Fiscal Nowcasting

Jacopo Cimadomo† Domenico Giannone‡ Michele Lenza§

This version: 28th February 2017

Abstract

Government budget balance data are only available at quarterly frequency and their release is particularly delayed coming, in general, later than national quarterly accounts. However, cash monthly data on the government borrowing requirement are published by official sources with a very limited delay. Though very timely, due to a different accounting methodology compared to the one for assessing the budget balance, monthly cash flows are a noisy indicator of the budget balance. This paper proposes a Bayesian Mixed Frequency VAR model aimed at extracting information on the budget balance while, at the same time, discounting the noisy content of monthly cash-flows. The proposed model allows to produce a monthly forecast of the annual budget balance, while forecasts from the government and official institutions are generally released only annually or bi-annually. An application based on Italian data shows that our parsimonious VAR model is characterized by a good forecasting accuracy, also in comparison with institutional forecasts such as the European Commission’s one.

JEL Classification: C11, E62, H68.

Keywords: Nowcasting, Mixed-frequency, Government budget balance, Cash data.

*Preliminary version, do not quote without authors’ permission. We would like to thank the participants at the International Conference on Macroeconomic Analysis and International Finance, at the Annual Conference of the International Association for Applied Econometrics (IAAE), at the International Symposium on Forecasting (ISF), and at Philadelphia Federal Reserve Conference on Real-Time Data Analysis, Methods, and Applications, for their useful comments. In particular, we would like to thank Dario Caldara, Peter Claeys, Dean Cronshoe, Thorsten Drautzburg, George Monokrousos, Keith Sill and Simon van Norden for their suggestions. Domenico Giannone gratefully acknowledges the Fiscal Policies Division of the ECB for its hospitality. The opinions expressed herein are those of the authors and do not necessarily reflect those of the ECB, the Eurosystem or the Federal Reserve Bank of New York.

†European Central Bank, Fiscal Policies Division, Sonnemannstrasse 20, D-60314 Frankfurt am Main, Germany. E-mail: jacopo.cimadomo@ecb.europa.eu

‡Federal Reserve Bank of New York, Macroeconomic and Monetary Studies Function, 33 Liberty Street New York, NY 10045. E-mail: domenico.giannone@ny.frb.org

§European Central Bank, Monetary Policy Division, Sonnemannstrasse 20, D-60314 Frankfurt am Main, Germany. E-mail: michele.lenza@ecb.europa.eu
1 Introduction

The deficit to GDP ratio is a synthetic indicator of state of public finances in one country, and it has a core role in the surveillance process in the context of the EU fiscal framework. Timely monitoring the tendency of such ratio is of fundamental importance, especially for countries that exceeded the 3% of GDP threshold and are subject to an “Excessive Deficit Procedure” (EDP). This paper describes and evaluates a new methodology to implement this task. In this paper, we focus on Italian data as an illustration, but the methodology we propose is more general and can be applied to other cases, as well.

Budget revenues and expenditures, the two constituencies of the budget balance, are generally quarterly variables and are released with a considerable delay. For example, in Italy, these variables are released only on the first business day of the fourth month after the end of the reference quarter, e.g., the budget balance for the fourth quarter of 2016 will be only released at the beginning of April 2017. However, two business days after the end of, say, month $t_m$, the Italian Treasury publishes its cash flow in month $t_m$. The sum of the cash flows in the quarter do not generally exactly sum to the budget balance of that quarter, due to different accounting methods. In fact, the ESA2010 (previously, ESA95) data on the budget balance, which are those relevant for fiscal surveillance, are not reported in terms of cash flows but are rather characterized by the accrual recording method. However, the cash flows should still reflect a large part of the items included in the evaluation of the Italian budget balance data relevant for fiscal surveillance.

This paper proposes a methodology, based on a mixed frequency Bayesian vector autoregressive model (VAR), which aims to reap the benefits of the timeliness in the releases of monthly cash data while, at the same time, trying to filter out the noise in the relationship with quarterly budget balance data induced by the different accounting procedures. The
methodology is similar to the BVAR model by Schorfheide and Song (2015), which develop a mixed-frequency Bayesian VAR model for GDP nowcasting. However, differently from these authors, we focus on the nowcasting of fiscal variables, which has been largely unexplored in the literature.

This paper is structured as follows: Section 2 briefly review the related literature on GDP nowcasting and fiscal forecasting; Section 3 presents the nowcasting problem, Section 5 the dataset used in this analysis; Section 5 outlines the the estimation methodology; Section 6 presents our preliminary empirical results. Finally, Section 7 concludes and discusses the next steps in this project.

2 Related literature

This paper lies at the intersection between the literature on nowcasting (i.e., current-period forecast), in particular GDP nowcasting, and fiscal forecasting. The literature on GDP nowcasting has developed massively over the last years. In particular, Giannone et al. (2008) evaluate the marginal impact that intra-monthly data releases have on the GDP growth nowcast. The proposed factor model allows to track the real-time flow of information monitored by central banks because it can handle large data sets with staggered data-release dates. More recently, Banbura et al. (2013) survey the literature on economic nowcasting with a special focus on those models that formalize key features of how market participants and policy makers read macroeconomic data releases in real-time.

This paper also connects with the fiscal forecasting literature. The latter is quite limited, and has developed mainly in Europe. This is probably due to the fact that fiscal surveillance (which implies practices of fiscal nowcasting) is particularly relevant in the euro zone, where the Stability and Growth Pact binds (for a survey of this literature, see Leal et al. (2008)).
Few papers in this literature have highlighted that - while accrual data on government deficits are only available with a relatively long time lag - monthly or quarterly intra-annual data are available with much shorter time lags, and can be used to derive accurate forecasts for end-of-year fiscal outcomes. (see e.g. Pérez (2007); Pedregal and Pérez (2010); Onorante et al. (2010)).

In particular, exploiting a Mixed Data Sampling approach (MiDaS), Asimakopoulou et al. (2013) assess the news content of quarterly fiscal data releases and their implications for the annual outturn of those series. Focusing on a sample of EU countries, they show that quarterly information is indeed very important to estimate annual outcomes. Hughes Halllett et al. (2010) focus on monthly cash data. They evaluate the relevance of such data as instruments for constructing early warnings indicators for future deficit deviating from targets. They also examine and compare two different strategies for correcting excessive.

Our modeling approach is based on a mixed-frequency Bayesian VAR model, which treats the low frequency variables as the result of aggregation of a high frequency latent process. This approach has been used in previous work. Namely, this approach has been followed by Giannone et al. (2009) and by Kuzin et al. (2011), based based on Maximum Likelihood (or under flat priors) estimation. In this paper, we follow more recent work which use informative priors (see, in particular, Schorfheide and Song (2015) and Brave et al. (2016).

1Earlier applications include Zadrozny (1990) and Mittnik and Zadrozny (2004).

2These papers use informative priors that have been widely used in traditional, single frequency, VARs (Doan et al. (1983), Busrara et al. (2010a), Giannone et al. (2015).
3 The now-casting problem

Our target variable is the annual budget balance to GDP ratio \((b_{ta})\) of the Italian general government in each specific year \(ta\), i.e.,

\[
b_{ta} = \frac{\sum_{t=ta,Q_1}^{ta,Q_4} D_t}{\sum_{t=ta,Q_1}^{ta,Q_4} Y_t P_t},
\]

In this application, we focus on monthly now-casts of \(b_{ta}\), i.e., on the evaluation of the budget balance to GDP ratio for the whole year \(ta\) conducted in each of the twelve months of the same year. Notice that this is not a trivial problem, given that, especially in the first months of the year, it implies to forecast the path of the budget balance to GDP ratio - hence, the path of the difference between revenues and expenditures, GDP and the GDP deflator - several months ahead.

In order to provide a realistic assessment of the challenges faced in now-casting the state of Italian public finances, we should also take into account the real time data availability faced by practitioners. However, at this stage we cannot address issues related to data revisions, given that we have only ex-post revised data, for the time being. Nevertheless, we fully address the issue of the end-sample data imbalance caused by the staggered nature of data releases. In order to mimic the data availability at the time of the now-cast production, which we assume to be the 15th of each month, we have reconstructed the data availability at the end of the sample that a practitioner would face in each of the twelve months of each year. In particular, Table 1 reports the data availability on the 15th of each month for the four variables used in our empirical application (as described in Section 5).

Table 1 reflects both the different timeliness of the variables (different dates of data releases) and their different sample frequency. In column 2 and 3, we report the available
Table 1: Data availability for Italy in the dates of the now-cast production.

<table>
<thead>
<tr>
<th>Date of now-cast</th>
<th>GDP</th>
<th>GDP Deflator</th>
<th>Budget balance</th>
<th>Cash balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-Jan</td>
<td>ta-1.Q3</td>
<td>ta-1.Q3</td>
<td>ta-1.Q3</td>
<td>ta-1.December</td>
</tr>
<tr>
<td>15-Feb</td>
<td>ta-1.Q3</td>
<td>ta-1.Q3</td>
<td>ta-1.Q3</td>
<td>ta.January</td>
</tr>
<tr>
<td>15-Mar</td>
<td>ta-1.Q4</td>
<td>ta-1.Q4</td>
<td>ta-1.Q3</td>
<td>ta.February</td>
</tr>
<tr>
<td>15-Apr</td>
<td>ta-1.Q4</td>
<td>ta-1.Q4</td>
<td>ta-1.Q4</td>
<td>ta.March</td>
</tr>
<tr>
<td>15-May</td>
<td>ta-1.Q4</td>
<td>ta-1.Q4</td>
<td>ta-1.Q4</td>
<td>ta.April</td>
</tr>
<tr>
<td>15-Jun</td>
<td>ta.Q1</td>
<td>ta.Q1</td>
<td>ta-1.Q4</td>
<td>ta.May</td>
</tr>
<tr>
<td>15-Jul</td>
<td>ta.Q1</td>
<td>ta.Q1</td>
<td>ta.Q1</td>
<td>ta.June</td>
</tr>
<tr>
<td>15-Aug</td>
<td>ta.Q1</td>
<td>ta.Q1</td>
<td>ta.Q1</td>
<td>ta.July</td>
</tr>
<tr>
<td>15-Sep</td>
<td>ta.Q2</td>
<td>ta.Q2</td>
<td>ta.Q1</td>
<td>ta.August</td>
</tr>
<tr>
<td>15-Oct</td>
<td>ta.Q2</td>
<td>ta.Q2</td>
<td>ta.Q2</td>
<td>ta.September</td>
</tr>
<tr>
<td>15-Nov</td>
<td>ta.Q2</td>
<td>ta.Q2</td>
<td>ta.Q2</td>
<td>ta.October</td>
</tr>
<tr>
<td>15-Dec</td>
<td>ta.Q3</td>
<td>ta.Q3</td>
<td>ta.Q2</td>
<td>ta.November</td>
</tr>
</tbody>
</table>

Releases of GDP and GDP deflator at each mid-month now-casting round. National accounts are released with a quarterly frequency and around mid-month, in the third month after the end of the reference quarter. Hence, say, the now-casts of the budget balance to GDP ratio produced in January and February are based on GDP and GDP deflators data until the third quarter of the previous year ($ta - 1$). At mid-March, instead, the release of the fourth quarter for the previous year becomes available. Successive national account releases follow the same path just described discussed for the first quarter.

Government accounts are released with a few weeks delay compared to Quarterly National accounts, generally at the beginning of the fourth month after the end of the reference quarter. Hence, differently from the case of GDP and the GDP deflator, even in March the now-casts are still based on budget balance data only until Q3 of the previous year and the fourth quarter release will only be factored in the now-casts from April onward. Again, the same pattern of releases then follows in the successive months. Cash data, instead, are released with monthly frequency and for, say, month $tm$, right at the beginning of the successive month (second business day after the end of the month). Hence, at the date of each now-cast, we have cash data releases ranging until the previous month.
4 Data

Our database includes quarterly data for GDP and the GDP deflator over the period 1985Q1 until 2016Q3. Quarterly data for the government revenue and expenditure are available since 1999Q1 and cover the period until 2016Q3. All quarterly data are collected from the European Commission’s database. The cash data for the government borrowing requirement are from January 1985 until December 2016, and are published by Banca d’Italia.3

Figure 1 offers a visual impression of the relationship between the quarterly budget balance data (revenues minus expenditures) and cash flow data for the Italian economy. In order to plot the data on the same time scale, we derive quarterly cash data by summing the three consecutive monthly values in each quarter.

![Graph showing quarterly budget balance and cash data](chart1)

**Figure 1: Quarterly budget balance and cash data.** The sample is from 1999Q1 to 2016Q1. Units are millions of euro.

Chart 1 shows that the medium-low frequency developments in the data on cash flows are definitely in line with the medium-low frequency in budget balance data. Hence, the very timely releases of cash data can be a very important asset in order to predict the budget

---

balance. However, cash data are also quite noisier than budget balance data and modelling devices should be used in order to appropriately filter out such noise without eliminating too much of their informative content.

5 The now-casting methodology

The now-casting problem defined above requires the solution of two issues of missing data. First, the variables are sampled with different frequency, quarterly and monthly. We assume that quarterly variables are monthly variables with missing observations in the first two months of the quarter.

Second, due to the staggered nature of data releases highlighted in table 1, several observations of the quarterly variables are missing at the end of the sample. This section briefly sketches the methodology we use in order to address these issues.

We assume that the levels of our $N (=4)$ variables (collected in the $N$-dimensional vector $X_{tm}$) are described by the following monthly vector autoregressive process with $p (=13)$ lags:

$$X_{tm} = A_0 + A_1 X_{tm} + \cdots + A_p X_{tm-p} + e_{tm}, \quad (1)$$

where $A_p$ is the $N \times N$ matrix collecting the coefficients of the $p$-th lag and $e_{tm}$ is a normally distributed multivariate white noise with covariance matrix $\Sigma$.

The choice of accounting for rich dynamics ($p = 13$ lags) is motivated by two main considerations. First, we want a general and flexible model which does not a-priori constraints the dynamic interrelationships among our variables. Second, the data are not seasonally adjusted and this dynamic specification is able to account for the seasonal fluctuations in the variables.
The rich dynamics we want to allow for in our VAR model imply that we face an issue of over-fitting, owing to the large number of parameters (the so-called “curse of dimensionality”). We address this issue by shrinking the model’s coefficients toward those of the naïve and parsimonious random walk with drift model, \( X_{i,t,m} = \delta_i + X_{i,t,m-1} + u_{i,t,m} \). De Mol et al. (2008) and Banbura et al. (2010b) have shown that this approach reduces estimation uncertainty without introducing substantial bias. This is achieved thanks to the tendency for macroeconomic time series to co-move over the business cycle, which creates scope for the data to point “massively” in the same direction against a naïve prior model that does not allow for any dynamic interaction. The resulting model offers a parsimonious but reliable estimate of the complex dynamic interactions among the macro, monetary and financial variables included in the data set.

More specifically, we use a Normal-Inverted Wishart prior centred on a random walk model. For \( \Sigma \), the covariance matrix of the residuals, we use an inverted Wishart with scale parameter given by a diagonal matrix \( \Psi \) and \( d = N + 2 \) degrees of freedom. This is the minimum number of degrees of freedom that guarantees the existence of the prior mean of \( \Sigma \), which is equal to \( \frac{\Psi}{(d-N-1)} = \Psi \). For the constant \( A_0 \) term, we use a flat prior. For the autoregressive coefficients \((A_1...A_p)\), we use the Minnesota prior, as originally proposed by Litterman (1980). As regards the Minnesota prior, conditional on the covariance matrix of the residuals, the prior distribution of the autoregressive coefficients is normal with the following means and variances:

\[
E(A_1) = I_N, E(A_2) = ... = E(A_p) = 0_{N,N},
\]

\[
\text{Cov}[(A_s)_{ij}, (A_r)_{hm} | \Sigma] = \lambda^2 \frac{\Sigma_{ih}}{s^2 \Psi_{ii}} \text{ if } m = j \text{ and } r = s, \text{ zero otherwise.}
\]
Notice that the variance of this prior distributions decays with the lag, and that coefficients associated with the same variables and lags in different equations are allowed to be correlated. The key hyperparameter is \( \lambda \), which controls the scale of all the prior variances and covariances, and effectively determines the overall tightness of this prior. For \( \lambda = 0 \) the posterior equals the prior and the data do not influence the estimates. If \( \lambda \to \infty \), on the other hand, posterior expectations coincide with the Ordinary Least Squares (OLS) estimates. The factor \( \frac{1}{\sigma^2} \) is the rate at which the prior variance decreases with increasing lag length and \( \frac{\sum u_i}{\Psi_{jj}} \) accounts for the different scale and variability of the data.

Summing up, the setting of these priors depends on the hyperparameter \( \lambda \), which reflects the informativeness of the prior distribution for the model’s coefficients. This parameter is usually set on the basis of subjective considerations or rules of thumb. For the sake of simplicity, at this stage, we set the value of this hyperparameter to 0.2, as it suggested in Sims and Zha (1998).

If we did not face the issue of missing data, the Bayes rule would allow us to easily draw parameters from the posterior distributions implied by the likelihood and the prior set-up just described. Then, the algorithm to produce conditional forecasts developed in Banbura et al. (2015) based on the simulation smoother of Carter and Kohn (1994), could be employed in order to produce the out-of-sample forecasts of the budget balance and nominal GDP. Notice that the need of an algorithm to produce conditional forecasts is due to the end-of-sample imbalance in our panel caused by the staggered data releases. The idea here is that we treat more timely data releases as future “conditions” on which we condition the other forecasts.

However, as described above, we have to tackle also a further issue of missing data in this set-up, due to the mixed frequency of the variables. We tackle the issue of missing
data by setting up a recursive procedure that, first, balances the database by providing a
draw of the missing data conditional on a draw from the posterior of the model parameters
and, then, provides another draw of the parameters conditional on the previous draw of the
variables.

A schematic way of representing our recursive algorithm for the panel available in month
\(tm\) and for a forecast horizon \(h\), is the following.

1) Initialization: \(X(0)_{tm}\) is obtained by interpolating the unbalanced panel by means of
standard univariate non-parametric interpolation techniques.

2) First draw of the parameters from their posterior distribution, conditional on initialization of the variables: \(A(1)_{0}...A(1)_{p}\).

3) First draw of the past, present and future of the variables from the distribution of
their conditional expectation, conditional on \(A(1)_{0}...A(1)_{p} : X(1)_{0}...X(1)_{tm}...X(1)_{tm+h}\) by

4) Second draw of parameters from their posterior distribution, conditional on previous
draw of the variables conditional on \(X(1)_{0}...X(1)_{tm} : A(2)_{0}...A(2)_{p}\).

5) Second draw of the past, present and future of the variables from the distribution of
their conditional expectation, conditional on \(A(2)_{0}...A(2)_{p} : X(2)_{0}...X(2)_{tm}...X(2)_{tm+h}\) by

6) Iterate 4 and 5 \(M\) times. on Data
6 Empirical results: now-cast of the annual budget balance-to-GDP ratio

In this section, we report the results of our preliminary analysis of now-cast accuracy. In particular, we produce annual forecasts - in the twelve months of the reference year - of the budget balance to GDP ratio for Italy. We report these now-casts in the years from 2004 to 2016. We limit our analysis to this sample because the quarterly observations for the budget balance only start in 1999 and we use roughly one fourth of the sample in order to estimate the model for the first now-casts of 2004. For this exercise, we have used quarterly data as available until 2016Q3 and monthly cash data until December 2016.

Chart 2 reports the results. The green line indicates the twelve now-casts while the blue straight lines indicate the outcomes for the budget balance ratio in a specific year. We plot point forecasts, which are given by the median of the predictive distribution produced by our model.

Chart 2 shows that, in spite of some volatility, our model provides a quite accurate account of the annual budget balance to GDP ratio. In particular, nowcasts produced around mid-year, i.e. now-casts produced about nine months before the release of the annual budget balance ratio and only based on the knowledge of the budget balance in the first quarter of the current year, are already pretty close to the outcomes. It is also generally the case that further releases of cash data improve the quality of the now-casts, pushing them closer to the final outcomes. This very informal evaluation of the model performance reveals that, in spite of the noisy nature of cash flow data, our model is able to extract information from the latter in order to inform our view on the state of public finances in Italy.
Figure 2: Now-casts of budget balance to GDP ratio. The budget balance ratio is expressed in percent of GDP. Point forecasts are given by the median of the predictive distribution. The black line indicates the twelve nowcast successively produced, month-by-month, for the same year, i.e., from left to right the nowcast factors in increasingly more information. The red straight lines indicates the outcome for the budget balance to GDP ratio in a particular year, as published in the European Commission’s Winter 2017 Forecasts. The blue dots are the May and November issues of the European Commission’s forecasts. The quarterly sample is from 1999Q1 to 2016Q3, the cash data sample is from January 1985 until December 2016.

We also compare the root mean square error (RMSE) from our forecast with the one from the European Commission’s forecasts. The European Commission publishes its forecasts twice per year, i.e., in May and in November, for a number of variables including the budget balance for Italy (blue dots in Figure 2). The European Commission’s forecast extends over a time horizon of at least two years and cover about 180 variables. The forecasts are are made by Commission’s experts using a variety of models and experts' judgement. It turns out that, despite the difference in the size and complexity of the two forecasting approaches, our forecast performs relatively well as reflected in RMSE ratio of 1.23 for the forecasts released in May and of 0.32 for the forecast released in November. It should be stressed that -

differently from the Commission’s forecast - our model is very parsimonious (including only four variables) and does not include any judgement. In addition, our model allows to obtain an monthly indicator of the budget balance, thus also in the months between the May and November editions of the Commission’s forecast. As such, it could be effectively employed as a tool for monitoring developments in public finances in real-time.

7 Conclusions and ongoing work

This paper describes a methodology to extract information from monthly cash data in order to now-cast the annual budget balance ratio to GDP in Italy. The methodology we propose is able to handle both staggered data releases and missing data in the estimation sample in a unified framework and its outcome is the predictive distribution of the budget balance ratio.

Our empirical application, in this paper, is on Italian data. In particular, our Mixed Frequency VAR model is estimated on quarterly data for GDP and the GDP deflator over the period 1985Q1 until 2016Q3, quarterly data for the government revenue and expenditure over the period 1999Q1 and 2016Q3. In addition, we use cash data for the period from January 1985 until December 2016. The results show that our very parsimonious forecasting model, including only 4 variables, produces a quite accurate account of the Italian budget balance to GDP ratio, also as compared to the European Commission’s forecast which includes 180 variables, judgement, and which is based on several models.

Ongoing work is devoted to:

- Evaluation of density forecasts; extend evaluation also to forecasts and back-casts;
- Extension of the cross-section of data in order to improve forecast accuracy (for ex-
ample, including monthly surveys to better forecast GDP) and extend the possible applications of the model (which, with a suitable variable choice, can be also used in order to provide scenario analysis).

- Comparison with forecasts from professional forecasters and other international organizations.
References


15


