Nowcasting and aggregation: Why small Euro area countries matter

Andrii Babii
Department of Economics, University of North Carolina – Chapel Hill
and
Luca Barbaglia
European Commission, Joint Research Centre (JRC)
and
Eric Ghysels
Department of Economics and Kenan-Flagler Business School, University of North Carolina – Chapel Hill
and
Jonas Striaukas
Department of Finance, Copenhagen Business School

July 26, 2024

Abstract

The paper studies the nowcasting of Euro area Gross Domestic Product (GDP) growth using mixed data sampling machine learning panel data regressions with both standard macro releases and daily news data. Using a panel of 19 Euro area countries, we investigate whether directly nowcasting the Euro area aggregate is better than weighted individual country nowcasts. Our results highlight the importance of the information from small- and medium-sized countries, particularly when including the COVID-19 pandemic period. The analysis is supplemented by studying the so-called Big Four—France, Germany, Italy, and Spain—and the value added of news data when official statistics are lagging.

Keywords: hierarchical nowcasting, high-dimensional panels, mixed-frequency data, text data

1 Introduction

A much-researched example of nowcasting is that of US real GDP growth. Traditional methods rely on dynamic factor models which treat quarterly GDP growth as a latent process and use so-called standard (monthly) macroeconomic data releases to obtain within-quarter estimates. There are several limitations to this approach. First, in the era of big data, it is challenging to expand dynamic factor models to include many high-frequency non-traditional data increasingly used by macroeconomists to gauge the state of the economy. Second, going beyond the nowcasting of a single series, and looking at the GDP growth for many countries modeled simultaneously or nowcasting earnings of many firms, which has many similarities with nowcasting GDP, is also challenging with the traditional models/methods. To that end, Babii, Ghysels & Striaukas (2022) introduced machine learning mixed data sampling (MIDAS) regressions for single series nowcasting using potentially large data sets, and Babii, Ball, Ghysels & Striaukas (2022*a*) extend these methods to machine learning panel data MIDAS regressions.

In this paper, we are interested in nowcasting Euro area 19 (EA-19) GDP growth (e.g., see Marcellino et al. 2003), which means that we aim to potentially nowcast GDP growth for multiple countries simultaneously. One can think of several ways to proceed, namely:

(a) nowcast aggregate EA-19 GDP growth using only aggregate Euro area data,

(b) nowcast aggregate EA-19 GDP growth using all available data from the individual countries,

(c) nowcast each of the 19 countries separately and then aggregate across countries to obtain an aggregate EA-19 GDP growth nowcast,

(d) nowcast only the GDP of the large constituents, ignoring the small countries, and use those to nowcast EA-19 GDP growth.

Option (a) is similar to the case of US GDP growth, with typically around 30 monthly series used to produce nowcasts.[1] Even for this case, Babii, Ghysels & Striaukas (2022) show that machine learning methods outperform traditional dynamic factor models using exactly the same data. Moving to option (b) the challenge of dealing with large data sets emerges. If we keep the same 30 series, but collect them for each individual country, we have potentially $30 \times 19 = 570$ predictors. Still, the target is a single series, i.e. aggregate EA-19 GDP growth. The more interesting options are (c) and (d), which are the novel contributions of the paper. In both cases, the combination of nowcasting and aggregation come into play. The former involves nowcasting all countries, regardless of their size and importance in the overall economic outlook of the EA. Option (d) has been entertained by Cascaldi-Garcia et al. (2023), among others, who propose a multi-country nowcasting model to simultaneously predict the GDP of the Euro area and its three largest countries—namely Germany, France, and Italy—using up to sixteen predictor variables per country.

Cases (c) and (d) pertain to nowcasting in a data-rich environment using high-dimensional mixed-frequency panel data models. This is a relatively new and unexplored research area. Khalaf et al. (2021) consider low-dimensional mixed-frequency panel data but not for nowcasting or forecasting. Fosten & Greenaway-McGrevy (2019) consider nowcasting with a mixed-frequency Vector AutoRegressive (VAR) panel data model, but not in the context of a high-dimensional data-rich environment as is the case in our application. Babii, Ghysels & Striaukas (2022) introduce structured machine learning for heavy-tailed dependent panel data regressions sampled at various frequencies. They focus on the sparse-group LASSO (sg-LASSO) regularization which they show can incorporate the mixed-frequency panel data structures. They derive oracle inequalities for the sparse-group LASSO panel model estimators with dependent fat-tailed data. Babii, Ball, Ghysels & Striaukas (2022*b*) apply

[1] We use a larger set of monthly and weekly standard data for the EA-19 compared to the series used for the US, as will be explained in Section 3.1.

the method to nowcast a large cross-section of earnings data of the US firms. In this paper, we apply the same framework for nowcasting panels of Euro area GDP growth based on real-time data vintages for standard macro data and news series which are extracted from a large set of newspaper articles.

Our paper uses standard macroeconomic series, as well as so-called non-standard series, in particular textual data. In recent years, the use of newspaper data for nowcasting has been explored by various authors.[2] Scotti (2016), Baker et al. (2016), Barbaglia et al. (2024), and Ashwin et al. (2024) are recent examples that use newspaper data for economic analysis in multiple countries. Although the construction of country-specific indicators brings an additional level of complexity, the presumption is that the resulting indexes provide an early signal of the local economic conditions that is more precise than a news-based general indicator. We follow Barbaglia et al. (2024) who propose country-specific text-based indicators for five European countries using local news translated to the English language. In this paper, we extend their analysis to all EA-19 countries and show how the inclusion of news-based indexes about smaller countries can improve the nowcasting performance.

We use the MIDAS sparse-group LASSO regression-based approach of Babii, Ghysels & Striaukas (2022) amenable to large dimensional data environments to test whether adding different levels of heterogeneity in panel models improves the quality of nowcasts. Moreover, we test whether specific data sources are informative when used in nowcasting settings. In particular, we look at whether news data can improve nowcasts over more traditionally used macro and financial series. Lastly, we apply several weighting schemes to compute the Euro area aggregate nowcast based on country-level GDP nowcasts. We show that

[2]See for example Larsen et al. (2021), Barbaglia et al. (2023), and Ellingsen et al. (2022) among others. Thorsrud (2020) uses daily business newspapers to construct a business cycle index based on quarterly GDP growth using a time-varying dynamic factor model with sparsity. While his work shows the feasibility of traditional state space models, the challenges grow when one starts thinking about adding potentially large-dimensional traditional data sets in a multi-country setting.

this strategy leads to more accurate nowcasts irrespective of the weighting scheme, suggesting that information in all European countries contributes to overall predictions. In addition, we test whether smaller panels involving only the Big Four countries—France, Germany, Italy, and Spain—suffice to compute the Euro area nowcasts. Our results suggest that smaller countries indeed matter and models that incorporate all EA-19 produce higher quality nowcasts. The efficient estimation of panel models using a large data set is compounded by the timely inclusion of information about the current state of the economy from smaller countries.

In sum, the paper makes multiple contributions. First, the topic of nowcasting and aggregation was to the best of our knowledge never addressed in the literature. Second, we study different aggregation schemes and find that a projection method appears to work best. Third, we investigate the advantages of machine learning panel data models versus single regression ones in the context of GDP nowcasting, extending earlier findings on the topic by Babii, Ball, Ghysels & Striaukas (2022*b*) who looked at corporate earnings. Fourth, we showcase the use of non-traditional data in nowcasting, more specifically news related data, a topic which received much attention recently. Fifth, we highlight to role played by small Euro area countries due to their timely release of data and the connectedness of their economies with the larger countries. Finally, our sample includes the pandemic and we analyze its impact on nowcasting and model performance. The latter leads us to investigate nowcast combination schemes across different models.

The paper is organized as follows. Section 2 introduces the models and estimators. Section 3 describes the Euro area GDP nowcasting application, whereas Section 4 reports and analyzes the empirical results. Section 5 concludes.

2 Mixed-frequency panel data machine learning nowcasting models

Our method builds on Babii, Ghysels & Striaukas (2022), who introduced machine learning MIDAS regressions with an application to single series nowcasting, and Babii, Ball, Ghysels & Striaukas (2022*a*), who extended the method to panel data settings. Moreover, Babii, Ball, Ghysels & Striaukas (2022*b*) consider an application of nowcasting price-earnings ratios for a large set of US firms.

Our goal is to nowcast a set of variables $y_{i,t+h}$ with $i \in [N]$ at horizon h measured at some low frequency, e.g. quarterly, $t \in [T]$, where $[p] = \{1, 2, \ldots, p\}$ for a positive integer p. For simplicity, we assume equally spaced data at different frequencies. In particular, n_k^H denotes the total number of high-frequency observations for the k^{th} covariate for each low-frequency period t, and n_k^L is the number of low-frequency periods used as lags. For example, in our application a quarter of high-frequency lags used as covariates corresponds to $n_k^L = 1$ while $n_k^H = 3$ indicates that three months of data are used in each quarter. Note that we can have a mix of quarterly, monthly, and weekly data, and for each covariate indexed by k, n_k^H represents different high-frequency sampling frequencies and associated lags $n_k^L n_k^H$. The information set consists of K predictors, i.e. $\left\{x_{i,t-j/n_k^H,k} :\ i \in [N], t \in [T], j = 0, \ldots, n_k^L n_k^H - 1, k \in [K]\right\}$, measured potentially at some higher frequency, e.g., monthly/weekly/daily with real-time updates into the quarter.

Consider the following panel data regression for the low-frequency panel target $y_{i,t|\tau}$, using information up to τ:

$$y_{i,t|\tau} = \alpha + \sum_{q=1}^{Q} \rho_{i,q} y_{i,t-q} + \sum_{k=1}^{K} \psi(L^{1/n_k^H}; \beta_k) x_{i,\tau,k} + u_{i,t|\tau},$$

where

$$\psi(L^{1/n_k^H};\beta_k)x_{i,\tau,k} = \frac{1}{k_{\max}}\sum_{j=0}^{k_{\max}-1}\beta_{j,k}L^{j/n_k^H}x_{i,\tau-j/n_k^H,k}, \tag{1}$$

and

- α is the intercept constant across all i;[3]
- $\rho_{i,q}$ are autoregressive lag coefficients possibly different across i;
- $k_{\max}$ is the maximum lag length which could depend on the covariate k, and for each high-frequency covariate $x_{i,\tau,k}$ we have the most recent information at time τ.[4] Note that this also implies the most recent vintage of data is used, i.e. past data revisions are incorporated as well.

When $\tau \leq t-1$, we are dealing with a forecasting situation, hence, our analysis applies to forecasting as well. Note that $k_{\max}$ is the maximum lag length which may depend on the covariate k, and for each high-frequency covariate $x_{i,\tau,k}$ we have the most up-to-date information available at time τ. For some high-frequency regressors that have not been updated yet, this could be stale information; see Babii, Ball, Ghysels & Striaukas (2022*b*) for further discussion. In our empirical analysis, we consider the following two variations of this general formulation. For $i, j \in [N]$:

- Pooled panel, denoted *pooled*, restriction: $\rho_{i,j} = \rho_j$.
- Country-specific AR lag coefficients, denoted *HetAR*.

In the pooled panel data case, the sample size is NT which is much larger than in the case of a single-country model. Therefore, we may expect a more accurate estimation compared to individual time series regressions at the expense of losing heterogeneity. In the HetAR case, we model country-specific effects, therefore we investigate whether more

[3]All parameters depend on τ, but we suppress this detail to simplify notation.
[4]To simplify the notation, we will not keep track of the dependence of x_t, $k_{\max}$, and K on τ.

flexible models can improve the quality of nowcasts relative to the pooled panel model by including country-specific autoregressive coefficients in the model. Lastly, in the pooled, we group pooled autoregressive lags in time series dimension rather than the cross-section.

The large number of predictors K with potentially large number of high-frequency measurements $k_{\max}$ can be a rich source of predictive information, yet at the same time, estimating $N \times Q + 1 + \sum_{k=1}^{K} k_{\max}$ parameters, where N is the size of the cross-section, Q is number of autoregressive lags (constant across i) and $k_{\max}$ is the max of high-frequency lags for each covariate (constant across i), is costly and may reduce the predictive performance in small samples. In addition, observing predictors at different frequencies leads to the frequency mismatch problem due to the missing data. To solve both problems, we follow the MIDAS (machine learning) literature and instead of estimating a large number of individual slope coefficients, we consider a weight function[5] $\omega : [0, q] \to \mathbf{R}$, parameterized by $\beta_k \in \mathbf{R}^L$

$$\psi(L^{1/n_k^H}; \beta_k) x_{i,\tau,k} = \frac{1}{k_{\max}} \sum_{j=0}^{k_{\max}-1} \omega\left(\frac{j}{n_k^H}; \beta_k\right) x_{i,\tau,k},$$

where

$$[0, q] \ni s \mapsto \omega(s; \beta_k) = \sum_{l=0}^{L-1} \beta_{l,k} w_l(s),$$

and $(w_l)_{l \geq 0}$ is a set of L functions, called the *dictionary*. One can use either splines or Legendre polynomials as a dictionary; see Babii, Ball, Ghysels & Striaukas (2022*a*) for further discussion on different choices of the dictionary. The attractive feature of this approach is that we can map the MIDAS regression to the simple linear regression framework. To that end, assuming that $k_{\max}$ is the same for all covariates and $\mathbf{x}_i = (X_{i,1}W_k, \ldots, X_{i,K}W_k)$, where for each $k \in [K]$, $X_{i,k} = \{x_{i,\tau - j/n_k^H, k}, j = 0, \ldots, k_{\max} - 1\}_{\tau \in [T]}$ is a $T \times k_{\max}$ matrix of predictors and $W_k k_{\max} = (w_l(j/n_k^H; \beta_k))_{0 \leq l \leq L-1, 0 \leq j \leq k_{\max}}$ is a $k_{\max} \times L$ matrix corresponding

[5] We use $q = 1$ for nowcasting without low-frequency lags. More generally, $q - 1$ denotes the number of low-frequency lags.

to the dictionary. Note that the matrices $(W_k)_{k\in[K]}$ solve the frequency mismatch problem.

Define $\mathbf{y}_i = (y_{i,1+h}, \dots, y_{i,T+h})^\top$, $\tilde{\mathbf{y}}_{i,q} = (y_{i,1-q}, \dots, y_{i,T-q})^\top$ for $q \in [Q]$, $\tilde{\mathbf{y}}_i = (\tilde{\mathbf{y}}_{i,1}, \dots, \tilde{\mathbf{y}}_{i,Q})$, and $\mathbf{u}_i = (u_{i,1}, \dots, u_{i,T})^\top$. Stacking time series observations in vectors, the regression equation for each $i \in [N]$ is

$$\mathbf{y}_i = \alpha\iota_T + \tilde{\mathbf{y}}_i\rho_i + \mathbf{x}_i\beta + \mathbf{u}_i,$$

where ι_T is a size T vector of ones, $\rho_i \in \mathbf{R}^Q$ coefficients of autoregressive lags and $\beta \in \mathbf{R}^{LK}$ is a vector of slopes of regressors. Lastly, let $\mathbf{y} = (\mathbf{y}_1^\top, \dots, \mathbf{y}_N^\top)^\top$, $\mathbf{X} = (\mathbf{x}_1^\top, \dots, \mathbf{x}_N^\top)^\top$, and $\mathbf{u} = (\mathbf{u}_1^\top, \dots, \mathbf{u}_N^\top)^\top$. Then the regression equation after stacking all cross-sectional observations is

$$\mathbf{y} = \alpha\iota_{NT} + \tilde{\mathbf{Y}}\rho + \mathbf{X}\beta + \mathbf{u},$$

where $\tilde{\mathbf{Y}}$ is a diagonal matrix with elements $\tilde{\mathbf{y}}_i, i = 1, \dots, N$, and $\rho = (\rho_1^\top, \dots, \rho_N^\top)^\top$.

The MIDAS approach reduces efficiently the dimensionality of high-frequency lag coefficients. An alternative approach, known as the U-MIDAS, see Foroni et al. (2015), would estimate all individual coefficients associated with high-frequency covariate lags hoping that machine learning would pick up relevant lags. This strategy is not appealing because we always pay a price for the model selection which can be substantial with heavy-tailed data time series data, typically leading to worse predictions compared to regularized MIDAS schemes; see Babii, Ghysels & Striaukas (2022), Babii, Ball, Ghysels & Striaukas (2022*a*) for further discussion and details.

Since the number of predictors K can be large, the regularization improves the predictive performance in small samples. In this paper, we work with the sg-LASSO regularization that was shown to be attractive for individual time series machine learning regressions in Babii, Ghysels & Striaukas (2022). The pooled panel data estimator with heterogeneous autoregressive dynamic and a variation of sparse-group regularization solves

$$\min_{(a,b,c)\in\mathbf{R}^{1+NQ+LK}} \|\mathbf{y} - a\iota_{NT} - \tilde{\mathbf{Y}}b - \mathbf{X}c\|_{NT}^2 + 2\lambda\Omega_\gamma(b,c), \tag{2}$$

where $\|.\|_{NT}^2 = |.|^2/(NT)$ is the scaled ℓ_2 norm squared and

$$\Omega_\gamma(b,c) = \gamma\left[|b|_1 + |c|_1\right] + (1-\gamma)\left[\|b\|_{2,1} + \|c\|_{2,1}\right],$$

is a regularizing functional, which is a linear combination of LASSO and group LASSO penalties. The parameter $\gamma \in [0,1]$ determines the relative weights of the ℓ_1 (sparsity) and the $\ell_{2,1}$ (group sparsity) norms, while the amount of regularization is controlled by the regularization parameter $\lambda \geq 0$. Recall also that for a group structure $\mathcal{G}$ described as a partition of $[p] = \{1,2,\ldots,p\}$, the group LASSO norm is computed as $\|u\|_{2,1} = \sum_{G\in\mathcal{G}} |u_G|_2$ for some generic vector u. The groups are assumed to be known to the econometrician, which in our setting corresponds to: a) country-specific autoregressive lags, and b) time series lags of covariates. Combining covariates of a similar nature in groups is also feasible.

3 Nowcasting EU output

As noted before, let $y_{i,t|t}$ the GDP growth nowcast for country i for the quarter t given the high-frequency information up to and including the quarter t. We use the following schemes to combine country-level nowcasts to compute the aggregate Euro area level GDP nowcasts, denoted $y_{\text{ea},t|t}$:

1. We compute the weights based on historical proportions of absolute GDP growth.

Each country weight $W^{(1)}_{i|t}$ at quarter t is

$$W^{(1)}_{i|t} = \frac{\sum_{j=1}^{t-1} |y_{ij}|}{\sum_{i=1}^{N} \sum_{j=1}^{t-1} |y_{ij}|}.$$

2. We compute the weights based on the most recent proportions of absolute GDP growth. Each country weight $W^{(2)}_{i|t}$ at quarter t is

$$W^{(2)}_{i|t} = \frac{|y_{it-1}|}{\sum_{i=1}^{N} |y_{it-1}|}.$$

3. We compute the weights based on the most recent proportions of the GDP level. Each country weight $W^{(3)}_{i|t}$ at quarter t is

$$W^{(3)}_{i|t} = \frac{\text{GDP}_{it-1}}{\sum_{i=1}^{N} \text{GDP}_{it-1}}.$$

4. We project the weights onto the most recent Euro area GDP growth series $y_{ea,t-1}$. Each country weight $W^{(4)}_{i|t}$ at quarter t is

$$y_{ea,t-1} = \sum_{i=1}^{N} W^{(4)}_{i|t} y_{it-1} + \epsilon_{t-1},$$

where $W^{(4)}_{i|t}$ are constrained to be positive and sum to 1. We estimate $W^{(4)}_{i|t}$ by constrained least squares.

For all the weighting schemes we use $t-1$ GDP data since at quarter t this is the information that is available in real-time. The Euro area aggregate nowcast is computed as $y_{\text{ea},t|t} = \sum_{i=1}^{N} W^{(q)}_{i|t} y_{i,t|t}$, for $q \in \{1, 2, 3, 4\}$ set of weights.

Finally, we also consider what we call the Euro area model. In this model, we nowcast Euro area GDP growth based on aggregate Euro area data. For this we use the machine learning MIDAS setup of Babii, Ghysels & Striaukas (2022).

3.1 Standard macro releases

We use real-time vintages of standard macro monthly releases from several sources (see the Appendix for the details of each series). GDP growth is quarterly and is available in real-time in our sample. The first vintage is January 2015, for which we have GDP vintages for all EA-19 countries. As predictor variables, we collect 64 traditional macroeconomic series: 57 are monthly series, 3 weekly, and 4 daily. Monthly series are hard data such as unemployment rate and industrial production; 47 are available at a country level while the remaining 10 are Euro area aggregates; weekly series are oil products which are available at a country level; 4 daily series are financial markets data covering stock market, gold, foreign exchange, and interest rates data.

It is worth mentioning that some macro series are available with a month, two, or even three months of delay relative to the nowcasting month. Since we use real-time data vintages, we naturally take into account such delays. This is particularly important when we analyze the additional gain in nowcasting accuracy when using news data—which we describe below—since such data is timely and available without delays.

3.2 News data

We collect news data from Dow Jones Factiva. The data set contains daily printed and online full-text articles from three different sources dealing with economic and financial issues, namely The Economist, Reuters News, and The Wall Street Journal. The final data set consists of approximately 2.5 million articles and 1 billion words from January 2005 to December 2022. We construct news-based indicators following the fine-grained aspect-based sentiment (FiGAS) by Consoli et al. (2022). This is a rule-based algorithm that provides sentence-level sentiment scores for textual information in the English language. The sentiment is (i) *aspect-based*, meaning that it analyzes only the words in a sentence that relate to a specific topic of interest based on linguistic dependencies, and (ii) *fine-grained*, that is, the sentence-level sentiment score comes from a humanly annotated dictionary tailored for economic and financial applications and is defined in $[-1, +1]$.

We compute news-based indicators for the following three topics covering different aspects of economic and financial activity: *economy*, *employment*, and *inflation*. Each topic is associated with a set of additional keywords that we look for in the articles: for instance, the *economy* topic includes also related terms, like GDP, output, or economic growth. We compute country-specific indicators for all EA-19 member states by filtering only on sentences where there is a direct mention of the country in the analysis. We refer to the online appendix for additional details about the news-based sentiment indicators, the selection of the topics' keywords, and the construction of country-specific measures.

The output of FiGAS consists of daily news-based indicators of *sentiment* and *volume* for each topic and country. The sentiment measures are obtained by summing the sentence-level sentiment scores for all articles within that day, while the volume corresponds to the number of sentences containing an explicit mention of the topic. Most importantly, the news-based sentiment indicators are real-time and we include them as additional regressors with no publication delay. In the remainder, our baseline model will include news-based indicators about *economy*, *employment*, and *inflation*.

4 Empirical results

We apply the machine learning regressions described in Section 2 to assess whether the aggregate Euro area GDP growth nowcasts are more accurate compared to panel data models which are based on individual country-level data. These models exploit panel data structures and several weighting schemes of individual country nowcasts to construct an aggregate nowcast.

4.1 Aggregate versus panel data regressions

Table 1 reports the nowcast comparison between the (aggregate) Euro area model and the panel data models, namely the pooled and the HetAR panel models, and the country-specific MIDAS regressions. The forecast accuracy is measured as root-mean-squared errors (RMSEs) at three nowcast horizons (i.e., 2- and 1-month ahead as well as at the end of the quarter). The first row in Panel A of Table 1 reports the RMSEs of the aggregate Euro area benchmark model, while the other rows report the RMSEs of the proposed panel data models relative to the benchmark: values below unity signal a better performance of the proposed model relative to the benchmark. For each panel data model, we document the performance of the four different weighting schemes ($W^{(1)}$-$W^{(4)}$) to aggregate individual country forecasts as described in Section 3. Columns 1 to 3 report the results on the full sample, while columns 4 to 6 consider only the pre-COVID period.

Panel A reports the results for the time series regressions when we nowcast directly Euro area aggregates. While the first row regressions include only aggregated information about the Euro area as an explanatory variable (option (a) in the Introduction), the other rows in the top panel add respectively the Aruoba et al. (2009) (ADS) index and information from individual EA-19 countries as additional regressors (option (b) in the Introduction).[6] A few patterns are worth highlighting. First, the model's accuracy is largely impacted by the inclusion of the COVID period observations, with the full sample results showing much larger errors than the pre-COVID sample. Second, the inclusion of timely information about the US business cycle proxied by the ADS index does not bring any systematic

[6] The ADS index is updated at the daily frequency and is based on a small set of macroeconomic variables available at the weekly, monthly, and quarterly frequency; see Aruoba et al. (2009) for more details.

	2-month	1-month	EoQ	2-month	1-month	EoQ
		Full sample			*Pre-COVID*	
			Panel A. *Time series*			
EA-19	4.960	3.642	3.366	0.185	0.215	0.176
+ ADS	0.830	1.043	0.886	0.961	1.146	1.031
+ other countries	1.501	2.136	1.685	1.653	1.050	0.753
+ countries & ADS	1.501	2.136	1.536	1.653	1.039	0.753
			Panel B. *Pooled*			
$W^{(1)}$	0.693	0.814	0.750	2.420	1.880	2.551
$W^{(2)}$	0.681	0.775	0.704	3.201	2.154	2.973
$W^{(3)}$	0.680	0.729	0.664	1.802	1.880	2.171
$W^{(4)}$	0.661	0.722	0.663	1.640	1.883	2.168
			Panel C. *HetAR*			
$W^{(1)}$	0.677	0.817	0.756	2.393	1.640	2.268
$W^{(2)}$	0.679	0.789	0.729	3.664	2.074	2.893
$W^{(3)}$	0.669	0.738	0.663	1.792	1.588	1.858
$W^{(4)}$	0.684	0.743	0.665	1.684	1.584	1.835
			Panel D. *Country-specific regressions*			
$W^{(1)}$	0.743	0.984	0.933	3.229	1.576	1.800
$W^{(2)}$	0.803	1.063	0.932	5.981	2.037	2.277
$W^{(3)}$	0.777	0.996	0.870	1.035	0.557	0.755
$W^{(4)}$	0.799	1.046	0.835	0.882	0.548	0.743

Table 1: Nowcast comparison table — root-mean-squared errors (RMSEs). Nowcast horizons are 2- and 1-month ahead, as well as the end-of-the-quarter (EoQ). The first row reports the absolute RMSE for the EA-19 model using only aggregate data. All remaining rows are relative RMSEs vis-à-vis the first row. Panel A displays results for time series regressions in which we directly nowcast Euro area GDP growth. Panels B and C report results for panel regression models and different weighting schemes. Rows $W^{(1)}$-$W^{(4)}$ denote different weights used to compute the Euro area aggregate based on individual country nowcasts. For rows $W^{(1)}$-$W^{(4)}$, we report RMSEs relative to the Euro area model: values smaller than unity indicate an improvement in prediction with respect to the benchmark. Panel B reports results for the *Pooled* panel model for 19 countries, while Panel C for the *HetAR* model. Panel D reports results using country-specific MIDAS regression nowcasts aggregated based on weighting schemes $W^{(1)}$-$W^{(4)}$. We use 5-fold cross-validation adjusted for panel data to compute both sg-LASSO tuning parameters. Out-of-sample periods: 2016 Q1 - 2019 Q4 (pre-COVID) and 2016 Q1 - 2022 Q4 (Full sample).

improvement with respect to the benchmark model. Third, an even worse performance results from the inclusion of information about individual EA-19 countries into the time series regression model compared to the benchmark model only using aggregate EA data. Hence, option (a) is better than (b), and therefore using only the aggregate data in machine learning models suffices.

Turning to the performance of the panel data machine learning models in Panels B and C, we obtain mixed results. For the pre-COVID sample, we note that the panel data models under perform vis-à-vis the benchmark. In contrast, the panel data models always outperform the benchmark when looking at the full sample. This result is robust with respect to the choice of the weighting scheme and the nowcasting horizon. The proposed models exploit the additional country-specific information included in the panel data structure: this information turns out to be redundant during normal times, while it proves relevant during the COVID pandemic. Comparing the two panel data models, the HetAR model generally achieves better performances than the pooled models: the inclusion of heterogeneity in the form of country-specific lags seems to be a better choice than pooling all coefficients.

Panel D of Table 1 reports the results with country-specific MIDAS regressions – hence not exploiting the panel data structure, but instead estimating single regressions per country – which are then aggregated using again the same weighting schemes. Hence, Panel D differs from Panels B and C only with respect to the panel structure, while the information set, the MIDAS structure and the weighting schemes are the same. Country-specific regressions aggregated following weighting schemes $W^{(3)}$ and $W^{(4)}$ attain good results both when looking at the full sample as well as pre-COVID. On the one hand, Panel D shows gains with respect to the Euro area benchmark during the pre-COVID sample, when panel models perform poorly. The gains attained by country-specific regressions are large and up to 45 percentage points. On the other hand, the performance in the full-sample, although still better than the benchmark, does not improve with respect to panel models which perform best when including the COVID pandemic in the analysis.

4.2 The Big Four

We test whether nowcasting the four largest Euro area countries—France, Germany, Italy, and Spain (i.e., the "Big Four")—separately may give an advantage in producing more accurate predictions for the Euro area GDP.[7] We compare the performance of the HetAR and pooled panel models, as well as, the country-specific regressions as in the previous section. To compute the aggregated Euro area nowcasts we use weighting scheme $W^{(4)}$ modified to include only the Big Four.

Table 2 reports the results of the four-country models relative to the benchmark model appearing in the first row of Table 1. All point to a better performance of the model involving nowcasting only the largest European economies. This result is robust across horizons, model specifications and, interestingly, the full and pre-COVID samples. Indeed, panel data models with the Big Four attain a more accurate performance even when considering the pre-COVID sample, whereas that was not the case in Table 1. However, looking at the full sample, panel data models with all Euro area countries always outperform the Big Four models. Hence, rather than focusing only on the indicators from the largest European countries, our evidence suggests that there is a value-added in the nowcasting model for all EA-19 countries when looking at the full sample and including the COVID in our model. Looking at the pre-COVID sample, the Big Four models are outperformed at all horizons by the country-specific regressions using all Euro area countries reported in Table 1. Overall, this suggests the best results are obtained with the inclusion of information of all EA-19 countries both in the full sample (with panel models) and pre-COVID (with

[7] See also Ashwin et al. (2024), Barbaglia et al. (2022), Cascaldi-Garcia et al. (2023), among others.

	2-month	1-month	EoQ	2-month	1-month	EoQ
		Full sample			*Pre-COVID*	
			Panel A. *Big four — Pooled*			
$W^{(4)}$	0.698	0.871	0.808	0.877	0.978	0.820
			Panel B. *Big four — HetAR*			
$W^{(4)}$	0.692	0.743	0.734	0.810	1.025	0.884
			Panel C. *Big four — country-specific regressions*			
$W^{(4)}$	0.813	1.046	0.820	0.965	0.557	0.807

Table 2: Nowcast comparison table — root-mean-squared errors (RMSEs). Nowcast horizons are 2- and 1-month ahead, as well as the end-of-the-quarter (EoQ). All entries are relative RMSEs to the first row of Table 1. Panel A and B display results for Pooled and HetAR panel regression models, respectively, based on France, Germany, Italy, and Spain (i.e. the "Big Four") countries. Panel C displays results for the Big Four country-specific regressions aggregated based on weighting scheme $W^{(4)}$. We use 5-fold cross-validation adjusted for panel data to compute both sg-LASSO tuning parameters. Out-of-sample periods: 2016 Q1 - 2019 Q4 (pre-COVID) and 2016 Q1 - 2022 Q4 (Full sample).

country-specific regressions).

To explain the relative improvement in nowcasting performance achieved by using the full panel of Euro area countries as opposed to just the Big Four, several factors comes into play. First, from a modeling perspective, machine learning panel regression models are estimated more accurately when N, i.e. the number of countries, is large (Babii, Ball, Ghysels & Striaukas 2022*a*). More accurate parameter estimates lead to higher quality nowcasts. Second, the real-time flow of data and information varies among countries. For instance, Belgium, a nation falling outside the Big Four category, consistently boasts superior and more timely survey data, significantly contributing to the accuracy of Euro area GDP predictions (Basselier et al. 2018). From an economic perspective, given Belgium's strong economic ties with Germany and France due to its geographical proximity, its economic news serves as a reliable signal for the broader economies, therefore influencing the overall Euro area GDP projection. By incorporating the entire panel of countries, we can effectively capture and harness these effects.

4.3 Nowcast aggregation and combination

We now focus on how to aggregate and combine individual nowcasts, starting from the nowcasting performance of the four proposed weighting schemes. Looking at the full sample results of Table 1, we observe that $W^{(4)}$ achieves the lowest RMSEs for the pooled panel model. This result does not hold for the HetAR model, where the $W^{(3)}$ weighting scheme produces slightly more accurate forecasts than $W^{(4)}$. As it regards country-specific regressions, $W^{(3)}$ and $W^{(4)}$ attain a similar performance both full sample and pre-COVID, and substantially outperform the other weighting schemes. Figure 1 illustrates the four estimated weighting schemes, with $W^{(3)}$ and $W^{(4)}$ notably exhibiting denser characteristics than the other schemes. Consequently, nowcasts employing denser weighting schemes yield more accurate results, further reinforcing the argument in favor of utilizing the entire panel of European countries to improve nowcasting precision.

To further investigate the nowcasting performance across weighting schemes, we turn to Figure 2 where we compare the weights obtained by $W^{(4)}$ (i.e., projections on GDP) and by $W^{(3)}$ (i.e., proportion of GDP level). We take the difference between $W^{(4)}$ and $W^{(3)}$ weights: positive values indicate a larger weight given by $W^{(4)}$ with respect to $W^{(3)}$. Note that by construction the $W^{(3)}$ weights are a direct measure of the size of each national economy within the Euro area. Compared to $W^{(3)}$, we observe that $W^{(4)}$ assigns smaller weights to Germany and to a lesser extent the Netherlands, while it inflates the relative importance of some small- and medium-sized countries, namely Austria, Belgium and Luxembourg. Although the size of these economies within the Euro area is relatively small, information about economic developments in those countries play a relevant role in attaining more accurate nowcasts.

Finally, we explore the relative importance of smaller countries against the Big Four by plotting in Figure 3 the sum of weights of the 15 small Euro area countries (*Small15*)

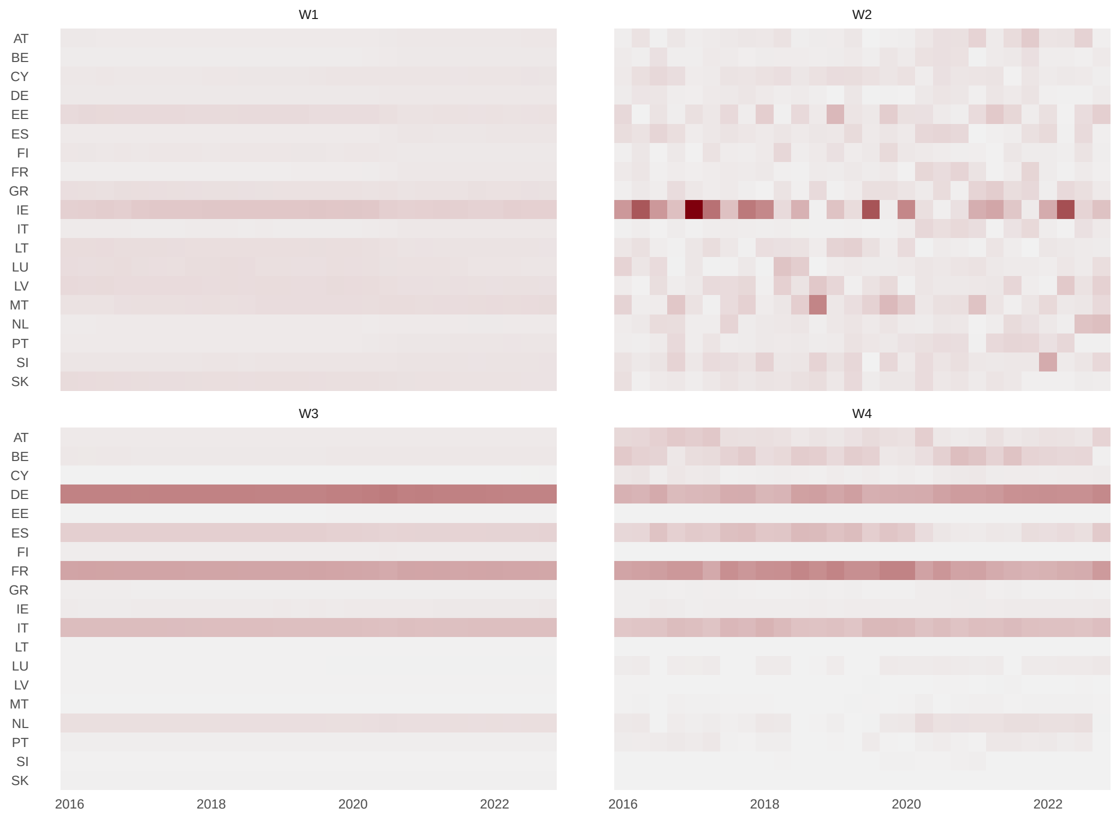

Figure 1: Estimated weights for the weighting schemes $W^{(1)}$-$W^{(4)}$ by country. Each tile corresponds to a quarter. The darker the color, the larger the weight.

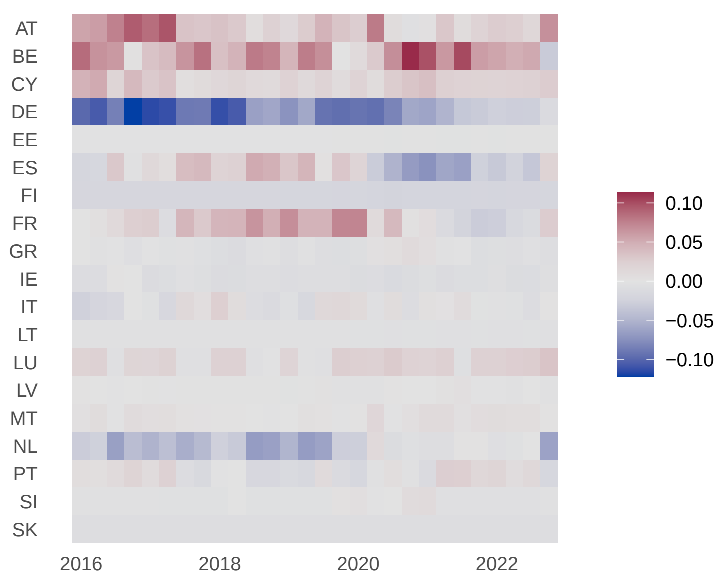

Figure 2: Difference between $W^{(4)}$ and $W^{(3)}$ weights by country. If difference is larger (smaller) than zero, $W^{(4)}$ weights are larger (smaller) than $W^{(3)}$ ones and are reported in red (blue). Each tile corresponds to a quarter. The darker the color, the larger the difference between weighting schemes.

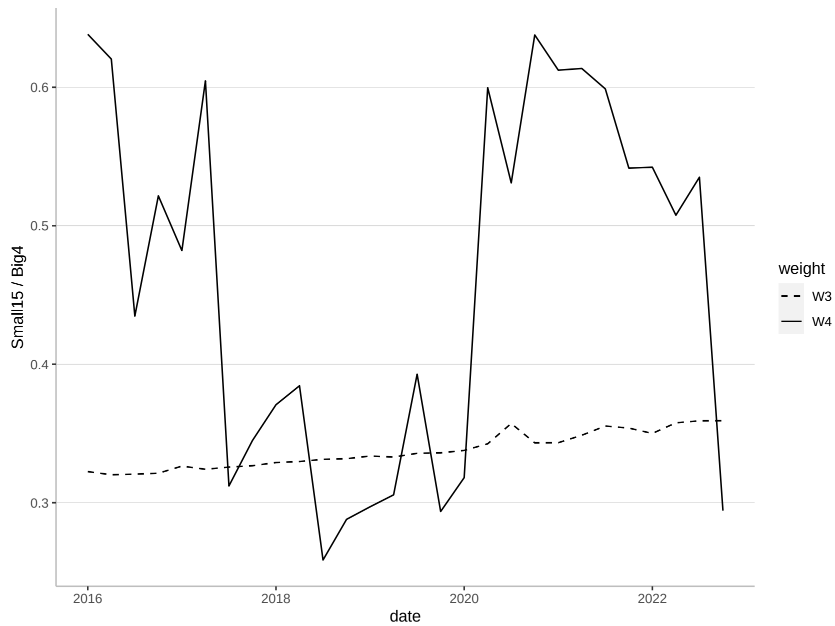

Figure 3: Sum of weights of the 15 small Euro area countries (*Small15*) relative to the sum of the weights of the Big Four (*Big4*) for weighting schemes $W^{(3)}$ and $W^{(4)}$.

relative to the sum of the weights of the Big Four for weighting schemes $W^{(3)}$ (dashed line) and $W^{(4)}$ (solid line). While the relative importance of smaller countries with respect to the Big Four is stable when looking at the $W^{(3)}$ weights (indeed, proportions of GDP level are steady in time and vary only at a slow pace), we observe large variability when considering $W^{(4)}$. Smaller countries are relatively more important in 2016-17 and, most notably, after 2020, when the total weight assigned to smaller countries doubles, going from 0.3 to approximately 0.6. The information coming from smaller Euro area countries plays an important role in attaining an accurate nowcast in the years following the COVID-19 pandemic, hence explaining the good performance of panel models on all EA-19 countries in the full sample reported in Table 1.

4.3.1 Nowcast combinations

Overall, our results highlight a heterogeneous nowcast performance of the analyzed models: panel models perform best when considering the full sample and information from all Euro area countries, country-specific MIDAS regressions outperform all other models in the pre-COVID sample, while nowcasting models on the Big Four attain a good performance across all time samples, although not the best one. Given this heterogeneity in performance, we combine forecasts obtained from individual models following the linear combination method by Stock & Watson (2004). In particular, we combine the forecast obtained following models: the *Pooled* panel model and *Country-specific regressions* on all EA-19 countries (Panels B and D, respectively, of Table 1), and the *HetAR* model on the Big Four (Panel B of Table 2). For all individual forecasts, we consider the aggregated forecasts obtained with $W^{(4)}$ weights.[8]

The nowcast combination results are reported in Table 3. We start with Panel A which covers models with news data. Combing the nowcasts consistently delivers large gains which range between a 20 to 30% reduction with respect to the benchmark RMSEs. Moreover, the gains are robust across horizons and, importantly, full versus pre-COVID samples. Indeed, combinations attain more accurate nowcasts both full sample as well pre-COVID, reaching a performance that is close to the best-performing individual model. Panel B of Table 3 explores one final aspect of our analysis, namely the added values of the news indicators. Compared to Panel A showing the nowcasting performance of models including news indicators, Panel B excludes them. The inclusion of news delivers nowcasting gains, even though these gains are not large, ranging between 1-5%. Therefore, although only marginally, the inclusion of news indicators positively impacts the accuracy of the nowcasts.

[8]We select forecasts aggregated with $W^{(4)}$ weights since they provide the best nowcasting performance overall. Regarding the selection of the individual nowcasting models, we include all model types analyzed in the paper (*Pooled* and *HetAR* panels, as well as country-specific MIDAS regressions) taking into account the time period and information set (i.e., all EA-19 or only Big Four countries) where they achieve their best performance. We have also experimented with other model subsets for which we obtained similar results.

	2-month	1-month	EoQ	2-month	1-month	EoQ
		Full sample			*Pre-COVID*	
		Panel A. *Combination – with news*				
$W^{(4)}$	0.693	0.815	0.717	0.831	0.770	0.733
		Panel B. *Combination – without news*				
$W^{(4)}$	0.701	0.832	0.725	0.890	0.801	0.773

Table 3: Nowcast comparison table — root-mean-squared errors (RMSEs). Nowcast horizons are 2- and 1-month ahead, as well as the end-of-the-quarter (EoQ). All entries are relative RMSEs to the first row of Table 1. Panel A reports results for forecast combination method combining models *Pooled* (Table 1 Panel B), *Country-specific regressions* (Table 1 Panel D), and *Big four — HetAR* (Table 2 Panel B) based on $W^{(4)}$ weights. Panel B reports forecast combination results for the same models as in Panel A but excluding news data from the predictor variable set. We use 5-fold cross-validation adjusted for panel data to compute both sg-LASSO tuning parameters. Out-of-sample periods: 2016 Q1 - 2019 Q4 (pre-COVID) and 2016 Q1 - 2022 Q4 (Full sample).

5 Conclusion

The paper studies the Euro area nowcasting using MIDAS machine learning panel data regression models. These models offer the flexibility needed to analyze extensive datasets gathered from diverse sources, sampled at varying frequencies, and available in both cross-sectional and time series dimensions. Through this research, we introduce several innovative empirical findings that carry significant relevance for policymakers. Our findings reveal a significant enhancement in the accuracy of nowcasts for the Euro area aggregate GDP when incorporating data from smaller countries. These improvements are substantial, reaching up to 30%, and remain robust across nowcasting horizons. We attribute these gains to two primary factors.

Firstly, in highly turbulent times such as the COVID-19 pandemic, the parameter estimates of the machine learning panel data regression models are estimated more precisely when a broader cross-sectional dataset is employed. Secondly, smaller countries, such as Austria, Belgium, or Luxembourg, in contrast to larger Euro-area countries like Germany or France, tend to possess higher-quality and more timely survey data. Given the high level of economic inter connectivity, especially among neighboring countries in the Euro area, the inclusion of data from these smaller nations results in more accurate signals, ultimately leading to enhanced model predictions.

In addition to our primary findings, we also contribute to the expanding body of literature concerning the application of alternative data, such as information extracted from newspaper articles, to enhance the accuracy of economic forecasts. The incorporation of such data is particularly advantageous due to its real-time availability and no publication delay of these data. Our research demonstrates that news data improve the accuracy of Euro area GDP nowcasts.

Acknowledgments

We received helpful comments from Peter Reusens, Wouter Van der Veken, and Raf Wouters, as well as participants at a National Bank of Belgium seminar, the Nowcasting workshop at the Paris School of Economics, and the 5^{th} Conference on "Nontraditional Data, Machine Learning, and Natural Language Processing in Macroeconomics." The views expressed are purely those of the authors and should not, in any circumstances, be regarded as stating an official position of the European Commission. Eric Ghysels and Jonas Striaukas gratefully acknowledge the financial support of the National Bank of Belgium. Jonas Striaukas also acknowledges the financial support of the F.R.S.–FNRS PDR under project Nr. PDR T.0044.22.

References

Aruoba, S. B., Diebold, F. X. & Scotti, C. (2009), 'Real-time measurement of business conditions', *Journal of Business & Economic Statistics* **27**(4), 417–427.

Ashwin, J., Kalamara, E. & Saiz, L. (2024), 'Nowcasting euro area gdp with news sentiment: A tale of two crises', *Journal of Applied Econometrics* **forthcoming**.

Babii, A., Ball, R. T., Ghysels, E. & Striaukas, J. (2022*a*), Machine learning panel data regressions with heavy-tailed dependent data: Theory and application. *Journal of Econometrics* (forthcoming).

Babii, A., Ball, R. T., Ghysels, E. & Striaukas, J. (2022*b*), Panel data nowcasting in a data-rich environment: The case of price-earnings ratios. *Journal of Applied Econometrics* (forthcoming).

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

Baker, S. R., Bloom, N. & Davis, S. J. (2016), 'Measuring Economic Policy Uncertainty*', *The Quarterly Journal of Economics* **131**(4), 1593–1636.

Barbaglia, L., Consoli, S. & Manzan, S. (2023), 'Forecasting with economic news', *Journal of Business & Economic Statistics* **41**(3), 708–719.

Barbaglia, L., Consoli, S. & Manzan, S. (2024), 'Forecasting gdp in europe with textual data', *Journal of Applied Econometrics* **39**(2), 338–355.

Barbaglia, L., Frattarolo, L., Onorante, L., Tiozzo Pezzoli, L., Pericoli, F. M. & Ratto, M. (2022), 'Testing big data in a big crisis: Nowcasting under covid-19', *International Journal of Forecasting* .

Basselier, R., de Antonio Liedo, D. & Langenus, G. (2018), 'Nowcasting real economic activity in the euro area: Assessing the impact of qualitative surveys', *Journal of Business Cycle Research* **14**, 1–46.

Cascaldi-Garcia, D., Ferreira, T. R., Giannone, D. & Modugno, M. (2023), 'Back to the present: Learning about the euro area through a now-casting model', *International Journal of Forecasting* .

Consoli, S., Barbaglia, L. & Manzan, S. (2022), 'Fine-grained, aspect-based sentiment analysis on economic and financial lexicon', *Knowledge-Based Systems* **247**, 108781.

Ellingsen, J., Larsen, V. H. & Thorsrud, L. A. (2022), 'News media versus fred-md for macroeconomic forecasting', *Journal of Applied Econometrics* **37**(1), 63–81.

Foroni, C., Marcellino, M. & Schumacher, C. (2015), 'Unrestricted mixed data sampling (midas): Midas regressions with unrestricted lag polynomials', *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **178**(1), 57–82.

Fosten, J. & Greenaway-McGrevy, R. (2019), 'Panel data nowcasting', *available at SSRN 3435691* .

Khalaf, L., Kichian, M., Saunders, C. J. & Voia, M. (2021), 'Dynamic panels with midas covariates: nonlinearity, estimation and fit', *Journal of Econometrics* **220**(2), 589–605.

Larsen, V. H., Thorsrud, L. A. & Zhulanova, J. (2021), 'News-driven inflation expectations and information rigidities', *Journal of Monetary Economics* **117**, 507–520.

Marcellino, M., Stock, J. H. & Watson, M. W. (2003), 'Macroeconomic forecasting in the euro area: Country specific versus area-wide information', *European Economic Review* **47**(1), 1–18.

Scotti, C. (2016), 'Surprise and uncertainty indexes: Real-time aggregation of real-activity macro-surprises', *Journal of Monetary Economics* **82**, 1–19.

Stock, J. H. & Watson, M. W. (2004), 'Combination forecasts of output growth in a seven-country data set', *Journal of Forecasting* **23**(6), 405–430.

Thorsrud, L. A. (2020), 'Words are the new numbers: A newsy coincident index of the business cycle', *Journal of Business and Economic Statistics* **38**(2), 393–409.