 2005		M	around the end of the previous month
	Business climate indicator	Germany	IFO	✓	✓	surveyed manufacturers, builders, wholesalers, and retailers	Survey of about 7,000 businesses that asks respondents to rate the relative level of current business conditions and expectations for the next 6 months	April 2001		M	around 3 weeks into the current month
	Business expectations	Germany	IFO	✓	×	idem		February 2002		M	idem
	Current assessment	Germany	ZEW	✓	×	surveyed German institutional investors and analysts	Survey of up to 350 German institutional investors and analysts that asks respondents about the general state of the economy as it relates to businesses	February 2004	January 2000	M	on the second or third Tuesday of the current month
	Economic sentiment	Germany	ZEW	✓	✓	idem	Survey of up to 350 German institutional investors and analysts that asks respondents to rate the relative 6-month economic outlook for Germany	December 2001	January 2000	M	idem

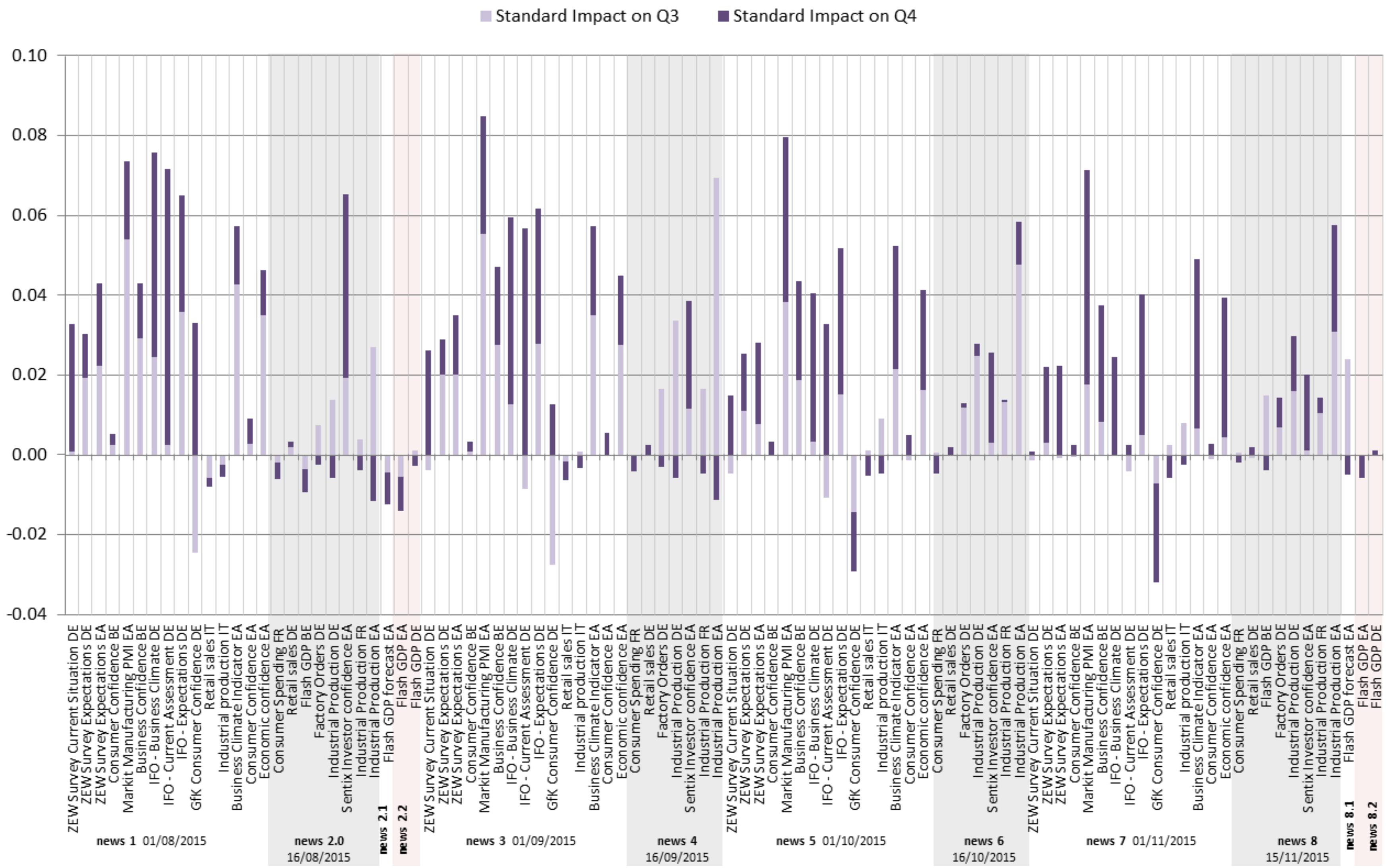
	Indicator	Country or Region	Source	Available on Bloomberg	Available on Forex Factory	Definition	Start of our sample	Sample extension	Frequency	Released
hard data	Industrial production m/m	Euro area	EC	✓	✓	Change in the total inflation-adjusted value of output produced by manufacturers, mines, and utilities	July 2001		M	about 45 days after the month ends
	Industrial production y/y	Euro area	EC	✓	×	idem	July 2001		M	idem
	Industrial production m/m	Germany	Destatis	✓	✓	idem	January 2006		M	about 40 days after the month ends
	Industrial production y/y	Germany	Destatis	✓	×	idem	July 2013	January 2006	M	idem
	Factory orders m/m	Germany	Destatis	✓	✓	Change in the total value of new purchase orders placed with manufacturers	April 2001		M	about 35 days after the month ends
	Factory orders y/y	Germany	Destatis	✓	×	idem	April 2001		M	idem
	Retail sales m/m	Germany	Destatis	✓	✓	Change in the total value of inflation-adjusted sales at the retail level, excluding automobiles and gas stations	June 2003		M	about 30 days after the month ends
	Retail sales y/y	Germany	Destatis	✓	×	idem	June 2003		M	idem
	Retail sales y/y	Spain	INE	✓	×	idem	June 2013		M	about 30 days after the month ends
	Retail sales m/m	Italy	Istat	✓	✓	Change in the total value of sales at the retail level	October 2003		M	about 55 days after the month ends
	Retail sales y/y	Italy	Istat	✓	×	idem	October 2003		M	idem
	Industrial production m/m	Italy	Istat	✓	✓	Change in the total inflation-adjusted value of output produced by manufacturers, mines, and utilities	May 2001		M	about 40 days after the month ends
	Industrial production y/y	Italy	Istat	✓	×	idem	May 2001		M	idem
	Industrial production m/m	France	INSEE	✓	✓	idem	April 2001		M	about 40 days after the month ends
	Industrial production y/y	France	INSEE	✓	×	idem	April 2001		M	idem
	Consumer spending m/m	France	INSEE	✓	✓	Change in the inflation-adjusted value of all goods expenditures by consumers	April 2001		M	about 27 days after the month ends
	Consumer spending y/y	France	INSEE	✓	×	idem	April 2001		M	idem
GDP	GDP flash	Belgium	NAI	✓	✓	Change in the total inflation-adjusted value of all goods and services produced by the economy	September 2003		Q	about 30 days after the quarter ends
	GDP flash forecast	Euro area	Bloomberg	✓	×	idem	June 2001		Q	about 44 days after the quarter ends
	GDP flash	Germany	Destatis	✓	✓	idem	June 2003		Q	about 45 days after the quarter ends
	GDP flash	Euro area	EC	✓	✓	idem	March 2001		Q	idem

Figure A.2: Detailed View of the Real-Time Dataflow



This figure represents the real-time dataflow. All data releases published between June 2015 and October 2015 are represented chronologically. Most of the surveys for each month or reference period are published before the end of the month. This means that the publication dates marked with triangles often fall inside the bar representing reference periods. In the extreme, some variables' with a strong expectations components are released prior to the reference period. Conversely, variables subject to publication lags, such as GDP or industrial production, will have the triangle way above the reference period. Hard data such as industrial production and retail sales are included in both month-on-month and year-on-year growth rates, except for Spanish retail sales, which are only included in month-on-month growth rates.

Figure A.3: Standard Impacts for euro area GDP Flash



This figure represents the weigths associated to the real-time newsflow multiplied by the standard deviation of the news.

NATIONAL BANK OF BELGIUM - WORKING PAPERS SERIES

The Working Papers are available on the website of the Bank: http://www.nbb.be.

276. "How do exporters react to changes in cost competitiveness?", by S. Decramer, C. Fuss and J. Konings, *Research series*, January 2015.
277. "Optimal monetary policy response to endogenous oil price fluctuations", by A. Stevens, *Research series*, January 2015.
278. "Comparing fiscal multipliers across models and countries in Europe", by J. Kilponen, M. Pisani, S. Schmidt, V. Corbo, T. Hledik, J. Hollmayr, S. Hurtado, P. Júlio, D. Kulikov, M. Lemoine, M. Lozej, H.Lundvall, J. R. Maria, B. Micallef, D. Papageorgiou, J. Rysanek, D. Sideris, C. Thomas and G. de Walque, *Research series*, March 2015.
279. "Assessing European competitiveness: The new CompNet micro-based database", by P. Lopez-Garcia, F. di Mauro and the CompNet Task Force, *Research series*, April 2015.
280. "FloGARCH: Realizing long memory and asymmetries in returns volatility", by H. Vander Elst, *Research series*, April 2015.
281. "Does education raise productivity and wages equally? The moderating roles of age, gender and industry", by F. Rycx, Y. Saks and I. Tojerow, *Research series,* April 2015.
282. "Assessing European firms' exports and productivity distributions: The CompNet trade module", by A. Berthou, E. Dhyne, M. Bugamelli, A.-M. Cazacu, C.-V. Demian, P. Harasztosi, T. Lalinsky, J. Merikül, F. Oropallo and A. C. Soares, *Research series,* May 2015.
283. "Economic importance of the Belgian ports: Flemish maritime ports, Liège port complex and the port of Brussels - Report 2013", by F. Van Nieuwenhove, *Document series*, June 2015.
284. "Crisis-proof services: Why trade in services did not suffer during the 2008-2009 collapse", by A. Ariu, *Research series,* July 2015.
285. "The labour market position of second-generation immigrants in Belgium", by V. Corluy, J. Haemels, I. Marx and G. Verbist, *Research series*, September 2015.
286. "The implications of household size and children for life-cycle saving", by B. Capéau and B. De Rock, *Research series*, September 2015.
287. "Monetary policy effects on bank risk taking", by A. Abbate and D. Thaler, Research series, September 2015.
288. "The Belgian production network 2002-2012", by E. Dhyne, G. Magerman and S. Rubínová, Research series, October 2015.
289. "Portfolio choice and investor preferences: A semi-parametric approach based on risk horizon", by G. Hübner and T. Lejeune, *Research series*, October 2015.
290. "Predicting Belgium's GDP using targeted bridge models", by Ch. Piette, *Research series*, January 2016.
291. "Did export promotion help firms weather the crisis?", by J. Van Biesebroeck, J. Konings and C. Volpe Martincus, *Research series*, January 2016.
292. "On the role of public policies and wage formation for business investment in R&D: A long-run panel analysis", by T. Buyse, F. Heylen and R. Schoonackers, *Research series*, January 2016.
293. "Unraveling firms: Demand, productivity and markups heterogeneity", by E. Forlani, R. Martin, G. Mion and M. Muûls, *Research series*, February 2016.
294. "Unemployment risk and over-indebtedness: A micro-econometric perspective", by Ph. Du Caju, F. Rycx and I. Tojerow, *Research series*, February 2016.
295. "International shocks and domestic prices: How large are strategic complementarities?", by A. Amiti, O. Itskhoki and J. Konings, *Research series,* March 2016.
296. "The supplier network of exporters: Connecting the dots", by E. Dhyne and S. Rubínová, *Research series*, May 2016.
297. "Does one size fit all at all times? The role of country specificities and state dependencies in predicting banking crises" by S. Ferrari and M. Pirovano, *Research series,* May 2016.
298. "Competition and product mix adjustment of multi-product exporters: Evidence from Belgium", by K. Breemersch, *Research series*, June 2016.
299. "Flemish maritime ports, Liège port complex and the port of Brussels – Report 2014", by G. Van Gastel, *Document series*, June 2016.
300. "Misalignment of productivity and wages across regions? Evidence from Belgian matched panel data", by F. Rycx, Y. Saks and I. Tojerow, *Research series*, July 2016.
301. "The European Payments Union and the origins of Triffin's regional approach towards international monetary integration", by I. Maes and I. Pasotti, *Research series,* September 2016.
302. "The transmission mechanism of credit support policies in the Euro Area", by J. Boeckx, M. de Sola Perea and G. Peersman, *Research series*, October 2016.

303. “Bank capital (requirements) and credit supply: Evidence from pillar 2 decisions”, by O. De Jonghe, H. Dewachter and S. Ongena, *Research series*, October 2016.
304. “Monetary and macroprudential policy games in a monetary union”, by R. Dennis and P. Ilbas, *Research series*, October 2016.
305. “Forward guidance, quantitative easing, or both?, by F. De Graeve and K. Theodoridis, *Research series*, October 2016.
306. “The impact of sectoral macroprudential capital requirements on mortgage loan pricing: Evidence from the Belgian risk weight add-on”, by S. Ferrari, M. Pirovano and P. Rovira Kaltwasser, *Research series*, October 2016.
307. “Assessing the role of interbank network structure in business and financial cycle analysis”, by J-Y Gnabo and N.K. Scholtes, *Research series*, October 2016.
308. “The trade-off between monetary policy and bank stability”, by M. Lamers, F. Mergaerts, E. Meuleman and R. Vander Vennet, *Research series,* October 2016.
309. “The response of euro area sovereign spreads to the ECB unconventional monetary policies”, by H. Dewachter, L. Iania and J-C. Wijnandts, *Research series*, October 2016.
310. “The interdependence of monetary and macroprudential policy under the zero lower bound”, by V. Lewis and S. Villa, *Research series*, October 2016.
311. “The impact of exporting on SME capital structure and debt maturity choices”, by E. Maes, N. Dewaelheynes, C. Fuss and C. Van Hulle, *Research series*, October 2016.
312. “Heterogeneous firms and the micro origins of aggregate fluctuations”, by G. Magerman, K. De Bruyne, E. Dhyne and J. Van Hove, *Research series*, October 2016.
313. “A dynamic factor model for forecasting house prices in Belgium”, by M. Emiris, *Research series*, November 2016.
314. “La Belgique et l’Europe dans la tourmente monétaire des années 1970 – Entretiens avec Jacques van Ypersele”, by I. Maes and S. Péters, *Research series*, December 2016.
315. “Creating associations to substitute banks’ direct credit. Evidence from Belgium”, by M. Bedayo, *Research series*, December 2016.
316. “The impact of export promotion on export market entry”, by A. Schminke and J. Van Biesebroeck, *Research series*, December 2016.
317. “An estimated two-country EA-US model with limited exchange rate pass-through”, by G. de Walque, Ph. Jeanfils, T. Lejeune, Y. Rychalovska and R. Wouters, *Research series,* March 2017.
318. Using bank loans as collateral in Europe: The role of liquidity and funding purposes”, by F. Koulischer and P. Van Roy, *Research series*, April 2017.
319. “The impact of service and goods offshoring on employment: Firm-level evidence”, by C. Ornaghi, I. Van Beveren and S. Vanormelingen, *Research series*, May 2017.
320. “On the estimation of panel fiscal reaction functions: Heterogeneity or fiscal fatigue?”, by G. Everaert and S. Jansen, *Research series*, June 2017.
321. “Economic importance of the Belgian ports: Flemish maritime ports, Liège port complex and the port of Brussels - Report 2015”, by C. Mathys, *Document series*, June 2017.
322. “Foreign banks as shock absorbers in the financial crisis?”, by G. Barboni, *Research series*, June 2017.
323. “The IMF and precautionary lending: An empirical evaluation of the selectivity and effectiveness of the flexible credit line”, by D. Essers and S. Ide, *Research series*, June 2017.
324. “Economic importance of air transport and airport activities in Belgium – Report 2015”, by S. Vennix, *Document series*, July 2017.
325. “Economic importance of the logistics sector in Belgium”, by H. De Doncker, *Document series*, July 2017.
326. “Identifying the provisioning policies of Belgian banks”, by E. Arbak, *Research series*, July 2017.
327. “The impact of the mortgage interest and capital deduction scheme on the Belgian mortgage market”, by A. Hoebeeck and K. Inghelbrecht, *Research series,* September 2017.
328. “Firm heterogeneity and aggregate business services exports: Micro evidence from Belgium, France, Germany and Spain”, by A. Ariu, E. Biewen, S. Blank, G. Gaulier, M.J. González, Ph. Meinen, D. Mirza, C. Martín and P. Tello, *Research series*, September 2017.
329. “The interconnections between services and goods trade at the firm-level”, by A. Ariu, H. Breinlichz, G. Corcosx, G. Mion, *Research series,* October 2017.
330. “Why do manufacturing firms produce services? Evidence for the servitization paradox in Belgium”, by P. Blanchard, C. Fuss and C. Mathieu, *Research series,* November 2017.
331. “Nowcasting real economic activity in the euro area: Assessing the impact of qualitative surveys”, by R. Basselier, D. de Antonio Liedo and G. Langenus, *Research series,* December 2017.

National Bank of Belgium
Limited liability company
RLP Brussels – Company's number: 0203.201.340
Registered office: boulevard de Berlaimont 14 – BE-1000 Brussels
www.nbb.be

Editor

Jan Smets

Governor of the National Bank of Belgium

Layout: Analysis and Research Group
Cover: NBB AG – Prepress & Image

Published in December 2017