De Economist (2025) 173:665–704
https://doi.org/10.1007/s10645-025-09462-w

ORIGINAL PAPER

Designing a Nowcasting Model for GDP Growth: A Practical Approach

Mick van Rooijen[1] · **Dorinth W. van Dijk**[1] · **Jasper M. de Winter**[1]

Received: 15 May 2025 / Accepted: 24 October 2025 / Published online: 19 November 2025

Abstract
This paper presents a practical guide for developing a nowcasting model for GDP growth. We employ a dynamic factor model to generate backcasts, nowcasts and forecasts of Dutch GDP in pseudo real-time. We evaluate forecast errors across a range of modeling alternatives, including data choice, transformations, outlier correction and model specification. The optimal combination of these alternatives outperforms standard benchmarks. Additionally, we demonstrate how to derive forecast contributions and assess the impact of new data releases, offering policymakers valuable insights for interpreting GDP nowcasts. Finally, we collect GDP nowcasts from professional forecasters and show they were relatively accurate during the COVID crisis.

Keywords Nowcasting · Dynamic factor model · Model selection

JEL Classification C33 · C53 · E37

1 Introduction

Policy institutes such as central banks face significant challenges in assessing the economy using Gross Domestic Product (GDP) as the primary indicator. First, GDP is reported quarterly, which is a low frequency to make informed policy decisions.

We gratefully acknowledge comments from Peter van Els and Maurice Bun. The views expressed herein are those of the authors and do not necessarily reflect the position of De Nederlandsche Bank. All remaining errors are our own.

✉ Mick van Rooijen
m.van.rooijen@dnb.nl

1 De Nederlandsche Bank, Amsterdam, The Netherlands

Second, GDP is usually published with a significant delay. In the Netherlands, for instance, the initial GDP estimate is released 30 days after the end of each quarter.[1]

Nowcasting models use high-frequency data to estimate the current state of the economy, offering a way to bridge the gap between the low-frequent delayed GDP releases and real-time policy needs. The literature on nowcasting is evolving rapidly, with various econometric models and techniques being discussed; see Stundziene et al. (2024) for a survey. While many studies compare models and recommend preferred techniques, practitioners must also make decisions on data selection, transformation, outlier correction, and model specification.

This paper presents a guide for practitioners to develop a nowcasting model for GDP growth. We use a dynamic factor model and analyze forecast errors of specific modeling alternatives regarding data choice, transformations, outlier correction and model specification. The analysis of forecast errors identifies an optimal combination of these modeling choices. Additionally, we outline how to compile a novel dataset incorporating a broad range of economic indicators. Forecast accuracy is compared across small, medium and large datasets. Moreover, we analyze which types of variables should be included in the datasets and how they should be transformed.

We design an experiment featuring a pseudo real-time, out-of-sample forecasting competition that generates forecast errors to evaluate model features. Using a rolling window analysis, 360 models are (re-)estimated producing 123,192 forecast errors over the period from 2013 to 2023. Inspired by Coulombe et al. (2022), we construct a model to assess the marginal gain (or loss) in forecast accuracy of each modeling alternative. The results indicate that a multi-factor model with a parsimonious lag structure trained on a medium sized dataset including survey data is optimal for predicting GDP growth.

We compare the forecasts produced by our optimal model with forecasts made by a group of professional forecasters. The private firm Consensus Economics collects GDP growth forecasts from professional forecasters in the Netherlands (referred to as *Consensus forecasts*). We show that our model outperforms the Consensus forecasts under normal conditions.

Given that our sample includes the volatile COVID period, we account for it in our analysis. We present results for the periods excluding COVID, up to COVID and from COVID onward. As anticipated, our findings show that the COVID period is crucial in assessing the accuracy of model forecasts. We observe that the Consensus forecasts were highly accurate during the COVID crisis and significantly outperformed statistical models.

In addition to evaluating the forecast accuracy of the nowcasting model employed in this paper, we apply two concepts that help practitioners interpret GDP nowcasts. First, we derive the contribution of each individual variable to the model forecasts to understand to role of specific variables in a nowcast episode. Second, we illustrate how the impact of new data releases on the GDP nowcast can be calculated. This enables practitioners to better explain revisions to the nowcast when new data

[1] Statistics Netherlands reduced the publication delay from 45 to 30 days starting in the first quarter of 2025. This paper focuses on the period prior to 2025, during which the publication delay was 45 days.

become available. The latter is important as nowcasts are usually updated frequently (e.g. monthly, weekly or even daily).

The contribution of this paper is twofold. First, we provide practical guidance for central banks and policymakers on improving and interpreting nowcasting models, including how to derive variable contributions and monitor new data releases. Second, we compare our nowcasts with the Consensus Forecast for GDP growth in the Netherlands, a source not previously studied. The exceptional performance of expert forecasts during the COVID crisis emphasize the need for a balanced approach that combines statistical models with human judgment in extreme crisis periods.

The remainder of the paper is organized as follows. Section 2 reviews the existing nowcasting literature. Section 3 provides a description of the dynamic factor model. Section 4 outlines the dataset that is used in the forecast competition described in Sect. 5. Section 6 compares the forecast accuracy of the dynamic factor model to benchmark models as well as the Consensus forecasts. Section 7 illustrates how the nowcasting model can be used practice. Section 8 concludes.

2 Related Literature

The systematic analysis of economic data to gauge the current state of the economy dates back to Burns et al. (1946), who first introduced business cycle analysis in academia. Central banks have since built on academic foundations, developing models to forecast key variables related to the business cycle. Sims (2002) provides a comprehensive review and evaluation of forecasting practices at several major central banks at that time.

More recently, research shifted focus to current-quarter forecasts (nowcasts) to predict quarterly time series such as GDP. Giannone et al. (2008) developed a framework that tracks the real-time flow of information to update nowcasts. Using a two-step estimator to combine principal components with Kalman filtering techniques, monthly data releases are bridged to quarterly GDP estimates. Doz et al. (2012) derive the theoretical properties of the two-step dynamic factor model and establish the viability of this approach.

Bańbura and Rünstler (2011) build on the framework of Giannone et al. (2008) and apply a DFM to nowcast euro area GDP. Using a pseudo real-time dataset that replicates data availability within each month, Bańbura and Rünstler (2011) show that including survey data improves forecast accuracy. Moreover, the authors apply a method proposed by Koopman and Harvey (2003) to calculate the contribution of each individual variable to the nowcast.

Bańbura and Modugno (2014) extend the two-step DFM by incorporating the expectation maximization (EM) algorithm in the estimation process. The advantage of their modified EM algorithm is that the model can be applied to mixed frequency datasets with any pattern of missing data. This is convenient when including series of different sample lengths, but also addresses the problem of having incomplete data patterns at the end of the sample (ragged edges).

Nowcasting economic time series remains actively researched by central banks. Luciani and Ricci (2014) use a Bayesian dynamic factor model to nowcast Norwegian GDP. Research done at the central bank of Italy shows that payment system data can improve forecast accuracy (Aprigliano et al., 2019). Recently, machine learning methods are introduced as an alternative to the classical nowcasting methods (see e.g. Babii et al. 2022). Buckmann et al. (2023) provide a generic workflow for the use of machine learning models at policy institutes such as central banks.

We choose to stick to the DFM popularized by Bańbura and Modugno (2014) in this paper and focus on modeling alternatives. This method remains among the most cited in the nowcasting literature (Stundziene et al., 2024). Moreover, DFMs continue to be widely used for nowcasting by major central banks such as the ECB (Linzenich & Meunier, 2024) and the New York Fed (Almuzara et al., 2023).

3 Dynamic Factor Model

In this Section we briefly introduce the dynamic factor model and its estimation procedure applied in this study. Furthermore, we derive the contributions of each variable to the nowcast and we show how new data releases impact the nowcast.

Let $y_t = [y_{1,t}, \dots, y_{n,t}]'$, $t = 1, \dots, T$, denote the n-dimensional vector of stationary variables, standardized to mean zero and unit variance. We assume that y_t is driven by a few unobserved factors, that can be described using the following factor model representation:

$$y_t = \Lambda f_t + \epsilon_t \tag{1}$$

where y_t contains $i = 1, ..., n$ variables that potentially have mixed frequencies. The $r \times 1$ vector f_t contains the unobserved common factors, Λ is an $n \times r$ matrix of time-invariant factor loadings and the idiosyncratic component is captured in $\epsilon_t = [\epsilon_{1,t}, \dots, \epsilon_{n,t}]'$. The common factors f_t are modeled as a vector autoregressive (VAR) process of order p

$$f_t = A_1 f_{t-1} + \cdots + A_p f_{t-p} + u_t, \quad u_t \overset{i.i.d.}{\sim} \mathcal{N}(0, Q) \tag{2}$$

where $A_1, \dots, A_p$ are $r \times r$ matrices of autoregressive coefficients for the factors, and Q represents the covariance matrix of the errors.

The idiosyncratic component ϵ_t captures movements that are specific to individual series. In many applications, the idiosyncratic component is assumed to be multivariate white noise (e.g. Baǹbura and Runstler 2011) as factor estimates are asymptotically consistent even when this assumption in violated (Doz et al., 2012). However, modeling the idiosyncratic component can improve forecasts for two reasons. First, most macroeconomic variables are serially correlated and therefore it would be inappropriate to assume uncorrelated idiosyncratic shocks (Shapiro et al. 2002). In that case, imposing an assumption of no serial correlation is quite restrictive because it is only valid asymptotically. Second, the efficiency of factor estimates could be improved in (small) real-time samples with ragged edges.[2] Hence, we choose to put structure on the idiosyncratic component and assume it to follow an autoregressive (AR) process of order 1:

$$\epsilon_{i,t} = \rho_i \epsilon_{i,t-1} + \varepsilon_{i,t}, \quad \varepsilon_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_i^2) \tag{3}$$

where ρ_i is an autoregressive coefficient for indicator *i* and we assume that the process is univariate.

The mixed frequency dynamic factor model with serial correlation in the idiosyncratic component can be estimated using the Expectation Maximization (EM) model popularized by Bańbura and Modugno (2014). To do so, we cast Eqs. 1-3 into state space form.

3.1 State Space Representation

To be able to cast Eqs. 1–3 into state space form, we first need to decompose the AR(1) process for the idiosyncratic component. We assume that $\epsilon_{i,t}$ can be written as

$$\begin{aligned} \epsilon_{i,t} &= \tilde{\epsilon}_{i,t} + \xi_{i,t}, \quad \xi_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}(0, \kappa) \\ \tilde{\epsilon}_{i,t} &= \rho_i \tilde{\epsilon}_{i,t-1} + e_{i,t}, \quad e_{i,t} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_i^2) \end{aligned} \tag{4}$$

where both $\xi_t = [\xi_{1,t}, \ldots, \xi_{n,t}]'$ and $\tilde{\epsilon}_t = [\tilde{\epsilon}_{1,t}, \ldots, \tilde{\epsilon}_{n,t}]'$ are cross-sectionally uncorrelated and κ is a small number.[3]

Combining Eqs. 1, 2 and 4 results in the following state space representation:

$$\begin{aligned} y_t &= \tilde{\Lambda} \tilde{f}_t + \xi_t, \quad \xi_t \sim \mathcal{N}(0, \tilde{R}) \\ \tilde{f}_t &= \tilde{A} \tilde{f}_{t-1} + \tilde{u}_t, \quad \tilde{u}_t \sim \mathcal{N}(0, \tilde{Q}) \end{aligned} \tag{5}$$

where $\tilde{f}_t = \begin{bmatrix} f_t \\ \tilde{\epsilon}_t \end{bmatrix}, \quad \tilde{u}_t = \begin{bmatrix} u_t \\ e_t \end{bmatrix}, \quad \tilde{\Lambda} = [\Lambda \quad I], \quad \tilde{A} = \begin{bmatrix} A & 0 \\ 0 & \mathrm{diag}(\rho_1, \ldots, \rho_n) \end{bmatrix},$
$\tilde{Q} = \begin{bmatrix} Q & 0 \\ 0 & \mathrm{diag}(\sigma_1^2, \ldots, \sigma_n^2) \end{bmatrix},$

$e_t = [e_{1,t}, \ldots, e_{n,t}]'$ and $\tilde{R}$ is a fixed diagonal matrix with κ on the diagonal.

Note that y_t contains monthly and quarterly observed variables. We apply the method developed by Mariano and Murasawa (2003) and for each quarterly variable we construct an unobserved monthly counterpart. Hence, we assume the frequency of the model to be monthly. At least four lags of the factors are needed to be able to estimate the unobserved monthly counterpart of quarterly data. See Appendix A for further details on the mixed frequency model specification and the estimation procedure.

[2] Ragged edges are patterns of incomplete data at the end of the sample caused by publication delays of certain variables.

[3] We arbitrarily set κ to 1e−4. A value for κ that is too small makes converging difficult, whereas a large κ makes the factor extraction unreliable.

3.2 Contribution of Individual Variables

Policy makers may not only be interested in the nowcast for GDP itself, but also want to understand which variables drive a certain nowcast. The contribution of each variable to the GDP nowcast can be derived using an algorithm described in Koopman and Harvey (2003) and applied to dynamic factor models in Bańbura and Rünstler (2011). This Section briefly recapitulates the idea of assigning implicit weights to individual variables in making GDP nowcasts.

The vector y_t defined in Eq. 1 contains GDP and all other variables from which factors are extracted. In the most generalized state space form, y_t can be expressed as:

$$\begin{aligned} y_t &= W(\theta)\alpha_t + \xi_t, \quad \xi_t \sim \mathcal{N}(0, \Sigma_\xi(\theta)) \\ \alpha_t &= T(\theta)\alpha_{t-1} + \tilde{u}_t, \quad \tilde{u}_t \sim \mathcal{N}(0, \Sigma_{\tilde{u}}(\theta)) \end{aligned} \tag{6}$$

where $\theta = (\tilde{\Lambda}, \tilde{A}, \tilde{R}, \tilde{Q})$ denotes the set of model parameters and errors ξ_t and $\tilde{u}_t$ are defined in Eqs. 4 and 5. The Kalman smoother provides the smoothed estimate of the state vector $a_{t|T} = \mathbb{E}[\alpha_t | y_T]$ conditional on the data. Each individual element of the smoothed state vector $a_{t|T}$ can be decomposed into a weighted sum of all observations through time. The weighted sum is given by:

$$a_{t|T} = \sum_{j=1}^{T} w_j(a_{t|T}) y_j \tag{7}$$

where the weight matrices $w_j(a_{t|T})$ are computed using the algorithm from Koopman and Harvey (2003). Note that the sum in Eq. 7 has subscript *j*, which indicates that the weights are computed after the Kalman smoother has been applied for $t = 1, ..., T$ to get $a_{t|T}$. Hence, for each point in time *t* we sum all observations in y_j given weight w_j for $j = 1, ..., T$. Naturally, when *j* gets closer to *t*, the weight of observations in y_j tends to get larger.

Now that we have split up each element of smoothed state vector $a_{t|T}$ into a weighted sum of observables, we can link the weighted sum to the GDP nowcasts $\bar{y}_t^{GDP} \in \bar{y}_t^Q$ defined in Eq. A.9 in Appendix A. To do so, we plug in the smoothed state vector in Eq. A.9 and define the nowcast contributions as:

$$\bar{y}_t^{GDP} = \lambda a_{t|T} = \lambda \sum_{j=1}^{T} w_j(a_{t|T}) y_j \tag{8}$$

where λ is the loading vector associated to GDP which is part of the loading matrix Λ_Q from Eq. A.9. Since $\bar{y}_t^{GDP}$ is a normalized value, we multiply the terms in Eq. 8 by the standard deviation of GDP and add the mean to revert the GDP forecast to the original units.

Note that the state vector α_t contains both the factors and the idiosyncratic component, as denoted in Eq. A.9. Hence, the estimated smoothed state vector $a_{t|T}$ (and its equivalent of weighted observables $\sum_{j=1}^{T} w_j(a_{t|T})y_j$) includes both the factors and the idiosyncratic component. The contribution of a variable to the GDP forecast is the sum of the individual elements in the state vector, scaled by the loading vector for GDP.

3.3 News in Nowcast Revisions

Policy makers continuously monitor new data releases and update their beliefs on the GDP nowcast. Intuitively, when new data come in, only the unexpected component of the data release should revise the nowcast. Following Bańbura et al. (2011), we update our GDP nowcast on a regular basis and assess the impact of new data releases on the nowcast. The state space framework provides forecasts for all other variables than GDP. Hence, we can extract the unexpected component from the data release and its effect on the GDP nowcast, which we define as *news*. The framework is used to understand the changes in GDP nowcasts when new data come in.

We consider two data vintages Ω_v and Ω_{v+1} of the same set of variables retrieved at dates v and $v+1$. The data vintage Ω_{v+1} contains new observations for variables $i = 1, ..., z$ that are not available in Ω_v.[4] The new observations are denoted by $y_{i,t}^{v+1}$ and tend to contain at least some new information. We assume that data are not revised and that the incoming new part of the information set is orthogonal to the current information set. Hence, we can write:

$$\mathbb{E}[\bar{y}_t^{GDP}|\Omega_{v+1}] = \mathbb{E}[\bar{y}_t^{GDP}|\Omega_v] + \mathbb{E}[\bar{y}_t^{GDP}|I_{v+1}] \tag{9}$$

where

$I_{v+1} = [I_{v+1,1}, ..., I_{v+1,z}]', \;\; I_{v+1,i} = y_{i,t}^{v+1} - \mathbb{E}[y_{i,t}^{v+1}|\Omega_v], \; i = 1, \ldots, z.$

The vector I_{v+1} represents the part of the release in each variable $y_{i,t}^{v+1}$ that is unexpected based on the information in Ω_v. This unexpected part is referred to as the surprise in the data release. Note that if a new observation in Ω_{v+1} is exactly equal to what is expected based on Ω_v, there is no surprise and the GDP nowcast will not change.

The aim is to compute the second term of Eq. 9, which is the forecast revision caused by the incoming data. We define a vector $B_{v+1} = [b_{v+1,1}, ..., b_{v+1,z}]$ that links the surprise in the data release of each variable to the new GDP forecast. The nowcast revision can be written as:

$$\underbrace{\mathbb{E}[\bar{y}_t^{GDP}|\Omega_{v+1}] - \mathbb{E}[\bar{y}_t^{GDP}|\Omega_v]}_{\text{nowcast revision}} = B_{v+1}I_{v+1} = \sum_{i=1}^{z} b_{v+1,i}(\underbrace{y_{i,t}^{v+1} - \mathbb{E}[y_{i,t}^{v+1}|\Omega_v]}_{\text{surprise}}). \tag{10}$$

[4] Note that we have new observations in Ω_{v+1} for variables $i = 1, ..., z$ which are not necessarily all n variables in Eq. 1. Furthermore, it is irrelevant at which day the new observation is observed, as long as it is between dates v and $v+1$.

As showed by Bańbura et al. (2011), both matrices B_{v+1} and I_{v+1} can be computed using the Kalman filter and smoother.

Note that the forecast of $\bar{y}_t^{GDP}$ not only depends on a given data vintage Ω_v or Ω_{v+1}, but also on a parameter vector θ that depends on one of the two data vintages. To identify the news impact of each variable, we should keep the parameter vector θ constant. We estimate the parameter vector θ using Ω_{v+1} and apply those parameter estimates to compute $\mathbb{E}[\bar{y}_t^{GDP}|\Omega_v]$.

In conclusion, the revision of the GDP forecast is decomposed as a weighted sum of surprises in the most recent data vintage. The news in the GDP forecast depends on the size of the surprises captured in I_{v+1}, as well as on the relevance of the variables in forecasting GDP indicated by B_{v+1}.

The news and contribution methods are particularly relevant for policymakers when applying the model in practice. This topic is discussed in greater detail in Sect. 7.

4 Data

4.1 Dataset

We construct a small, medium and large dataset for our analysis by meticulously examining all indicators from the publicly available data sources: Statistics Netherlands, Eurostat, the ECB Data portal and the European Commission. In addition to publicly available data, the database also includes information from proprietary sources: Refinitiv (e.g. PMI manufacturing), Betaalvereniging Nederland for debit card payments, and Het Financieele Dagblad (FD) for tone-adjusted news topics based on previous research (see Van Dijk and De Winter 2023). The small, medium, and large datasets contain 24, 67 and 129 monthly series, respectively. Additionally, we have 4, 9 and 9 quarterly variables in the small, medium and large datasets, respectively. A comprehensive list of all variables can be found in Appendix B.

The variable selection procedure to define the three databases is inspired by Bańbura and Modugno (2014), Alvarez and Perez-Quiros (2016) and Barigozzi & Luciani (2024) and based on availability of the data.[5] We categorize variables in seven distinct groups following the FRED-MD database developed by McCracken and Ng (2016). Furthermore, the data are transformed analogously to the FRED-MD database. In short, the different databases can be summarized as follows:

- ***Small***: This dataset comprises the main "market-moving indicators" of real activity for the Dutch economy.[6] These include industrial production, the stock of new

[5] We only select series that have a start date no later than March 2000.

[6] Following Bańbura et al. (2013) we used the Bloomberg terminal to identify the "market moving indicators" for the Netherlands by filtering all macroeconomic information that is relevant to the country. We thank one of the anonymous referees for raising this point. The filtering process yielded a total of 13 indicators, each accompanied by a relevance score (shown in brackets): CPI Netherlands (94.7), S&P global/Nevi Netherlands manufacturing PMI (90), unemployment rate (84.2), gross domestic product (79.0), retail sales (52.6), consumer spending (47.4), manufacturing production (42.1), producer confi-

orders, retail sales, the unemployment rate, the economic sentiment indicator, the purchasing manager index and confidence levels reported in newspapers. Additionally, it includes financial series such as stock price indices or raw material prices as well as the headline FD-sentiment index. This dataset encompasses a total of 24 monthly series and 4 quarterly series;

- ***Medium***: In addition to the series covered in the small specification, provides more disaggregated information on industrial production, survey data, and national accounts. It nearly encompasses all the essential real economic indicators for the Netherlands as reported by Statistics Netherlands and Eurostat. Furthermore, it includes four financial newspaper sentiment indices, each representing the sentiment of one of the four main topics that make up the headline index and the European stock market index. This dataset consists of 67 monthly series and 9 quarterly series;
- ***Large***: Apart from the indicators in the medium-sized dataset, this dataset incorporates more series also regarding the Euro area. It offers greater sectoral granularity for industrial production, the services sector, and retail trade. Additionally, it includes sub-indices from sentiment surveys, sixteen more granular newspaper sentiment indices, and PMIs. The large dataset comprises 129 monthly series and 9 quarterly series.

4.2 Consensus Forecasts

We also compare the forecasts of our dynamic factor model to the forecasts of professional forecasters. The quarterly GDP forecasts by professional forecasters were collected from the monthly *Consensus forecasts*, published by Consensus Economics. This publication offers a survey of private sector analysts' expectations for a set of key macroeconomic variables for a broad range of countries. Once a quarter, the publication contains the averaged forecasts for quarterly GDP over a horizon of six quarters, starting with the nearest quarter for which no officially released figure is available. New quarterly Consensus forecasts become available in the second week of the last month of the quarter. The survey date (deadline for respondents) is typically the second Monday of the third month of a quarter; publication is usually three days later on Thursday. The timing of the survey is therefore broadly in line with the timing of the monthly data vintages we collected. In our information set Consensus forecasts are not updated in the first and second months in a quarter, while monthly indicators are updated every month. Moreover, at the time analysts form their expectations they have official information on GDP growth in the preceding quarter. To the best of our knowledge, the Consensus forecasts have not been used before to assess model-based forecast accuracy, except in a recent case study for the Netherlands (Jansen and de Winter 2018).

dence index (36.8), industrial sales (26.3), House price index (21.15), consumer confidence index (10.5), exports (5.3) and imports (0). We added another 13 indicators based on the variables and size of the small dataset in Bańbura et al. (2013). These include the AEX stock market index, loans to households and Eurozone gross domestic product.

5 Empirical Application

In this Section we describe the design of our forecast competition and show the results. Also, we indicate how we evaluate modeling alternatives. Based on the out-of-sample forecast errors of models with different features, we propose an optimal specification for a nowcasting model for GDP growth.

5.1 Forecasting Design

GDP is only observed at the quarterly frequency and had a publication lag of approximately 45 days in the Netherlands.[7] Hence, we produce forecasts, nowcasts and backcasts as the model yields an estimated value for GDP growth each month.

A sequence of eight estimates for GDP growth is constructed for each quarter, obtained for consecutive months. Table 1 illustrates the timing of the forecasting exercise in more detail, taking the estimates for the third quarter of 2013 (2023Q3) as an example. We make the first forecast on April 1, 2023, which is a one-quarter-ahead forecast in month one. Subsequently, we produce monthly forecasts for the next seven months up to and including November. The last estimate is produced just two weeks before the (first) release of GDP in mid-November. Following the conventional terminology, forecasts refer to one or more quarter-ahead forecasts, nowcasts refer to current quarter estimates, and backcasts refer to estimates for the preceding quarter, before official GDP figures become available.

All experiments are conducted in a pseudo real-time setting. The first forecasts are produced based on model estimations with data used up until April 2013, and we re-estimate the model in each consecutive month to produce new forecasts up until the last model estimation in November 2023. We estimate the parameters of all models recursively in a 10-year rolling window, using only the information that was available at the time of the forecast. More specifically, starting from April 2013, we reconstruct pseudo-real-time vintages by replicating the data availability pattern as implied by a stylized release schedule. This is done by recursively removing observations from the full dataset according to a fixed schedule.

Table 1 Timing of forecast exercise for third quarter

No	Forecast type	Month	Forecast on the 1^{st} of
1	One-quarter-ahead forecast	1	April
2		2	May
3		3	June
4	Nowcast	1	July
5		2	August
6		3	September
7	Backcast	1	October
8		2	November

[7] Statistics Netherlands changed the publication lag from 45 to 30 days, but only as of the first quarter of 2025. For the period analyzed in this paper, the publication lag was 45 days.

Given the unavailability of real-time historic data vintages, only final revised data downloaded on January 26, 2024 is used in the forecasting exercise.[8] As a result, the role of historic data revisions is ignored and we can only perform a pseudo real-time out-of-sample exercise. However, factor models are known to be robust to data revisions since revision errors are idiosyncratic by nature and may cancel out; see for example Bernanke and Boivin (2003) for the United States and Schumacher and Breitung (2008) for Germany. For similar approaches, see Giannone et al. (2008), Jansen and de Winter (2018) and Kant et al. (2025), among others. Another limitation is that we assume the publication calendar to stay the same throughout the entire sample period. In reality, the release delays might change over time.

In the forecast exercise we consider the period starting from 2013Q3 up until 2023Q3. This period holds three sub-periods of economic upswing and just as many downturns, giving a view on the forecast accuracy of the models over the business cycle. Moreover, the evaluation period is long enough to determine the statistical significance reliably.[9]

The analyzed period is also special as it includes the aftermath of the European debt crisis, the COVID pandemic, and a surge in energy prices due to the Russian invasion of Ukraine. The swings in GDP during the COVID crisis have no precedence in terms of size. To assess the robustness of our regression results for the COVID crisis, we also estimate regression models over the period 2013Q3-2023Q3, excluding the COVID crisis. The latter is defined as the period 2020Q1-2020Q3. We run separate regressions for all models during the period leading up to the COVID crisis (2013Q3-2019Q4) to assess the forecasting performance of the different model specifications in more tranquil times. The statistical outcomes for the sub-periods are somewhat less reliable compared to those over the full sample period, because the number of observations is smaller.

5.2 Evaluating Modeling Alternatives

In the pseudo real-time forecasting exercise several modeling alternatives are considered. We analyze the impact of modeling alternatives on the forecasting performance along four dimensions:

- ***Model specification***: Based on out-of-sample forecasting performance, we determine the number of factors and lags included in our model (Eq. 5).
- ***Dataset***: We investigate if it matters how many data series we use, by estimating the models with the small, medium, and large datasets extracted from the database described in Sect. 4. Besides the size of the dataset, we also investigate the merits of including/excluding survey data and quarterly data. Furthermore, we determine the added value of data that are only available with a subscription (Refinitiv) or that are confidential (payments data or the FD sentiment indicator).

[8] Statistics Netherlands provides real-time data revisions for GDP across all quarters. However, for our explanatory variables, only the revised observations are available. Therefore, our analysis is based solely on revised data.

[9] Periods of economic upswing and downturn defined on the basis of the DNB business cycle indicator.

- ***Outliers & transformation***: We show the impact on forecast accuracy by comparing the performance when we do not treat the data for outliers, do a mild outlier correction, or a more stringent correction.[10] Additionally, we examine how to transform survey data as there is an ongoing debate among practitioners whether to include surveys in levels or in first differences.
- ***Timing***: Forecast accuracy tends to increase as more monthly data are available. We empirically assess this along three axes: the quarter in which the forecast is made (i.e. is it a backcast, nowcast or one-quarter-ahead forecast), the month within the quarter (month 1, month 2 and month 3), and the day within each month (first day of the month, middle of the month or end of the month).

To evaluate the modeling alternatives, we stack the forecast errors from all treatments in one vector and tease out the impact of the different model features on the forecast accuracy by defining separate dummy variables for each model feature. The regression model to evaluate model alternatives is given by:

$$e^2_{m,h,t} = \delta + \beta' D_m + \tau_{y(t)} + u_{m,h,t} \tag{11}$$

where $e^2_{m,h,t}$ is the log squared forecast error produced by model *m* for horizon *h* at time *t*, δ is a constant, $\tau_{y(t)}$ captures yearly time fixed effects and D_m is a vector of dummy variables that represent the features of model *m*. The vector of coefficients β captures the improvement/deterioration of features corresponding to model *m*. In this log-level specification the percentage impact of the model features on the squared forecast error can be calculated as $(\exp(\beta) - 1) \times 100$. The squared forecast error puts a higher weight on large forecast errors, in line with the loss function of a central bank (i.e. making large forecast errors is much more costly than small forecast errors). In total we analyze 167,076 forecast errors.

In each regression, the null hypothesis is that there is no predictive accuracy gain with respect to the base specification. The base specification is estimated using the small dataset and has the following model features: estimation with data available on the 1st day of the month, surveys in first difference, both public and restricted data, no quarterly data, mild outlier correction and a DFM with a single factor and one lag.

5.3 Forecasting Competition

This Section presents the results of a forecasting competition among the modeling alternatives described in the previous Section. The whole sample period from 2013Q3 up until 2023Q3 is considered. If, however, a certain result is strongly impacted by the COVID period, we highlight separate results for the pre-COVID period as well as for the period as of COVID. All figures related to these alternative evaluation periods are presented in Appendix C.

[10] We define a mild outlier correction as a winsorization at the median ± three times the empirical interquartile range. For a strong outlier correction we winsorize at the median ± two times the empirical interquartile range.

The figures in this Section illustrate the forecasting gain (or loss) relative to the base specification. The bars correspond to regression estimates of β in Eq. 11. All effects are relative to a benchmark model with one factor, one lag, estimated using the small dataset (including both public and restricted data), exluding quarterly variables, surveys in first difference and only extreme outliers are deleted. Our evaluation of model specifications involves two approaches:

- ***Economically meaningful***: If a bar falls within the shaded region, the squared errors deviate by more than 5% from the base model specification.[11] This ±5% threshold serves as a rough, informal gauge of the economic significance of forecast accuracy gains resulting from this model feature, as previously employed by Jansen and de Winter (2018);
- ***Statistical significance***: Coefficient significance is formally tested at the 5%-level, following methodologies by Coulombe et al. (2022) and Carriero et al. (2019). Colored bars indicate statistically significant coefficients:
  - ***Green bars***: Represent an improvement in forecast accuracy;
  - ***Red bars***: Represent a deterioration in forecast accuracy.

5.3.1 Results Model Specification

One of the messages conveyed by Fig. 1 is that a single lag in the dynamic structure of the model is sufficient. However, to better capture the variance in the monthly dataset, the model requires additional factors. The number of factors could potentially increase the forecast accuracy up to 17% for backcasting. The gains in forecast accuracy are largest for backcasts, followed by the nowcasts.

Overall, the best-performing specifications have three factors.[12] This outcome is strongly driven by the period post-COVID. Pre-COVID, the only statistically and economically significant impact on the forecast accuracy stems from the third factor when backcasting (see Fig. 11 in Appendix C). This indicates that more model features are required to capture the more complicated dynamics since the COVID crisis.

5.3.2 Results Dataset

In comparing the different datasets described in Sect. 4, we find that the small dataset performs as good as the larger datasets in terms of forecast accuracy as showed in

[11] The base model specification includes 1 factor and 2 lags, and all coefficients are estimated using a 10 year rolling window. The model is estimated using a small dataset including both public & non-public data, excluding series with a quarterly frequency. All series are corrected for extreme outliers and all survey data are included in first difference. The ragged edges in the data are constructed as if the data where downloaded on the first day of the month.

[12] This outcome might raise the question if there is any increment from going from three to four factors. This increment is (very) small and is in line with the statistical tests conducted to determine the number of factors, following Bai and Ng (2002), and scree plot tests Catell (1966). Moreover, including too many factors can yield non-negligible estimation errors, or overfitting of the model (see e.g., Miranda et al 2022 and Barigozzi and Cho 2020).

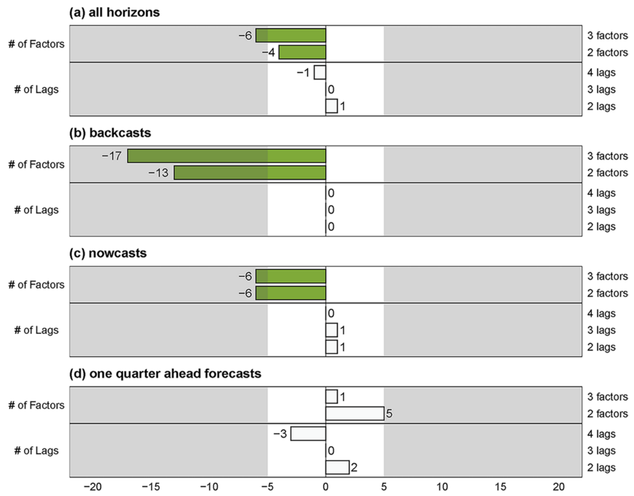

Fig. 1 Impact of model-specification on forecast accuracy, full sample *Note:* Impact of features on out-of sample mean squared forecast, compared to base model, in percent

Fig. 2. Hence, a model that only includes series that measure total economy concepts performs equally well compared models estimated using larger datasets.[13] This may be attributed to the challenges of extracting a relevant signal when dealing with indicators of varying quality, as highlighted by Boivin and Ng (2006) and Bańbura and Modugno (2014). Our findings align with the existing literature (e.g., Caggiano et al., 2011 and Havrlant et al.,2016) suggesting that small to medium sized datasets (typically containing 10-30 variables) perform just as well as models with larger datasets (containing over 100 variables). The results remain consistent across different time periods, including the tranquil pre-COVID period and the period without the COVID crisis.

When considering the type of data to include, Fig. 2 indicates that incorporating surveys strongly and significantly improves forecast accuracy. This results is apparent in both the pre-COVID period as well as the period without COVID. Furthermore, adding quarterly data (such as production capacity and sub-components of GDP) does not lead to improved forecast accuracy in our sample. Measured over the total sample period, adding series from restricted data sources (Refinitiv, FD) does not impact the forecast accuracy, and this results is quite persistent.[14] There is one exception: Adding the restricted series lowers the average RMSFE of the backcasts during the no-COVID period by 14%. This implies that including restricted series increases the forecast accuracy of backcast after the COVID crisis.

[13] The gains from using a large dataset are comparatively small. Measured over the total sample the deterioration is statistically significant, but economically, the deterioration is not sizable (3%).

[14] The FD-indicators significantly improves the forecast accuracy in a nowcasting model over a longer sample. This is also one of the reasons why we opt to include FD-indicators, see Van Dijk & De Winter (2023).

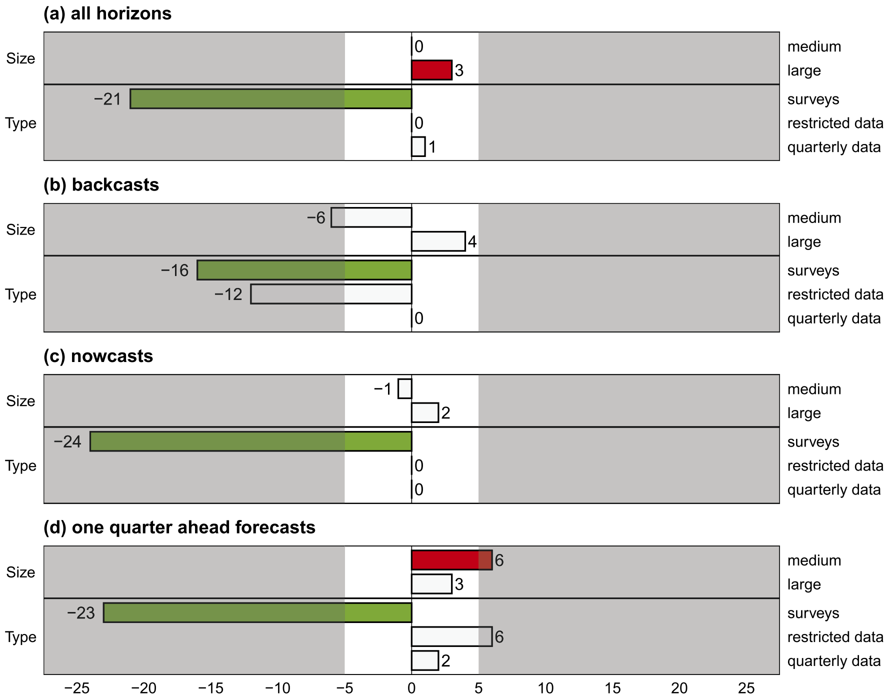

Fig. 2 Impact of choice of data on forecast accuracy *Note:* Impact of features on out-of-sample mean squared forecast, compared to base model, in percent

Based on these findings, we conclude that expanding the dataset size and incorporating quarterly data do not necessarily enhance forecast accuracy. However, policymakers might still prefer a medium-sized or large dataset over a small one, for the purpose of interpreting the information conveyed by their releases. Notably, the inclusion of survey indicators significantly improves forecast accuracy and should be considered. Although adding restricted data to the dataset does not increase the forecast accuracy, it also does not hurt. Some indicators, such as the PMI and financial market data are strongly favored by policy makers and we therefore include them in our model. We decide to opt for a medium-sized dataset including survey indicators and series from restricted data sources. We do not include quarterly data.

5.3.3 Results Outliers & Transformation

Fig. 3 shows the impact of outlier correction and transformation of survey variables. We find that a mild correction for outliers appears to be sufficient.[15] The forecast accuracy is not adversely affected by using a more stringent outlier correction. Using

[15] As stated in Sect. 5.2, we define a mild outlier correction as a winsorization at the median ± three times the empirical interquartile range. For a strong outlier correction we winsorize at the median ± two times the empirical interquartile range. Note that the outlier correction is conducted over the estimation window. We calculate the forecast accuracy using the true (non-outlier corrected) GDP growth.

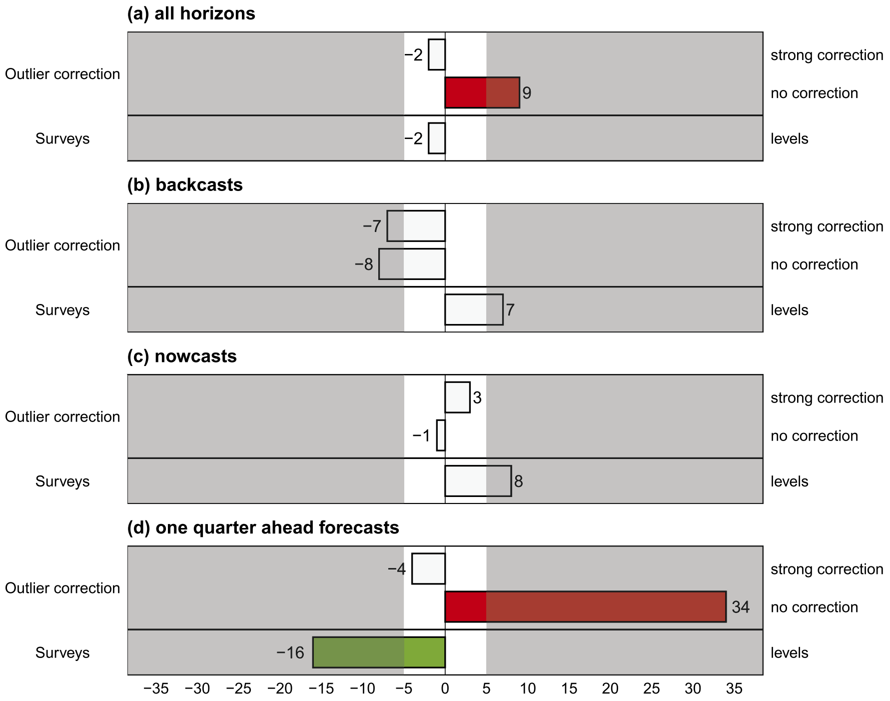

Fig. 3 Impact of outlier correction & transformation on forecast accuracy *Note:* Impact of features on out-of sample mean squared forecast, compared to base model, in percent

no outlier correction strongly worsens accuracy, especially in one-quarter-ahead forecasts. It can lead to a 37% lower accuracy for forecasts in the period excluding the COVID crisis and 27% lower forecast accuracy in the pre-COVID period.

Given the debate in the literature on inclusion of surveys in level or first difference, we tested both specifications. As shown in Fig. 3, there is no significant difference between including surveys in levels and in first difference when all horizons are considered. The improved forecast accuracy in the one-quarter-ahead forecasts is offset by poor forecasting performance in the other forecast horizons. Hence, we do not find convincing evidence for including surveys in first difference or in levels. This claim is robust to analyzing the periods before and without the COVID crisis.

5.3.4 Results Timing

The results presented in Fig. 4 show that when time passes by and more information gets available, the forecast accuracy increases. Measured over the full period, the nowcasts and backcasts are 25% and 39% more accurate than the one-quarter-ahead forecasts, respectively. The same logic applies when forecasting later in the quarter: forecasting in the third month of a quarter improves the forecast accuracy by 9% compared to the first month. Measured over the complete sample, this effect is much more pronounced for the nowcasts than the one-quarter-ahead forecast. Note that for backcasts we do not find a significant improvement in the second month compared

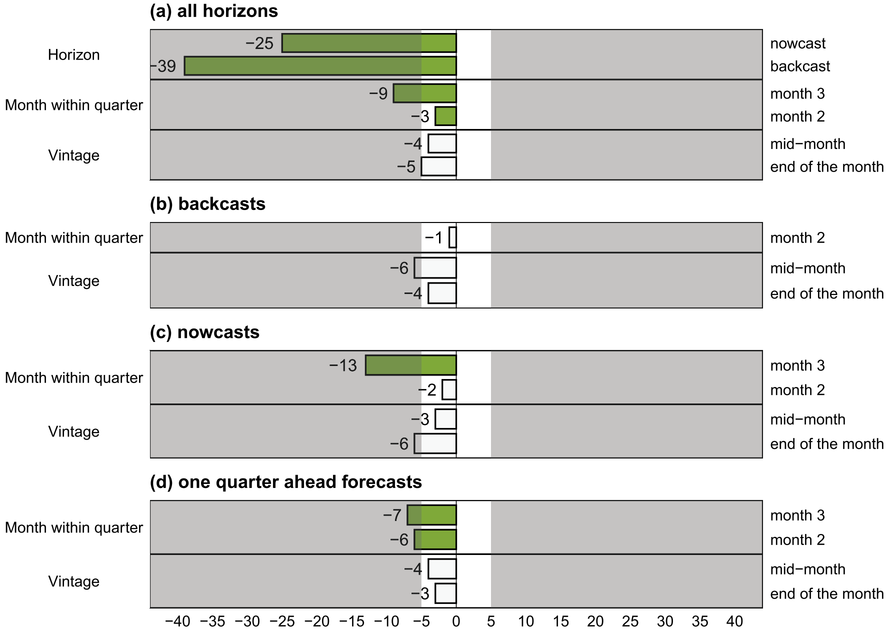

Fig. 4 Impact of timing on forecast accuracy *Note:* Impact of features on out-of sample mean squared forecast, compared to base model, in percent

to the first month. This indicates that the most important information for backcasting a quarter is already available at the first month. More details on when data become available is provided in Appendix C.1.

The differences in forecasting at the beginning, middle, or end of the month are neither economically meaningful nor statistically significant. This result implies that it is much more important whether you are backcasting, nowcasting, or forecasting a quarter, and which specific month you are in within a quarter, rather than the specific day of the month.

5.3.5 Final Model Specification

Based on our empirical results we construct an optimal model to produce backcasts, nowcasts and one-quarter-ahead forecasts. The model has three dynamic factors, each with a single lag. The medium database is chosen to estimate the model parameters and we apply a mild outlier correction method to all series. We include survey data and data from restricted sources, but we abstain from including quarterly data. As we do not find clear evidence for including survey data in first differences or levels, we estimate the model twice using both options. We take the simple average of the two model forecasts as our final forecast. Optimizing this simple average using techniques like Bayesian model averaging could open up an avenue for further research.

6 Comparison to Benchmark Models

This Section compares the forecast accuracy of our optimal model based on Bańbura and Modugno (2014) (hereafter *BM*) with a number of benchmark models.[16]. The benchmark models considered are a random walk model, an autoregressive model with one lag and a popular alternative dynamic factor model specification (Bańbura and Runstler, 2011). Additionally, we report the relative forecast accuracy of BM against the forecasts of professional analysts, the Consensus Forecasts.

6.1 Comparing Forecast Accuracy

Table 2 presents the main outcomes of the comparison. Panel (a) shows the out-of-sample relative Root Mean Squared Forecast Error (rRMSFE) of the benchmark models and the Consensus Forecasts over the total evaluation period, i.e. the period 2013Q3–2023Q3. All RMSFEs are in relative terms (rRMSFE) against BM, i.e.: $\mathrm{rRMSFE}_{\mathrm{ALT}} = \frac{\mathrm{RMSFE}_{\mathrm{ALT}}}{\mathrm{RMSFE}_{\mathrm{BM}}}$. Here, ALT denotes the forecast of the alternative, i.e. one of the benchmark models or the Consensus forecasts. A value higher (lower) than 1 indicates that the RMSFE of the alternative forecast is higher (lower) than the RMSFE of BM. Bold cells indicate the cases where the alternative model RMSFE is at least 5% higher than the RMSFE of BM. Starred entries (*, **, ***) indicate that the one-sided Diebold and Mariano (1995) test indicate the difference is statistically significant at the 10%, 5%, and 1% levels,

Table 2 Relative RMSFE of benchmark models and Consensus forecasts versus BM

	Backcast		Nowcast			1Q Ahead forecast		
	M2	M1	M3	M2	M1	M3	M2	M1
(a) Total evaluation period: 2013Q3 - 2023Q3 (*N*= 41)								
Random walk	**1.21****	**1.21****	**1.20****	**1.19*****	**1.14*****	**1.07****	0.96	1.00
Autoregressive model	**1.20****	**1.20****	**1.19***	**1.09**	1.04	0.98	0.94	0.98**
Dynamic factor model B&R	0.99	1.03	**1.05**	1.00	**1.05**	1.02	0.96	0.97
Consensus forecast	0.68	0.68	0.67	0.87	0.84	0.79	0.91	0.95
(b) Pre-COVID: 2013Q3 - 2019Q4 (*N*= 26)								
Random walk	**1.30***	**1.28***	**1.23***	**1.18***	**1.20****	**1.06**	**1.31****	**1.30**
Autoregressive model	**1.11**	**1.10**	**1.06**	1.00	1.02	0.90	1.00	0.99
Dynamic factor model B&R	**1.06**	**1.08**	**1.06**	**1.08***	**1.11****	0.98	1.02	1.02
Consensus forecast	**1.35****	**1.33****	**1.27****	1.00	1.02	0.90	0.84	0.84
(c) No-COVID: 2013Q3 - 2023Q3, excluding 2020Q1 -2020Q3 (*N* = 38)								
Random walk	**1.40****	**1.21****	**1.19****	**1.19*****	**1.14*****	**1.07****	0.95	1.00
Autoregressive model	**1.27***	**1.27***	**1.22**	**1.24***	**1.17***	1.00	0.98	0.95
Dynamic factor model B&R	**1.05**	1.04	1.01	1.01	1.01	0.88	0.88	0.88
Consensus forecast	**1.49****	**1.48****	**1.43****	0.98	0.92	0.79	0.81	0.78

Bold cells indicate the RMSFE is at least 5% worse than BM. Starred entries (*, **, ***) indicate that the one-sided Diebold-Mariano test (alternative is worse than the baseline) is significant at the 10%, 5%, and 1% levels, respectively

[16] This model is actually used by the DNB to produce GDP nowcasts for the Dutch economy. The estimates are published on a monthly frequency on the website of DNB.

respectively.[17] Panel (b) shows the outcomes for the same key figures for the Pre-COVID period, whilst Panel (c) shows the outcomes for the entire period excluding COVID. The outcomes in Table 2 point to several interesting results.

First, BM backcasts and nowcasts are generally more accurate than the random walk and autoregressive models, further the "naive benchmark models". This conclusion specifically holds in the total evaluation period and the period without the volatile COVID quarters. Although we find somewhat lower significance levels in the DM test regarding the pre-COVID, the rRMFSE still points towards an advantage of BM over the naive benchmark models. When forecasting one-quarter-ahead BM does not always beat the naive benchmark models. This is a well-known phenomenon in the literature: DFMs perform well when information on the quarter to forecast is partly known, but does not necessarily have the competitive edge when the forecast horizon is longer and no information on the quarters is available, see e.g. Giannone et al. (2008) and Jansen and de Winter (2018). The added value of BM increases when more data for the forecasted quarter arrive and can increase up to 40% for the period excluding the COVID period.

Second, comparing the outcomes in panel (a) of Table 2 with the outcomes in panel (b) and (c) of Table 2 it is evident that the competitive edge of BM over the naive benchmark models diminished during the COVID crisis. This is not surprising, as the onset and severity of the crisis were largely unapparent in the monthly data releases, partly because all series were adjusted for extreme outliers.

Third, BM outperforms the popular dynamic factor model of Bańbura and Rünstler (2011) in the pre-COVID period, but only by a small margin. We observe a moderate, but insignificant, improvement of 7% on average when backcasting. The nowcasts of BM perform up to 11% better in the pre-COVID period compared to the model proposed by Bańbura and Rünstler (2011). However, measured over the whole sample as well as excluding the COVID crisis, both models have a rather similar forecast accuracy. Note that BM has some practical advantages over the model introduced in Bańbura and Rünstler (2011), such as the ability to cope with missing values in an elegant manner.[18]

Fourth, BM outperforms the Consensus forecasts when nowcasting and backcasting in normal times, but Consensus forecasts are more accurate in times of large distress. This result has been documented before in Jansen and de Winter (2018), Liebermann (2014) and Lundquist and Stekler (2012). It reflects the inability of mechanical statistical models to incorporate expert knowledge. Professional forecasters are very responsive to the latest information about the state of the economy that is not captured in the monthly indicators and can adjust their forecasts quickly. Strikingly, professional analysts fail to produce accurate back- and nowcasts during tranquil times, and the forecast accuracy is the mirror image of the good forecast performance during the COVID crisis. Over the whole evaluation period excluding the three COVID quarters the professional analysts are beaten by a 46% margin by BM when backcasting.

[17] The RMSFEs are based on the most recent realization of GDP growth, the outcomes are robust to using the "first estimate" of GDP growth of CBS at the time of release.

[18] Besides the arbitrary pattern of missing values, BM employs the EM algorithm accounting for uncertainty in factor extraction. Furthermore, Bańbura and Modugno (2014) provide a comparison between their model and the model of Bańbura and Rünstler (2011) and find that both methods perform similar in terms of forecast accuracy.

6.2 Comparing the Evolution of Forecast Accuracy in Time

Fig. 5 casts the outcomes in Table 2 into the time dimension, showing the *cumulative* sum of squared error difference (CSSED) moving forward in time, calculated as the cumulative sum of squared errors of the alternative model *minus* the cumulative squared error of BM. A CSSED *below* zero indicates that the alternative model's forecasts have a *lower* CSSE up until that that point in time, and are therefore more accurate than BM. If the CSSED is *above* zero this indicates the reverse and the alternative model has a a lower forecast accuracy at

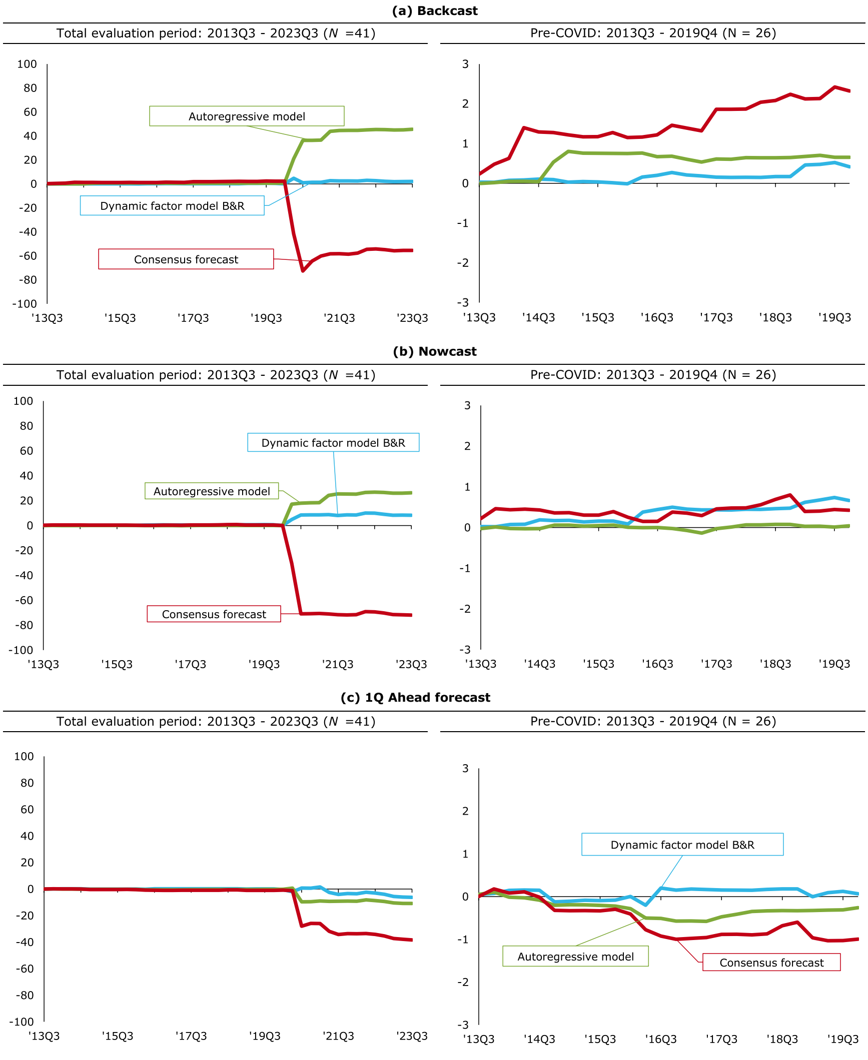

Fig. 5 Cumulative sum of squared error difference (CSSED) *Note:* CSSED calculated for the average errors in the backcast (M1 and M2), nowcast (M1, M2 and M3) and 1 quarter ahead forecast (M1, M2 and M3). A CSSED above zero indicates better performance of BM versus a benchmark

that point in time. Furthermore, an increase in the CSSED indicates that the model performance of the alternative model is decreasing vis-a-vis BM. A decline indicates the opposite.

The left-hand graphs in panel a, b and c of Fig. 5 describe the evolution of the CSSED's over the entire evaluation period for the backcasts, nowcasts and one-quarter-ahead forecasts, respectively. The left-hand graphs of Fig. 5 obscure the development prior to the COVID crisis due to significant shifts in the relative forecast accuracy of the models in COVID. To gain a clearer understanding of the pre-COVID CSSED evolution, the right-hand graphs display the CSSED up to the COVID period for the backcast, nowcast, and one-quarter-ahead forecast, respectively.

Panel a of Fig. 5 shows that when backcasting, the forecasting advantage of the Consensus forecast solely stems from the COVID period. There is a huge and sudden decline in the CSSE of the Consensus forecast vis-a-vis BM. After COVID, there is an upward trend in the red line, indicating that BM regained some of its forecasting advantage. The COVID period caused a significant decline in the forecast accuracy of the autoregressive model, as evidenced by the large increase in the green line. The alternative dynamic factor model (Bańbura and Runstler, 2011) showed no noticeable change, suggesting that its forecast accuracy during backcasting remained roughly the same as BM.

The evolution of the CSSED for the nowcasts (Panel b of Fig. 5) is similar to that for the backcasts, though by a lesser extent. A notable difference is the significant deterioration in the forecast accuracy of the dynamic factor model of Bańbura and Rünstler (2011). When

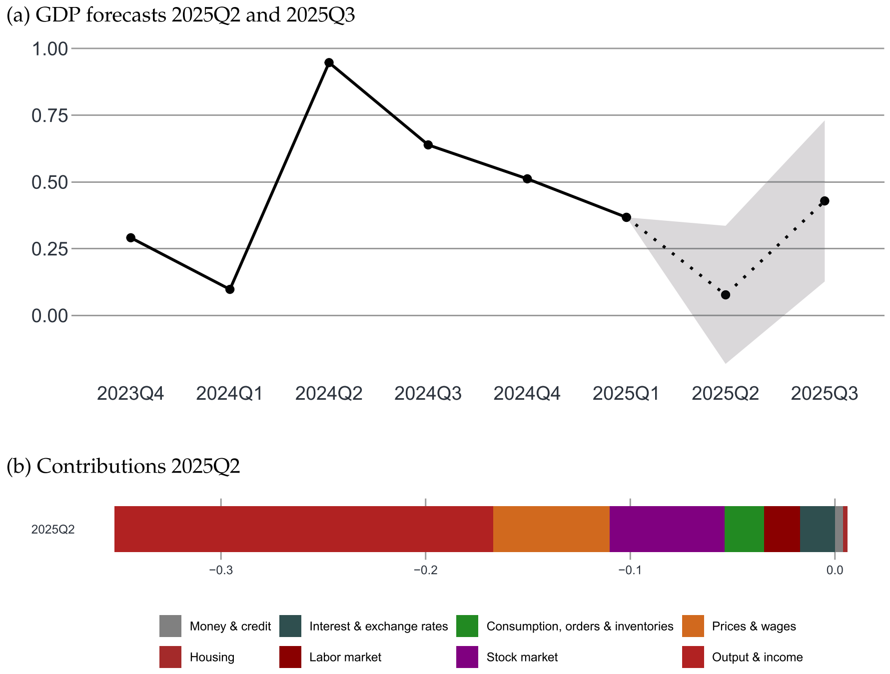

Fig. 6 GDP forecasts and contributions *Note*: Panel (a) displays the GDP backcast for 2025Q2 and the GDP nowcast for 2025Q3, both made on 15-07-2025. The forecasts represent quarter-on-quarter growth rates in percentages. Panel (b) shows the contribution of each group to the 2025Q2 backcast in percentage points, as described in Sect. 3.2

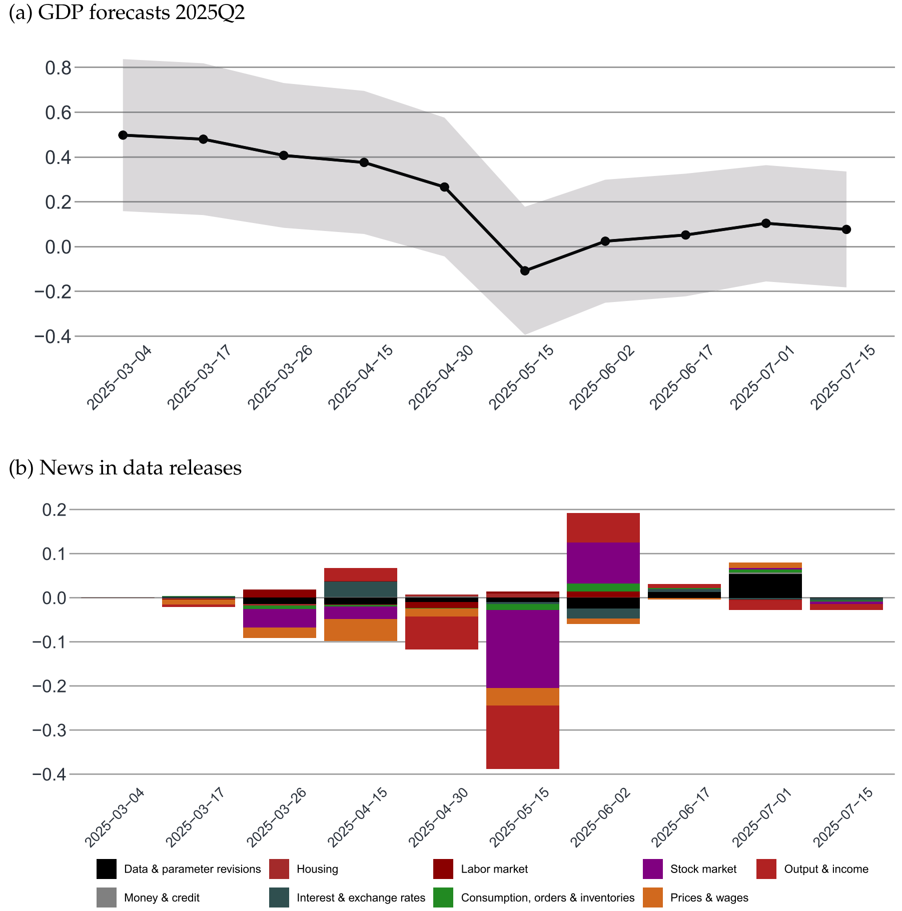

Fig. 7 Nowcast revisions and news *Note*: Panel (a) presents all GDP forecasts made between 04-03-2025 and 15-07-2025, along with the 68% confidence interval. The forecasts represent quarter-on-quarter growth rates in percentages. Panel (b) Illustrates how new data releases impact the GDP forecasts, with the sum of each bar in (b) corresponding to the change in the GDP forecasts in (a). The news concept is described in Sect. 3.3

forecasting one-quarter-ahead (panel c of Fig. 5), the models are slightly more accurate than BM measured over the entire period. In line with the outcomes in Table 2, the differences are quite small, and are mainly caused by deviating forecast performance during the COVID crisis.

Interestingly, the forecast accuracy of the Consensus forecasts was declining compared to BM in the period leading up to the crisis, for both the backcasts (panel a) and the nowcasts (panel b). Note that the scale of this decline is significantly smaller than that shown in the left-hand graphs. Additionally, there was a noticeable deterioration in the forecast accuracy of the dynamic factor model by Bańbura and Rünstler (2011) compared to BM, starting in the second half of 2015 and continuing right up to the COVID crisis. This trend is observed in backcasts, nowcasts, as well as one-quarter-ahead forecasts.

7 Nowcasting in Practice

In this Section we illustrate how we use the nowcasting model described in this paper in practice. The model is updated bi-weekly, providing policymakers with timely GDP estimates that are used in making economic projections. We take the nowcast episode for 2025Q2 as an example.

Figure 6a shows the backcast for 2025Q2 and the nowcast for 2025Q3, as estimated on 15 July 2025. The shaded area indicates the 68% confidence interval, based on the smoothed state variance obtained from the Kalman smoother. Policymakers typically require more than a point estimate and a confidence interval to interpret economic developments. Therefore, we compute the contribution of each variable to the nowcast, as described in Sect. 3.2. Since the data are divided into eight groups (see Sect. 4), we sum the contributions of series within a group.

Figure 6b shows the contribution of each group to the GDP backcast for 2025Q2. Most groups contribute negatively, resulting in a GDP forecasts below its long-term average. In particular, the output and income group exhibits a substantial negative contribution. Although some individual variables contribute positively, as shown in Appendix D, the negative contributions dominate resulting in a GDP forecasts of 0.08% for 2025Q2.

The contributions provide a breakdown of the nowcast at a certain point in time. However, policymakers monitor new data releases and adjust their expectations for the GDP forecasts. We use the procedure outlined in Sect. 3.3 to extract the news from data releases that lead to revisions in the GDP nowcast.

Figure 7a show the revisions made to the 2025Q2 GDP forecasts, from the initial estimate on 04-03-2025 to the final estimate on 15-07-2025. The sum of each bar in Fig. 7b corresponds to the change in the GDP forecasts shown in Fig. 7a. This figure illustrates how turbulence in the stock market and negative surprises in output and income variables for April pushed GDP forecasts into negative territory, followed by a partial rebound after the release of May data. Ultimately, the realized GDP growth rate for 2025Q2 was 0.2%, well within the confidence interval.

Note that Fig. 7b includes an additional group labeled *Data & parameter revisions*. This residual group accounts for changes in GDP forecasts resulting from revisions in published data releases. Additionally, since the model is re-estimated at each point in time, GDP forecasts are also influenced by parameter revisions. Further research could investigate modeling data and parameter revisions similar to Anesti et al. (2022).

8 Conclusion

This paper provides a comprehensive guide for practitioners on designing a nowcasting model for GDP growth. We demonstrate how to compile a novel database that includes a broad range of economic indicators. Using a dynamic factor model proposed by Bańbura and Modugno (2014), we conduct a forecast competition and assess the forecast errors resulting from various modeling choices. We find that a multi-factor model with a parsimonious lag structure, estimated on a medium sized dataset including survey data, delivers the most accurate forecasts for Dutch GDP growth. Moreover, our model outperforms the Consensus Forecasts in the non-COVID period.

We illustrate how to use our nowcasting model in practice, offering policymakers valuable insights beyond point estimates of GDP growth. The methodologies outlined in this paper can easily be adopted by other central banks to design nowcasting models tailored to their economies.

Appendix

State Space Representation of Mixed Frequency DFM

A notable challenge for nowcasting models is the discrepancy between quarterly GDP growth figures and the monthly frequency of macroeconomic time series. Restricting our dataset to quarterly variables would render the model inadequate for real-time forecasting. Real-time forecasting requires regularly updated information, much of which is received monthly. Fortunately, mixed frequency datasets can be integrated into the factor model by treating lower frequency series as high-frequency indicators with missing data. As a result, information from indicators collected at a lower frequency (mainly GDP) can still be used to estimate the factors. The model can also be employed to forecast the lower frequency series or enhance their interpolation. To achieve this, we represent the quarterly variables in our model as partially observed monthly variables. Initially, we use the quarterly GDP in volume terms to define its monthly counterparts as follows:

$$GDP_t^Q = GDP_t^M + GDP_{t-1}^M + GDP_{t-2}^M \quad \text{for } t = 3, 6, 9, \ldots \tag{A.1}$$

Additionally, we specify the transformations:

$$Y_t^Q = 100 \times \log\left(GDP_t^Q\right) \tag{A.2}$$

$$Y_t^M = 100 \times \log\left(GDP_t^M\right) \tag{A.3}$$

Using Eqs. A.1 and A.3, we can define the monthly and quarterly GDP growth rates as follows:

$$y_t^Q = Y_t^Q - Y_{t-3}^Q \tag{A.4}$$

$$y_t^M = Y_t^M - Y_{t-1}^Q \tag{A.5}$$

Next, we define the monthly representation of the quarterly growth rate ($\bar{y}_t^Q$) as:

$$\bar{y}_t^Q = \begin{cases} Y_t^Q - Y_{t-3}^Q & \text{if } t = 3, 6, 9, \ldots \\ \text{NA} & \text{otherwise} \end{cases} \tag{A.6}$$

In $\bar{y}_t^Q$ the quarterly growth rates are in fact are 'assigned' to the third month of each quarter, following the usual convention in the nowcasting literature. To bridge $\bar{y}_t^Q$ to y_t^Q we follow the approximation developed by Mariano and Murasawa (2003), i.e.:

$$\bar{y}_t^Q = Y_t^Q - Y_{t-3}^Q = (Y_t^M + Y_{t-1}^M + Y_{t-2}^M) - (Y_{t-3}^M + Y_{t-4}^M + Y_{t-5}^M) \quad \text{(A.7)}$$

$$= y_t^M + 2y_{t-1}^M + 3y_{t-2}^M + 2y_{t-3}^M + y_{t-4}^M \quad \text{(A.8)}$$

where y_t^M and $\bar{y}_t^Q$ denote the $n_M \times 1$ and $n_Q \times 1$ vectors of monthly and quarterly data, respectively. Further, let Λ_M and Λ_Q denote the corresponding factor loadings for the monthly y_t^M, and for the unobserved monthly growth rates of the quarterly y_t^Q, respectively. We can then cast the model in the following state space representation. The measurement equation is defined as:

$$\begin{bmatrix} y_t^M \\ \bar{y}_t^Q \end{bmatrix} = \begin{bmatrix} \Lambda_M & 0 & 0 & 0 & 0 & I_n & 0 & 0 & 0 & 0 & 0 \\ \Lambda_Q & 2\Lambda_Q & 3\Lambda_Q & 2\Lambda_Q & \Lambda_Q & 0 & I_{n_Q} & 2I_{n_Q} & 3I_{n_Q} & 2I_{n_Q} & I_{n_Q} \end{bmatrix} \begin{bmatrix} f_t \\ f_{t-1} \\ f_{t-2} \\ f_{t-3} \\ f_{t-4} \\ \varepsilon_t^M \\ \varepsilon_t^Q \\ \varepsilon_{t-1}^Q \\ \varepsilon_{t-2}^Q \\ \varepsilon_{t-3}^Q \\ \varepsilon_{t-4}^Q \end{bmatrix} + \begin{bmatrix} \xi_t^M \\ \xi_t^Q \end{bmatrix} \quad \text{(A.9)}$$

and the transition equation:

$$\begin{bmatrix} f_t \\ f_{t-1} \\ f_{t-2} \\ f_{t-3} \\ f_{t-4} \\ \tilde{\epsilon}_t^M \\ \tilde{\epsilon}_t^Q \\ \tilde{\epsilon}_{t-1}^Q \\ \tilde{\epsilon}_{t-2}^Q \\ \tilde{\epsilon}_{t-3}^Q \\ \tilde{\epsilon}_{t-4}^Q \end{bmatrix} = \begin{bmatrix} A_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ I_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho_M & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \rho_Q & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & I_{n_Q} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I_{n_Q} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & I_{n_Q} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & I_{n_Q} & 0 \end{bmatrix} \begin{bmatrix} f_{t-1} \\ f_{t-2} \\ f_{t-3} \\ f_{t-4} \\ f_{t-5} \\ \tilde{\epsilon}_{t-1}^M \\ \tilde{\epsilon}_{t-1}^Q \\ \tilde{\epsilon}_{t-2}^Q \\ \tilde{\epsilon}_{t-3}^Q \\ \tilde{\epsilon}_{t-4}^Q \\ \tilde{\epsilon}_{t-5}^Q \end{bmatrix} + \begin{bmatrix} u_t \\ 0 \\ 0 \\ 0 \\ 0 \\ e_t^M \\ e_t^Q \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Here, $\rho_M = \text{diag}(\rho_{M,1}, \ldots, \rho_{M,n_M})$ and $\rho_Q = \text{diag}(\rho_{Q,1}, \ldots, \rho_{Q,n_Q})$ collect the AR(1) coefficients of the idiosyncratic component of monthly and quarterly data. As described above ξ_t^M and ξ_t^Q have fixed and small variances.[19]

The state-space form allows inference using the Kalman filter and smoother. The state space framework also provides a convenient framework for handling the irregularities of the data in real time (i.e., mixed frequencies and non-synchronicity of the data releases) and updating the forecasts. The Kalman filter processes incoming data

[19] Notice that the size of the state-vector quickly expands when the VAR order increases. Notice that this expansion comes at exponentially increasing computational time.

in a clear and intuitive manner. It updates model forecasts recursively by weighting the innovation components of new data based on their timeliness and quality. Additionally, because the model generates forecasts for all variables simultaneously, analyzing the data flow does not require combining multiple unrelated models.

Model Estimation

We estimate the dynamic factor model using an algorithm developed by Bańbura and Modugno (2014), which is a modification of the original Expectation Maximization (EM) algorithm first introduced by Dempster et al. (1977). Since the factors f_t are unobserved, the maximum likelihood estimators of the model parameters do not have a closed form solution. The idea of the EM algorithm is to first compute the expectation of the (log) likelihood conditional on the data and a set of parameters (from the previous iteration or from the initialization). The second step is to maximize the expected (log) likelihood by re-estimating the parameters. This two-step procedure continues until convergence.

The initial values for parameters are obtained similarly to Giannone et al. (2008). First, $F = [f_t, ..., f_T]$ and Λ from Eq. 1 are estimated using principal components derived from the covariance matrix of the standardized data $Y = [y_t, ..., y_T]$. Next, we initialize the coefficients $A = [A_1, ..., A_p]$ and the covariance matrix Q from Eq. 2 by applying a VAR to $\hat{F}$. Finally, the parameters from the AR(1) process of the idiosyncratic component ρ_i and σ_i^2 are initialized using $\hat{\epsilon}_t = y_t - \hat{\Lambda}\hat{f}_t$.

Given these initializations, the Kalman smoother provides the expected log likelihood for the first step of the EM algorithm. In the second step, EM algorithm updates the parameters such that the expected log likelihood is maximized. For the algorithm to converge, we use the following metric as stopping criterion:

$$c_m = \frac{l(\Omega_T; \theta(m)) - l(\Omega_T; \theta(m-1))}{\frac{1}{2}(|l(\Omega_T; \theta(m))| + |l(\Omega_T; \theta(m-1))|)} \tag{A.10}$$

where $l(\Omega_T; \theta(m))$ denotes the log-likelihood of the data Ω_T conditional on the parameter vector θ at iteration m. The algorithm is stopped at iteration M if $c_m < 10^{-5}$ or $M = 500$, whichever condition is met first.

Dataset

Additional Results

Availability of Data Within the Quarter

Figure 8a shows the average RMSFE across all models and forecasted quarters discussed in Sect. 5. Figure 8b shows the share of observations available at each forecasts, based on the publication lags of the variables. The total number of observations per quarter equals the number of monthly variables multiplied by three, plus the number of quarterly variables, all listed in Table 3. The timing of forecasts, nowcasts and backcast is also explained in Table 1.

Table 3 Overview of variables in small, medium-sized and large datasets

Nr.	Series	Group	Freq.	Dataset size			Trans.		Source	Start	Publ.lag (days)	Type
				S	M	L	Ln.	Diff.				
1	Unemployment rate: 15-75 years	Labor market	M	X	X	X	O	X	CBS	Jan-'83	17	hard
2	Collectively agreed wages: all sectors	Prices & wages	M	O	X	X	X	X	CBS	Jan-'00	7	hard
3	Bankruptcies: total	Output & income	M	O	O	X	X	X	CBS	Jan-'81	12	hard
4	Construction: value added	Output & income	M	O	X	X	X	X	CBS	Jan-'00	48	hard
5	Building permits: firms	Output & income	M	O	X	X	X	X	CBS	Jan-'90	54	hard
6	Building permits: housing	Housing	M	O	X	X	X	X	CBS	Jan-'95	54	hard
7	House price index	Housing	M	X	X	X	X	X	CBS	Jan-'95	22	hard
8	Houses sold	Housing	M	X	X	X	X	X	CBS	Jan-'95	22	hard
9	Consumption: gas	Cons., orders & inv.	M	O	X	X	X	X	CBS	Jan-'82	16	hard
10	Industrial production: manufacturing	Output & income	M	X	X	X	X	X	CBS	Jan-'95	39	hard
11	Industrial production: food, bev. & tobacco	Output & income	M	O	X	X	X	X	CBS	Jan-'95	39	hard
12	Industrial production: chemical (products)	Output & income	M	O	X	X	X	X	CBS	Jan-'95	39	hard
13	Industrial production: machinery & equipment	Output & income	M	O	X	X	X	X	CBS	Jan-'95	39	hard
14	Consumption households: domestic	Cons., orders & inv.	M	X	X	X	X	X	CBS	Jan-'95	38	hard
15	Consumer confidence: headline	Cons., orders & inv.	M	X	X	X	O	X	CBS	Apr-'86	-8	soft
16	Consumer confidence: economic sit. < 12 m.	Cons., orders & inv.	M	O	O	X	O	X	CBS	Apr-'86	-8	soft
17	Consumer confidence: economic sit. > 12 m.	Cons., orders & inv.	M	O	O	X	O	X	CBS	Apr-'86	-8	soft
18	Consumer confidence: unemployment > 12 m.	Labor market	M	O	X	X	O	X	CBS	Apr-'86	-8	soft
19	Consumer conf.: purchase of dur. goods > yr.	Cons., orders & inv.	M	O	X	X	O	X	CBS	Apr-'86	-8	soft
20	Cons. confidence: buying/build.a home > 2 yr.	Housing	M	O	O	X	O	X	CBS	Apr-'86	-8	soft
21	Import of goods	Output & income	M	X	X	X	X	X	CBS	Jan-'95	42	hard
22	Export of goods	Output & income	M	X	X	X	X	X	CBS	Jan-'95	42	hard
23	Terms of trade	Prices & wages	M	O	O	X	X	X	CBS	Jan-'95	42	hard
24	Consumer price index: headline excl. energy	Prices & wages	M	O	X	X	X	X	CBS	Jan-'96	1	hard
25	Harmonised index of consumer prices: total	Prices & wages	M	X	X	X	X	X	CBS	Jan-'96	1	hard
26	Producer prices: manufacturing	Prices & wages	M	X	X	X	X	X	CBS	Jan-'81	30	hard
27	Flights: tonnes of goods	Output & income	M	O	O	X	X	X	CBS	Jan-'99	32	hard
28	Industrial confidence: headline	Output & income	M	X	X	X	O	X	Eurostat	Jan-'80	0	soft
29	Industrial confidence: employment > 3 m.	Labor market	M	O	X	X	O	X	Eurostat	Apr-'82	0	soft
30	Industrial confidence: export order books	Cons., orders & inv.	M	O	X	X	O	X	Eurostat	Feb-'90	0	soft
31	Industrial confidence: order-book	Cons., orders & inv.	M	O	O	X	O	X	Eurostat	Jan-'80	0	soft
32	Industrial confidence: production > 3 m.	Output & income	M	O	X	X	O	X	Eurostat	Jan-'80	0	soft
33	Industrial confidence: production < 3 m.	Output & income	M	O	O	X	O	X	Eurostat	Jan-'80	0	soft
34	Industrial confidence: stocks finished products	Cons., orders & inv.	M	O	O	X	O	X	Eurostat	Jan-'80	0	soft
35	Construction confidence: headline	Housing	M	O	X	X	O	X	Eurostat	Jan-'80	0	soft
36	Construction confidence: employment > 3 m.	Labor market	M	O	O	X	O	X	Eurostat	Jan-'80	0	soft
37	Construction confidence: evolution order books	Housing	M	O	O	X	O	X	Eurostat	Jan-'80	0	soft
38	Construction confidence: building activity < 3m.	Housing	M	O	O	X	O	X	Eurostat	Jan-'80	0	soft
39	Retail confidence: stock of finished products	Cons., orders & inv.	M	O	O	X	O	X	Eurostat	Jan-'86	0	soft
40	Retail confidence: headline	Output & income	M	O	X	X	O	X	Eurostat	Jan-'86	0	soft
41	Retail confidence: activity > 3 m.	Output & income	M	O	O	X	O	X	Eurostat	Jan-'86	0	soft
42	Retail confidence: employment > 3 m.	Labor market	M	O	O	X	O	X	Eurostat	Jan-'86	0	soft
43	Retail confidence: activity < 3 m.	Output & income	M	O	O	X	O	X	Eurostat	Jan-'86	0	soft
44	Economic sentiment: headline	Output & income	M	X	X	X	O	X	Eurostat	Jan-'80	0	soft
45	Services confidence: business situation < 3 m.	Output & income	M	O	X	X	O	X	Eurostat	Apr-'93	0	soft
46	Services confidence: demand > 3 m.	Output & income	M	O	O	X	O	X	Eurostat	Jan-'96	0	soft
47	Services confidence: demand < 3 m.	Output & income	M	O	O	X	O	X	Eurostat	Apr-'93	0	soft
48	Services confidence: headline	Output & income	M	X	X	X	O	X	Eurostat	Jan-'96	0	soft
49	Services confidence: employment > 3 m.	Labor market	M	O	X	X	O	X	Eurostat	Apr-'93	0	soft
50	Services confidence: employment < 3 m.	Labor market	M	O	O	X	O	X	Eurostat	Apr-'93	0	soft
51	Unemployment rate: < 25 years	Labor market	M	O	X	X	O	X	Eurostat	Jan-'83	31	hard
52	10-year government bond yield	Interest & exch. rates	M	X	X	X	O	X	Eurostat	Apr-'86	10	hard
53	Exchange rate eur/usd	Interest & exch. rates	M	O	X	X	X	X	Eurostat	Jan-'99	2	hard
54	Nominal effective exchange rate	Interest & exch. rates	M	X	X	X	X	X	Eurostat	Jan-'99	30	hard
55	Real effective exchange rate	Interest & exch. rates	M	O	O	X	X	X	Eurostat	Jan-'99	30	hard
56	Retail sales: mail orders & online stores	Output & income	M	O	X	X	X	X	Eurostat	Jan-'00	39	hard
57	Retail sales: total	Output & income	M	X	X	X	X	X	Eurostat	Jan-'00	39	hard
58	Consumer confidence: Belgium headline	Output & income	M	O	O	X	O	X	Eurostat	Jan-'85	0	soft
59	Consumer confidence: Germany headline	Output & income	M	O	O	X	O	X	Eurostat	Jan-'85	0	soft
60	Consumer confidence: eurozone headline	Output & income	M	O	X	X	O	X	Eurostat	Jan-'85	0	soft
61	Consumer confidence: Spain headline	Output & income	M	O	O	X	O	X	Eurostat	Jun-'86	0	soft
62	Consumer confidence: France headline	Output & income	M	O	O	X	O	X	Eurostat	Jan-'85	0	soft
63	Consumer confidence: Italy headline	Output & income	M	O	O	X	O	X	Eurostat	Jan-'85	0	soft
64	Producer confidence: Belgium headline	Output & income	M	O	O	X	O	X	Eurostat	Jan-'85	0	soft
65	Producer confidence: Germany headline	Output & income	M	O	O	X	O	X	Eurostat	Jan-'80	0	soft
66	Producer confidence: eurozone headline	Output & income	M	O	O	X	O	X	Eurostat	Jan-'80	0	soft
67	Producer confidence: Spain headline	Output & income	M	O	O	X	O	X	Eurostat	Apr-'87	0	soft
68	Producer confidence: France headline	Output & income	M	O	O	X	O	X	Eurostat	Feb-'85	0	soft
69	Producer confidence: Italy headline	Output & income	M	O	O	X	O	X	Eurostat	Jan-'80	0	soft
70	Passenger car registration	Cons., orders & inv.	M	O	X	X	X	X	ECB	Jan-'90	48	hard
71	Industrial production: comp. & electronics	Output & income	M	O	O	X	X	X	ECB	Jan-'90	41	hard
72	Credit to the domestic private sector	Money & credit	M	O	X	X	X	X	ECB	Mar-'70	26	hard
73	Loans: households	Money & credit	M	X	X	X	X	X	ECB	Sep-'97	26	hard

Freq. frequency of series (M=monthly, Q= Quarterly), *Trans.* transformation of series, *Ln* take logarithm (X= yes, O= no), *Diff.* take first difference of series (X= Yes, O= No), *source* of series (*CBS*: Statistics Netherlands, *BVNL* Betaalvereniging Nederland, *EC* European Commission, *ECB* European Central Bank, *Eurostat* Eurostat, *FD* Financieele Dagblad, *Refinitiv* Refinitiv), *Start* start mont/quarter of series, *Publ. Lag* publication lag of series in days; negative publ. lag implies serie is released n days before the end of the month. *Type* soft (survey-based indicators) or hard (objective economic measures)

Table 4 Overview of variables in small, medium-sized and large datasets

Nr.	Series	Group	Freq.	Dataset size			Trans.		Source	Start	Publ.lag (days)	Type
				S	M	L	Ln.	Diff.				
74	Loans: non-financial corporations	Money & credit	M	O	X	X	X	X	ECB	Sep-'97	26	hard
75	M1	Money & credit	M	O	O	X	X	X	ECB	Jan-'80	26	hard
76	M3	Money & credit	M	X	X	X	X	X	ECB	Jan-'70	26	hard
77	Int.% new loans nfc's: < 1 y. & <= eur 1 mil.	Interest & exch. rates	M	O	X	X	O	X	ECB	Jan-'80	33	hard
78	Int. % new loans hh. house purchase: < 1 y.	Interest & exch. rates	M	O	X	X	O	X	ECB	Jan-'80	33	hard
79	Interest rate new debt securityr: 10 years	Interest & exch. rates	M	O	X	O	O	X	ECB	Jan-'93	12	hard
80	3-month interest rate	Interest & exch. rates	M	O	X	X	O	X	ECB	Oct-'72	1	hard
81	country level index of financial stress	Stock market	M	O	O	X	O	X	ECB	Jan-'70	25	hard
82	Crude oil price: brent spot free on board	Prices & wages	M	O	X	X	X	X	ECB	Feb-'99	1	hard
83	1-year euribor interest rate	Interest & exch. rates	M	O	O	X	O	X	ECB	Jan-'94	1	hard
84	Industrial production: eurozone	Output & income	M	O	X	X	X	X	ECB	Jan-'70	45	hard
85	Retail trade: eurozone	Output & income	M	O	X	X	X	X	ECB	Jan-'00	39	hard
86	Unemployment rate: eurozone	Labor market	M	O	O	X	O	X	ECB	Apr-'98	34	hard
87	Gold price	Stock market	M	O	O	X	X	X	ECB	Jan-'70	5	hard
88	composite indicator of systemic stress	Stock market	M	O	O	X	O	X	ECB	Jan-'99	1	hard
89	Dow Jones Euro Stoxx: consumer services	Stock market	M	O	X	X	X	X	ECB	Dec-'91	1	hard
90	Commodity Price index: non-energy comm.	Prices & wages	M	X	X	X	X	X	ECB	Jan-'96	5	hard
91	PMI manufacturing: eurozone headline	Output & income	M	O	O	X	O	X	Refinitiv	Jun-'97	3	soft
92	PMI services: eurozone headline	Output & income	M	O	O	X	O	X	Refinitiv	Jul-'98	5	soft
93	PMI manufacturing: headline	Output & income	M	X	X	X	O	X	Refinitiv	Mar-'00	1	soft
94	PMI manufacturing: production	Output & income	M	O	O	X	O	X	Refinitiv	Mar-'00	1	soft
95	PMI manufacturing: new orders	Cons., orders & inv.	M	X	X	X	O	X	Refinitiv	Mar-'00	1	soft
96	PMI manufacturing: employment	Labor market	M	O	O	X	O	X	Refinitiv	Mar-'00	1	soft
97	AEX-midkap	Stock market	M	O	O	X	X	X	Refinitiv	Jan-'83	1	hard
98	Dow Jones Euro Stoxx: 50	Stock market	M	O	X	X	X	X	Refinitiv	Jan-'87	1	hard
99	Dow Jones Euro Stoxx: basic materials	Stock market	M	O	O	X	X	X	Refinitiv	Jan-'87	1	hard
100	Dow Jones Euro Stoxx: technology	Stock market	M	O	O	X	X	X	Refinitiv	Jan-'87	1	hard
101	Dow Jones Euro Stoxx: industrials	Stock market	M	O	O	X	X	X	Refinitiv	Jan-'87	1	hard
102	AEX	Stock market	M	X	X	X	X	X	Refinitiv	Jan-'83	1	hard
103	PMI composite: world (headline)	Output & income	M	X	X	X	O	X	Refinitiv	Jul-'98	5	soft
104	World steel production	Output & income	M	O	X	X	X	X	Refinitiv	Jan-'80	21	hard
105	Baltic exchange dry index	Prices & wages	M	O	O	X	X	X	Refinitiv	May-'85	1	hard
106	VIX Europe	Stock market	M	O	O	X	X	X	Refinitiv	Feb-'99	1	hard
107	Serv. conf.: ev. of de. < 3 m.; pc. progr. & conslt.	Output & income	M	O	X	X	O	X	EC	Apr-'97	0	soft
108	Serv. conf.: ev. of de. < 3 m.: conslt. & head off.	Output & income	M	O	X	X	O	X	EC	Apr-'97	0	soft
109	Serv. conf.: ev. of de. < 3 m.: empl. activities	Output & income	M	O	X	X	O	X	EC	Apr-'97	0	soft
110	Global supply chain pressure index	Output & income	M	O	O	X	O	X	Refinitiv	Jan-'98	4	soft
111	Debit card payments: total	Cons., orders & inv.	M	O	X	X	X	X	BVNL	Jan-'95	10	hard
112	Debit card payments: services	Cons., orders & inv.	M	O	O	X	X	X	BVNL	Jan-'98	10	hard
113	US high yield bond spread	Stock market	M	O	O	X	X	X	Refinitiv	Jan-'97	0	hard
114	Eurozone high yield bond spread	Stock market	M	O	X	X	X	X	Refinitiv	Jan-'98	0	hard
115	FD: headline	Output & income	M	X	O	O	O	O	FD	Jan-'85	0	soft
116	FD: financial markets (FM)	Stock market	M	O	X	O	O	O	FD	Jan-'85	0	soft
117	FD: companies (COM)	Output & income	M	O	X	O	O	O	FD	Jan-'85	0	soft
118	FD: economics (ECO)	Output & income	M	O	X	O	O	O	FD	Jan-'85	0	soft
119	FD: politics (POL)	Output & income	M	O	X	O	O	O	FD	Jan-'85	0	soft
120	FD FM: financial markets	Stock market	M	O	O	X	O	O	FD	Jan-'85	0	soft
121	FD FM: financials	Stock market	M	O	O	X	O	O	FD	Jan-'85	0	soft
122	FD FM: news	Stock market	M	O	O	X	O	O	FD	Jan-'85	0	soft
123	FD FM: financial indices	Stock market	M	O	O	X	O	O	FD	Jan-'85	0	soft
124	FD COM: infrastructure	Output & income	M	O	O	X	O	O	FD	Jan-'85	0	soft
125	FD COM: multinationals	Output & income	M	O	O	X	O	O	FD	Jan-'85	0	soft
126	FD COM: construction & energy	Output & income	M	O	O	X	O	O	FD	Jan-'85	0	soft
127	FD COM: demography	Output & income	M	O	O	X	O	O	FD	Jan-'85	0	soft
128	FD ECO: elections	Output & income	M	O	O	X	O	O	FD	Jan-'85	0	soft
129	FD ECO: indicators	Output & income	M	O	O	X	O	O	FD	Jan-'85	0	soft
130	FD ECO: raw materials	Output & income	M	O	O	X	O	O	FD	Jan-'85	0	soft
131	FD ECO: EU	Output & income	M	O	O	X	O	O	FD	Jan-'85	0	soft
132	FD POL: parliament	Output & income	M	O	O	X	O	O	FD	Jan-'85	0	soft
133	FD POL: national	Output & income	M	O	O	X	O	O	FD	Jan-'85	0	soft
134	FD POL: lower government	Output & income	M	O	O	X	O	O	FD	Jan-'85	0	soft
135	FD POL: social partners	Output & income	M	O	O	X	O	O	FD	Jan-'85	0	soft
136	Consumption	Cons., orders & inv.	Q	O	X	X	X	X	CBS	Mar-'96	45	hard
137	Gross private investment	Output & income	Q	O	X	X	X	X	CBS	Mar-'96	45	hard
138	Exports	Output & income	Q	O	X	X	X	X	CBS	Mar-'96	45	hard
139	Imports	Output & income	Q	O	X	X	X	X	CBS	Mar-'96	45	hard
140	Hours worked	Labor market	Q	X	X	X	X	X	CBS	Mar-'95	45	hard
141	Labour productivity	Labor market	Q	O	X	X	X	X	Eurostat	Mar-'96	45	hard
142	Capacity utilization	Output & income	Q	X	X	X	O	X	Eurostat	Mar-'80	-3	soft
143	Gross domestic product: eurozone	Output & income	Q	X	X	X	X	X	Eurostat	Mar-'95	45	hard
144	Gross domestic product	Output & income	Q	X	X	X	X	X	CBS	Mar-'96	45	hard
	N (monthly)			24	67	129						
	N (quarterly)			4	9	9						
	N (total)			28	76	138						

Freq. frequency of series (M=monthly, Q= Quarterly), *Trans.* transformation of series, *Ln* take logarithm (X= yes, O= no), *Diff.* take first difference of series (X= yes, O= NO), *source* of series (*CBS* Statistics Netherlands, *BVNL* Betaalvereniging Nederland, *EC* European Commission, *ECB* European Central Bank, Eurostat: Eurostat, *FD* Financieele Dagblad, *Refinitiv* Refinitiv), *Start* start mont/quarter of series, *Publ. Lag* publication lag of series in days; negative publ. lag implies serie is released n days before the end of the month. *Type* soft (survey-based indicators) or hard (objective economic measures)

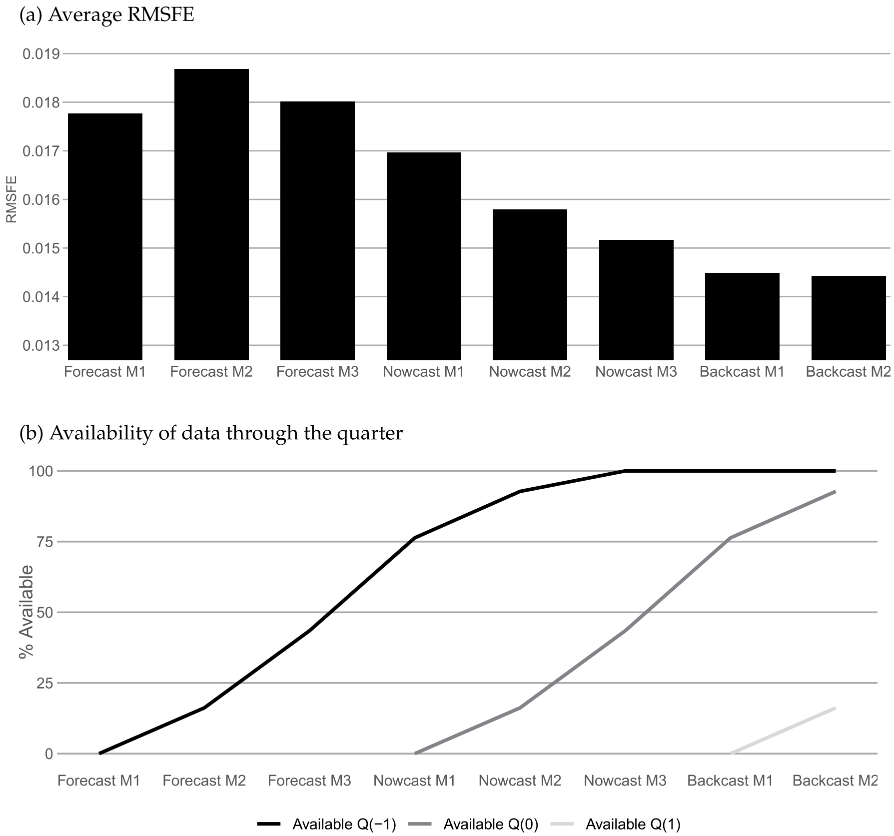

Fig. 8 RMSFE and data availability *Note:* Panel (a) reports the average Root Mean Squared Forecast Error across all DFM model specifications evaluated in Sect. 5 in the period 2013Q3-2023Q3, broken down by forecasting horizon as defined in Table 1. Panel (b) Shows the cumulative percentage of observations that become available over time, based on the publication lag of each variable reported in Table 3. Q(-1), Q(0), and Q(1) denote the previous, current, and next quarter, respectively. For instance, by the third month of the current quarter (Nowcast M3), 50% of its observations are available, while all observations from the previous quarter have been released

Overall, we observe that forecast accuracy improves as more information becomes available.[20] Given that forecasts are produced at the first day of the month, information regarding a quarter starts to appear from the second month onward. Moreover, by the third month, approximately 50% of the observations are available for the current quarter and all observations for the previous quarter. Note that Q(0) never reaches full availability, even when backcasting in the second month after the quarter ends, because some variables are released alongside GDP and thus cannot be used when forecasting the current quarter.

[20] The exception to the monotonic improvement in RMSFE is Forecast M1, which is based on the full sample. However, this exception disappears when the COVID period is excluded.

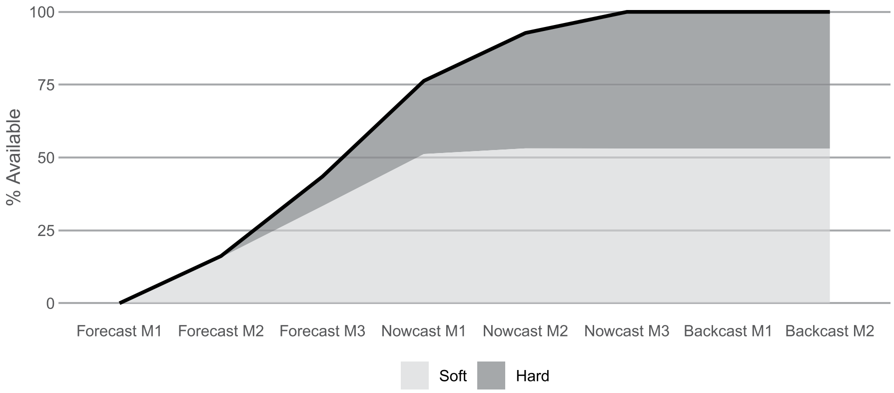

Fig. 9 Availability Soft and Hard data Q(-1)

Fig. 9 illustrates the availability of different types of observations for the previous quarter Q(-1) per forecasting horizon. Soft data refers to survey-based indicators, while hard data consists of objective economic measures. Table 3 classifies each variable as either soft or hard. We observe that soft data are typically released earlier, whereas hard data tend to have longer publication lags. It remains difficult to disentangle whether the reduction in RMSFE is primarily driven by the timely release of soft indicators or the delayed availability of hard indicators. Bańbura and Rünstler (2011) address this issue and conclude that both types of indicators are important, once publication lags are appropriately considered.

Results Pre-COVID Period: 2013Q3 – 2019Q4

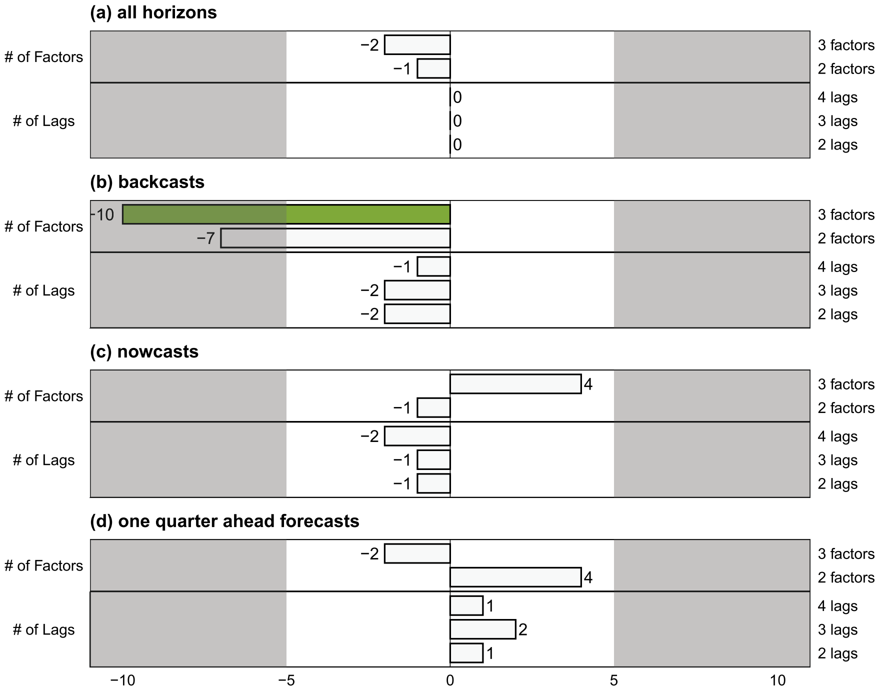

Fig. 10 Impact of model-specification on forecast accuracy: pre-COVID period *Note:* Impact of features on out-of sample mean squared forecast, in percent

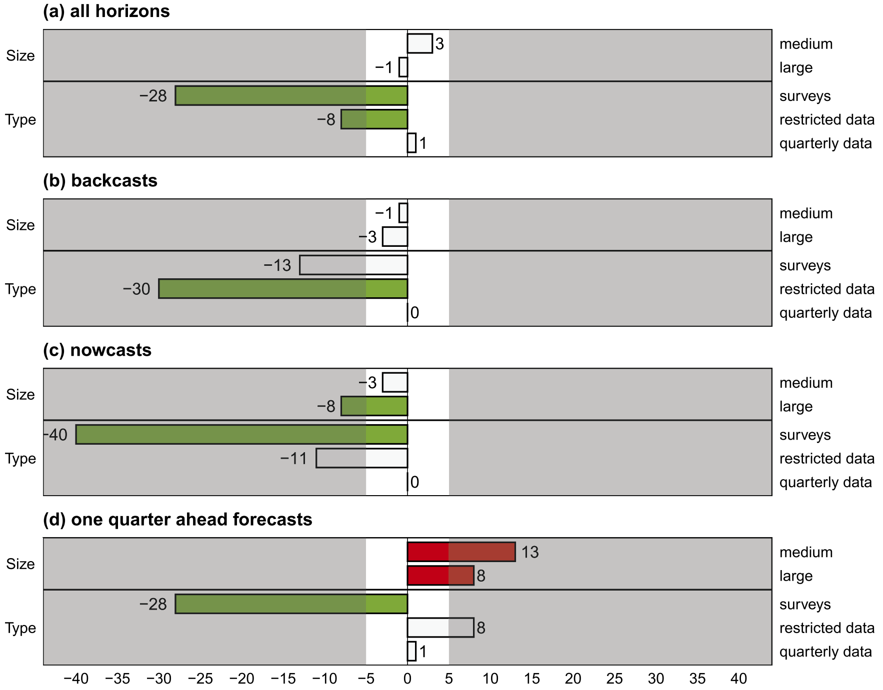

Fig. 11 Impact of choice of data on forecast accuracy: pre-COVID period *Note:* Impact of features on out-of-sample mean squared forecast, in percent

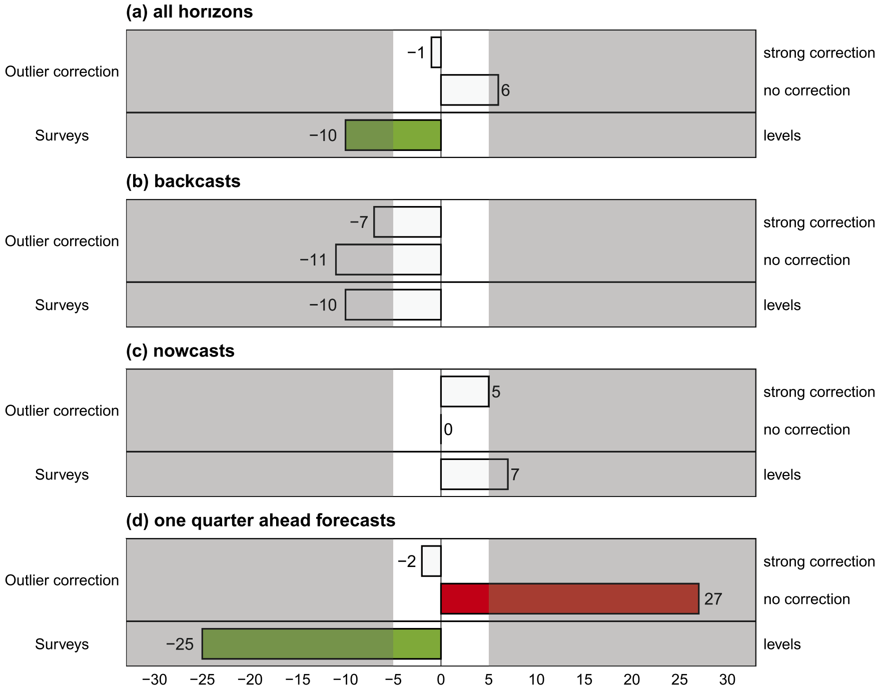

Fig. 12 Impact of estimation & transformation on forecast accuracy: pre-COVID period *Note:* Impact of features on out-of sample mean squared forecast, in percent

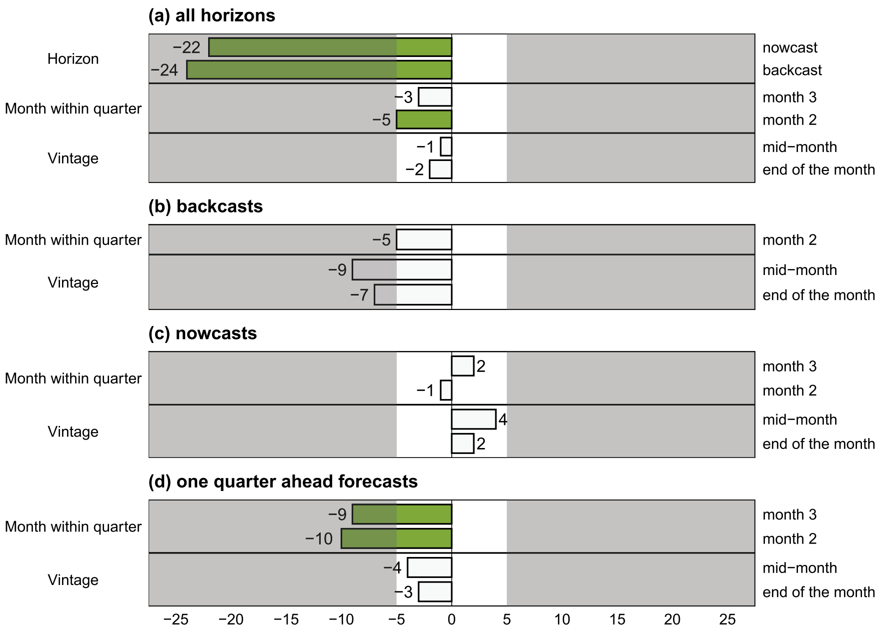

Fig. 13 Impact of timing on forecast accuracy: pre-COVID period *Note:* Impact of features on out-of sample mean squared forecast, in percent

Results No-COVID Period: 2013Q3 – 2019Q4 & 2020Q4-2023Q3

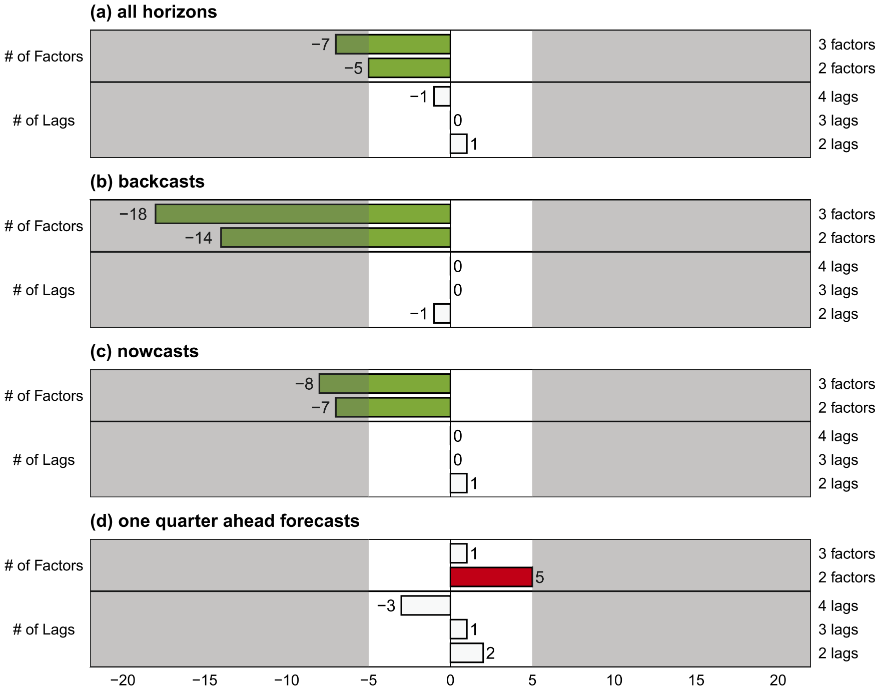

Fig. 14 Impact of model-specification on forecast accuracy: no-COVID period *Note:* Impact of features on out-of sample mean squared forecast, in percent

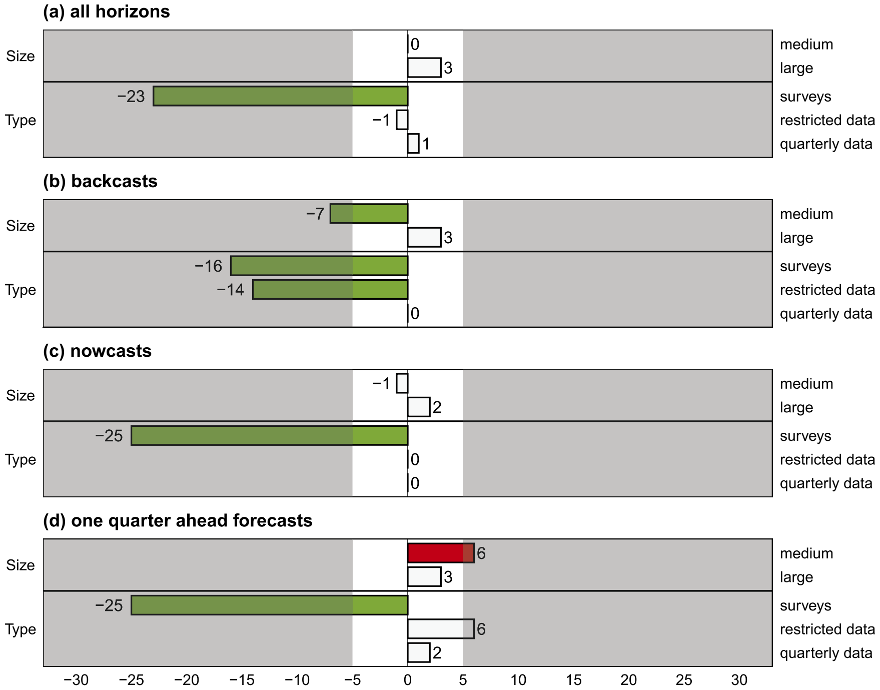

Fig. 15 Impact of choice of data on forecast accuracy: no-COVID period *Note:* Impact of features on out-of-sample mean squared forecast, in percent

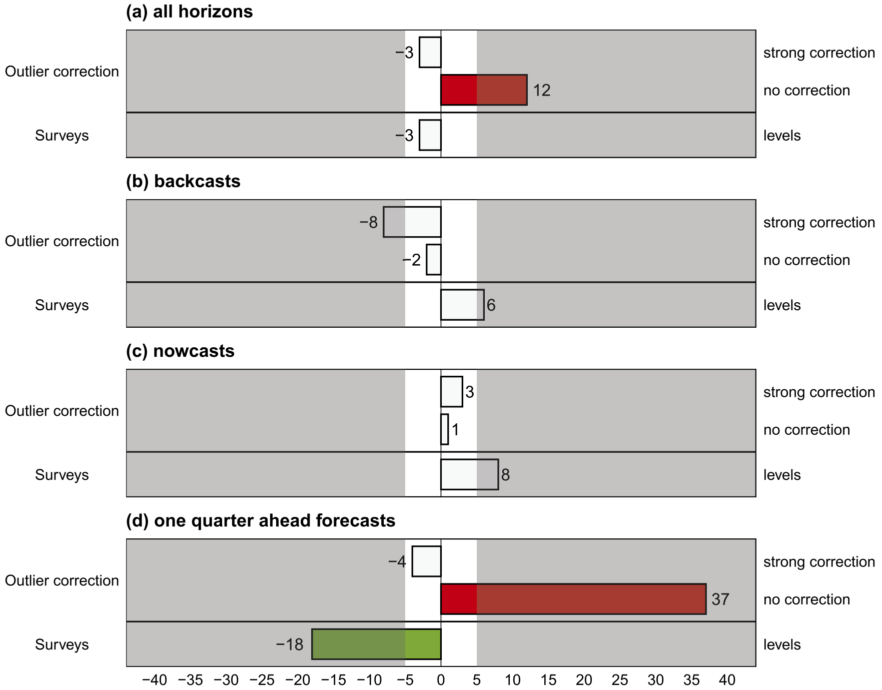

Fig. 16 Impact of estimation & transformation on forecast accuracy: no-COVID period *Note:* Impact of features on out-of sample mean squared forecast, in percent

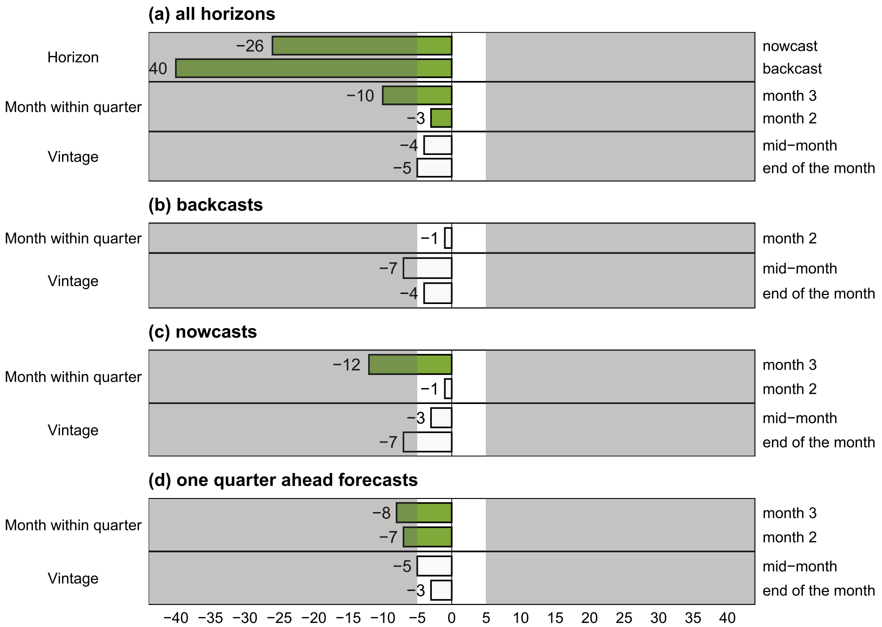

Fig. 17 Impact of timing on forecast accuracy: no-COVID period *Note:* Impact of features on out-of sample mean squared forecast, in percent

Contributions of Individual Variables

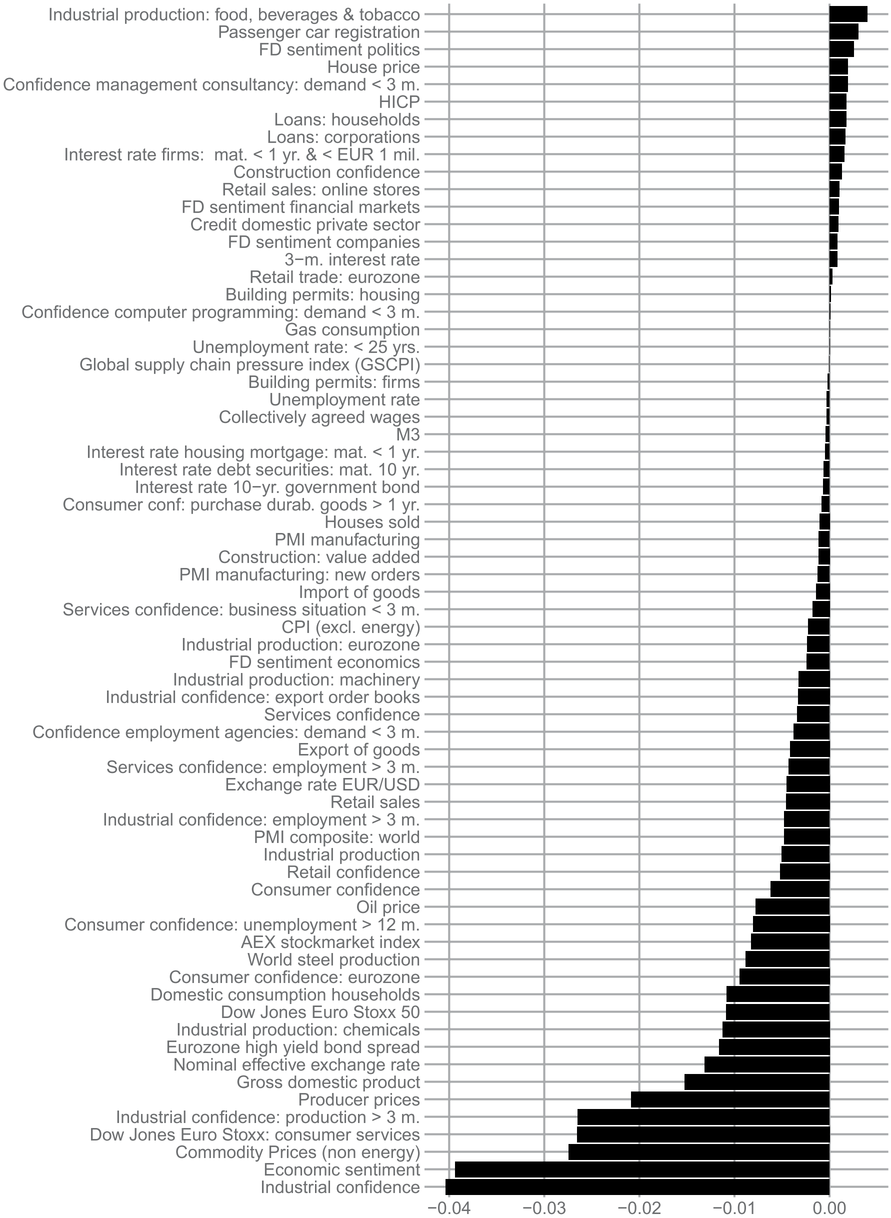

Fig. 18 Contributions of individual variables *Note*: Contributions of all variables to the 2025Q2 backcast made at 15-07-2025 displayed in Fig. 6a

Declarations

Conflict of interest The authors have no financial or non-financial interests to disclose. The paper fully complies with the ethical standards of the journal.

References

Almuzara, M., Baker, K., O'Keeffe, H., & Sbordone, A. (2023). The new york fed staff nowcast 2.0. Staff nowcast technical paper, New York Fed. https://www.newyorkfed.org/medialibrary/media/research/blog/2023/NYFed-Staff-Nowcast_technical-paperlink.

Alvarez, R., & Perez-Quiros, G. (2016). Aggregate versus disaggregate information in dynamic factor models. *International Journal of Forecasting, 32*, 680–694. https://doi.org/10.1016/j.ijforecast.2015.10.006. link.

Anesti, N., Galvão, A. B., & Miranda-Agrippino, S. (2022). Uncertain Kingdom: Nowcasting gross domestic product and its revisions. *Journal of Applied Econometrics, 37*(1), 42–62. https://doi.org/10.1002/jae.2845. link.

Aprigliano, V., Ardizzi, G. & Monteforte, L. (2019). Using payment system data to forecast economic activity. *International Journal of Central Banking 60*, 55–80. https://www.ijcb.org/journal/ijcb19q4a2.pdflink.

Babii, A., Ghysels, E., & Striaukas, J. (2022). Machine learning time series regressions with an application to nowcasting. *Journal of Business & Economic Statistics, 40*, 1094–1106. https://doi.org/10.1080/07350015.2021.1899933. link.

Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica, 70*(1), 191–221. https://doi.org/10.1111/1468-0262.00273. link.

Bańbura, M., Giannone, D., Modugno,M., & Reichlin, L. (2013). Now-casting and the real-time data flow. In G. Elliott and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 2, Part A, pp. 195–237. Elsevier. https://doi.org/10.1016/B978-0-444-53683-9.00004-9link.

Bańbura, M., Giannone,D., Reichlin, L. (2011). Nowcasting. In M. P. Clements and D. F. Hendry (Eds.), *Oxford Handbook on Economic Forecasting*, pp. 63–90. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195398649.013.0008link.

Bańbura, M., & Modugno, M. (2014). Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data. *Journal of Applied Econometrics, 29*(1), 133–160. https://doi.org/10.1002/jae.2306. link.

Bańbura, M., & Rünstler, G. (2011). A look into the factor model black box: Publication lags and the role of hard and soft data in forecasting GDP. *International Journal of Forecasting, 27*(2), 333–346. https://doi.org/10.1016/j.ijforecast.2010.01.011. link.

Barigozzi, M., & Cho, H. (2020). Consistent estimation of high-dimensional factor models when the factor number is over-estimated. *Electronic Journal of Statistics, 14*(2), 2892–2921. https://doi.org/10.1214/20-EJS1741. link.

Barigozzi, M. & Luciani,M. (2024). Quasi maximum likelihood estimation and inference of large approximate dynamic factor models via the EM algorithm. Working paper 1910.03821, arXiv. https://doi.org/10.48550/arXiv.1910.03821 link.

Bernanke, B. S., & Boivin, J. (2003). Monetary policy in a data-rich environment. *Journal of Monetary Economics, 50*(3), 525–546. https://doi.org/10.1016/S0304-3932(03)00024-2. link.

Boivin, J., & Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics, 1*, 169–194. https://doi.org/10.1016/j.jeconom.2005.01.027. link.

Buckmann, M., Joseph, A., & Robertson, H. (2023). An interpretable machine learning workflow with an application to economic forecasting. *International Journal of Central Banking, 19*(4), 449–552. https://www.ijcb.org/journal/ijcb23q4a10.pdf.

Burns, A. F. & Mitchell, W. C. (1946). *Measuring Business Cycles*. NBER. https://www.nber.org/books-and-chapters/measuring-business-cycleslink.

Caggiano, C., Kapetanios, G., & Labhard, V. (2011). Are more data always better for factor analysis? Results for the euro area, the six largest euro area countries and the UK. *Journal of Forecasting, 30*, 736–753. https://doi.org/10.1002/for.1208. link.

Carriero, A., Galvão, A. B., & Kapetanios, G. (2019). A comprehensive evaluation of macroeconomic forecasting methods. *International Journal of Forecasting, 35*, 1226–1239. https://doi.org/10.1016/j.ijforecast.2019.02.007. link.

Catell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10. link.

Coulombe, P. G., Leroux, M., Stevanovic, D., & Suprenant, S. (2022). How is machine learning useful for macroeconomic forecasting? *Journal of Applied Econometrics, 37*, 920–964. https://doi.org/10.1002/jae.2910. link.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B B (Statistical Methodology), 39*(1), 1–22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x. link.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics, 13*(3), 134–144. https://doi.org/10.1198/073500102753410444. link.

Doz, C., Giannone, D., & Reichlin, L. (2012). A quasi-maximum likelihood approach for large, approximate dynamic factor models. *The Review of Economics and Statistics, 94*(4), 1014–1024. https://doi.org/10.1162/REST_a_00225. link.

Giannone, D., Reichlin, L., & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics, 55*(4), 665–676. https://doi.org/10.1016/j.jmoneco.2008.05.010. link.

Havrlant, D., Tóth, P., & Wörz, J. (2016). On the optimal number of indicators – nowcasting GDP growth in CESEE. Focus on European Economic Integration 4, Oesterreichische Nationalbank. https://www.oenb.at/dam/jcr:1b28450b-edab-4f2a-a71a-c521a022e52f/feei_2016_q4_studies_havrlant_tlink.

Jansen, W. J., & de Winter, J. M. (2018). Combining model-based near-term GDP forecasts and judgmental forecasts: A real-time exercise for the G7 countries. *Oxford Bulletin of Economics and Statistics, 80*(6), 1213–1242. https://doi.org/10.1111/obes.12250. link.

Kant, D., Pick, A., & de Winter, J. M. (2025). Nowcasting GDP using machine learning methods. *AStA Advances in Statistical Analysis, 109*, 1–24. https://doi.org/10.1007/s10182-024-00515-0. link.

Koopman, S. J., & Harvey, A. (2003). Computing observation weights for signal extraction and filtering. *Journal of Economic Dynamics and Control, 27*(7), 1317–1333. https://doi.org/10.1016/S0165-1889(02)00061-1. link.

Liebermann, J. (2014). Real-time nowcasting of gdp: A factor model vs. professional forecasters. *Oxford Bulletin of Economics and Statistics, 76*(6), 783–811. https://doi.org/10.1111/obes.12047. link.

Linzenich, J. & Meunier, B. (2024). Nowcasting made easier: a toolbox for economists. Working Paper Series 3004, European Central Bank. https://www.ecb.europa.eu/pub/pdf/scpwps/ecb.wp3004~3ce9d0d8ca.en.pdflink.

Luciani, M., & Ricci, L. (2014). Nowcasting Norway. *International Journal of Central Banking, 37*, 215–248. https://www.ijcb.org/journal/ijcb14q4a7.pdf.

Lundquist, K., & Stekler, H. (2012). Interpreting the performance of business economists during the great recession. *Business Economics, 47*, 148–154. https://doi.org/10.1057/be.2012.2. link.

Mariano, R., & Murasawa, Y. (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics, 18*(4), 427–443. https://doi.org/10.1002/jae.695. link.

McCracken, M., & Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics, 34*(3), 574–589. https://doi.org/10.1080/07350015.2015.1086655. link.

Miranda, K., Poncela, P., & Ruiz, E. (2022). Dynamic factor models: Does the specification matter? *SERIEs Journal of the Spanish Economic Association, 13*(1), 397–428. https://doi.org/10.1007/s13209-021-00248-2. link.

Schumacher, C., & Breitung, J. (2008). Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data. *International Journal of Forecasting, 24*(3), 386–98. https://doi.org/10.1016/j.ijforecast.2008.03.008. link.

Shapiro, A. H., Sudhof, M., & Wilson, D. J. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association, 97*, 1167–1179. https://doi.org/10.1198/016214502388618960. link.
Sims, C. A. (2002). The role of models and probabilities in the monetary policy process. *Brookings Papers on Economic Activities, 2002*(2), 1–40. https://doi.org/10.1353/eca.2003.0009. link.
Stundziene, A., Pilinkiene, V., Bruneckiene, J., Grybauskas, A., Lukauskas, M., & Pekarskiene, I. (2024). Future directions in nowcasting economic activity: A systematic literature review. *Journal of Economic Surveys, 38*, 1199–1233. https://doi.org/10.1111/joes.12579. link.
Van Dijk, D. & De Winter J.(2023). Nowcasting GDP using tone-adjusted time varying news topics: Evidence from the financial press. Working Paper 766, De Nederlandsche Bank. https://ideas.repec.org/p/dnb/dnbwpp/766.htmllink.