BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working Papers)

Nowcasting Italian GDP growth: a Factor MIDAS approach

by Donato Ceci, Orest Prifti and Andrea Silvestrini

March 2024

Number 1446

BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working Papers)

Nowcasting Italian GDP growth: a Factor MIDAS approach

by Donato Ceci, Orest Prifti and Andrea Silvestrini

Number 1446 - March 2024

ISSN 2281-3950 (online)

Designed by the Printing and Publishing Division of the Bank of Italy

NOWCASTING ITALIAN GDP GROWTH: A FACTOR MIDAS APPROACH

by Donato Ceci*, Orest Prifti** and Andrea Silvestrini*

Abstract

This paper examines the role of weekly financial data in nowcasting the quarterly growth rate of Italian real GDP, with a specific focus on the impact of the COVID-19 pandemic. It combines factor models and MIxed DAta Sampling (MIDAS) regression models to set up Factor MIDAS specifications, which leverage a large set of higher-frequency financial variables to exploit the information flow within the quarter. The analysis is performed using a comprehensive dataset that includes monthly macroeconomic data and weekly financial data. The predictive accuracy is assessed by conducting a pseudo out-of-sample nowcast exercise and evaluating the performance of the models with and without the inclusion of factors derived from financial indicators. Financial variables improve the nowcast of real GDP growth in Italy, particularly in the first month of the quarter, when few macroeconomic data are available. The models incorporating financial variables consistently exhibit high nowcasting accuracy throughout the quarter.

JEL Classification: C22, C43, C53, C55, E32, E37.
Keywords: nowcasting, mixed frequency, factor models, variable selection, financial markets, factor MIDAS.

---

* Bank of Italy, Directorate General for Economics, Statistics and Research.

** Università degli Studi di Roma “Tor Vergata”.

1 Introduction[1]

The capacity to efficiently gather and interpret information on the state of the economy is pivotal for informed decision-making and adaptive policy formulation. With the Covid-19 shock in 2020 the importance of delivering timely forecasts of real economic activity became increasingly evident, see Delle Monache et al. (2020) and Aprigliano et al. (2021). Such forecasts can help policy-makers respond to potential recessions or mitigate their effects.

An effective approach to achieve timely forecasts involves the integration of financial market variables into the forecasting framework. Financial data, such as stock prices, the yield curve, and credit spreads, frequently act as leading indicators of real economic activity, see Estrella and Hardouvelis (1991), Estrella and Mishkin (1998), Stock and Watson (2003), Ang et al. (2006) and Gilchrist and Zakrajšek (2012), among others. These variables often exhibit changes prior to those observed in broader macroeconomic indicators, thereby providing economists with potentially useful insights into the current and future state of the economy, while standard macroeconomic data often suffer from a lack of timeliness. An extensive body of literature has also demonstrated the suitability of financial variables in tracking the tail growth rate of real GDP, see Adrian et al. (2019) and Adams et al. (2021). In addition, owing to their high-frequency availability, the incorporation of financial market variables as covariates in econometric models allows more frequent updates of forecasts. Consequently, high-frequency financial variables have gained extensive usage in econometric models and a substantial amount of research has been dedicated to investigating the informative value of these variables; see recent contributions by Babii et al. (2022), Jardet and Meunier (2022), and Cimadomo et al. (2022).

In this context, the primary objective of this paper is to provide a nowcast for the

[1]The views expressed herein are those of the authors and do not necessarily reflect those of the Bank of Italy or the Eurosystem. The authors are grateful to two anonymous reviewers, Valentina Aprigliano, Lorenzo Braccini, Fabio Busetti, Davide Delle Monache, Antonio Di Cesare, Simone Emiliozzi, Raffaela Giordano, Nicolò Gnocato, Marco Flaccadoro, Marco Fruzzetti, Stefano Neri, Federico Nucera, Luigi Palumbo, Marcello Pericoli, Tiziano Ropele, Gwyneth Schaefer, Stefania Villa and Gabriele Zinna for useful comments and suggestions on a previous draft.

growth rate of Italian real gross domestic product (GDP) prior to its official release. This nowcasting exercise is conducted within a data-rich environment (Bernanke and Boivin, 2003), employing an extensive dataset encompassing standard real economic indicators along with higher-frequency financial market variables. The underlying assumption that is intended to be tested is that the information contained within financial variables enhances the precision of nowcasts, allowing for more accurate predictions of real GDP. This assumption cannot be taken for granted in Italy. In contrast to the United States, Italy exhibits a bank-based economy where financial markets have a relatively less important role in facilitating private sector financing (Ceci and Silvestrini, 2023). In addition, according to Ferrara and Marsilli (2013), the predictive power of well-known financial variables to anticipate GDP growth differs across the four main euro-area countries (Germany, France, Italy and Spain). As a result, the extent to which financial market indicators possess predictive capabilities for real economic activity in Italy remains uncertain and warrants more research. This paper attempts to further investigate this issue.

Incorporating financial market variables into a nowcasting framework is not straightforward. Several relevant issues need to be dealt with. First, real economic indicators and financial data are often published at different time frequencies. For instance, macroeconomic variables such as GDP are typically reported quarterly, while financial market data, such as stock prices or exchange rates, are available on a weekly or daily basis. As a result, when constructing a nowcasting model, economists frequently encounter mixed-frequency data, which presents challenges in terms of appropriate statistical techniques that can account for the differences in observation frequencies. A second problem concerns data unbalancedness at the end of the sample, which is referred to as the 'ragged-edge' problem of the dataset (Wallis, 1986). Ragged-edge data often results in missing observations, which mainly occur due to different publication lags (i.e., asynchronous data publication): for instance, GDP is released by statistical agencies with a substantial delay with respect to the reference period and, consequently, is unknown during the current quarter and becomes available only during the following quarter; on the other hand, many monthly business-cycle indicators such as the industrial production index or the purchasing managers' indices (PMI) are available throughout each quarter, while financial market data are released with no delay and therefore are immediately accessible for analysis (Moneta, 2005).

To address all these challenges, various econometric approaches have been developed, including temporal aggregation (Silvestrini and Veredas, 2008), temporal disaggregation techniques (Chow and Lin, 1971), interpolation methods (Dagum and Cholette, 2006), state-space factor models estimated with the Kalman filter (Harvey, 1990; Mariano and Murasawa, 2003; Camacho and Perez-Quiros, 2010), and mixed data sampling (MIDAS) models, introduced in a regression context in a number of papers including Ghysels et al. (2004), Ghysels et al. (2006), Ghysels et al. (2007), Ghysels and Wright (2009), Andreou et al. (2010), Marcellino and Schumacher (2010), Foroni and Marcellino (2014), and Audrino et al. (2019), among others. MIDAS models combine variables sampled at different time frequencies in a parsimonious way, without any prior time aggregation or interpolation (Guérin and Marcellino, 2013). Compared with conventional state-space models, which are also widely used in econometrics due to their ability to handle irregular and incomplete data patterns, MIDAS models are easier to estimate and less prone to parameter estimation errors and/or specification errors; see Bai et al. (2013).

Another significant aspect that is highly relevant for this paper is the availability of large-scale datasets (Stock and Watson, 2002a), which has experienced a substantial growth in recent years. With advancements in technology and data collection methods, researchers nowadays have access to an extensive array of economic data from various sources, providing new opportunities for enhancing economic forecasting models. Data-rich environments have been found to be helpful for delivering useful coincident indicators of economic activity and particularly valuable for nowcasting purposes, when the goal is to generate up-to-date estimates of macroeconomic variables before official statistics are released (Castle et al., 2017; Bok et al., 2018).

This paper takes advantage of the availability of large datasets and evaluates the contribution of weekly financial data in nowcasting the Italian GDP growth rate over the period 2015-2022, focusing on the impact of the Covid-19 pandemic. A novel mixed-frequency dataset is specially assembled for this purpose, consisting of 118 monthly macroeconomic and 109 weekly financial time series sampled from 2001 onwards. The monthly series are basically those gathered by Aprigliano and Bencivelli (2013) for constructing Ita-coin, a coincident indicator for the Italian economy, while the weekly series are those used by Ceci and Silvestrini (2023) to produce short-term forecasts of the expansion and recession phases of the Italian business cycle.[2]

In terms of the modelling approach, two widely established strands of the econometric literature are merged in this paper: (i) first, the factor models introduced by Stock and Watson (2002a,b), which are a very useful tool in this context due to their ability to handle a large number of variables and extract relevant information from complex datasets; (ii) second, the MIDAS regression models proposed by Ghysels et al. (2006). By combining these two approaches, as proposed by Marcellino and Schumacher (2010), two similar econometric specifications can be obtained, the Factor MIxed DAta Sampling (Factor-MIDAS) and the Factor AutoRegressive MIxed DAta Sampling (FAR-MIDAS) models, where the second one differs only in the addition of the autoregressive term, as suggested by Clements and Galvão (2008). These two models are estimated in a two-stage process, starting with a principal component analysis (PCA) to extract the common factors that capture the correlation structure among a pre-selected set of variables. Subsequently, these factors are employed as covariates in the MIDAS regressions.

In order to appreciate the contribution of financial market variables to improve forecast accuracy, two different specifications of the Factor-MIDAS and FAR-MIDAS models are considered: (i) one including both monthly real and weekly financial factors (Model 1); (ii) another incorporating only the monthly real factor (Model 2). Their comparison allows us to evaluate the informative contribution of financial variables in nowcasting the growth rate of Italian real GDP. Given the large amount of available data, the Least Angle Regression (LARS) pre-selection technique of Efron et al. (2004) is employed prior

[2]Only a small number of discontinued series have been excluded from the present analysis, as our dataset begins on 5 January 2001, while the sample period in Ceci and Silvestrini (2023) starts from 7 February 2003.

to factor estimation in order to improve the forecasting accuracy (Boivin and Ng, 2006) following the “soft thresholding rule” Bai and Ng (2008), which consists in keeping the most targeted predictors in the variables’ subset.[3]

In principle, nowcasts generated by Factor-MIDAS and FAR-MIDAS models can be computed several times during the reference quarter owing to the flow of information from progressively available data. Specifically, real-time forecasts can be produced at the end of the first, second, and third months of the quarter. Structuring the nowcasting exercise in this manner allows for a precise assessment of the impact of various data releases on the predictive ability of the different models. For instance, focusing on the information flow within the reference quarter, at the end of the first month of the quarter only survey data for that month are available. At the end of the second month, price data for the first month are added, while the industrial production index is only included at the end of the third month. Financial data, on the other hand, are always available in real-time.

As a preview of the results, financial information turns out to be particularly useful in nowcasting the growth rate of Italian real GDP, especially in the first month of the current quarter, when most macroeconomic data have not yet been released. As business cycle indicators become progressively available, the higher predictive accuracy of Models (Factor-MIDAS and FAR-MIDAS) 1, which incorporate factors extracted from the financial dataset, tends to diminish, and the differences in predictive accuracy compared to Models 2, which include only factors from the real dataset, are no longer statistically significant. Nevertheless, the predictive accuracy of Models 1 remains unaffected in the second and third months of the reference quarter and continues to improve as long as the flow of new information progresses, similarly to Models 2. Among the financial variables included in the dataset, corporate credit risk, price earnings/earnings per share, commodity futures, government bond yields, and stock indices exhibit the highest information content and contribute the most to out-of-sample predictive accuracy.

[3]As a robustness test, the t-statistics pre-selection technique is also considered, see Bai and Ng (2008).

The rest of the paper proceeds as follows. Section 2 presents the dataset and the information flow within the quarter. Section 3 outlines the econometric methodology. Section 4 illustrates the *pseudo* real-time nowcasting exercise and the baseline out-of-sample forecasting results, together with some robustness checks implemented to validate these results. Finally, Section 5 provides concluding remarks and proposes some lines of future research. Additional tables, charts and detailed robustness checks are included in the Appendix.

2 Data description and the information flow

This Section offers an overview of the multi-frequency database collected for the empirical analysis and the nowcast design. Section 2.1 provides a comprehensive description of the dataset, as well as details regarding the sources from which the data were obtained. Section 2.2 illustrates the real-time flow of information within the quarter.

2.1 The dataset

Both macroeconomic and financial data cover the period from January 2001 to December 2022. There are 88 quarterly observations for the growth rate of real GDP in Italy; for each monthly variable, there are 264 observations available and for each weekly variable, 1144 observations. Monthly and weekly data, along with their respective publication delays, are summarized in Table 1.

The 118 monthly macro variables are those used by Aprigliano and Bencivelli (2013) to build Ita-coin, a coincident indicator of the Italian business cycle, and they are representative of various aspects and sectors of the economy. These variables provide a broad picture of the macroeconomic environment and belong to several different blocks, including industrial production indices, international trade, consumer and producer sentiment, purchase managers' indices (PMI), inflation on both the consumer and producer sides, and the labour market. Among these macro variables, consumer and producer sentiment indicators along with PMIs are particularly important as they are already published at the end of each month, providing timely information for the nowcasts.[4] All the monthly series are standardized and seasonally adjusted. The data sources are Istat and Refinitiv Eikon.

The 109 weekly financial variables are those used by Ceci and Silvestrini (2023) to nowcast recession probabilities in Italy, and include stock market indices, price earning ratios, earnings per share, forward yields, corporate credit risk indicators for banks and firms, money market rates, sovereign bond yields, commodity futures, oil prices and exchange rates. The data sources are Bloomberg, Refinitiv Eikon and ICE Bofa Merrill Lynch.

All the macro and financial time series are preserved at their original frequency without any temporal aggregation and undergo standard transformations to ensure their stationarity (either first-differencing or first log-differencing), as indicated in Table 1.[5]

The timeliness of the indicators varies substantially. Official quarterly Italian GDP data are published by the Italian National Institute of Statistics (Istat) at the beginning of the third month of the following quarter, with a publication delay of approximately 60 days, even if a preliminary GDP estimate (the so-called 'flash estimate') is released 30 days after the end of the quarter. Monthly variables also have different publication lags: while the PMIs and the surveys are collected and released by the end of the reference month, the publication of the other monthly variables is subject to delays, typically ranging from 1 to 2 months. On the other hand, weekly financial data, which originate from the markets, are promptly available at the end of the working week (i.e., every Friday) without any delay.

Therefore, the dataset is unbalanced by construction, due to asynchronous publication lags. This issue, referred to as 'ragged-edge' problem in the literature, is addressed using

[4] In order to accurately evaluate the contribution of financial information through weekly data, the financial block in Aprigliano and Bencivelli (2013) (composed of 20 time series) is excluded from the monthly dataset. This step ensures that the estimated monthly factor does not incorporate any financial information, thus allowing the weekly financial variables to maintain their distinct contribution.

[5] This paper does not address the issue of seasonality for weekly variables. The primary reason for this decision is that all the weekly variables are derived from financial markets, where the presence of seasonality is controversial.

Table 1: Data description

Category	Number of series	Frequency	Transformation	Publication delay
Macroeconomic dataset				
Consumers' price indices	9	Monthly	Δ log	1 month
Producers' price indices	31	Monthly	Δ log	1 month
Consumers surveys	14	Monthly	None	None
Industry surveys	21	Monthly	None	None
Industrial production	20	Monthly	Δ log	2 months
Wages	2	Monthly	Δ log/None	1 month
Demand indicators	8	Monthly	Δ log	2 months
External trade	6	Monthly	Δ/Δ log	2 months
Coincident/leading indicators	7	Monthly	Δ/Δ log	1 month
Financial dataset				
Commodity futures	3	Weekly	Δ log	None
Oil price	1	Weekly	Δ log	None
Money Market rates	1	Weekly	Δ	None
Sovereign yields	20	Weekly	Δ	None
Exchange rates	13	Weekly	Δ	None
Corporate credit risk (banks and firms)	12	Weekly	Δ/Δ log	None
Earnings per share	17	Weekly	Δ/Δ log	None
Forward yields (EPS and PE)	27	Weekly	Δ/Δ log	None
Price earnings	6	Weekly	Δ/Δ log	None
Stock market indices	9	Weekly	Δ/Δ log	None

Notes: the transformations indicated in the fourth column of this table are applied to the time series prior to the selection step.

the 'vertical re-alignment' procedure of Altissimo et al. (2010). The vertical re-alignment method consists in shifting downwards to the end of the dataset the series with missing observations, so that the last available observation is considered as the contemporaneous value.[6] Although this method assumes a constant structure of cross-sectional correlations among variables, Marcellino and Schumacher (2010) find no substantial differences as compared to other alternative procedures proposed in the literature for handling ragged-edge data.

2.2 The information flow

The nowcasting exercise has been structured to mimic the real-time process to predict the dynamics of GDP, carefully replicating the flow of economic information available to the econometrician. Unfortunately, given that the real-time data vintages for both the GDP growth rate and some monthly variables are not available over the entire sample

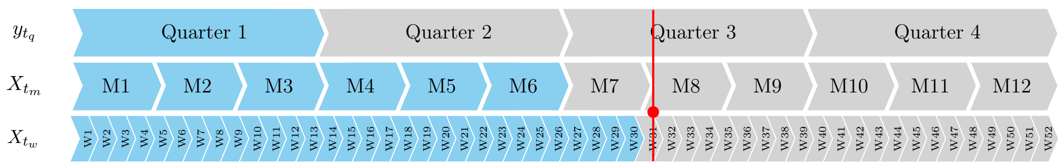
[6] Formally, given a series $\tilde{x}_t$ with publication lag q, the series becomes $x_t \equiv \tilde{x}_{t-q}$.

period, the nowcasting exercise can be considered only as a *pseudo* real-time one.[7]

The information flow collected during the reference quarter in a given year is presented in Figure 1, Figure 2 and Figure 3, which reproduce the information sets available to the forecaster at the end of each month in terms of quarterly (y_{t_q}), monthly (X_{t_m}) and weekly data (X_{t_w}). Three nowcasting origins per quarter to be nowcast are considered, at the end of each month within the quarter. The nowcast at a monthly frequency has been conventionally chosen in order to appreciate the cumulative effect of information gathering over the respective quarter to be nowcast.

The information set represented in Figure 1 is related to the nowcasts produced at the end of the first month of the reference quarter (i.e., January, April, July and October)[8] and can be labelled as the 'small' information set. Indeed, only the soft indicators (consumer and industry surveys) and the financial data are available at the end of the first month of the reference quarter, whereas the other indicators are not yet available and must be filled in with the vertical re-alignment procedure. The preliminary (*flash*) GDP estimate for the previous quarter is also released at the end of the first month of the reference quarter (i.e., at t+30 days).

Figure 1: The 'small' information set available in the first month



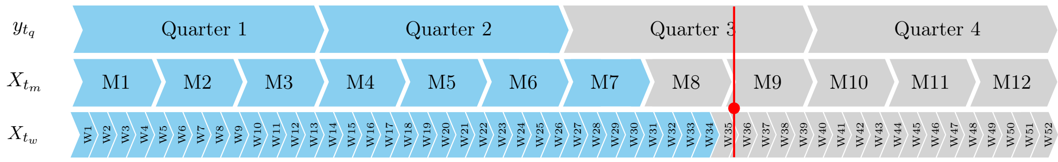
The information set available at the end of the second month of the quarter (i.e., February, May, August and November) is defined as the 'medium' information set: along with the weekly financial data and soft indicators for the first and second month of the reference quarter, the remaining monthly data (mainly hard indicators) for the first

[7]On the other hand, financial market variables are not subject to revisions by construction.

[8]In order to incorporate as much information as possible pertaining to the current month and, at the same time, preserve the timely nature of nowcasting, the prediction associated to each month is assumed to be generated on the second day of the subsequent month.

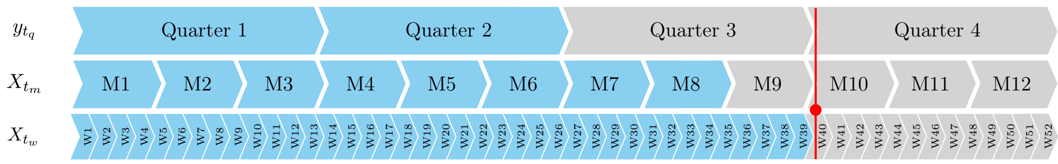
month are also available, as illustrated in Figure 2. In addition, the official GDP estimate for the previous quarter is released.

Figure 2: The 'medium' information set available in the second month



Finally, as shown in Figure 3, the 'large' information set becomes available at the end of the third month (i.e., March, June, September and December), with a complete set of weekly financial data and soft indicators for all three months of the reference quarter plus hard monthly indicators for the first two months.

Figure 3: The 'large' information set available in the third month



Overall, these charts illustrate the close alignment between the most recent financial market data and the timing of nowcasting, offering a considerable lead-time compared to monthly macro data. This temporal proximity may enhance the capacity to capture and incorporate current economic dynamics into the nowcasting process.

3 The econometric model

Since the seminal paper of Ghysels et al. (2006), MIDAS models have been widely used to handle time series data sampled at different frequencies. The fundamental idea behind MIDAS is to leverage the predictive power of high frequency data, such as monthly and weekly indicators, in order to obtain more accurate predictions. The main characteristic of this class of models is that their formulation is flexible since it is regression-based while being parsimonious in the estimation of the parameters.

Consider the variable y_{t_q}, where $t_q = 1, 2, 3, \dots T_q$ is the quarterly time index, representing the Italian GDP quarter-on-quarter growth rate. Nowcasting variable y in quarter t_q based on all the available information at monthly and weekly frequencies can be easily handled by a MIDAS model defined as follows:

$$y_{t_q} = \beta_0 + \beta_1\, b(L_m, \bar{\theta}^m) X^0_{t_m} + \beta_2\, b(L_w, \bar{\theta}^w) X^0_{t_w} + u_{t_q} \tag{1}$$

where $X^0_{t_m}$ and $X^0_{t_w}$ are, respectively, the N_m- and N_w- dimension vectors of monthly (macroeconomic) and weekly (financial) regressors ($t_m = 1, 2, 3, \dots T_m = 3T_q$ and $t_w = 1, 2, 3, \dots T_w = 52T_q$) and u_{t_q} is the nowcasting error. Equation (1) is characterised by the two-parameter exponential Almon lag polynomial $b(\cdot)$, in which L_m and L_w are lag operators at the monthly and weekly frequency, respectively, and $\bar{\theta}^m$ and $\bar{\theta}^w$ are the corresponding parameters. Specifically, the exponential Almon polynomial is defined by the following expressions (the frequency indicator $\ell = m, w$ is suppressed for ease of notation):

$$b(L, \bar{\theta}) = \sum_{j=0}^{J} \omega(j, \bar{\theta}) L^j \,,$$
$$\omega(j, \bar{\theta}) = \frac{exp(\theta_1 j + \theta_2 j^2)}{\sum_{j=0}^{J} exp(\theta_1 j + \theta_2 j^2)}$$

with $\bar{\theta} = (\theta_1, \theta_2)$. The presence of the exponential Almon polynomial represents the main difference between MIDAS and ARDL regression: while the latter requires the estimation of as many coefficients as the number of lags of the regressors, the former relies only on two parameters (θ_1 and θ_2) to characterise the entire lag structure. As opposed to Bridge Models introduced by Baffigi et al. (2004), high frequency data are not temporally aggregated so as to exploit as much as possible the time granularity of the variables. This feature of the MIDAS regression in (1) appears to be particularly relevant in nowcasting the dynamics of the dependent variable: the MIDAS model possesses the capability to track the real-time data progression of monthly and weekly variables, enabling the generation of nowcasts promptly upon the release of new data.

Note that model (1) can easily accommodate different lag lengths for the two data frequencies. The results presented in the following sections are obtained with $J = 2$ for the monthly frequency and $J = 11$ for the weekly frequency, but they are robust to different values of J. The model is estimated using nonlinear least squares in a regression of y_{t_q} on the high frequency variables and their lagged values.

In a data-rich environment, it is possible to include in the MIDAS regression the factors extracted from a large dataset in order to obtain a parsimonious model that, at the same time, exploits the information contained in the estimated factors. The resulting Factor-MIDAS specification (Marcellino and Schumacher, 2010) consists of a regression of the low-frequency variable y_{t_q} on the factors extracted from the two different high-frequency datasets:

$$y_{t_q} = \beta_0 + \beta_1\, b(L_m, \theta^m) F^0_{t_m} + \beta_2\, b(L_w, \theta^w) F^0_{t_w} + u_{t_q} \tag{2}$$

where the new quantities $F^0_{t_m}$ and $F^0_{t_w}$ are, respectively, the factors extracted from the monthly (macroeconomic) and weekly (financial) variables X^0_m and X^0_w (see below for additional discussion related to the estimation of the factors). Marcellino and Schumacher (2010) find substantial improvement in using Factor-MIDAS to nowcast and forecast quarterly GDP growth.

The Factor AutoRegressive MIDAS (FAR-MIDAS) model is a further extension of (2) that incorporates autoregressive dynamics of the low frequency variable y_{t_q} together with the latent factors extracted from the higher-frequency variables. However, adding an autoregressive term may produce seasonal effects of the regressors on y_{t_q}.[9] Therefore, following Clements and Galvão (2008), the autoregressive component is introduced as a

[9]Taking the inverse of the autoregressive polynomial may generate a seasonal response of y_{t_q} to the regressors irrespective of whether or not the high-frequency variables display a seasonal pattern (Clements and Galvão, 2008).

common factor in the original Factor-MIDAS (Hendry and Mizon, 1978):

$$y_{t_q} = \beta_0 + \lambda y_{t_q-1} + \beta_1\, b(L_m, \theta^m) F^0_{t_m} + \beta_2\, b(L_w, \theta^w) F^0_{t_w} + u_{t_q}. \tag{3}$$

The autoregressive component captures the time dependence in the data and also conveys useful low-frequency information not contained in the high-frequency variables alone, which would be otherwise absent owing to the fact that, differently from the original paper of Altissimo et al. (2010), our dataset does not include variables observed at a quarterly frequency (besides real GDP).

Both Factor-MIDAS and FAR-MIDAS models provide flexibility in choosing relevant predictor variables, incorporating additional explanatory factors as well as selecting appropriate lags. This flexibility allows researchers to customize the model to their specific nowcasting needs and economic domain knowledge. Additionally, the inclusion of latent factors offers a more interpretable framework for understanding the underlying economic forces driving the low-frequency variable and enhances model interpretation.

In order to estimate the higher-frequency factors, it is assumed that both $X^0_{t_m}$ and $X^0_{t_w}$ can be represented as the sum of two orthogonal elements: a component driven by latent factors common to all the variables in the dataset (the so-called *common component*, χ_t) and an idiosyncratic component (ξ_t). The factor structure of the common component is represented by the r-dimensional vector of factors F^0_t and a loadings matrix Λ linking the observables to the factors, i.e.,

$$X^0_{t_\ell} = \chi_{t_\ell} + \xi_{t_\ell} = \Lambda F^0_{t_\ell} + \xi_{t_\ell} \tag{4}$$

with $\ell = m, w$. Equation (4) represents an approximate factor model, which allows for limited cross-correlation among the idiosyncratic elements in ξ_{t_ℓ}. Following Stock and Watson (2002b), the estimation of factors is conducted via Principal Component Analysis. The results presented in the following sections are obtained with $r = 1$, as in Jardet and Meunier (2022).

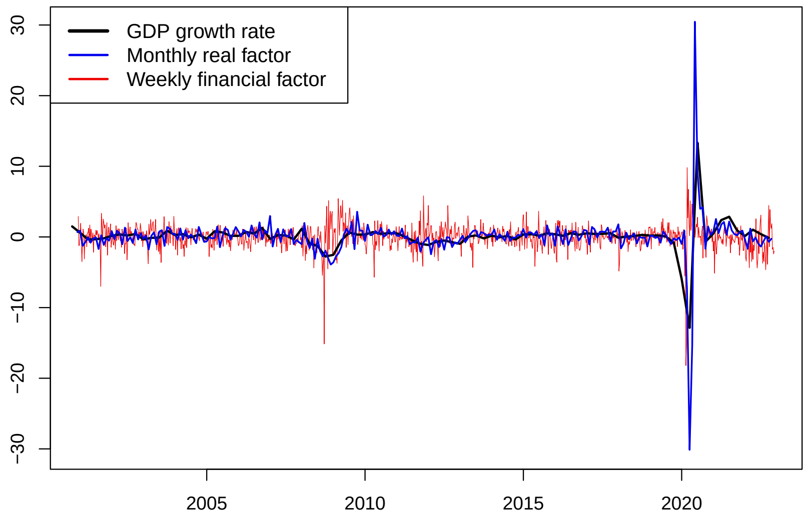
Factors are very useful to synthesize relevant information from large datasets. However, they represent statistical artefacts whose extraction is unsupervised with respect to the target variable, i.e., they are estimated taking into account only the covariance among the variables in the dataset. The forecasting literature has highlighted that the prediction accuracy of models incorporating factors can be improved through the introduction of a pre-selection step. This step intends to select a subset of the original dataset comprising only the more informative predictors (Boivin and Ng, 2006; Bai and Ng, 2008; Schumacher, 2010). Variable pre-selection aims to mitigate issues such as model complexity, multicollinearity, and overfitting, while improving the interpretability and predictive performance of factors. In the spirit of Bai and Ng (2008), the pre-selection step implemented in this paper follows a soft thresholding procedure under which the predictors are ranked following a particular technique, and only the $n^*_\ell \leq N_\ell$ ($\ell = m, w$) top-ranked variables are included in the dataset from which the factors are extracted. In particular, two alternative techniques are considered to rank the variables: (i) the Least Angle Regression (LARS) of Efron et al. (2004), implemented as in Bai and Ng (2008); and (ii) the t-statistic associated with the coefficient in the univariate regression of the target variable on each predictor aggregated at the quarterly frequency (similar to Bai and Ng 2008 and Jurado et al. 2015). While the latter is based on a univariate relationship of each regressor with the low-frequency variable, the former ranks the variables taking into account also the other predictors.[10] The optimal number n^*_ℓ of variables to include in the final dataset from which the factors are extracted is chosen by testing over a grid of values and picking the one optimizing the forecast performance in a *pseudo* out-of-sample exercise (see Section 4.1 for further details). As an illustration, Figure 4 shows the monthly and weekly factors estimated in-sample and using the top $n^*_m = n^*_w = 20$ variables out of the variable ordering produced by the LARS pre-selection.[11] As a matter of comparison, the quarterly GDP growth rate is also displayed in addition to these two

[10]This approach has been extensively used in nowcasting GDP growth (Schumacher, 2010; Bulligan et al., 2015; Jardet and Meunier, 2022). The algorithm adds one regressor at each step: therefore, when the number of iterations is taken as the total number of variables (N_ℓ), it provides an ordering of the regressors in terms of their relative performance to predict the target variable.

[11]As opposed to the *pseudo* real-time exercise described in the following sections, the factors are extracted from the entire sample covering the period January 2001 – December 2022.

factors.

Figure 4: Factors and GDP growth



Define X_{t_m} and X_{t_w} the datasets of size n_m^* and n_w^* obtained by applying the pre-selection to the original data $X_{t_m}^0$ and $X_{t_w}^0$, and F_{t_m} and F_{t_w} the corresponding factors. In Section 4 the prediction performance of four different specifications of MIDAS regressions will be evaluated:

$$
\begin{aligned}
\text{FAR-MIDAS Model 1:} \quad & y_{t_q} = \beta_0 + \lambda y_{t_q-1} + \beta_1 b(L_m, \theta^m) F_{t_m} + \beta_2 b(L_w, \theta^w) F_{t_w} + u_{t_q} \\
\text{FAR-MIDAS Model 2:} \quad & y_{t_q} = \beta_0 + \lambda y_{t_q-1} + \beta_1 b(L_m, \theta^m) F_{t_m} + u_{t_q} \\
\text{Factor-MIDAS Model 1:} \quad & y_{t_q} = \beta_0 + \beta_1 b(L_m, \theta^m) F_{t_m} + \beta_2 b(L_w, \theta^w) F_{t_w} + u_{t_q} \\
\text{Factor-MIDAS Model 2:} \quad & y_{t_q} = \beta_0 + \beta_1 b(L_m, \theta^m) F_{t_m} + u_{t_q}
\end{aligned}
$$

The comparison between Model 1 and Model 2 of the two specifications (FAR-MIDAS and Factor-MIDAS) will allow us to appreciate whether the inclusion of weekly financial variables improves the model's nowcasting performance.

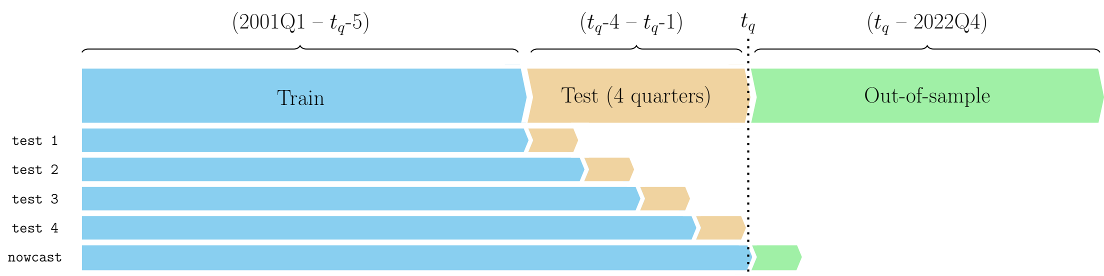
4 Empirical analysis

This Section evaluates the performance of the Factor MIDAS models described in Section 3 by their ability to nowcast the Italian GDP growth rate. The *pseudo* real-time nowcast exercise is presented in Section 4.1, while Section 4.2 discusses the out-of-sample results and Section 4.3 illustrates a robustness analysis.

4.1 The *pseudo* real-time exercise

To mimic the real-time setup, each model is recursively estimated following the procedure shown in Figure 5.

Figure 5: Real-time setup for the nowcasting exercise



In each quarter t_q the ranking of the regressors is performed over the in-sample period (composed by the train set and the test set), which represents the information available to the forecaster in date t_q. The optimal number of regressors n^*_ℓ ($\ell = m, w$) also is recalibrated at each date t_q. The calibration of the hyperparameters n^*_ℓ is obtained estimating each model in the train set ($2001Q1 - t_q$-5) and producing a *pseudo* out-of-sample nowcast in the test set (rolling from $t_q - 4$ to $t_q - 1$). Note that, being a rolling window, the test set selects the model with the best temporal proximity to the nowcast. This process is repeated across a predetermined grid of values n^*_ℓ ($\ell = m, w) \in \{20, 40, 60, 80, 100\}$.[12] The optimal number of regressors is chosen with the aim

[12]The choice to rely on a small grid of values is made in order to obtain computational efficiency and stability of the hyperparameter n^*_ℓ. However, Section 4.3.3 presents also the results obtained using a wider grid of values for n^*_ℓ.

of minimizing the root mean squared error in the test sample, and the resulting n_m^* and n_w^* are allowed to be different. Factor-MIDAS and FAR-MIDAS based on n_m^* and n_w^* are then used to nowcast the Italian GDP growth rate in t_q, which represents the first period in the out-of-sample. The in-sample is then expanded to include t_q, moving forward the test set (t_q-3 $-$ t_q) and adding one observation at the end of the train set (2001$Q1$ $-$ t_q-4); the ranking of the regressors as well as n_m^* and n_w^* are re-estimated and a new nowcast for quarter $t_q + 1$ is produced.

This procedure is separately repeated for the first, second and third months of the quarter,[13] assuming that the information relevant in the first month may be different from that in the second and third months, owing in particular to the ragged-edge nature of the data. An analysis of the variables in the pre-selection step confirms this assumption.

4.2 Nowcasting results

This Section illustrates the nowcasting results for the out-of-sample 2015$Q1$ $-$ 2022$Q4$ obtained with the LARS pre-selection method (results based on the t-stat ranking are reported in Section 4.3)[14]. Performance of model i is evaluated in terms of its Root Mean Squared Forecast Error (RMSFE), a commonly employed metric for assessing the accuracy of forecasting models defined as follows:

$$RMSFE^{(i)} = \sqrt{\frac{1}{n}\sum\nolimits_{t_q=2015Q1}^{2022Q4}(y_{t_q} - \hat{y}_{t_q}^{(i)})^2}, \tag{5}$$

[13] Depending on the month of the quarter in which the nowcast is produced, the test sample is composed by the corresponding months (the first, the second or the third) of the four preceding quarters. As an illustration, suppose the researcher wants to nowcast GDP growth in the second quarter, being in April 2020: given that April corresponds to the first month of the (second) quarter, the test sample is composed by January 2020, October 2019, July 2019 and April 2019. Therefore, the test sample is composed by four quarters. The use of an 8-quarter window (based on the average business-cycle expansion phase in Italy, which exceeds one year) leads to a smoother and less erratic variable selection process. However, this choice only marginally affects the forecast results, which are available from the authors upon request.

[14] Altering the initial year of the out-of-sample does not affect the results and the main findings. Additional results based on the out-of-samples 2014$Q1$ $-$ 2022$Q4$ and 2010$Q1$ $-$ 2022$Q4$ are available from the authors upon request.

where y_{t_q} and $\hat{y}_{t_q}^{(i)}$ are, respectively the observed official GDP growth rate and the corresponding nowcast obtained with model i.

In addition to the nowcast results of the factor MIDAS models discussed in Section 3, this paper evaluates the performance of five additional benchmark models:

1. A Bridge Model (Golinelli and Parigi, 2007) using industrial production as a monthly predictor;
2. An ARIMA(p, d, q) model with automatically selected lags for each nowcast;
3. A Random Walk model;
4. A simple AR(1) model;
5. The median projection in the monthly survey of economists released by Bloomberg Consensus.

For the sake of comprehensiveness and in order to provide insights into the nowcasting performance of the Factor MIDAS models before the publication of the *flash* GDP estimate, the first month involves two separate nowcasts: one with the inclusion of *flash* GDP and one without. Specifically, results obtained taking into account the *flash* estimate of GDP growth are presented as an additional nowcast produced in the first month. In this case, for both Factor MIDAS specifications, the *flash* GDP estimate is considered as a supplementary data point exclusively in the information set available at the end of the first month: therefore, variable pre-selection and estimation are carried out using a larger information set which also includes the monthly and weekly variables of the most recent quarter.

Table 2 reports the main nowcast results in terms of RMSFE ratios. Model comparison is based on the ratio between the RMSFE of each competing model and the RMSFE of the AR(1) process. Thus, values greater than one indicate that the RMSFE of a model is larger than the RMSFE of the AR(1), and vice versa. Figures A.1 and A.2 in Appendix A display the corresponding nowcasts, while Figure A.3 shows the variation of the optimal n_ℓ^* for factor estimation during the out-of-sample exercise. Finally, in order to shed additional light on the composition of factors in each Factor MIDAS model, Figures A.4, A.5, A.6, A.7 show the heat maps of factor loading values.

As expected, as the flow of information increases, i.e., moving from the first to the third month of the quarter, the nowcasting performance of all models improves. FAR-MIDAS and Factor-MIDAS models exhibit a substantial superiority over the classical univariate benchmark models. There is also a clear improvement in their nowcasting accuracy compared to a standard Bridge Model using the industrial production index as predictor, with a precision gain of up to almost 30%: looking solely at the contribution of the industrial sector turns out to be slightly misleading when nowcasting economic activity in general.

Overall, Model 1 outperforms Model 2 in almost all the nowcasting scenarios for both FAR-MIDAS and Factor-MIDAS specifications. The only exceptions are in the second and in the third months for Factor-MIDAS, although the results of Model 2 point only to a minor gain of 1.5%. More specifically, in the first month Model 1 reports a gain of approximately 4% in terms of accuracy with respect to Model 2 across both MIDAS specifications. The lowest RMSFE ratio is attained in the third month with FAR-MIDAS Model 1, a value which is very close to the well-known benchmark of Bloomberg Consensus based on prediction and, primarily, judgment.

The strong performance of Model 1 suggests including financial variables when producing nowcasts of the GDP growth rate, particularly at the beginning of the quarter, when limited information regarding the reference quarter is available.

Table 2: RMSFE relative to AR(1). Out-of-sample: $2015Q1 - 2022Q4$.

	FAR-MIDAS		Factor-MIDAS		BRIDGE	ARIMA	Random	Bloomberg
	Model 1	Model 2	Model 1	Model 2	(indpro)	(auto)	Walk	consensus
month 1	**0.696**	0.740	0.720	0.760	0.889	1.239	1.245	0.354
month 1 (*flash*)	0.525	0.560	**0.473**	0.507	0.596	1.167	1.108	0.303
month 2	0.372	0.417	0.385	**0.369**	0.658	1.196	1.092	0.298
month 3	**0.300**	0.432	0.327	0.314	0.528	1.196	1.092	0.274

Notes: results refer to the LARS pre-selection. Model comparison is based on the ratio between the RMSFE of each competing model and the RMSFE of the AR(1) process. Bold values indicate the best RMSFE ratio for a given month without considering Bloomberg consensus. The term "month 1 (*flash*)" indicates the nowcasting produced in the first month also using the information of the *flash* GDP estimate. "BRIDGE (indpro)" refers to a BRIDGE model built with the Industrial Production Index (IPI) and an autoregressive term of order one.

Table 3 compares the predictive accuracy of the models based on the tests of Diebold and Mariano (1995), Giacomini and White (2006) and Harvey et al. (1997).[15]
The Giacomini and White (2006) and Harvey et al. (1997) tests are alternative versions of the Diebold-Mariano test: while the first is designed to capture the effect of estimation uncertainty on relative forecast performance and to handle forecasts derived from both nested and nonnested models, the latter addresses the issue of limited sample sizes.

Note that while the 4% gain obtained by both Models 1 with respect to Models 2 in the first month of the quarter (with and without the *flash* estimate of GDP growth) is statistically significant, this is not the case for the differences in the second and third months for both Factor MIDAS models.

Table 3: Predictive accuracy tests. Out-of-sample: $2015Q1 - 2022Q4$.

	FAR-MIDAS						Factor-MIDAS					
	Model 1 vs. Model 2			Model 1 vs. AR(1)			Model 1 vs. Model 2			Model 1 vs. AR(1)		
	DM	GW	HLN	DM	GW	HLN	DM	GW	HLN	DM	GW	HLN
month 1	**0.01**	**0.01**	**0.01**	**0.03**	**0.02**	**0.03**	**0.01**	**0.01**	**0.01**	**0.03**	**0.02**	**0.03**
month 1 (*flash*)	**0.03**	**0.02**	**0.03**	0.14	0.14	0.14	**0.03**	**0.02**	**0.03**	0.15	0.15	0.15
month 2	0.25	0.25	0.25	0.14	0.13	0.14	0.81	0.82	0.81	0.13	0.13	0.13
month 3	0.20	0.19	0.20	0.14	0.13	0.14	0.74	0.74	0.74	0.14	0.14	0.14

Notes: results refer to the LARS pre-selection. Bold values indicate rejection of null hypothesis of equal predictive at 10% significance level. The acronyms DM, GW and HLN refer respectively to Diebold-Mariano, Giacomini-White and Harvey-Leybourne-Newbold tests.

In order to have a visual inspection of the baseline results presented in Table 2 and Table 3, it is possible to use the Cumulative Sum of Squared Forecast Error Difference (CSSFED), see Welch and Goyal (2007), allowing us to understand in which part of the out-of-sample financial variables matter the most in terms of predictive performance (e.g., FAR-MIDAS model 1 compared to FAR-MIDAS model 2 and AR(1) model). The CSSFED is thus used to offer insights into the most relevant variables for nowcasting GDP growth over time. The CSSFED is presented in Figure 6 and Figure 7 for the FAR-MIDAS and Factor-MIDAS models, respectively.

[15] In this paper, all the predictive accuracy tests comparing two competing models are structured as follows: the null hypothesis (H_0) assumes equal accuracy between the models, while the alternative hypothesis (H_1) posits that the second competing model exhibits lower predictive accuracy compared to the first model.

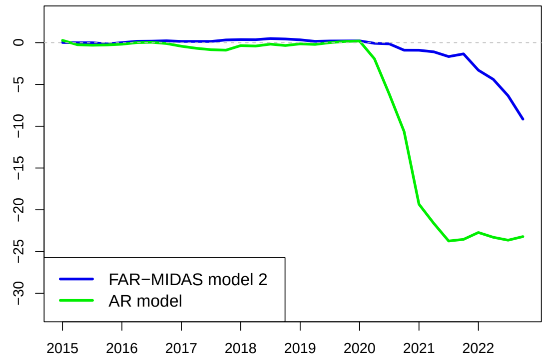
After examining the CSSFED charts alongside the heat maps in Appendix A, it becomes apparent that financial variables play a significant role starting from 2020/2021 onwards, particularly in the first month of the quarter (both with and without the flash GDP estimate). The most important financial categories during this period are:

- forward yield of earnings per share and price-earnings ratio
- options-adjusted spread and yield to maturity of banks and firms

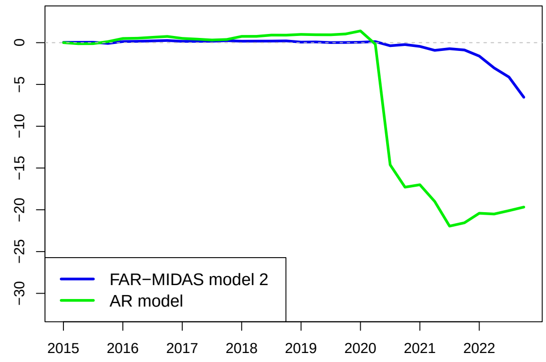
The most important macroeconomic categories in the first month of nowcasting are: Producer Price Index (PPI), coincident indicators and industrial production.

In the second and third months of nowcasting, when financial information becomes less relevant following the release of key macroeconomic variables for the current quarter, the main macroeconomic categories, in order of importance, become: industrial production, coincident indicators, and demand. It is worth noting that the Producer Price Index (PPI) is released with a one-month delay, while the industrial production with a two-month delay. This explains the shift in relevance from the Producer Price Index (PPI) to industrial production between the first and second month, as new information on industrial production for the current quarter is incorporated for the first time in the second month.

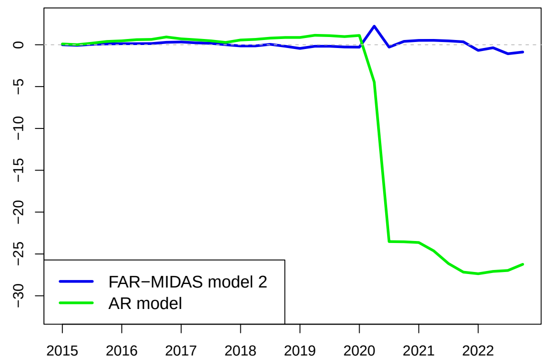
Figure 6: CSSFED of FAR-MIDAS model 1 compared to FAR-MIDAS model 2 and AR(1) model in $2015Q1 - 2022Q4$ out-of-sample.



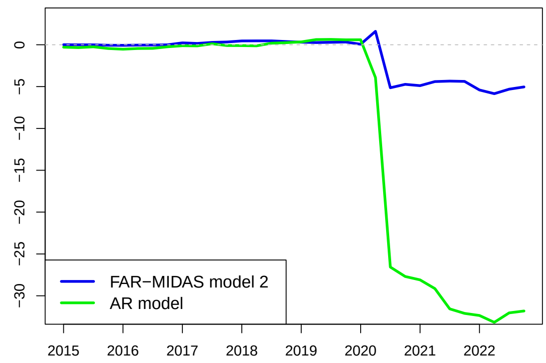
(a) month 1



(b) month 1 (*flash*)

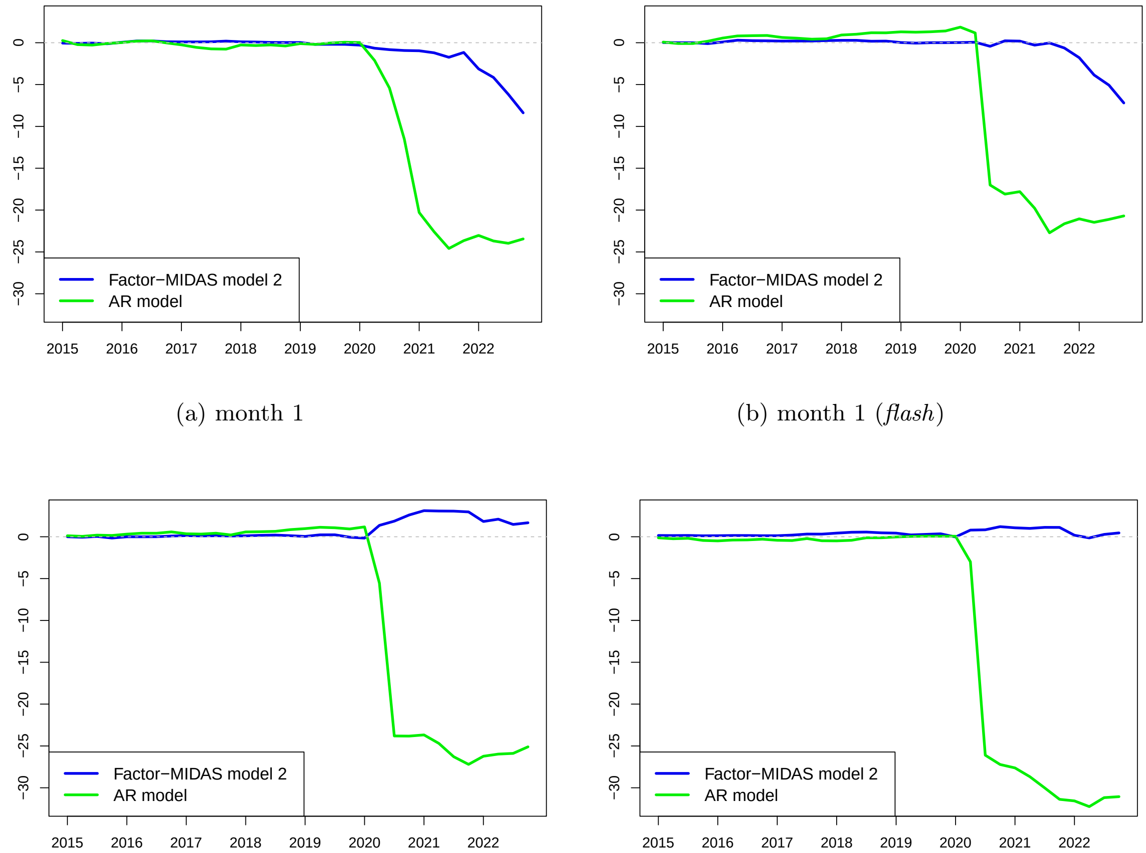


(c) month 2



(d) month 3

Figure 7: CSSFED of Factor-MIDAS model 1 compared to Factor-MIDAS model 2 and AR(1) model in $2015Q1 - 2022Q4$ out-of-sample.



(a) month 1

(b) month 1 (*flash*)

(c) month 2

(d) month 3

These results are also supported by an analysis based on the Model Confidence Set (Hansen et al., 2011) presented in Table 4. With the Model Confidence Set procedure one can evaluate the accuracy of a set of models and not only the performance of two models each time. With an 80% confidence level, in the first month FAR-MIDAS and Factor-MIDAS Models 1 are included in the set of superior models, while both Models 2 are excluded. In the second month the small difference in accuracy between FAR-MIDAS Model 1 and Factor-MIDAS Model 2 is instead not statistically significant. In fact, the p-values of these two models in the Model Confidence Set are almost equal (respectively, 0.966 and 1).

Table 4: Model Confidence Set (T_R) - I.C. 80%. Out-of-sample: $2015Q1 - 2022Q4$.

	FAR-MIDAS		Factor-MIDAS		AR(1)
	Model 1	Model 2	Model 1	Model 2	
month 1	1	×	0.260	×	×
month 1 (*flash*)	0.534	×	1	×	0.265
month 2	0.966	0.650	0.652	1	0.483
month 3	1	0.740	0.619	0.784	0.589

Notes: results refer to the LARS pre-selection and report p-values of T_R statistics of Hansen et al. (2011) for a set of superior models (SSM) with a confidence level at 80% (α=0.2). The symbol × indicates the exclusion of the model from the SSM.

Nowcasting with single financial categories

Taking into consideration the information heterogeneity across financial variables, the nowcasting exercise is now replicated separately for each of the eight financial categories. In doing so, the financial variables are not subject to pre-selection prior to factor estimation given the homogeneity of information within each category. The results of this exercise are reported in Table 5 and Table 6 for the FAR-MIDAS model. In general, Model 1 demonstrates superior performance compared to Model 2 across various financial categories. Note that, particularly in the first month of the quarter, corporate credit risk indicators ('oas & ytm') produce an improvement in nowcasting GDP growth, exhibiting a statistically significant difference compared to Model 2.

Table 5: RMSFE of FAR-MIDAS with single financial categories relative to AR(1). Out-of-sample: $2015Q1 - 2022Q4$.

	Model 1								Model 2
	forward yield (EPS and PE)	price earnings	earnings per share	oas & ytm (banks + firms)	exchange rates	stock indices	gov. bonds	commod. futures	
month 1	0.745	0.730	**0.693**	0.701	0.704	0.758	0.729	0.724	0.740
month 1 (*flash*)	0.549	0.542	0.548	0.538	0.559	0.546	**0.537**	0.543	0.560
month 2	0.401	0.400	0.412	0.389	0.423	**0.361**	0.420	0.437	0.417
month 3	0.370	0.325	0.476	0.324	0.464	0.365	0.449	**0.292**	0.432

Notes: the results refer to the LARS pre-selection. The columns represent the financial category used for the estimation of the financial factor in the FAR-MIDAS model, thus excluding all other financial categories from the dataset.

Table 6: Diebold-Mariano test of FAR-MIDAS with single financial categories. Out-of-sample: $2015Q1 - 2022Q4$.

	Model 1 vs. Model 2							
	forward yield (EPS and PE)	price earnings	earnings per share	oas & ytm (banks + firms)	exchange rates	stock indices	gov. bonds	commod. futures
month 1	0.70	0.21	0.15	**0.03**	0.91	0.77	0.17	**0.06**
month 1 (*flash*)	0.13	**0.01**	0.22	**0.04**	0.43	**0.07**	**0.04**	**0.08**
month 2	0.21	0.19	0.30	0.27	0.85	0.18	0.81	0.66
month 3	0.17	0.23	0.58	0.25	0.87	0.28	0.88	0.20

Notes: the results are related to the FAR-MIDAS models with LARS pre-selection. Bold values indicate rejection of null hypothesis of equal predictive ability at 10% significance level.

The enhancements achieved using specific financial categories are also statistically significant when using Factor-MIDAS models, as reported in Table A.1 and Table A.2 in Appendix A. Similar to the FAR-MIDAS scenario, Model 1 exhibits superior performance over Model 2 within the same financial categories.

The role of variable pre-selection

The importance of pre-selecting the most informative variables before extracting the factors is well resumed in Table 7, which presents the RMSFE of the FAR-MIDAS and Factor-MIDAS models using two pre-selection methods (LARS and t-statistics), as well as no pre-selection at all. The absence of pre-selection produces the highest RMSFE for all three months of the quarter throughout the out-of-sample period $2015Q1 - 2022Q4$. Furthermore, by closely examining the bold values which represent the best RMSFE for a specific month and pre-selection method, it becomes evident that they predominantly belong to Model 1.

LARS pre-selection achieves the lowest RMSFE in the third month (1.514), with a substantial difference compared to the t-statistics technique in the same month. In Table 7 the mentioned difference stands out as the largest between the two pre-selection methods considering bold values for a specific month. This is in line with Jardet and Meunier (2022) and Chinn et al. (2023), who find that the LARS is the best dimension-reduction technique compared with t-statistics and others methods, making it the recommended one.[16]

Table 7: RMSFE across pre-selection techniques. Out-of-sample: $2015Q1 - 2022Q4$.

technique	model	month 1	month 1 (*flash*)	month 2	month 3
LARS	FAR-MIDAS Model 1	**2.956**	2.605	1.873	**1.514**
	FAR-MIDAS Model 2	3.144	2.775	2.102	2.178
	Factor-MIDAS Model 1	3.058	**2.346**	1.941	1.649
	Factor-MIDAS Model 2	3.229	2.515	**1.861**	1.585
t-statistics	FAR-MIDAS Model 1	**2.297**	**1.726**	1.874	2.585
	FAR-MIDAS Model 2	2.420	1.734	2.194	2.609
	Factor-MIDAS Model 1	2.639	1.736	**1.818**	2.688
	Factor-MIDAS Model 2	2.549	1.832	1.866	**2.453**
none	FAR-MIDAS Model 1	3.450	4.493	2.870	2.987
	FAR-MIDAS Model 2	3.580	4.858	3.098	3.231
	Factor-MIDAS Model 1	**3.258**	**2.509**	**2.459**	**2.692**
	Factor-MIDAS Model 2	3.466	2.904	2.583	2.913

In conclusion, all the results so far presented convey two clear messages:

(i) financial variables play a crucial role in filling the information gap during the first month of the quarter, when there are significant delays in the publication of macroeconomic data;

(ii) the inclusion of financial information in Factor MIDAS models does not result in a deterioration of the nowcasts, but rather it tends to enhance their accuracy.

For the sake of completeness, additional nowcasting results for different out-of-sample periods are reported in Appendix B.

4.3 Sensitivity analysis

This Section examines the robustness of the nowcasting results presented in Section 4.2. Specifically, five types of sensitivity analysis are proposed: (i) sensitivity to the pre-

[16] Observing the results obtained through the t-statistics method, it can be noted that the RMSFE does not consistently improve throughout the quarter as more information is integrated: while pre-selection may aid in identifying more informative variables, in the present case when using the t-statistics method the incremental information gained from incoming variables is limited compared to outgoing ones.

selection method, (ii) sensitivity to the number of variables n_ℓ^* used for factor estimation, (iii) sensitivity to the grid of values for n_ℓ^*, (iv) sensitivity to the number of real factors, and (v) sensitivity to pre-selection in real-time.

4.3.1 Sensitivity to pre-selection method

The nowcasting results obtained using the t-statistics pre-selection method are shown in Table 8 and Table 9 as a test of robustness. The RMSFE ratios highlight an overperformance of Model 1 in the initial two months of the quarter. Although there is an edge in the first and second months as well, only the FAR-MIDAS model without the *flash* GDP reaches a statistical significance at the 10% level, as shown in Table 9. It is important to bear in mind the evidence supporting the message presented at the end of Section 4.2: also pre-selecting the variables with the t-statistics there is no deterioration of the nowcast results, with a potential gain when including financial variables.

Table 8: RMSFE relative to AR(1). Out-of-sample: $2015Q1 - 2022Q4$.

	FAR-MIDAS		Factor-MIDAS	
	Model 1	Model 2	Model 1	Model 2
month 1	**0.542**	0.570	0.622	0.601
month 1 (*flash*)	**0.349**	**0.349**	0.350	0.370
month 2	0.371	0.433	**0.361**	0.370
month 3	0.513	0.516	0.533	**0.486**

Notes: the results are related to the t-statistics pre-selection. Bold values indicate the best RMSFE ratio for a given month.

Table 9: Diebold-Mariano test. Out-of-sample: $2015Q1 - 2022Q4$.

	Model 1 vs. Model 2	
	FAR-MIDAS	Factor-MIDAS
month 1	**0.06**	0.85
month 1 (*flash*)	0.49	0.19
month 2	0.25	0.19
month 3	0.46	0.84

Notes: the results are related to the t-statistics pre-selection.

4.3.2 Sensitivity to optimal number of variables n_ℓ^* used for factor estimation

The sensitivity of the nowcasting results to the optimal number of variables n_ℓ^* ($\ell = m, w$) used for factor estimation is assessed in Table 10 and Table 11, which report the nowcasting performance of the FAR-MIDAS model when n_ℓ^* is fixed over the entire sample and it is not optimized via the rolling test set of one year (as described in Section 4.1). The tables report the results with several values of $n_\ell^* \in \{20, 40, 60, 80, 100\}$, each time kept constant over the whole sample. For comparative purposes, the first column presents the results of Table 2 obtained with the optimal n_m^* and n_w^* values.

It can be observed that an increase in the fixed n_ℓ^* leads to an almost linear rise in the average RMSFE. This result is in line with the findings of Boivin and Ng (2006), Bai and Ng (2008), Schumacher (2010), Jardet and Meunier (2022) and Chinn et al. (2023), confirming that more data are not always better for factor analysis.

Table 10: RMSFE of FAR-MIDAS with fixed n_ℓ^* relative to AR(1). Out-of-sample: $2015Q1 - 2022Q4$.

	n_m^*, n_w^*		n_m^*=n_w^*=20		n_m^*=n_w^*=40		n_m^*=n_w^*=60		n_m^*=n_w^*=80		n_m^*=n_w^*=100	
	Mod. 1	Mod. 2	Mod. 1	Mod. 2	Mod. 1	Mod. 2	Mod. 1	Mod. 2	Mod. 1	Mod. 2	Mod. 1	Mod. 2
month 1	**0.696**	0.740	0.736	0.745	0.744	0.731	0.826	0.827	0.773	0.810	0.808	0.867
month 1 (*flash*)	**0.525**	0.560	0.537	0.562	0.739	0.757	0.880	0.956	0.857	0.938	0.900	0.988
month 2	**0.372**	0.417	0.393	0.386	0.420	0.460	0.409	0.462	0.420	0.483	0.501	0.538
month 3	**0.300**	0.432	0.335	0.428	0.310	0.462	0.359	0.468	0.442	0.527	0.495	0.568
average	**0.473**	0.537	0.500	0.530	0.553	0.603	0.619	0.678	0.623	0.690	0.676	0.740

Notes: the results refer to the LARS pre-selection. Bold values indicate the best RMSFE ratio for a given month.

Table 11: Diebold-Mariano test of FAR-MIDAS with fixed n_ℓ^*. Out-of-sample: $2015Q1 - 2022Q4$.

	Model 1 vs. Model 2					
	n_m^*, n_w^*	n_m^*=n_w^*=20	n_m^*=n_w^*=40	n_m^*=n_w^*=60	n_m^*=n_w^*=80	n_m^*=n_w^*=100
month 1	**0.01**	**0.08**	0.68	0.50	0.23	0.17
month 1 (*flash*)	**0.03**	**0.08**	**0.09**	0.13	**0.10**	**0.09**
month 2	0.25	0.82	0.12	0.18	0.15	0.22
month 3	0.20	0.27	0.17	0.15	0.16	0.20

Notes: the results are related to the FAR-MIDAS models with LARS pre-selection. Bold values indicate rejection of null hypothesis of equal predictive ability at 10% significance level.

For the sake of completeness, the results obtained with the Factor-MIDAS model are available in Table A.3 and Table A.4 in Appendix A, confirming the same message as above.

4.3.3 Sensitivity to the grid of values for n_ℓ^*

All the previous results are obtained with the grid of values $n_\ell^* \in \{20, 40, 60, 80, 100\}$ mainly for computational efficiency reasons, as explained in Section 2.2. The robustness of the results is now evaluated with a wider grid: $n_w^* \in \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$ and $n_m^* \in \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110\}$.[17] Therefore, the rolling one-year test designed to determine the optimal number of variables n_ℓ^* for estimating the factors is carried out with 10/11 possible values instead of 5. The corresponding results are reported in Table 12 and Table 13, supporting the two conclusions at the end of Section 4.2 about the contribution of financial variables for nowcasting.

Table 12: RMSFE relative to AR(1). Out-of-sample: $2015Q1 - 2022Q4$.

	FAR-MIDAS		Factor-MIDAS	
	Model 1	Model 2	Model 1	Model 2
month 1	**0.691**	0.706	0.698	0.732
month 1 (*flash*)	0.470	0.512	**0.451**	0.491
month 2	0.377	0.382	**0.362**	0.374
month 3	**0.310**	0.449	0.456	0.344

Notes: results refer to the LARS pre-selection. Bold values indicate the best RMSFE ratio for a given month.

[17] The two grids have different ranges due to the size of their respective datasets. The monthly macroeconomic dataset consists of 118 variables, so the maximum reachable n_m^* is 110, while the weekly financial dataset consists of 109 variables, resulting in a maximum n_w^* of 100.

Table 13: Diebold-Mariano test. Out-of-sample: $2015Q1 - 2022Q4$.

	Model 1 vs. Model 2	
	FAR-MIDAS	Factor-MIDAS
month 1	**0.09**	**0.01**
month 1 (*flash*)	0.20	**0.03**
month 2	0.44	0.28
month 3	0.15	0.81

Notes: the table reports the p-values of the test with bold values indicating a significance level of 10%.

4.3.4 Sensitivity to the number of real factors

One might be concerned about the fact that the baseline specification of Model 2 includes only one factor ($r = 1$), limiting the information content of macroeconomic variables in nowcasting the GDP growth rate and, therefore, biasing the comparison between Model 1 and Model 2. To address this issue, the sensitivity analysis presented in this Section allows Model 2 to incorporate one, two or three real factors ($r \in (1, 2, 3)$). The number of real factors is determined using the rolling window test described in Section 4.1, which selects the best combination of n_m^* and, this time, also considers the best number of real factors that produces the lowest RMSFE during the one-year test. Model 1 remains unchanged with two factors, one real and the other financial.

In comparison to Table 2, the RMSFE ratios in Table 14 highlight a remarkable precision gain of Model 2 in the second month of the quarter (even outperforming the Bloomberg Consensus GDP nowcast with the FAR-MIDAS specification, see Table 2), but only a small gain in the first month. Despite these improvements, Model 2 does not outperform Model 1 in month 1 and month 1 (*flash*), as confirmed by the Diebold-Mariano tests in Table 15, which once again demonstrate the statistical superiority of Model 1 in the first month of the quarter.

Table 14: RMSFE relative to AR(1). Out-of-sample: $2015Q1 - 2022Q4$.

	FAR-MIDAS		Factor-MIDAS	
	Model 1	Model 2	Model 1	Model 2
month 1	**0.696**	0.765	0.720	0.767
month 1 (*flash*)	0.525	0.527	**0.473**	0.531
month 2	0.372	**0.295**	0.385	0.303
month 3	**0.300**	0.437	0.327	0.314

Notes: results refer to the LARS pre-selection. Bold values indicate the best RMSFE ratio for a given month.

Table 15: Diebold-Mariano test. Out-of-sample: $2015Q1 - 2022Q4$.

	Model 1 vs. Model 2	
	FAR-MIDAS	Factor-MIDAS
month 1	**0.07**	**0.06**
month 1 (*flash*)	0.44	**0.06**
month 2	0.96	0.95
month 3	0.19	0.75

Notes: the table reports the p-values of the test with bold values indicating a significance level of 10%.

Table A.5 and Table A.6 in Appendix A present corresponding results with a maximum of two real factors for Model 2, supporting our previous findings.

4.3.5 Sensitivity to pre-selection in real-time

One of the main features of the *pseudo* real-time exercise is the pre-selection of the variables conducted in each quarter based on the prediction performance in the test set. The inclusion of this step in the nowcasting procedure is motivated by the presumption that, although some variables are relevant over almost the entire sample, some indicators may gain relevance in specific periods of time while losing usefulness in others (e.g., during periods of financial turmoil or geopolitical tensions). On the other hand, a one-shot selection of the variables creates a fixed dataset and may produce a computational gain.

Table 16 and Table 17 report the results when LARS pre-selection is applied only once at the beginning of the out-of-sample, i.e., using the sample from 2001Q1 to 2014Q4. The obtained findings align with previous evidence and demonstrate an improvement in the nowcasting accuracy of Model 1 relative to Model 2. Moreover, these results further affirm the two conclusions at the end of Section 4.2 regarding the statistically significant contribution of financial variables in the first month of the quarter.

Table 16: RMSFE relative to AR(1). Pre-selection in $2014Q4$.

	FAR-MIDAS		Factor-MIDAS	
	Model 1	Model 2	Model 1	Model 2
month 1	0.628	0.686	**0.622**	0.698
month 1 (*flash*)	0.739	0.764	**0.433**	0.437
month 2	**0.396**	0.465	0.397	0.423
month 3	0.450	0.453	0.334	**0.326**

Notes: results refer to the LARS pre-selection. Bold values indicate the best RMSFE ratio for a given month.

Table 17: Diebold-Mariano test. Pre-selection in $2014Q4$.

	Model 1 vs. Model 2	
	FAR-MIDAS	Factor-MIDAS
month 1	**0.09**	**0.08**
month 1 (*flash*)	**0.03**	0.37
month 2	0.11	0.24
month 3	0.47	0.62

Notes: the table reports the p-values of the test with bold values indicating a significance level of 10%.

In Appendix A (Nowcasting with one-time pre-selection), the same exercise is described using the entire sample period $2001Q1 - 2022Q4$ for performing only once the pre-selection process and then using it for nowcasting over the out-of-sample $2015Q1 - 2022Q4$.

5 Concluding remarks

This paper provides a comprehensive analysis of the incorporation of weekly financial variables into Factor MIDAS models, alongside traditional monthly macroeconomic variables, with the objective of assessing the capacity of financial variables to enhance the accuracy of Italian GDP growth nowcasting. The use of weekly financial variables improves the nowcasting, particularly at the beginning of the quarter, due to the lags in the publication of the macroeconomic variables; there is a statistically significant improvement in the nowcasting performance.

The robustness of these results is tested across several dimensions, including the variable pre-selection method for factor estimation, the number of variables employed, the granularity of the grid used to select the optimal number of variables, the number of real factors considered, and the time-varying pre-selection ordering. Building upon the research of Boivin and Ng (2006), Bai and Ng (2008), Schumacher (2010), Jardet and Meunier (2022) and Chinn et al. (2023), our results align with their findings that a targeted selection of variables for factor estimation improves the accuracy of Factor MIDAS models.

The results suggest that not only does the inclusion of weekly financial variables in the Factor MIDAS model not compromise the accuracy of nowcasting, but rather it has the potential to enhance it, particularly during periods of major economic shocks, such as the Covid-19 pandemic. Consequently, the forecaster has the opportunity to update the nowcasts on a weekly basis to ensure timely monitoring of real GDP growth.

Further research might explore additional financial variables and alternative methodologies to effectively integrate high-frequency financial data into the modelling framework, thereby improving the nowcasting accuracy. Moreover, it would be interesting to extract common factors from blocks of similar variables, in the framework of hierarchical dynamic factor models as proposed by Moench et al. (2013). Lastly, based on findings presented by Ferrara and Marsilli (2013) regarding differences in the forecasting ability of financial variables across the four main euro-area countries, performing robustness tests across other countries would offer valuable insights into the generalizability of the results. All these issues are left for future investigations.

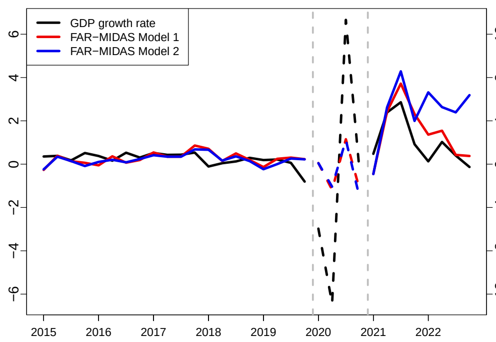
APPENDIX

A Further results

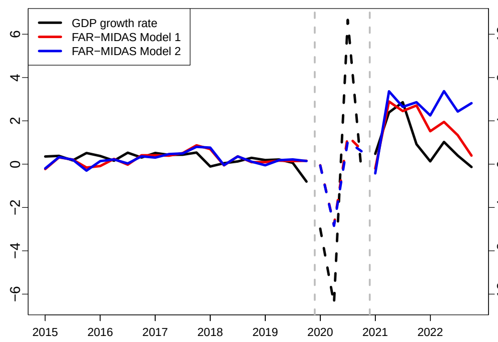
Nowcasting plots

Figure A.1 and Figure A.2 display the nowcasts of real GDP growth underlying Table 2.

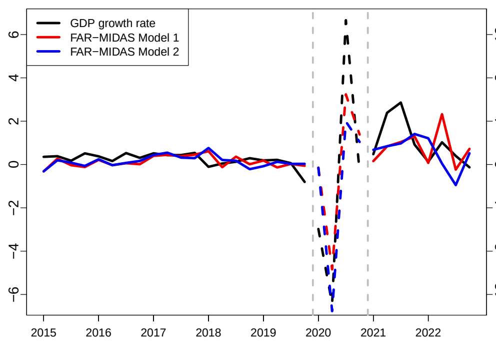
Figure A.1: FAR-MIDAS out-of-sample nowcasts against real quarterly GDP growth



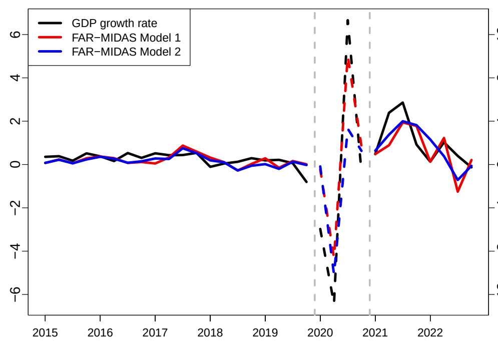
(a) month 1



(b) month 1 (*flash*)



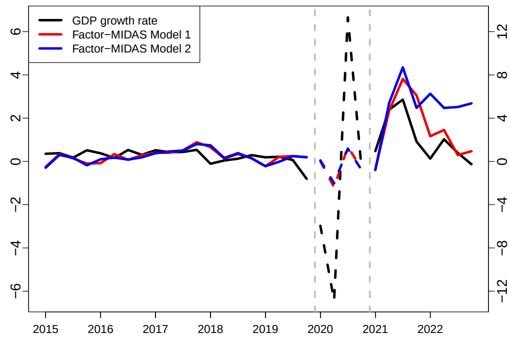
(c) month 2



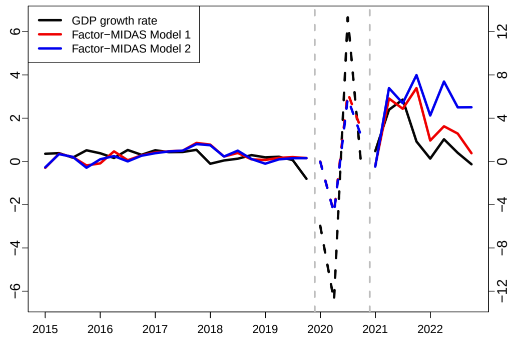
(d) month 3

Notes: the values in the period $2020Q1 - 2020Q4$, enclosed by the two dashed gray lines, have been halved to provide a more detailed view of the $2015 - 2019$ and $2021 - 2022$ out-of-sample periods. The scale on the right side of the graphs represents the original values for the period $2020Q1 - 2020Q4$.

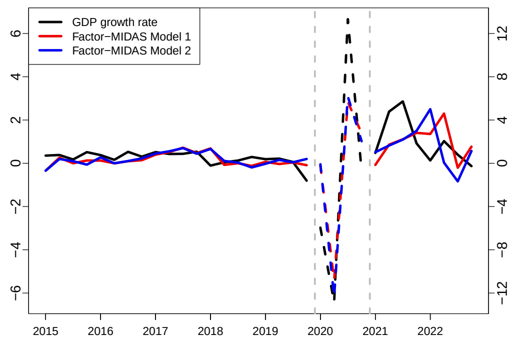
Figure A.2: Factor-MIDAS out-of-sample nowcasts against real quarterly GDP growth



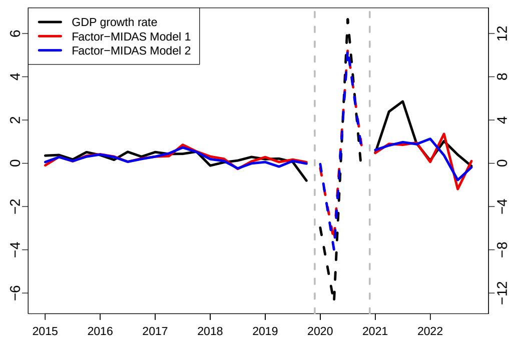
(a) month 1



(b) month 1 (*flash*)



(c) month 2



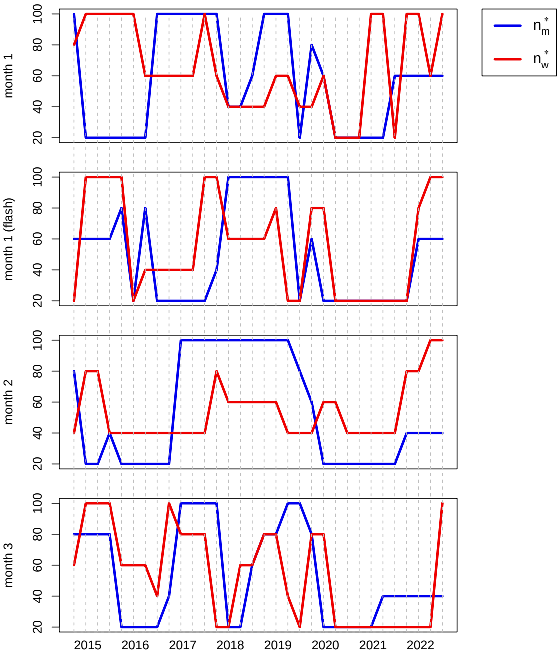
(d) month 3

Notes: the values in the period $2020Q1 - 2020Q4$, enclosed by the two dashed gray lines, have been halved to provide a more detailed view of the $2015 - 2019$ and $2021 - 2022$ out-of-sample periods. The scale on the right side of the graphs represents the original values for the period $2020Q1 - 2020Q4$.

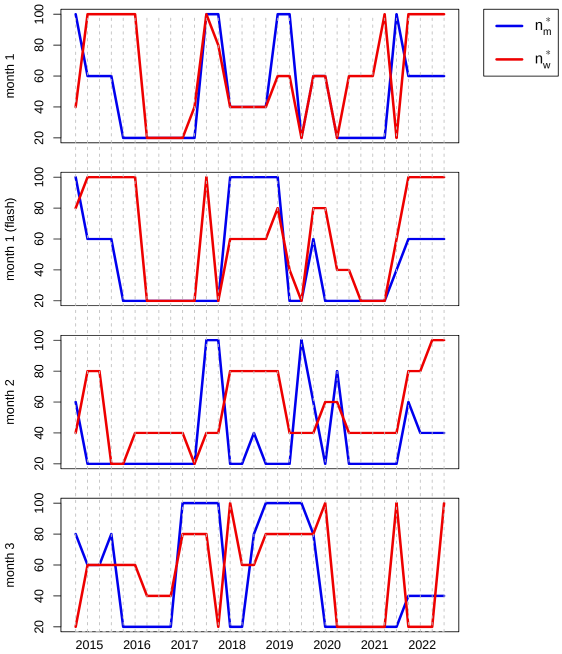
Optimal number of variables n_ℓ^* for factor estimation

Figure A.3 displays the plots of the optimal number of variables n_ℓ^* used for factor estimation, related to the results obtained from both Factor MIDAS models in Table 2 of Section 4.2. It can be observed that, starting from 2020 onwards, the optimal number of monthly real variables (n_m^*) tends to decrease and stabilize at relatively low values.

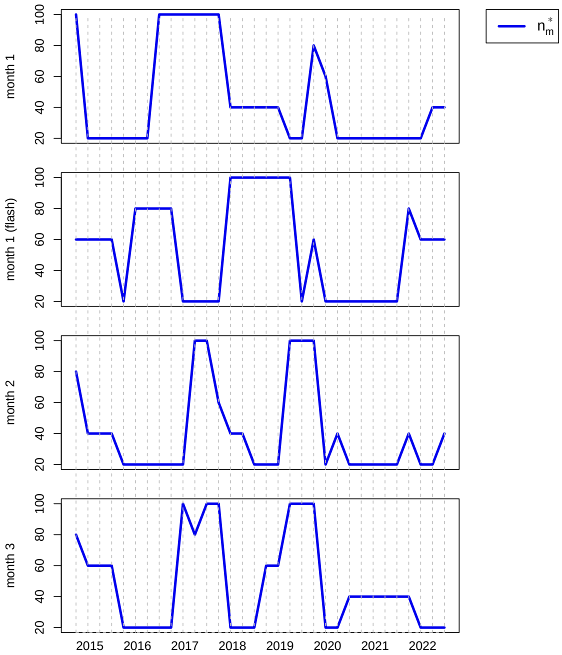
Figure A.3: Optimal n_ℓ^* for $2015Q1 - 2022Q4$ out-of-sample.



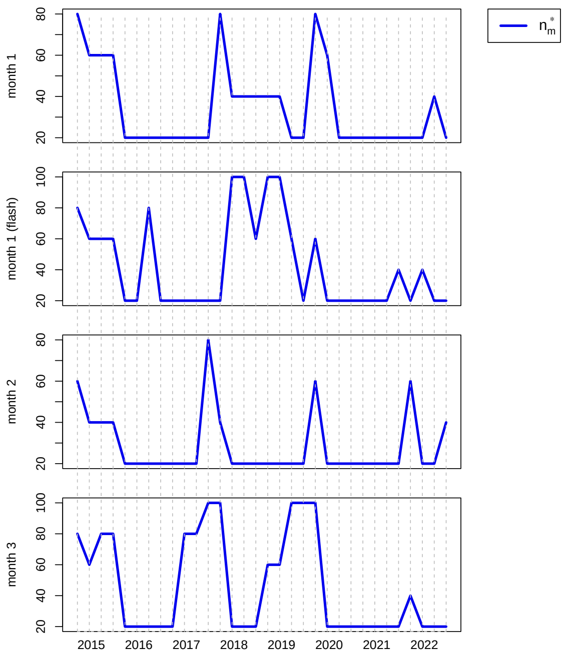
(a) FAR-MIDAS Model 1



(b) Factor-MIDAS Model 1



(c) FAR-MIDAS Model 2



(d) Factor-MIDAS Model 2

On the other hand, a relatively higher number of weekly financial data is selected during the Covid-19 pandemic crisis, as they contain forward-looking information that is useful to anticipate the downturn of economic activity. More specifically, it appears that the optimal number of weekly indicators is equal or greater than the number of monthly variables for both econometric setups.

Factor composition

Figure A.4 and Figure A.5 present, respectively, heat maps of estimated loading values for the real and financial factors in the FAR-MIDAS and Factor-MIDAS Model 1, which lead to the results of Table 2 in Section 4.2. The rows of each panel correspond to the different categories of real and financial variables, while the columns correspond to the quarters in the out-of-sample. The greater the factor loading associated with a specific variable category, the more pronounced the color intensity becomes.

Figure A.6 and Figure A.7 show conversely the heat maps of the real factor loadings of the FAR-MIDAS and Factor-MIDAS Model 2, respectively.

Figure A.4: Factor loading composition of FAR-MIDAS Model 1.

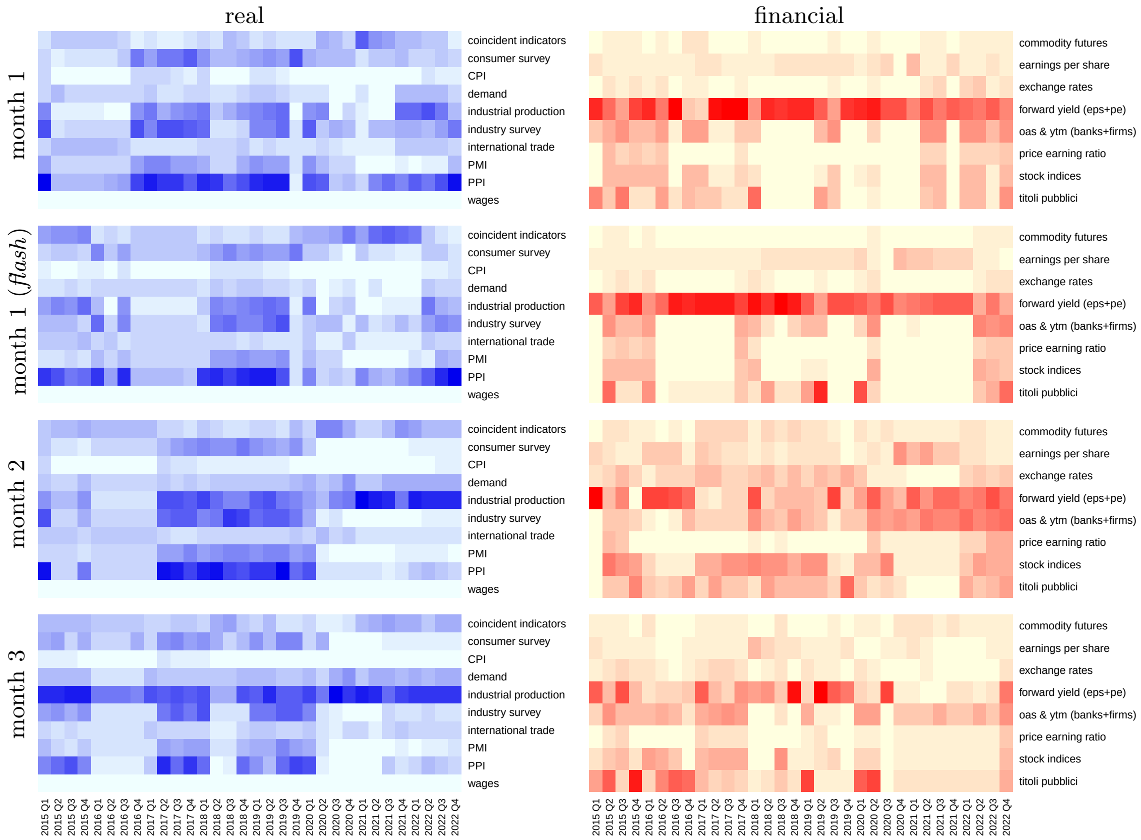


Figure A.5: Factor loading composition of Factor-MIDAS Model 1.

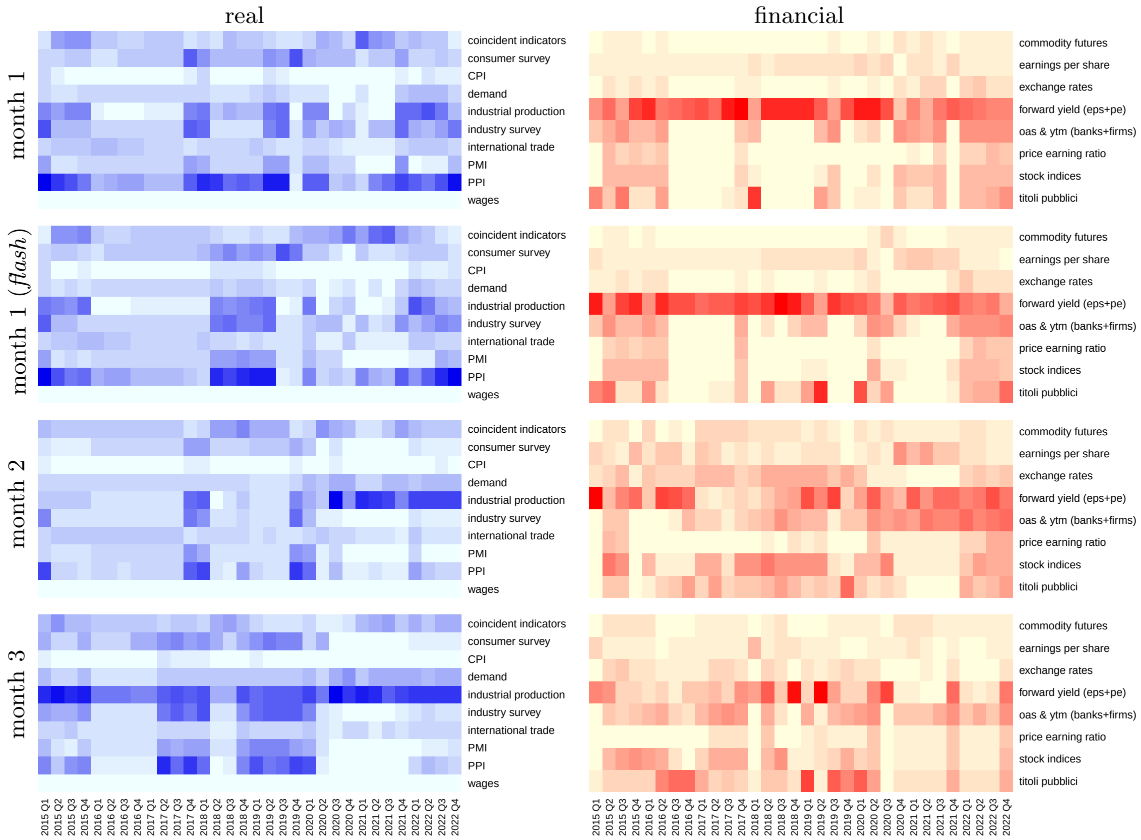


Figure A.6: Factor loading composition of FAR-MIDAS Model 2.

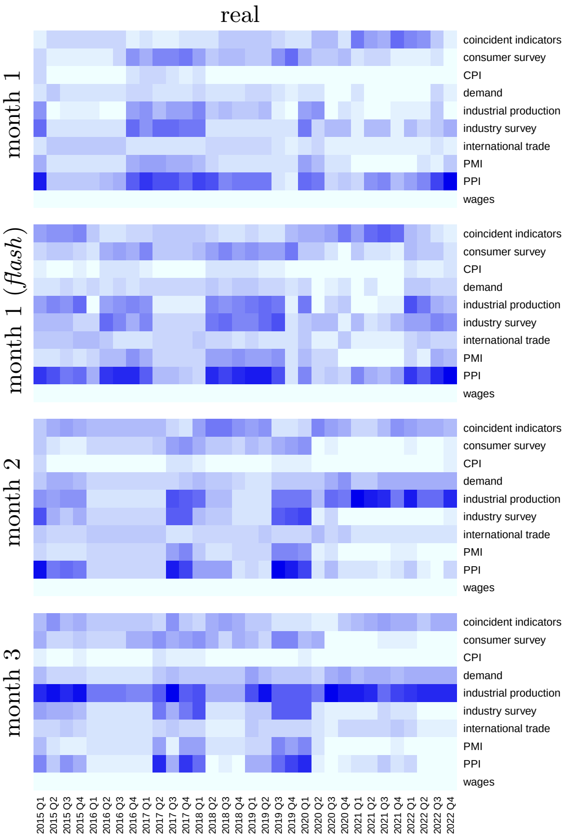
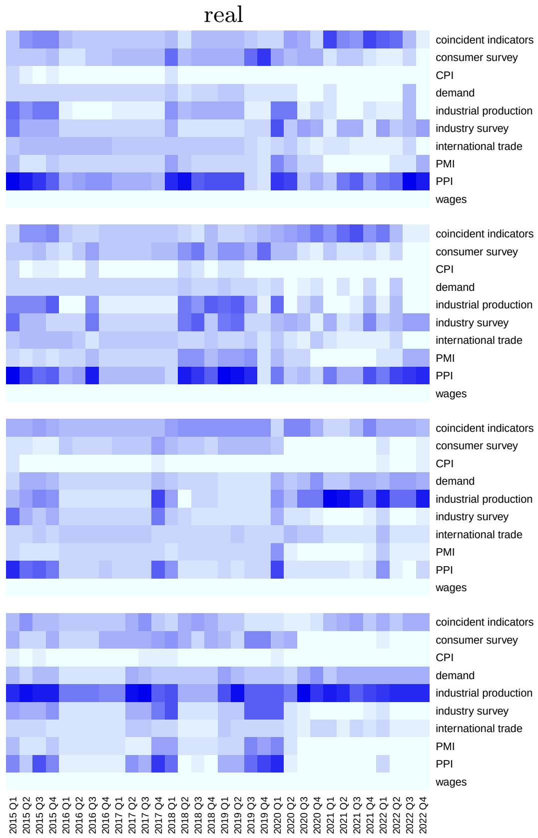


Figure A.7: Factor loading composition of Factor-MIDAS Model 2.



Nowcasting at weekly frequency

The purpose of the subsequent two charts (Figure A.8 and Figure A.9) is to illustrate the predictive capability of Model 1, for both FAR-MIDAS and Factor-MIDAS specifications, in nowcasting real GDP growth using all the available weekly financial data releases. The red line represents Model 1 and demonstrates its ability to generate 12 weekly nowcasts per quarter. In comparison, Model 2 (blue line) provides only 3 monthly nowcasts per quarter.

Figure A.8: FAR-MIDAS weekly nowcasts

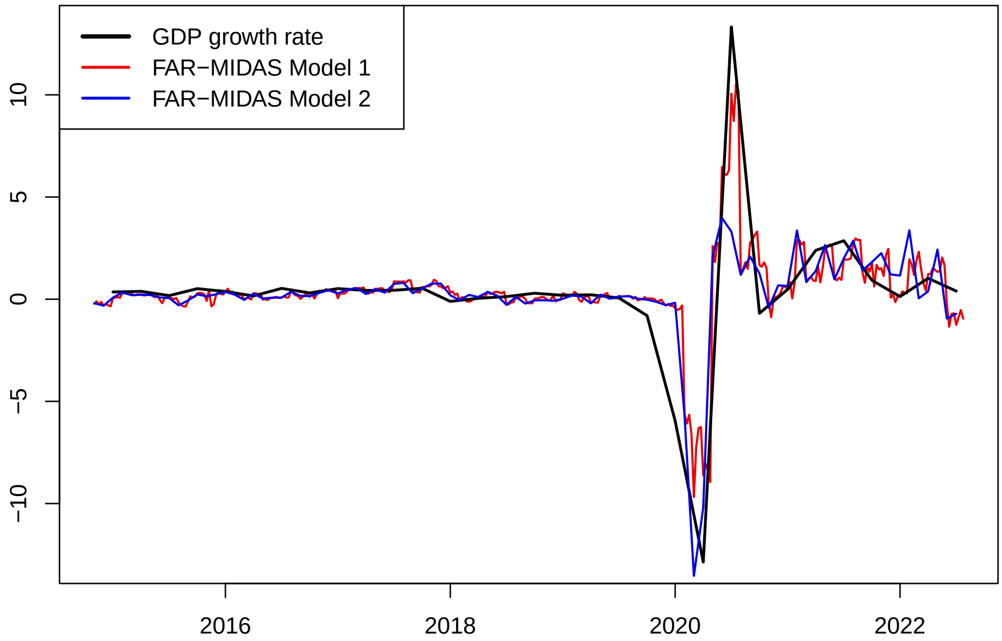
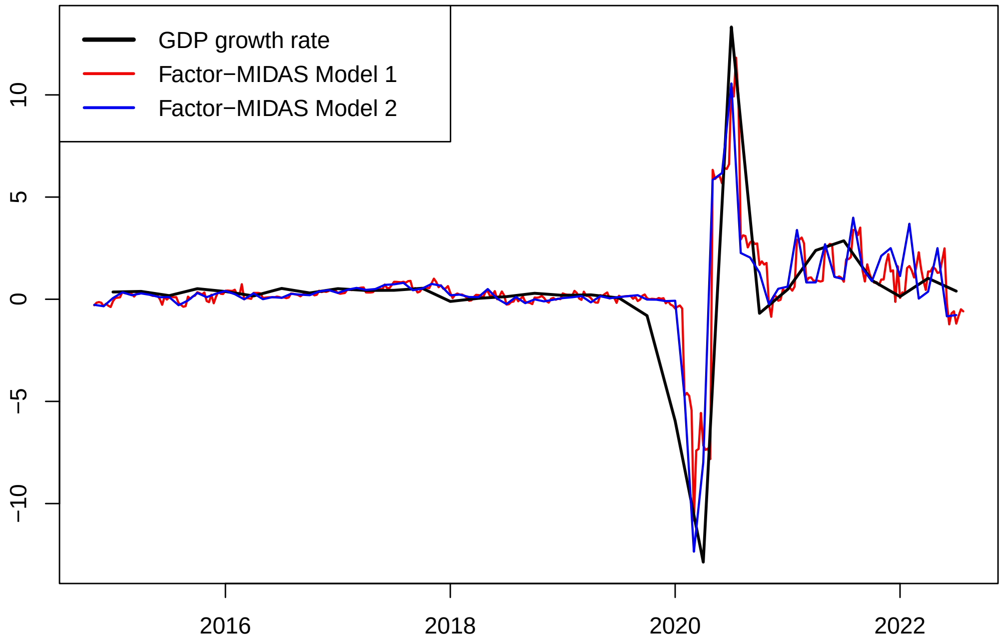


Figure A.9: Factor-MIDAS weekly nowcasts



Nowcasting with single financial categories

The results of the Factor-MIDAS model with single financial categories, similar to the approach used for the FAR-MIDAS model discussed in Section 4.2, are reported in Table A.1 and Table A.2 here below.

Table A.1: RMSFE of Factor-MIDAS with single financial categories relative to AR(1). Out-of-sample: $2015Q1 - 2022Q4$.

	Model 1								
	forward yield (EPS and PE)	price earnings	earnings per share	oas & ytm (banks + firms)	exchange rates	stock indices	gov. bonds	commod. futures	Model 2
month 1	0.758	0.748	**0.688**	0.723	0.771	0.755	0.756	0.750	0.760
month 1 (flash)	0.493	0.482	0.550	**0.480**	0.503	0.502	0.488	0.493	0.507
month 2	**0.354**	0.380	0.381	0.391	0.371	0.386	0.366	0.374	0.369
month 3	0.306	**0.303**	0.500	0.335	0.343	0.340	0.321	0.329	0.314

Notes: the results refer to the LARS pre-selection. The columns represent the financial category used for the estimation of the financial factor in the Factor-MIDAS model, thus excluding all other financial categories from the dataset.

Table A.2: Diebold-Mariano test of Factor-MIDAS with single financial categories. Out-of-sample: $2015Q1 - 2022Q4$.

	Model 1 vs. Model 2							
	forward yield (EPS and PE)	price earnings	earnings per share	oas & ytm (banks + firms)	exchange rates	stock indices	gov. bonds	commod. futures
month 1	0.37	**0.06**	**0.04**	**0.03**	0.90	0.29	0.31	0.15
month 1 (*flash*)	0.19	**0.07**	0.71	**0.04**	0.43	0.31	**0.07**	**0.01**
month 2	0.14	0.72	0.80	0.77	0.78	0.76	0.11	0.60
month 3	0.32	0.28	0.85	0.81	0.95	0.81	0.77	0.82

Notes: the results are related to the Factor-MIDAS models with LARS pre-selection. Bold values indicate rejection of null hypothesis of equal predictive ability at 10% significance level.

Nowcasting with fixed n_ℓ^*

The nowcast results of the Factor-MIDAS model with a fixed number of variables for factor estimation (n_ℓ^*), as conducted in Section 4.3.2 focusing on the FAR-MIDAS model, are reported in Table A.3 and Table A.4.

Table A.3: RMSFE of Factor-MIDAS with fixed n_ℓ^* relative to AR(1). Out-of-sample: $2015Q1 - 2022Q4$.

	n_m^*, n_w^*		n_m^*=n_w^*=20		n_m^*=n_w^*=40		n_m^*=n_w^*=60		n_m^*=n_w^*=80		n_m^*=n_w^*=100	
	Mod. 1	Mod. 2	Mod. 1	Mod. 2	Mod. 1	Mod. 2	Mod. 1	Mod. 2	Mod. 1	Mod. 2	Mod. 1	Mod. 2
month 1	0.720	0.760	0.755	0.767	0.731	**0.717**	0.806	0.809	0.780	0.814	0.776	0.826
month 1 (*flash*)	0.473	0.507	0.486	0.507	**0.459**	0.474	0.470	0.530	0.472	0.531	0.493	0.564
month 2	0.385	0.369	0.359	**0.358**	0.408	0.434	0.386	0.424	0.388	0.425	0.455	0.478
month 3	0.327	0.314	0.332	0.298	**0.285**	0.297	0.359	0.385	0.426	0.456	0.469	0.533
average	0.476	0.488	0.483	0.483	**0.471**	0.481	0.505	0.537	0.517	0.557	0.548	0.600

Notes: the results refer to the LARS pre-selection. Bold values indicate the best RMSFE ratio for a given month.

Table A.4: Diebold-Mariano test of Factor-MIDAS with fixed n_ℓ^*. Out-of-sample: $2015Q1 - 2022Q4$.

	Model 1 vs. Model 2					
	n_m^*, n_w^*	n_m^*=n_w^*=20	n_m^*=n_w^*=40	n_m^*=n_w^*=60	n_m^*=n_w^*=80	n_m^*=n_w^*=100
month 1	**0.01**	**0.07**	0.71	0.48	0.21	0.15
month 1 (*flash*)	**0.03**	0.18	0.12	**0.01**	**0.02**	**0.01**
month 2	0.81	0.55	0.13	0.12	0.12	0.17
month 3	0.74	0.96	0.22	**0.07**	**0.06**	**0.09**

Notes: the results are related to the Factor-MIDAS models with LARS pre-selection. Bold values indicate rejection of null hypothesis of equal predictive ability at 10% significance level.

Nowcasting with a maximum of two real factors

The following Table A.5 and Table A.6 report the nowcasting results with a maximum of two real factors for Model 2, as conducted in Section 4.3.4 with a maximum of three.

Table A.5: RMSFE relative to AR(1). Out-of-sample: $2015Q1 - 2022Q4$.

	FAR-MIDAS		Factor-MIDAS	
	Model 1	Model 2	Model 1	Model 2
month 1	**0.696**	0.740	0.720	0.751
month 1 (*flash*)	0.525	0.544	**0.473**	0.504
month 2	0.372	**0.302**	0.385	0.330
month 3	**0.300**	0.437	0.327	0.317

Notes: results refer to the LARS pre-selection. Bold values indicate the best RMSFE ratio for a given month.

Table A.6: Diebold-Mariano test. Out-of-sample: $2015Q1 - 2022Q4$.

	Model 1 vs. Model 2	
	FAR-MIDAS	Factor-MIDAS
month 1	**0.03**	**0.05**
month 1 (*flash*)	0.11	**0.03**
month 2	0.95	0.88
month 3	0.19	0.70

Notes: the table reports the p-values of the test with bold values indicating a significance level of 10%.

Nowcasting with one-time pre-selection

The following Table A.7 shows the RMSFE ratios obtained after applying the pre-selection only once, as in Section 4.3.5, but this time using all the available data, i.e., $2001Q1 - 2022Q4$. Table A.8 refers to the corresponding Diebold-Mariano test.

Table A.7: RMSFE relative to AR(1). Pre-selection in $2022Q4$.

	FAR-MIDAS		Factor-MIDAS	
	Model 1	Model 2	Model 1	Model 2
month 1	0.728	0.887	**0.675**	0.777
month 1 (*flash*)	0.708	0.872	**0.529**	0.658
month 2	0.376	0.365	**0.308**	0.339
month 3	0.358	0.347	0.385	**0.326**

Notes: results refer to the LARS pre-selection. Bold values indicate the best RMSFE ratio for a given month.

Table A.8: Diebold-Mariano test. Pre-selection in 2022Q4.

	Model 1 vs. Model 2	
	FAR-MIDAS	Factor-MIDAS
month 1	**0.01**	**0.02**
month 1 (*flash*)	**0.07**	**0.02**
month 2	0.72	**0.10**
month 3	0.56	0.80

Notes: the table reports the p-values of the test with bold values indicating a significance level of 10%.

Alternative MCS with different statistics and confidence intervals

The following Table A.9, Table A.10 and Table A.11 pertain to the MCS test results corresponding to Table 2 in Section 4.

Table A.9: Model Confidence Set (T_{max}) - I.C. 80%. Out-of-sample: 2015Q1 − 2022Q4.

	FAR-MIDAS		Factor-MIDAS		AR(1)
	Model 1	Model 2	Model 1	Model 2	
month 1	1	1	1	1	0.205
month 1 (*flash*)	1	1	1	1	0.343
month 2	1	0.593	1	1	×
month 3	1	0.505	1	1	×

Notes: results refer to the LARS pre-selection and report p-values of T_{max} statistics of Hansen et al. (2011) for a set of superior models (SSM), with a confidence level at 80% (α=0.2). The symbol × indicates the exclusion of the model from the SSM.

Table A.10: Model Confidence Set (T_{max}) - I.C. 95%. Out-of-sample: $2015Q1 - 2022Q4$.

	FAR-MIDAS		Factor-MIDAS		AR(1)
	Model 1	Model 2	Model 1	Model 2	
month 1	1	1	1	1	0.205
month 1 (*flash*)	1	1	1	1	0.355
month 2	1	1	1	1	0.170
month 3	1	1	1	1	0.169

Notes: results refer to the LARS pre-selection and report p-values of T_{max} statistics of Hansen et al. (2011) for a set of superior models (SSM), with a confidence level at 95% (α=0.05). The symbol $\times$ indicates the exclusion of the model from the SSM.

Table A.11: Model Confidence Set (T_R) - I.C. 95%. Out-of-sample: $2015Q1 - 2022Q4$.

	FAR-MIDAS		Factor-MIDAS		AR(1)
	Model 1	Model 2	Model 1	Model 2	
month 1	1	0.078	0.646	0.060	0.430
month 1 (*flash*)	0.716	0.052	1	0.095	0.535
month 2	0.962	0.654	0.656	1	0.497
month 3	1	0.743	0.664	0.782	0.594

Notes: results refer to the LARS pre-selection and report p-values of T_R statistics of Hansen et al. (2011) for a set of superior models (SSM), with a confidence level at 95% (α=0.05). The symbol $\times$ indicates the exclusion of the model from the SSM.

B Sensitivity to the out-of-sample

This Appendix presents the nowcasting results for different out-of-sample periods. The purpose is, first, to provide insights into the performance of the Factor MIDAS models excluding the Covid-19 pandemic and the Global Financial Crisis (GFC), i.e., in the out-of-sample $2010Q1 - 2019Q4$. The influence of the GFC is evaluated separately in the second part of this Appendix, while the third part of the Appendix presents the nowcasting performance post Covid-19, i.e., in the out-of-sample $2021Q1 - 2022Q4$. The last part of the Appendix also examines the influence of possible distorted estimations derived from the Covid-19 shock.

Nowcasting without the Global Financial Crisis and Covid-19

Table B.12 presents the RMSFE ratios for the out-of-sample period $2010Q1 - 2019Q4$, in which the GFC and the Covid-19 shock are not included. The main characteristic of this out-of-sample is the presence of the Sovereign Debt Crisis in $2011 - 2012$, which had a relatively minor impact on GDP growth compared to the Covid-19 and the GFC.

Table B.12: RMSFE relative to AR(1). Out-of-sample: $2010Q1 - 2019Q4$.

	FAR-MIDAS		Factor-MIDAS		BRIDGE	ARIMA	Random
	Model 1	Model 2	Model 1	Model 2	(indpro)	(auto)	Walk
month 1	0.832	0.828	**0.823**	0.844	0.948	0.977	1.061
month 1 (*flash*)	1.094	1.005	1.077	1.021	1.114	**0.999**	1.018
month 2	1.013	0.968	**0.933**	0.981	1.122	0.999	1.018
month 3	0.913	**0.799**	0.883	0.820	1.049	0.999	1.018

Notes: results refer to the LARS pre-selection. Model comparison is based on the ratio between the RMSFE of each competing model and the RMSFE of the AR(1) process. Bold values indicate the best RMSFE ratio for a given month. The term "month 1 (*flash*)" indicates the nowcasting produced in the first month also using the information of the *flash* GDP estimate. "BRIDGE (indpro)" refers to a BRIDGE model built with the Industrial Production Index (IPI) and an autoregressive term of order one.

The results show that, compared to the out-of-sample $2015Q1 - 2022Q4$, the Factor MIDAS models lose much of the accuracy gain with respect to the AR(1). Notably, Model 2 demonstrates superior performance in the third month, exhibiting the best overall RMSFE ratio, supported by a p-value of nearly 1 in the Diebold-Mariano test, which confirms that Model 1 does not possess the capability to outperform Model 2 (Table B.13). Moreover, there is no longer a statistically significant improvement in the first month for Model 1. This implies that the inclusion of financial variables does not offer considerable benefits during generally stable times, indicating that in these periods macroeconomic variables alone are adequate for accurate nowcasting.

Table B.13: Diebold-Mariano test. Out-of-sample: $2010Q1 - 2019Q4$.

	Model 1 vs. Model 2		Model 1 vs. AR(1)		Model 2 vs. AR(1)	
	FAR-MIDAS	Factor-MIDAS	FAR-MIDAS	Factor-MIDAS	FAR-MIDAS	Factor-MIDAS
month 1	0.55	0.24	**0.08**	**0.05**	**0.07**	**0.09**
month 1 (*flash*)	0.89	0.93	0.84	0.80	0.52	0.59
month 2	0.75	0.23	0.56	0.24	0.38	0.42
month 3	1.00	0.94	0.18	**0.10**	**0.02**	**0.03**

Notes: the table reports the p-values of the test with bold values indicating a significance level of 10%.

Nowcasting during the Global Financial Crisis

Analyzing the nowcasting performance during the GFC in Table B.14, Model 1 exhibits a less pronounced superiority in the first month compared to the period of Covid-19 and subsequent years, as previously presented in the paper. Nevertheless, using the *flash* GDP estimate in the first month leads to a statistically significant improvement at the 10% level, as indicated in Table B.15.

Table B.14: RMSFE relative to AR(1). Out-of-sample: $2008Q1 - 2009Q4$.

	FAR-MIDAS		Factor-MIDAS		BRIDGE	ARIMA	Random
	Model 1	Model 2	Model 1	Model 2	(indpro)	(auto)	Walk
month 1	0.740	**0.680**	0.736	0.730	1.124	0.964	1.049
month 1 (*flash*)	0.666	0.704	**0.645**	0.734	1.146	1.096	0.904
month 2	0.700	0.655	0.753	**0.629**	1.060	1.099	0.908
month 3	0.479	0.484	0.479	**0.446**	0.962	1.099	0.908

Notes: results refer to the LARS pre-selection. Model comparison is based on the ratio between the RMSFE of each competing model and the RMSFE of the AR(1) process. Bold values indicate the best RMSFE ratio for a given month. The term "month 1 (*flash*)" indicates the nowcasting produced in the first month also using the information of the *flash* GDP estimate. "BRIDGE (indpro)" refers to a BRIDGE model built with the Industrial Production Index (IPI) and an autoregressive term of order one.

Table B.15: Diebold-Mariano test. Out-of-sample: $2008Q1 - 2009Q4$.

	Model 1 vs. Model 2	
	FAR-MIDAS	Factor-MIDAS
month 1	0.91	0.58
month 1 (*flash*)	**0.10**	**0.08**
month 2	0.69	0.86
month 3	0.46	0.76

Notes: results refer to the LARS pre-selection.
Bold values indicate rejection of null hypothesis of equal predictive at 10% significance level.

Nowcasting excluding 2020

The results in Table B.16 and Table B.17 illustrate the nowcasting performance of both Factor MIDAS models (FAR-MIDAS and Factor-MIDAS) when applying LARS pre-selection over the $2021Q1 - 2022Q4$ out-of-sample period, which excludes the most acute phase of the Covid-19 pandemic in 2020.

Table B.16: RMSFE relative to AR(1). Out-of-sample: $2021Q1 - 2022Q4$.

	FAR-MIDAS		Factor-MIDAS		BRIDGE	ARIMA	Random
	Model 1	Model 2	Model 1	Model 2	(indpro)	(auto)	Walk
month 1	**0.226**	0.539	0.267	0.514	1.048	1.049	1.320
month 1 (*flash*)	**0.618**	1.161	0.659	1.261	1.196	0.794	0.759
month 2	**0.644**	0.708	0.696	0.829	0.982	0.795	0.724
month 3	0.563	**0.501**	0.651	0.646	0.950	0.795	0.724

Notes: results refer to the LARS pre-selection. Model comparison is based on the ratio between the RMSFE of each competing model and the RMSFE of the AR(1) process. Bold values indicate the best RMSFE ratio for a given month. The term "month 1 (*flash*)" indicates the nowcasting produced in the first month also using the information of the *flash* GDP estimate. "BRIDGE (indpro)" refers to a BRIDGE model built with the Industrial Production Index (IPI) and an autoregressive term of order one.

Table B.17: Diebold-Mariano test. Out-of-sample: $2021Q1 - 2022Q4$.

	Model 1 vs. Model 2	
	FAR-MIDAS	Factor-MIDAS
month 1	**0.03**	**0.04**
month 1 (*flash*)	**0.02**	**0.01**
month 2	0.20	0.18
month 3	0.72	0.53

Notes: results refer to the LARS pre-selection.
Bold values indicate rejection of null hypothesis of equal predictive at 10% significance level.

The results of Table 3 are overall confirmed as in month 1 (with or without *flash* GDP) Model 1 is significantly more accurate compared to Model 2.

Table B.18 and Table B.19 present the nowcast results using the t-statistics pre-selection technique rather than LARS. A similar message is obtained: overall, even excluding the effect of the pandemic and the extremely large errors in model-based forecasts in 2020, the model with factors extracted from real and financial variables is significantly more accurate in the first month of the quarter compared to the model that solely relies on the information from the real dataset.

These findings hold significant implications for GDP forecasting and highlight that integrating financial variables into a mixed-frequency regression framework improves nowcasting accuracy, especially in times of significant economic disruptions. Notably, these conclusions possess broader validity and are not solely attributable to the Covid-19 pandemic.

Table B.18: RMSFE relative to AR(1). Out-of-sample: 2021Q1 − 2022Q4.

	FAR-MIDAS		Factor-MIDAS		BRIDGE	ARIMA	Random
	Model 1	Model 2	Model 1	Model 2	(indpro)	(auto)	Walk
month 1	0.496	0.590	**0.492**	0.515	1.048	1.049	1.320
month 1 (*flash*)	0.833	0.857	0.872	1.114	1.196	0.794	**0.759**
month 2	**0.539**	0.666	0.552	0.665	0.982	0.795	0.724
month 3	0.748	**0.729**	0.818	0.777	0.950	0.795	0.724

Notes: results refer to the t-statistics pre-selection. Model comparison is based on the ratio between the RMSFE of each competing model and the RMSFE of the AR(1) process. Bold values indicate the best RMSFE ratio for a given month. The term "month 1 (*flash*)" indicates the nowcasting produced in the first month also using the information of the *flash* GDP estimate. "BRIDGE (indpro)" refers to a BRIDGE model built with the Industrial Production Index (IPI) and an autoregressive term of order one.

Table B.19: Diebold-Mariano test. Out-of-sample: 2021Q1 − 2022Q4.

	Model 1 vs. Model 2	
	FAR-MIDAS	Factor-MIDAS
month 1	**0.01**	0.34
month 1 (*flash*)	0.32	**0.07**
month 2	0.16	**0.09**
month 3	0.60	0.67

Notes: results are related to the t-statistics pre-selection. Bold values indicate rejection of null hypothesis of equal predictive at 10% significance level.

The Covid-19 shock and challenges for nowcasting

In 2020, as the pandemic unfolded worldwide, the variation exhibited by some macroeconomic time series was so extreme that it may have distorted the estimated coefficients since March 2020, posing serious challenges to forecasting models (Bobeica and Hartwig, 2023).

To mitigate the potential distortion in estimations caused by the Covid-19 shock, in this exercise the model parameters were estimated with information up to 2019Q4 and kept fixed for nowcasting the GDP growth rate in 2020-2021. These fixed estimated coefficients correspond to factors derived from a specific number of variables ordered by the LARS procedure. Consequently, both the optimal number of variables and their ordering were held constant as of 2019Q4.

The results are displayed in Table B.20 and Table B.21, where only the FAR-MIDAS and Factor-MIDAS models have fixed coefficients until the end of 2021. The enhancements in nowcasting the first month of the quarter achieved by the inclusion of financial information are no longer evident, leading to quasi-identical results between Model 1 and Model 2 in both the FAR-MIDAS and Factor-MIDAS specifications. A possible explanation for this outcome is that the financial factor had less weight in the model estimated in 2019Q4, and the order of variables used to extract the factors remained unchanged as of 2019Q4. As a result, when new data for 2020-2021 is added to the model, the financial factor plays a much smaller role in the nowcasting process, resulting in nearly identical nowcasts.

Finally, Table B.22 shows the results for the 2020-2021 out-of-sample without keeping the parameters fixed, demonstrating that they are overall better for both model 1 and 2 in both specifications in each month.

Table B.20: RMSFE relative to AR(1). Out-of-sample: 2020Q1 − 2021Q4

	FAR-MIDAS		Factor-MIDAS		BRIDGE	ARIMA	Random
	Model 1	Model 2	Model 1	Model 2	(indpro)	(auto)	Walk
month 1	0.729	0.727	0.777	**0.708**	0.887	1.237	1.245
month 1 (*flash*)	0.641	0.630	0.595	0.599	**0.592**	1.166	1.108
month 2	0.602	0.619	0.570	**0.547**	0.655	1.195	1.092
month 3	**0.523**	0.597	0.540	0.569	0.524	1.195	1.092

Notes: results refer to the LARS pre-selection. Model comparison is based on the ratio between the RMSFE of each competing model and the RMSFE of the AR(1) process. Bold values indicate the best RMSFE ratio for a given month. The term "month 1 (*flash*)" indicates the nowcasting produced in the first month also using the information of the *flash* GDP estimate. "BRIDGE (indpro)" refers to a BRIDGE model built with the Industrial Production Index (IPI) and an autoregressive term of order one.

Table B.21: Predictive accuracy tests. Out-of-sample: $2020Q1 - 2021Q4$

	FAR-MIDAS						Factor-MIDAS					
	Model 1 vs. Model 2			Model 1 vs. AR(1)			Model 1 vs. Model 2			Model 1 vs. AR(1)		
	DM	GW	HLN	DM	GW	HLN	DM	GW	HLN	DM	GW	HLN
month 1	0.92	0.95	0.92	0.46	**0.01**	**0.02**	0.78	0.80	0.78	0.58	**0.06**	**0.09**
month 1 (*flash*)	0.85	0.87	0.85	0.24	0.16	0.18	0.13	0.11	0.13	0.21	0.16	0.18
month 2	0.35	0.35	0.35	0.20	0.14	0.16	0.87	0.89	0.87	0.18	0.15	0.17
month 3	0.20	0.19	0.20	0.16	0.14	0.16	**0.06**	**0.04**	**0.06**	0.16	0.13	0.15

Notes: results refer to the LARS pre-selection. Bold values indicate rejection of null hypothesis of equal predictive at 10% significance level. The acronyms DM, GW and HLN refer respectively to Diebold-Mariano, Giacomini-White and Harvey-Leybourne-Newbold tests.

Table B.22: RMSFE relative to AR(1). Out-of-sample: $2020Q1 - 2021Q4$

	FAR-MIDAS		Factor-MIDAS		BRIDGE	ARIMA	Random
	Model 1	Model 2	Model 1	Model 2	(indpro)	(auto)	Walk
month 1	**0.692**	0.706	0.717	0.732	0.887	1.237	1.245
month 1 (*flash*)	0.605	0.620	**0.545**	0.555	0.592	1.166	1.108
month 2	0.432	0.483	0.445	**0.416**	0.655	1.195	1.092
month 3	**0.345**	0.507	0.379	0.364	0.524	1.195	1.092

Notes: results refer to the LARS pre-selection. Model comparison is based on the ratio between the RMSFE of each competing model and the RMSFE of the AR(1) process. Bold values indicate the best RMSFE ratio for a given month. The term "month 1 (*flash*)" indicates the nowcasting produced in the first month also using the information of the *flash* GDP estimate. "BRIDGE (indpro)" refers to a BRIDGE model built with the Industrial Production Index (IPI) and an autoregressive term of order one.

References

Adams, P. A., T. Adrian, N. Boyarchenko and D. Giannone, "Forecasting macroeconomic risks," International Journal of Forecasting 37 (2021), 1173–1191.

Adrian, T., N. Boyarchenko and D. Giannone, "Vulnerable growth," American Economic Review 109 (April 2019), 1263–89.

Altissimo, F., R. Cristadoro, M. Forni, M. Lippi and G. Veronese, "New Eurocoin: Tracking economic growth in real time," The Review of Economics and Statistics 92 (November 2010), 1024–1034.

Andreou, E., E. Ghysels and A. Kourtellos, "Regression models with mixed sampling frequencies," Journal of Econometrics 158 (2010), 246–261, specification Analysis in Honor of Phoebus J. Dhrymes.

Ang, A., M. Piazzesi and M. Wei, "What does the yield curve tell us about GDP growth?," Journal of Econometrics 131 (2006), 359–403.

Aprigliano, V. and L. Bencivelli, "Ita-coin: a new coincident indicator for the Italian economy," Temi di discussione (Economic working papers) 935, Bank of Italy, Economic Research and International Relations Area, October 2013.

Aprigliano, V., A. Borin, F. P. Conteduca, S. Emiliozzi, M. Flaccadoro, S. Marchetti and S. Villa, "Forecasting Italian GDP growth with epidemiological data," Questioni di Economia e Finanza (Occasional Papers) 664, Bank of Italy, Economic Research and International Relations Area, December 2021.

Audrino, F., A. Kostrov and J.-P. Ortega, "Predicting U.S. bank failures with MIDAS logit models," Journal of Financial and Quantitative Analysis 54 (2019), 2575–2603.

Babii, A., E. Ghysels and J. Striaukas, "Machine learning time series regressions with an application to nowcasting," Journal of Business & Economic Statistics 40 (2022), 1094–1106.

Baffigi, A., R. Golinelli and G. Parigi, "Bridge models to forecast the euro area GDP," International Journal of Forecasting 20 (2004), 447–460.

Bai, J., E. Ghysels and J. H. Wright, "State Space Models and MIDAS Regressions," Econometric Reviews 32 (2013), 779–813.

Bai, J. and S. Ng, "Forecasting economic time series using targeted predictors," Journal of Econometrics 146 (2008), 304–317, honoring the research contributions of Charles R. Nelson.

Bernanke, B. S. and J. Boivin, "Monetary policy in a data-rich environment," Journal of Monetary Economics 50 (April 2003), 525–546.

Bobeica, E. and B. Hartwig, "The COVID-19 shock and challenges for inflation modelling," International Journal of Forecasting 39 (2023), 519–539.

Boivin, J. and S. Ng, "Are more data always better for factor analysis?," Journal of Econometrics 132 (2006), 169–194.

Bok, B., D. Caratelli, D. Giannone, A. M. Sbordone and A. Tambalotti, "Macroeconomic nowcasting and forecasting with big data," Annual Review of Economics 10 (2018), 615–643.

Bulligan, G., M. Marcellino and F. Venditti, "Forecasting economic activity with targeted predictors," International Journal of Forecasting 31 (2015), 188–206.

Camacho, M. and G. Perez-Quiros, "Introducing the euro-sting: Short-term indicator of euro area growth," Journal of Applied Econometrics 25 (2010), 663–694.

Castle, J., D. Hendry and O. Kitov, Forecasting and nowcasting macroeconomic variables: A methodological overview (Eurostat, 2017).

Ceci, D. and A. Silvestrini, "Nowcasting the state of the Italian economy: The role of financial markets," Journal of Forecasting 42 (2023), 1569–1593.

Chinn, M. D., B. Meunier and S. Stumpner, "Nowcasting world trade with machine learning: a three-step approach," (June 2023).

Chow, G. C. and A.-l. Lin, "Best linear unbiased interpolation, distribution, and extrapolation of time series by related series," The Review of Economics and Statistics 53 (1971), 372–375.

Cimadomo, J., D. Giannone, M. Lenza, F. Monti and A. Sokol, "Nowcasting with large Bayesian vector autoregressions," Journal of Econometrics 231 (2022), 500–519, special Issue: The Econometrics of Macroeconomic and Financial Data.

Clements, M. P. and A. B. Galvão, "Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the United States," Journal of Business & Economic Statistics 26 (2008), 546–554.

Dagum, E. B. and P. A. Cholette, "Benchmarking, temporal distribution, and reconciliation methods for time series," (2006).

Delle Monache, D., S. Emiliozzi and A. Nobili, "Tracking economic growth during the Covid-19: a weekly indicator for Italy," Banca d'Italia, "Note Covid-19", 27 gennaio 2021. (2020).

Diebold, F. X. and R. S. Mariano, "Comparing predictive accuracy," Journal of Business & Economic Statistics 13 (1995), 253–263.

Efron, B., T. Hastie, I. Johnstone and R. Tibshirani, "Least Angle Regression," The Annals of Statistics 32 (2004), 407–451.

Estrella, A. and G. A. Hardouvelis, "The term structure as a predictor of real economic activity," The Journal of Finance 46 (1991), 555–576.

Estrella, A. and F. S. Mishkin, "Predicting U.S. recessions: financial variables as leading indicators," The Review of Economics and Statistics 80 (1998), 45–61.

Ferrara, L. and C. Marsilli, "Financial variables as leading indicators of GDP growth: Evidence from a MIDAS approach during the Great Recession," Applied Economics Letters 20 (February 2013), 233–237.

Foroni, C. and M. Marcellino, "A comparison of mixed frequency approaches for nowcasting Euro area macroeconomic aggregates," International Journal of Forecasting 30 (2014), 554–568.

Ghysels, E., P. Santa-Clara and R. Valkanov, "The MIDAS touch: Mixed data sampling regression models," Cirano working papers, CIRANO, 2004.

———, "Predicting volatility: getting the most out of return data sampled at different frequencies," Journal of Econometrics 131 (2006), 59–95.

Ghysels, E., A. Sinko and R. Valkanov, "MIDAS Regressions: Further results and new directions," Econometric Reviews 26 (2007), 53–90.

Ghysels, E. and J. H. Wright, "Forecasting professional forecasters," Journal of Business & Economic Statistics 27 (2009), 504–516.

Giacomini, R. and H. White, "Tests of conditional predictive ability," Econometrica 74 (2006), 1545–1578.

Gilchrist, S. and E. Zakrajšek, "Credit spreads and business cycle fluctuations," American Economic Review 102 (June 2012), 1692–1720.

Golinelli, R. and G. Parigi, "The use of monthly indicators to forecast quarterly GDP in the short run: an application to the G7 countries," Journal of Forecasting 26 (2007), 77–94.

Guérin, P. and M. Marcellino, "Markov-Switching MIDAS models," Journal of Business & Economic Statistics 31 (2013), 45–56.

Hansen, P. R., A. Lunde and J. M. Nason, "The model confidence set," Econometrica 79 (2011), 453–497.

Harvey, A. C., Forecasting, structural time series models, and the Kalman filter (Cambridge u.a.: Cambridge Univ. Press, 1990).

Harvey, D., S. Leybourne and P. Newbold, "Testing the equality of prediction mean squared errors," International Journal of Forecasting 13 (1997), 281–291.

Hendry, D. F. and G. E. Mizon, "Serial correlation as a convenient simplification, not a nuisance: A comment on a study of the demand for money by the Bank of England," The Economic Journal 88 (1978), 549–563.

Jardet, C. and B. Meunier, "Nowcasting world GDP growth with high-frequency data," Journal of Forecasting 41 (2022), 1181–1200.

Jurado, K., S. C. Ludvigson and S. Ng, "Measuring uncertainty," American Economic Review 105 (2015), 1177–1216.

Marcellino, M. and C. Schumacher, "Factor MIDAS for nowcasting and forecasting with ragged-edge data: A model comparison for German GDP," Oxford Bulletin of Economics and Statistics 72 (2010), 518–550.

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

Moench, E., S. Ng and S. Potter, "Dynamic hierarchical factor models," Review of Economics and Statistics 95 (2013), 1811–1817.

Moneta, F., "Does the yield spread predict recessions in the euro area?," International Finance 8 (2005), 263–301.

Schumacher, C., "Factor forecasting using international targeted predictors: The case of German GDP," Economics Letters 107 (2010), 95–98.

Silvestrini, A. and D. Veredas, "Temporal aggregation of univariate and multivariate time series models: A survey," Journal of Economic Surveys 22 (July 2008), 458–497.

Stock, J. H. and M. W. Watson, "Forecasting using principal components from a large number of predictors," Journal of the American Statistical Association 97 (2002a), 1167–1179.

———, "Macroeconomic forecasting using diffusion indexes," Journal of Business & Economic Statistics 20 (2002b), 147–162.

———, "Forecasting output and inflation: The role of asset prices," Journal of Economic Literature 41 (September 2003), 788–829.

Wallis, K. F., "Forecasting with an econometric model: The 'ragged edge' problem," Journal of Forecasting 5 (1986), 1–13.

Welch, I. and A. Goyal, "A comprehensive look at the empirical performance of equity premium prediction," The Review of Financial Studies 21 (03 2007), 1455–1508.

RECENTLY PUBLISHED "TEMI" (*)

N. 1419 – *Temperatures and search: evidence from the housing market*, by Michele Cascarano and Filippo Natoli (July 2023).

N. 1420 – *Flight to climatic safety: local natural disasters and global portfolio flows*, by Fabrizio Ferriani, Andrea Gazzani and Filippo Natoli (July 2023).

N. 1421 – *The effects of the pandemic on households' financial savings: a Bayesian structural VAR analysis*, by Luigi Infante, Francesca Lilla and Francesco Vercelli (October 2023).

N. 1422 – *Decomposing the monetary policy multiplier*, by Piergiorgio Alessandri, Fabrizio Venditti and Oscar Jordà (October 2023).

N. 1423 – *The short and medium term effects of full-day schooling on learning and maternal labor supply*, by Giulia Bovini, Nicolò Cattadori, Marta De Philippis and Paolo Sestito (October 2023).

N. 1424 – *Subsidizing business entry in competitive credit markets*, by Vincenzo Cuciniello, Claudio Michelacci and Luigi Paciello (October 2023).

N. 1425 – *Drivers of large recessions and monetary policy responses*, by Giovanni Melina and Stefania Villa (October 2023).

N. 1426 – *The performance of household-held mutual funds: evidence from the euro area*, by Valerio Della Corte and Raffaele Santioni (November 2023).

N. 1427 – *Trade in the time of COVID-19: an empirical analysis based on Italian data*, by Gianmarco Cariola (November 2023).

N. 1428 – *Natural gas and the macroeconomy: not all energy shocks are alike*, by Piergiorgio Alessandri and Andrea Giovanni Gazzani (November 2023).

N. 1429 – *Inflation is not equal for all: the heterogenous effects of energy shocks*, by Francesco Corsello and Marianna Riggi (November 2023).

N. 1430 – *Labor market dynamics and geographical reallocations*, by Gaetano Basso, Salvatore Lo Bello and Francesca Subioli (November 2023).

N. 1431 – *Monetary and fiscal policy responses to fossil fuel price shocks*, by Anna Bartocci, Alessandro Cantelmo, Pietro Cova, Alessandro Notarpietro and Massimiliano Pisani (December 2023).

N. 1432 – *Do female leaders choose women? Evidence from visible and hidden appointments*, by Andrea Cintolesi and Edoardo Frattola (December 2023).

N. 1433 – *Monetary policy tightening in response to uncertain stagflationary shocks: a model-based analysis*, by Anna Bartocci, Alessandro Cantelmo, Alessandro Notarpietro and Massimiliano Pisani (December 2023).

N. 1434 – *Inflation, capital structure and firm value*, by Andrea Fabiani and Fabio Massimo Piersanti (December 2023).

N. 1435 – *Announcement and implementation effects of central bank asset purchases*, by Marco Bernardini and Antonio M. Conti (December 2023).

N. 1436 – *Connecting the dots: the network nature of shocks propagation in credit markets*, by Stefano Pietrosanti and Edoardo Rainone (December 2023).

N. 1437 – *Inflation expectations and misallocation of resources: evidence from Italy*, by Tiziano Ropele, Yuriy Gorodnichenko and Olivier Coibion (December 2023).

N. 1438 – *Women in economics: the role of gendered references at entry in the profession*, by Audinga Baltrunaite, Alessandra Casarico and Lucia Rizzica (February 2024).

N. 1439 – *Procuring survival*, by Matilde Cappelletti, Leonardo M. Giuffrida and Gabriele Rovigatti (February 2024).

N. 1440 – *Estimating the returns to occupational licensing: evidence from regression discontinuities at the bar exam*, by Omar Bamieh, Andrea Cintolesi and Mario Pagliero (February 2024).

N. 1441 – *Household perceived sources of business cycle fluctuations: a tale of supply and demand*, by Clodomiro Ferreira and Stefano Pica (February 2024).

(*) Requests for copies should be sent to:
Banca d'Italia – Servizio Studi di struttura economica e finanziaria – Divisione Biblioteca e Archivio storico – Via Nazionale, 91 – 00184 Rome – (fax 0039 06 47922059). They are available on the Internet www.bancaditalia.it.

2022

ANDINI M., M. BOLDRINI, E. CIANI, G. DE BLASIO, A. D'IGNAZIO and A. PALADINI, *Machine learning in the service of policy targeting: the case of public credit guarantees,* Journal of Economic Behavior & Organization, v. 198, pp. 434-475, **WP 1206 (February 2019).**

ANGELICO C., J. MARCUCCI, M. MICCOLI and F. QUARTA, *Can we measure inflation expectations using twitter?,* Journal of Econometrics, v. 228, 2, pp. 259-277, **WP 1318 (February 2021).**

BARTOCCI A., A. NOTARPIETRO and M. PISANI, *Covid-19 shock and fiscal-monetary policy mix in a monetary union,* Economic challenges for Europe after the pandemic, Springer Proceedings in Business and Economics, Berlin-Heidelberg, Springer, **WP 1313 (December 2020).**

BOTTERO M., C. MINOIU, J. PEYDRÒ, A. POLO, A. PRESBITERO and E. SETTE, *Expansionary yet different: credit supply and real effects of negative interest rate policy,* Journal of Financial Economics, v. 146, 2, pp. 754-778, **WP 1269 (March 2020).**

BRONZINI R., A. D'IGNAZIO and D. REVELLI, *Financial structure and bank relationships of Italian multinational firms,* Journal of Multinational Financial Management, v. 66, Article 100762, **WP 1326 (March 2021).**

CANTELMO A., *Rare disasters, the natural interest rate and monetary policy,* Oxford Bulletin of Economics and Statistics, v. 84, 3, pp. 473-496, **WP 1309 (December 2020).**

CARRIERO A., F. CORSELLO and M. MARCELLINO, *The global component of inflation volatility,* Journal of Applied Econometrics, v. 37, 4, pp. 700-721, **WP 1170 (May 2018).**

CIAPANNA E. and G. ROVIGATTI, *The grocery trolley race in times of Covid-19. Evidence from Italy,* Italian Economic Journal / Rivista italiana degli economisti, v. 8, 2, pp. 471-498, **WP 1341 (June 2021).**

CONTI A. M., A. NOBILI and F. M. SIGNORETTI, *Bank capital requirement shocks: a narrative perspective,* European Economic Review, v.151, Article 104254, **WP 1199 (November 2018).**

FAIELLA I. and A. MISTRETTA, *The net zero challenge for firms' competitiveness,* Environmental and Resource Economics, v. 83, pp. 85-113, **WP 1259 (February 2020).**

FERRIANI F. and G. VERONESE, *Hedging and investment trade-offs in the U.S. oil industry,* Energy Economics, v. 106, Article 105736, **WP 1211 (March 2019).**

GUISO L., A. POZZI, A. TSOY, L. GAMBACORTA and P. E. MISTRULLI, *The cost of steering in financial markets: evidence from the mortgage market,* Journal of Financial Economics, v.143, 3, pp. 1209-1226, **WP 1252 (December 2019).**

LAMORGESE A. and D. PELLEGRINO, *Loss aversion in housing appraisal: evidence from Italian homeowners,* Journal of Housing Economics, v. 56, Article 101826, **WP 1248 (November 2019).**

LI F., T. MÄKINEN, A. MERCATANTI and A. SILVESTRINI, *Causal analysis of central bank holdings of corporate bonds under interference,* Economic Modelling, v.113, Article 105873, **WP 1300 (November 2020).**

LOBERTO M, A. LUCIANI and M. PANGALLO, *What do online listings tell us about the housing market?,* International Journal of Central Banking, v. 18, 4, pp. 325-377, **WP 1171 (April 2018).**

MIRENDA L., M. SAURO and L. RIZZICA, *The economic effects of mafia: firm level evidence,* American Economic Review, vol. 112, 8, pp. 2748-2773, **WP 1235 (October 2019).**

MOCETTI S., G. ROMA and E. RUBOLINO, *Knocking on parents' doors: regulation and intergenerational mobility,* Journal of Human Resources, v. 57, 2, pp. 525-554, **WP 1182 (July 2018).**

PERICOLI M. and M. TABOGA, *Nearly exact Bayesian estimation of non-linear no-arbitrage term-structure models,* Journal of Financial Econometrics, v. 20, 5, pp. 807-838, **WP 1189 (September 2018).**

ROSSI P. and D. SCALISE, *Financial development and growth in European regions,* Journal of Regional Science, v. 62, 2, pp. 389-411, **WP 1246 (November 2019).**

SCHIVARDI F., E. SETTE and G. TABELLINI, *Credit misallocation during the European financial crisis,* Economic Journal, v. 132, 641, pp. 391-423, **WP 1139 (September 2017).**

TABOGA M., *Cross-country differences in the size of venture capital financing rounds: a machine learning approach,* Empirical Economics, v. 62, 3, pp. 991-1012, **WP 1243 (November 2019).**

2023

APRIGLIANO V., S. EMILIOZZI, G. GUAITOLI, A. LUCIANI, J. MARCUCCI and L. MONTEFORTE, *The power of text-based indicators in forecasting Italian economic activity,* International Journal of Forecasting, v. 39, 2, pp. 791-808, **WP 1321 (March 2021).**

BARTOCCI A., A. NOTARPIETRO and M. PISANI, *Non-standard monetary policy measures in non-normal times,* International Finance, v. 26, 1, pp. 19-35, **WP 1251 (November 2019).**

CAPPELLETTI G. and P. E. MISTRULLI, *The role of credit lines and multiple lending in financial contagion and systemic events,* Journal of Financial Stability, v. 67, Article 101141, **WP 1123 (June 2017).**

CECI D. and A. SILVESTRINI, *Nowcasting the state of the Italian economy: the role of financial markets,* Journal of Forecasting, v. 42, 7, pp. 1569-1593, **WP 1362 (February 2022).**

CIAPANNA E, S. MOCETTI and A. NOTARPIETRO, *The macroeconomic effects of structural reforms: an empirical and model-based approach,* Economic Policy, v. 38, 114, pp. 243-285, **WP 1303 (November 2020).**

DAURICH D, S. DI ADDARIO and R. SAGGIO, *The macroeconomic effects of structural reforms: an empirical and model-based approach,* Review of Economic Studies, v. 90, 6, pp. 2880–2942, **WP 1390 (November 2022).**

DI ADDARIO S., P. KLINE, R. SAGGIO and M. SØLVSTEN, *The effects of partial employment protection reforms: evidence from Italy,* Journal of Econometrics,v. 233, 2, pp. 340-374, **WP 1374 (June 2022).**

FERRARI A. and V. NISPI LANDI, *Toward a green economy: the role of central bank's asset purchases,* International Journal of Central Banking, v. 19, 5, pp. 287-340, **WP 1358 (February 2022).**

FERRIANI F., *Issuing bonds during the Covid-19 pandemic: was there an ESG premium?,* International Review of Financial Analysis, v. 88, Article 102653, **WP 1392 (November 2022).**

GIORDANO C., *Revisiting the real exchange rate misalignment-economic growth nexus via the across-sector misallocation channel,* Review of International Economics, v. 31, 4, pp. 1329-1384, **WP 1385 (October 2022).**

GUGLIELMINETTI E., M. LOBERTO and A. MISTRETTA, *The impact of COVID-19 on the European short-term rental market,* Empirica, v. 50, 3, pp. 585-623, **WP 1379 (July 2022).**

LILLA F., *Volatility bursts: a discrete-time option model with multiple volatility components,* Journal of Financial Econometrics, v. 21, 3, pp. 678-713, **WP 1336 (June 2021).**

LOBERTO M., *Foreclosures and house prices,* Italian Economic Journal / Rivista italiana degli economisti, v. 9, 1, pp. 397-424, **WP 1325 (March 2021).**

LOMBARDI M. J., M. RIGGI and E. VIVIANO, *Worker's bargaining power and the Phillips curve: a micro-macro analysis, and wages,* Journal of the European Economic Association, v. 21, 5, pp. 1905–1943, **WP 1302 (November 2020).**

NERI S., *Long-term inflation expectations and monetary policy in the Euro Area before the pandemic*, European Economic Review, v. 154, Article 104426, **WP 1357 (December 2021).**

ORAME A., *Bank lending and the European debt crisis: evidence from a new survey*, International Journal of Central Banking, v. 19, 1, pp. 243-300, **WP 1279 (June 2020).**

RIZZICA L., G. ROMA and G. ROVIGATTI, *The effects of shop opening hours deregulation: evidence from Italy,* The Journal of Law and Economics, v. 66, 1, pp. 21-52, **WP 1281 (June 2020).**

TANZI G. M., *Scars of youth non-employment and labour market conditions,* Italian Economic Journal / Rivista italiana degli economisti, v. 9, 2, pp. 475-499, **WP 1312 (December 2020).**

2024

MORO A. and V. NISPI LANDI, *The external financial spillovers of CBDCs,* Journal of Economic Dynamics and Control, v. 159, Article 104801, **WP 1416 (July 2023).**

FORTHCOMING

BALTRUNAITE A., M. CANNELLA, S. MOCETTI and G. ROMA, *Board composition and performance of state-owned enterprises: quasi experimental evidence,* The Journal of Law, Economics, and Organization, **WP 1328 (April 2021).**

BUONO I, F. CORNELI and E. DI STEFANO, *Capital inflows to emerging countries and their sensitivity to the global financial cycle,* International Finance, **WP 1262 (February 2020).**

CORNELI F., *Sovereign debt maturity structure and its costs,* International Tax and Public Finance, **WP 1196 (November 2018).**

CUCINIELLO V. and N. DI IASIO, *Determinants of the credit cycle: a flow analysis of the extensive margin,* Journal of Money, Credit and Banking, **WP 1266 (March 2020).**

FERRARI A. and V. NISPI LANDI, *Whatever it takes to save the planet? Central banks and unconventional green policy,* Macroeconomic Dynamics, **WP 1320 (February 2021).**

FLACCADORO M., *Exchange rate pass-through in small, open, commodity-exporting economies: lessons from Canada,* Journal of International Economics, **WP 1365 (April 2022).**

GAUTIER E., C. CONFLITTI, R. FABER, B. FABO, L. FADEJEVA, V. JOUVANCEAU, J.-O. MENZ, T. MESSNER, P. PETROULAS, P. ROLDAN-BLANCO, F. RUMLER, S. SANTORO, E. WIELAND and H. ZIMMER, *New facts on consumer price rigidity in the euro area,* American Economic Journal: Macroeconomics, **WP 1375 (July 2022).**

MICHELANGELI V. and E. VIVIANO, *Can internet banking affect households' participation in financial markets and financial awarness?,* Journal of Money, Credit and Banking, **WP 1329 (April 2021).**

MISTRETTA A., *Synchronization vs transmission: the effect of the German slowdown on the Italian business cycle,* International Journal of Central Banking, **WP 1346 (October 2021).**

RAINONE E., *Reservation rates in interbank money markets,* Journal of Money, Credit and Banking, **WP 1160 (February 2021).**

RAINONE E., *Real-time identification and high frequency analysis of deposits outflows,* Journal of Financial Econometrics, **WP 1319 (December 2017).**