Université de Neuchâtel
Institut de recherches économiques

IRENE Working Paper 23-02

Do daily lead texts help nowcasting GDP growth?

Marc Burri

Do daily lead texts help nowcasting GDP growth?*

Marc Burri†

July 21, 2023

Abstract

This paper evaluates whether publicly available daily news lead texts help nowcasting Swiss GDP growth. I collect titles and lead texts from three Swiss newspapers and calculate text-based indicators for various economic concepts. A composite indicator calculated from these indicators is highly correlated with low-frequency macroeconomic data and survey-based indicators. In a pseudo out-of-sample nowcasting exercise for Swiss GDP growth, the indicator outperforms a monthly Swiss business cycle indicator if one month of information is available. Improvements in nowcasting accuracy mainly occur in times of economic distress.

JEL classification: C53, E32, E37

Keywords: Mixed-frequency data, composite leading indicator, news sentiment, recession, natural language processing, nowcasting

*I thank David Ardia, Daniel Kaufmann, Maximilian Schröder, Leif Anders Thorsrud, Mark Watson and Philipp Wegmüller for helpful comments and discussions and participants at the IAAE 2023 Conference, the IRENE PhD meeting and the Money-Macro-Finance Conference for valuable comments.

†University of Neuchâtel, Institute of Economic Research, Rue A.-L. Breguet 2, CH-2000 Neuchâtel, marc.burri@unine.ch

1 Introduction

Text data, such as news articles, provide a timely source of information that can be used by policymakers in their decision-making. Unlike traditional survey or hard data, news data is available at daily frequency, potentially offering timelier information on the state of the economy. Such timely information is particularly valuable during dynamic situations such as the Covid-19 pandemic. By incrementally examining the information provided by daily news texts, this paper shows whether high-frequency text data helps for nowcasting real Swiss Gross Domestic Product (GDP) growth.[1]

I propose a context-based method to create sentiment indicators for various hand-selected economic concepts, alongside recession indicators. Based on the lexical methodology (see e.g. Shapiro et al., 2020, for an overview), I develop context-based sentiment indicators, following a similar approach as suggested by Barbaglia et al. (2022). The recession indicators (also referred to as R-word indexes, see e.g. The Economist, 2011) rely on the counting of terms associated with economic recessions. By combining these indicators using a factor model, I develop a daily composite business cycle measure, referred to as the "Short Economic News Indicator" (SENI). I then evaluate the daily informational content of the SENI in a pseudo real-time out-of-sample nowcasting exercise using **mi**xed-**da**ta **s**ampling (MIDAS) and bridge models (Baffigi et al., 2004; Ghysels et al., 2004).

The main results show that daily textual data is at least as informative about the current state of the economy as existing low-frequency business cycle indicators based on survey data. In-sample, the SENI is highly correlated with these business cycle indicators. Therefore, textual data allows to track business cycle fluctuations in a timely and cost-effective manner. The out-of-sample analysis shows that nowcasts of GDP growth using the SENI are more accurate than those using existing business cycle indicators. Therefore, the SENI provides accurate and timely information about the state of the economy. I then investigate the role of the state of information, and the predictive performance over the business cycle. The SENI provides more accurate forecasts once one month of daily information of the current quarter is available. This reflects that daily fluctuations in text-based indicators are volatile and noisy. However, considering that monthly survey data for the current quarter are released at the start of the following month, the SENI still contains valuable information even after accounting

[1] Nowcasting refers to the problem of predicting the present, the very near future, and the very recent past. See Bańbura et al. (2012) and Bańbura et al. (2013) for extensive surveys.

for realistic publication lags. Furthermore, the findings emphasize that improvements in nowcasting performance mostly occur during turning points in the business cycle, underscoring the SENI's effectiveness in real-time recession detection.

Using news texts for nowcasting GDP growth poses various challenges. Two key challenges are the resource-intensive nature of acquiring relevant articles on a daily basis, as well as comprehensively reviewing a large volume of articles each day. To address these challenges, I propose simple solutions. First, I employ web scraping techniques to retrieve publicly available titles and lead texts from online archives and news websites.[2] This eliminates the need for multiple costly newspaper subscriptions. Second, I use text mining methods to extract relevant information. This approach automates the process of reading the articles and extracting the information. Moreover, it enables the identification of patterns that might be relevant to track economic activity, and may not be readily apparent to human observers.

One of the main benefits of using publicly available news data is its immediate availability. The data set can be usually updated with a delay of one day. In addition, news data is available over a longer time period than other high-frequency data sources (see e.g. Becerra et al., 2020; Wegmüller et al., 2021). Furthermore, text data is usually not revised ex-post, in contrast to many macroeconomic statistics. Finally, the process of extracting information from the data is straightforward, and since the data is publicly accessible, no newspaper subscription is needed to replicate and verify the results. While the public availability of news data is a positive aspect, it should be noted that typically only the titles, lead texts or specific passages of articles are available, as opposed to full articles. However, this should not be considered a significant drawback as these lead texts often succinctly convey the main message of the article and, therefore, may include less irrelevant information. In other words, I use the information where the newspapers may have already separated the signal from some of the noise in the data. In the end it is an empirical question whether short texts are sufficient for forecasting economic variables which is also addressed in this paper.

The Covid-19 pandemic has triggered a wealth of studies on high-frequency indicators measuring business cycle fluctuations. Daily measures using news texts have been

[2] Some of the data sources provide title and lead texts while one source provides small excerpts of newspaper articles. In what follows "lead texts" refer to all collected texts.

proposed by Shapiro et al. (2020) and Burri and Kaufmann (2020).[3] Using data on real economic activity, weekly indices to track economic activity have been developed for various countries (see e.g. Lewis et al., 2022; Wegmüller et al., 2021).[4] News texts are also increasingly used for forecasting economic variables.[5] Ardia et al. (2019) shows that news sentiments improve forecasts of U.S. industrial production growth. Ellingsen et al. (2021) find that news data are particularly informative for forecasting consumption developments. Moreover, Barbaglia et al. (2022) and Kalamara et al. (2022) demonstrate that it improves forecasts of macroeconomic variables such as GDP, inflation, and unemployment. Textual data has also been used to measure uncertainty (Baker et al., 2016; Larsen, 2021) and recession prevalence (The Economist, 2011). Iselin and Siliverstovs (2013) show that counting the term "recession" in newspapers improves predictions of GDP growth.

This paper makes four main contributions to the literature. First, it investigates the information content of daily news lead texts for measuring business cycle fluctuations. So far, most studies evaluated the information content of monthly indicators. Although Shapiro et al. (2020) also construct daily indicators, they do not investigate how many daily observations are needed to make an accurate GDP growth nowcast. Second, unlike previous studies that relied on expensive data sources, this research utilizes publicly available short news lead texts, making it more accessible and cost-effective. Third, an innovation of this study lies in the integration of recession indicators with sentiment indicators. While studies exist, evaluating the predictive value for both separately, their combined value has not been examined before. Finally, the methodology to create news-based indicators, as well as the aggregation process is simple and based on economic intuition. This makes it more accessible for policy makers and the public at large. Moreover, the news-based indicators are robust to data leakage, as opposed to approaches using topic models.

The paper proceeds as follows. In the next section, I describe the text data. Section 3 presents the methodology. Section 4 evaluates the text-based business cycle indicator in-

[3]Research on the use of textual data for measuring and forecasting economic activity has been conducted long before the Covid–19 pandemic (See Bybee et al., 2021; Larsen & Thorsrud, 2018; Shiller, 2019; Thorsrud, 2020, for a few examples).

[4]There are various other initiatives, that aim to provide reliable high-frequency information on the economy. These initiatives propose various novel approaches to construct composite indexes based on data sources such as search engine data, labor market data, debit and credit card data, traffic and mobility data, payments and cash withdrawals data (Becerra et al., 2020; Brown & Fengler, 2020; Eckert et al., 2020; Eckert & Mikosch, 2020).

[5]Short news texts are also used to predict asset prices. Li et al. (2022) and Y. Bai et al. (2022) show that news headlines can effectively be used to predict future prices.

and out-of-sample. The last section concludes.

2 Data

I use publicly and quickly available news data to create sentiment and recession indicators based on three Swiss newspapers. The newspaper data stem from the online archives of the *Tages-Anzeiger* (TA), the *Neue Zürcher Zeitung* (NZZ) and the *Finanz und Wirtschaft* (FUW)[6]. These newspapers are three of the most relevant German-language newspapers reporting on economic affairs in Switzerland and abroad. Their archives cover the period from 2000 at the latest until today.[7] For this paper, I use data from January 1, 2000 to December 31, 2021. Since the

The publicly available news data primarily comprises titles, lead texts, or specific passages of articles, which tend to concisely capture the main message of the article and, consequently, contain less irrelevant information. To further remove irrelevant information, which may reduce the signal-to-noise ratio, I utilize solely texts pertaining to the economy. Articles about subjects that are not related to the economy, like sports, may also express a sentiment, but it does not necessarily have any meaning for the economy. I assume that articles about the economy comprise specific German keywords such as *Wirtschaft*, *Konjunktur* and *Rezession* (which translate to economy, business cycle, and recession respectively).[8] As a small open economy, Switzerland is greatly affected by economic developments in other countries. To account for this, I create indicators that measure sentiments and recession prevalence for both Switzerland and foreign countries by using location-specific keywords. I use specific keywords to identify articles related to the Euro area, Germany and the USA as these countries are major trading partners of Switzerland. For example *Wirtschaft & Schweiz* or *Rezession & Deutschland* (economy &

[6]See https://www.tagesanzeiger.ch/zeitungsarchiv-930530868737, https://zeitungsarchiv.nzz.ch/archive and https://www.fuw.ch/archiv, .

[7]Sometimes the *Tages-Anzeiger* updates its archive with a relevant delay or not at all. Therefore, I additionally use lead texts from the *Tages-Anzeiger* website: https://www.tagesanzeiger.ch/wirtschaft. This ensures real-time feasibility of the approach.

[8]Why not using keywords such as *Wirtschaftsaufschwung* (economic recovery) as well? The research conducted by Becerra et al. (2020) using Google Trends data suggests that terms associated with positive sentiment do not align with changes in economic activity. This highlights that people's interest in the economy is not symmetrical. This is also reflected in the behavior of journalists, who tend to focus more on recessions than on periods of growth. This phenomenon, known as "negativity bias" is not exclusive to journalists and is well-documented in the literature, where it has been shown that people tend to pay more attention to and remember negative information over positive information (See e.g. Baumeister et al., 2001).

Table 1 — Descriptive statistics of the news data

Journal	#Texts	Avg. #Words	Avg. Sentiment	Coverage
Finanz und Wirtschaft	100’084	15.8	0.043	2000 - 2021
Neue Zürcher Zeitung	720’530	62.7	0.038	2000 - 2021
Tages Anzeiger	29’359	19.2	0.015	2000 - 2021
Tages Anzeiger Webpage	54’288	16.2	0.003	2008 - 2021

Notes: The total number of texts is not a unique count of articles. It is the total count of all articles satisfying the search queries represented in Table 5 in the appendix. The average number of words is calculated from the cleaned texts as outlined in Section 3. The average sentiment is calculated as the total number of positive minus the total number of negative words as defined by Remus et al. (2010), divided by the total number of words.

switzerland, recession & germany).[9]

Table 1 presents an overview of the web-scraped text data. Over all sources, roughly 900,000 lead texts, text passages, or titles were collected. Note that these 900’000 texts are not unique articles, but rather the results of multiple search queries. Moreover, the average number of words is higher than 20 only for the *NZZ*, as it is the only newspaper that provides short passages from articles instead of just titles and lead texts.

3 Methodology

In this section, I will describe the method of extracting information from the data, discuss the creation of various sub-indices covering different areas of the economy, and explain the process of aggregating them into a business cycle indicator. Additionally, I will provide an explanation of the models used for the out-of-sample nowcasting exercise.

3.1 Creating text-based indicators

To convert the high-dimensional and unstructured newspaper texts into time series, they have to be preprocessed (cleaned). I, therefore, filter out irrelevant information, as is common in the natural language processing (NLP) literature. I remove Hyper Text Markup Language (HTML) tags, punctuation, numbers, and stopwords, that is

[9]The & means both terms must be contained in the article. For more information on the search queries used to filter out relevant articles, see Table 5 in the Appendix.

words that are not informative, typically conjunctions such as "or" and "if". The stop words are provided by Feinerer and Hornik (2019). Finally, I transform all letters to lowercase. Many NLP applications then stem the words, which is a process of removing and replacing word suffixes to arrive at a common root form of the word. However, this is not necessary because I use the sentiment lexicon developed by Remus et al. (2010) that is not stemmed. Furthermore, unlike many NLP applications, I refrain from identifying frequently occurring phrases (known as bigrams) like "labor market" in the text corpus. This method necessitates setting a threshold based on the overall frequency of the bigram in the text corpus, which introduces a look-ahead bias by utilizing future information to assess the importance of a specific bigram (see Kalamara et al., 2022, for a more detailed discussion).

With the cleaned data, I then construct text-based indicators that capture various economic concepts in Switzerland and abroad. This is achieved by using different sets of keywords, which define these contexts. A detailed list of the topic defining keywords is shown in Table 2. The indicators are based on two methods.[10] First, I create recession indicators by simply counting the occurrence of keywords related to recessions and sum them up to a daily time series. The recession index, also known as the R-word index, was proposed by The Economist (2011) in the early 90's. Iselin and Siliverstovs (2013) create an R-word index for Switzerland and find that it has predictive power to forecast Swiss GDP growth. The created recession indicators are highly correlated with uncertainty indicator (Baker et al., 2016). Therefore, they do not only measure the perception of a potential recession, but also are related to economic uncertainty. [11] Hence, these indicators measure uncertainty in the economy and recession prevalence and are therefore negatively correlated with the business cycle.

[10] The procedure is documented in detail in algorithm 1 in the Appendix.

[11] In a robustness check, I additionally considered uncertainty indicators. However, they did not comprise additional predictive power to the recession and sentiment indicators. For brevity, I do not report the results.

Table 2 — Keywords for economic topics

Topic	Keywords	English	Method
Recession	rezession, krise	recession, crisis	Count
Labor market	arbeit, job, beschäftigung	labor, job, employment	KWIC
Financial market	stock, asset, anlage, aktionär, aktie, dividend, börse, finanz, \bsmi\b, dax, \bspi\b, nasdaq, msci, wechselkurs	stock, asset, investment, share, dividend, financial,\bsmi\b, dax, \bspi\b, nasdaq, msci, exchange rate	KWIC
Government	regierung, staat, minister, govern, \bbund\b, steuer, politik	government, state, minister, federal, tax, policy	KWIC
Investment	invest	invest	KWIC
Economy	wirtschaft, konjunktur, industrie, handel, import, export	economy, business cycle, industry, trade, import, export	KWIC
Inflation	inflation, teuerung, preis	inflation, price	KWIC

Notes: The column 'English' lists contextual translation of the German words. The queries use wildcard operators (i.e "spi" also matches spillovers). The symbol \b reverses the wildcard operator (i.e \bspi\b doesn't match spillovers). The indices are created with two different methods. "Count" means simply counting all occurring keywords, "KWIC" calculates sentiment indices based on the keyword-in-context method as explained in Section 3.

The second approach, known as the keyword-in-context (KWIC) method (Luhn, 1960), utilizes word co-occurrences to calculate topic-specific sentiment indicators. This is done by creating new sets of documents by screening the texts for topic defining keywords. Whenever a keyword is found, the keyword, along with the ten preceding and ten following words, is extracted into a new document, d.[12] A sentiment score is then calculated for each of these documents. Thus, the sentiment score is local in the sense that it considers only the text related to a topic of interest. Denote by $\mathcal{P}$ and $\mathcal{N}$ the list of phrases identified as pertaining to positive and negative sentiment derived by Remus et al. (2010). The sentiment score subtracts the counts of words in $\mathcal{N}$ from the counts of terms in $\mathcal{P}$ in document d, and scales it by the number of total terms in document d. This is also referred to as the lexical methodology (see, e.g., Ardia et al., 2019; Shapiro et al., 2020; Thorsrud, 2020). More formally, let $w_{t,d,i,j} = (w_{t,d,i,j,1}, w_{t,d,i,j,2}, ...w_{t,d,i,j,N_{t,d,i,j}})$ be the list of terms in document d at date t for topic j. i is either "domestic" or "foreign". For simplicity, I drop the subscript i in what follows. The document-level news sentiment is hence given by

[12] This means I use a context window of ten words. I have also tested context windows of five or fifteen words. However, this does not significantly change the results.

Figure 1 — Document-level sentiment score

Lead text of FUW from March 6, 2020

Before cleaning
Das Coronavirus trifft die Schweizer Wirtschaft ab dem Frühjahr mit voller Kraft. Volkswirte stimmen auf ein schwaches zweites Quartal ein. Konsum, Tourismus und Exportindustrie leiden bereits. Der Bundesrat sollte die Kurzarbeit ausweiten.

After cleaning
coronavirus trifft schweizer wirtschaft frühjahr voller kraft volkswirte stimmen schwaches zweites quartal konsum tourismus exportindustrie leiden bereits bundesrat kurzarbeit ausweiten

---

In English

Before cleaning
The corona virus is hitting the Swiss economy with full force from the spring. Economists are predicting a weak second quarter. Consumption, tourism and the export industry are already suffering. The Federal Council should extend short-time work.

After cleaning
corona virus hitting swiss economy spring full force economists predicting weak second quarter consumption tourism export industry already suffering federal council extend short time work

Notes: Example of how document-level sentiment scores for two topics are calculated based on a newspaper lead from FUW. For the general economy topic that is defined by the keyword (in blue) *wirtschaft*, the number of negative words (in red) is subtracted from the number of positive words (in green) within the ten preceding and following words from the keyword, and this result is divided by the total number of words. In this case, the sentiment score is $S_{t,d,\text{economy}} = (1-1)/14 = 0$. Note that there are only 14 words in the denominator because the keyword is close to the beginning of the text. The sentiment score for the second matching keyword is given by $S_{t,d,\text{economy}} = (1-2)/19 = -0.05$.

$$S_{t,d,j} = \frac{\sum_n \mathbb{1}(w_{t,d,j,n} \in \mathcal{P}) - \sum_n \mathbb{1}(w_{t,d,j,n} \in \mathcal{N})}{N_{t,d,j}}$$

where $N_{t,d,j}$ is the number of terms in the document. Figure 1 provides a more intuitive example of how the document-level sentiment score is being calculated. Finally, daily news sentiment indicators, $S_{t,j}$, for the domestic and foreign economy and for a given topic j are calculated as a simple average of the sentiment scores.

A number of studies have used a probabilistic topic model to classify articles into topics (see, e.g., Ellingsen et al., 2021; Hansen et al., 2018; Thorsrud, 2020). Using a topic model, articles can be classified into distinct topics based on their word content, enabling the assignment of similar-worded articles to the same topics. However, due to the possibility of articles discussing multiple topics, assigning the article-level sentiment score to a specific topic becomes challenging. In comparison, my approach offers greater specificity by focusing solely on the topic-specifying keyword and calculating the sentiment score based on a few surrounding words. Moreover, the brevity of the texts and occasional absence of complete passages pose difficulties in accurately estimating a topic model (Yan et al., 2013). Finally, employing the proposed approach, using keywords not tied to any economic events, helps mitigate potential data leakage issues, which could adversely impact the nowcasting exercise (see Kalamara et al., 2022).[13]

3.2 Creating aggregate indicators

The objective of this paper is to assess the informational content of daily news-based text indicators for nowcasting quarterly Swiss GDP growth. To accomplish this, models allowing for the inclusion of time series of varying frequencies in the same regression are necessary. **Mi**xed-**da**ta **s**ampling (MIDAS) and bridge models are commonly used in literature for this purpose (see Baffigi et al., 2004; Ghysels et al., 2004). However, including a large number of explanatory variables in the MIDAS model can lead to parameter proliferation. Additionally, news indicators can be quite volatile and correlated with one another.[14] To effectively summarize the information content of the data and eliminate idiosyncratic noise, while avoiding parameter proliferation, I estimate a factor model in static form:

$$X = F\Lambda + e$$

The model comprises N variables and T daily observations. Therefore, the data matrix X is $(T \times N)$, the common factors F are $(T \times r)$, the factor loadings Λ are $(r \times N)$, and the unexplained error term e is $(T \times N)$. The advantage of using a factor model is that it allows for summarizing the information in a large data matrix X with a small

[13]Nevertheless, I have tested two topic models specifically designed for short texts. The first is an algorithm that models co-occurrences of bi-terms (bi-terms are pairs of words appearing together in a text) and the second is a structural topic model, that is a general framework for topic modeling with document-level covariate information (see Roberts et al., 2014; Yan et al., 2013). The results were significantly worse than with the method used here.

[14]The news indicators are rather volatile (see Figures 8 and 9 in the Appendix). I, therefore, compute a one-sided ten-day moving average before including them in the factor model. Comparable studies smooth their news sentiments with a sixty-day or higher moving average (see e.g. Shapiro et al., 2020; Thorsrud, 2020). The choice of the moving average time window is a trade-off between less volatility and more timeliness.[15]

number of common factors r. Factors and loadings can be estimated through principal components, under the assumption that the idiosyncratic components are only weakly serially and cross-sectionally correlated (J. Bai & Ng, 2013; Stock & Watson, 2002). [16]

Given that the construction of the indicators is based on economic reasoning, the first principal component of the static factor model can be interpreted as a coincident business cycle indicator. Moreover, an information criterion to determine the number of factors in approximate factor models proposed by J. Bai and Ng (2002) confirms that one factor is representing the data well enough.[17] I use the recommended information criterion BIC_3. In what follows, I refer to this first principal component as short economic news indicator (SENI).

3.3 Pseudo out-of-sample evaluation models

How reliable is the SENI and what is the informational content of the daily frequency? To answer these questions I perform a daily pseudo-real-time forecast evaluation using mixed frequency methods.

The variable of interest is quarterly GDP growth, which is denoted as y_{t_q}, where t_q is the quarterly time index $t_q = 1, 2, ..., T_y$, with T_y being the last quarter for which GDP figures are available. I use the real-time data set for quarterly GDP vintages by Indergand and Leist (2014). Thereby taking into account the ragged-edge structure as a result of different publication dates of official quarterly GDP figures. The aim is to now- and forecast quarterly GDP growth, y_{Ty+H+1} with a horizon of $H = 0, 1$ quarters. I use this notation to emphasize that a horizon of $H = 0$ corresponds to a nowcast, whereas $H = 1$ is a forecast.

Similar to Kuzin et al. (2011) and Schumacher (2016), I assume that the information set for now- and forecasting includes one stationary daily indicator x_{t_d} in addition to the real-time GDP observations. For simplicity I assume every quarter to have $D = 60$ days, reflecting approximately five working days per week and four weeks per month. Hence, the time index for the daily observations is defined as a fraction of the low-frequency quarter according to $t_d = 1 - 59/60, 1 - 58/60, ..., 1, 2 - 59/60, ..., T_x - 1/60, T_x$, where T_x is the last day for which the daily indicator is available. Nowcasts are predictions for

[16] I exclude weekends and holidays. Then, I interpolate additional missing values using an EM-algorithm (Stock & Watson, 2002), after standardizing the data to have zero mean and unit variance. I choose a relatively large number of factors for interpolating the data ($r = 4$). Finally, I use the first principal component of the interpolated data set.

[17] Nevertheless, an interesting extension would be to examine whether more than one factor comprises relevant information for Swiss economic activity. I leave this extension for future research.

horizons of $h = 0, ..., 59$ days, and forecasts are predictions for horizons of $h = 60, ...119$ days. The now- or forecast for GDP is conditional on information available in T_x, including all observations until T_x and the GDP observations up to T_y. The sample spans from January 1, 2000, to December 31, 2021. To compute the forecast errors, I use the release of quarterly GDP from December 2021.

To determine the informational content of the SENI, I forecast GDP growth on an expanding information window. The forecasts are based on three models that exploit the information contained in the high-frequency indicator and link it to the low-frequency variable. First, I estimate a MIDAS model introduced by Ghysels et al. (2004) and Ghysels et al. (2007). Second, I employ bridge equations following Baffigi et al. (2004). Third, I consider an iterative MIDAS model, which is a mixture of both as discussed by Schumacher (2016).

3.3.1 MIDAS model

The MIDAS approach is a direct multi-step forecasting tool. I use the following model for a forecast horizon of H quarters (using the terminology of Schumacher (2016))

$$y_{t_q+H+1} = \alpha + \sum_{p=0}^{P-1} \beta_p \sum_{k=0}^{K-1} b(k,\theta) L^{(pD+k)/D} x_{t_d+T_x-T_y} + \varepsilon_{t+H+1} \tag{1}$$

where α is a constant, P denotes the number of low-frequency lags and K is the number of high-frequency lags per low-frequency lag (both including zero). This modeling strategy is very flexible, allowing for different lag structures. I set $P = 2$ and $K = 60$, meaning the dependent variable depends on all 60 high-frequency values of the current and the last quarter. The daily lag operator is defined as $L^{1/60} x_{t_d} = x_{t_d-1/60}$. I determine the effect of the daily indicator $x_{t_d+T_x-T_y}$ on y_{t_q+H+1} by estimating regression coefficients β_p. Because x_{t_d} is sampled at a much higher frequency than y_{t_q}, I potentially have to include many high-frequency lags to achieve adequate modeling, which easily leads to overparameterization in the unrestricted linear case. To avoid parameter proliferation, I use a non-linear weighting scheme given by the polynomial $b(k, \theta)$. Note that I use the same polynomial specification for all low-frequency lags included in the model.[18]

[18] I also estimate a model with different polynomial specifications for every included low-frequency lag. However, this led to convergence issues in the NLS estimation for some periods, and hence, this deteriorates forecasting performance.

For the polynomial specification, I use the commonly employed exponential Almon lag of order two (see (Ghysels et al., 2007)):[19]

$$b(k,\theta) = \frac{exp(\theta_1 k + \theta_2 k^2)}{\sum_{j=0}^{K} exp(\theta_1 j + \theta_2 j^2)}$$

As it is shown by Ghysels et al. (2007), this functional form allows for many different shapes. The weighting scheme can for instance be hump-shaped, declining, or flat. By definition, they sum to one. Moreover, it parsimoniously represents the large number of predictors – with $P = 2$ I only have to estimate five parameters. The parameters are estimated by non-linear least squares (NLS) for each forecast horizon. Since MIDAS models are a direct forecasting tool and depend on the forecast horizon H, I have to estimate one model for every H and re-estimate them every time when new information becomes available (here every day).

3.3.2 Bridge equation

Bridge equations link the low-frequency variable and time-aggregated high-frequency indicators (See e.g. Baffigi et al., 2004; Diron, 2008; Foroni & Marcellino, 2013). This approach is a two-step procedure. In the first step, the high-frequency variable has to be forecasted to the end of the desired quarter and then aggregated over time to obtain values corresponding to the low-frequency. In the second step, the aggregated values are used in the bridge equation to forecast the low-frequency variable. I estimate a bridge model for a forecast horizon H of the following form:

$$y_{t_q+H+1} = \alpha + \sum_{p=0}^{P-1} \beta_p L^p x_{t_q+H+1} + \varepsilon_{t_q+H+1} \tag{2}$$

where α is a constant, P the number of lags and the lag operator is defined as $L^1 x_{t_q} = x_{t_q-1}$. Note that

$$x_{t_q} = \sum_{k=0}^{K-1} \omega(k) L^{k/D} x_{t_d} = \sum_{k=0}^{K-1} \omega_k L^{k/D} x_{t_d} \tag{3}$$

[19] For robustness I also use a Legendre polynomial proposed by Babii et al. (2021). The results are shown in Table 6 in the Appendix.

is the time aggregated high-frequency variable. The aggregation function depends on the nature of the indicator. I choose a simple average (i.e. $\omega_k = 1/D \quad \forall k$). The bridge equation in (2) can be estimated by OLS only on sample periods where all the high-frequency variables are available. To get a forecast of the low-frequency variable, I use an autoregressive model of order p (AR(p)), where the lag order is determined by the Bayesian Information Criterion (BIC). These predictions are then aggregated according to equation (3) and plugged into the estimated equation (2).

3.3.3 Iterative MIDAS

The iterative MIDAS (Midas-IT) model is an intermediate model between bridge and MIDAS (Schumacher, 2016). In the bridge model the aggregation function $\omega(k)$ is replaced with the restricted weighting polynomial $b(k, \theta)$. As for the bridge model I use an AR(p) model to forecast the indicator variable to the end of the desired quarters. I use the same polynomial specification as for the MIDAS model. Using these three model types allows me to identify the advantage of selected aspects of MIDAS and bridge models.

3.3.4 Benchmarks

The SENI forecasts are compared to three benchmarks. First, I use an AR(1) estimated on the corresponding real-time vintage for GDP growth. Second, using bridge equations, I forecast GDP growth using the KOF Economic Barometer, a well-known monthly composite leading indicator for Switzerland (Abberger et al., 2014). Because there is no real-time vintage of the KOF Barometer available, I suppose that the value for the current month is available three days before the month ends. This is a reasonable assumption since the Barometer is usually published towards the end of each month. Within the bridge equation, I use the mean of the KOF Barometer for a given quarter to project its value until the quarter's end. Third, I compare the forecasts to the preliminary quarterly GDP growth release for the respective quarter. Given that the quarterly GDP figures are revised after the initial release, I consider the initial quarterly GDP release to be a forecast of the final GDP outcome.

4 Evaluation of the SENI

This section first provides a descriptive analysis of the SENI and its underlying indicators. It then demonstrates its in-sample information content, highlighting that it is available earlier than most other leading indicators. In addition, it provides an evaluation of its pseudo out-of-sample performance for nowcasting (forecasting) real

GDP growth. While the focus stays on real GDP growth, it should be noted that the SENI is correlated with many key macroeconomic variables (See Figure 11 in the Appendix).

4.1 Descriptive analysis

Most of the indicators underlying the SENI are substantially correlated with GDP growth.[20] Figure 2 provides an overview of the cross-correlations of these indicators as well as the aggregate indicator. The absolute correlation coefficients range from 0.26 to 0.63. The domestic indicators tend to be more strongly correlated with GDP growth. Interestingly, the recession indicator displays the highest absolute coincident correlation. The largest correlations in absolute value appear for the coincident correlation with GDP growth. In addition, some weaker cross-correlations suggest that the text-based indicators are lagging GDP growth. As GDP growth is published with a significant delay, and is heavily revised, the coincident correlation may still comprise useful information for now- and forecasting GDP growth(see Indergand & Leist, 2014).[21] The aggregate SENI has a coincident correlation of 0.57, which is the second highest coefficient. This shows the usefulness of the aggregation step. Based on these observations, a SENI based on domestic indicators might perform better.

[20]To account for persistence in GDP growth and the SENI, the cross-correlations are shown with pre-whitened data (see Neusser, 2016, Ch. 12.1).

[21]GDP growth is published with a delay of about 9 weeks. See Table 4 in the Appendix for publication lags of relevant Swiss indicators.

Figure 2 — Cross-correlations of text based indicators with GDP growth

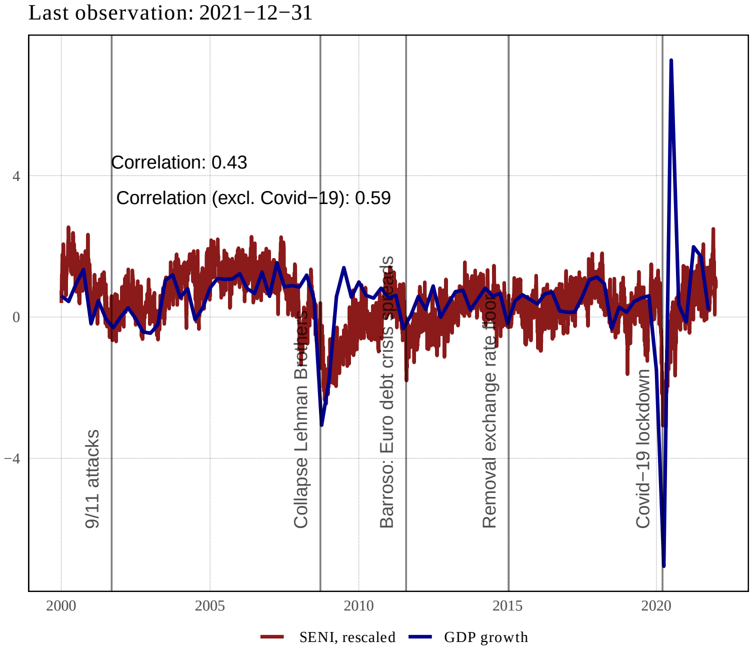
Notes: Cross-correlation between news-based sentiment indicators and real Swiss GDP growth for domestic (left) and foreign (right) topics. I aggregate all data to quarterly frequency. Only statistically significant correlations at displacement s given on x-axis are labelled. A significant correlation at $s > 0$ means the indicator is leading. Before computing the cross-correlation the series have been pre-whitened with an AR(p) model (see Neusser, 2016, Ch. 12.1). The lag order has been determined using the Bayesian Information Criterion.

Panel (a) of Figure 3 displays the SENI together with actual GDP growth, revealing the indicator closely follows economic crises. It declines during the downturn during the Global Financial Crisis, responds to the removal of the minimum exchange rate, and the euro area debt crisis.[22] The SENI also reacts strongly to the Covid-19 crisis, as seen in panel (b). The indicator begins to drop in late February, as the impact of the Covid-19 crisis on most European countries became evident. It reaches a low point shortly after the lockdown was implemented, then gradually increases with news about economic stimulus and the easing of lockdown measures. The low point of the Covid-19 crisis is similar to that of the Global Financial Crisis, but with a faster downturn. The crisis was also less persistent. By the end of July 2020, the SENI had improved to one-fourth of its low value during the lockdown. The indicator also reflects the health situation well, showing continuous improvement as restrictions are lifted, and rising just before tougher restrictions are imposed. The SENI is more volatile (because less smoothed) than other news-based sentiment indicators, making it difficult to visually assess the

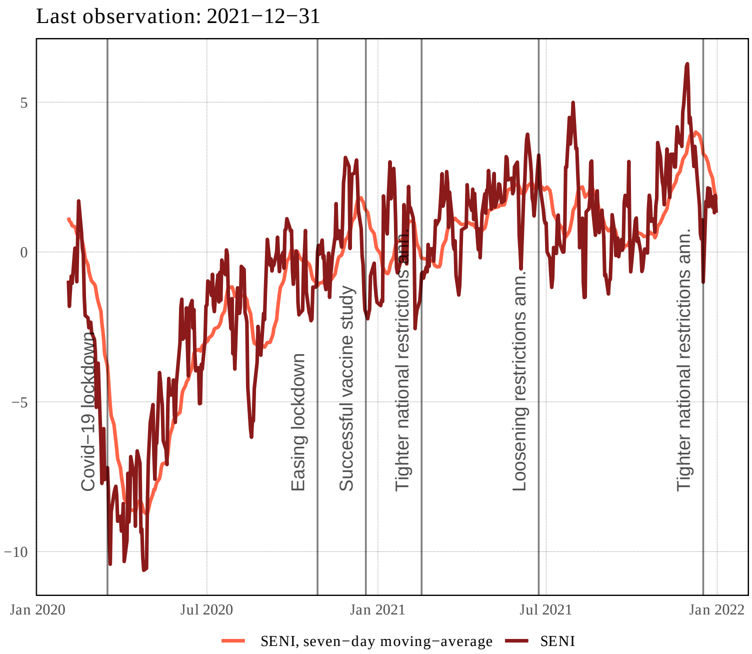
[22]On January 15, 2015, the Swiss National Bank abandoned its long-standing policy of maintaining a minimum exchange rate of 1.20 Swiss francs per euro. This move resulted in the Swiss franc appreciating dramatically against the euro and other major currencies. The decision had far-reaching implications for currency markets and led to significant volatility in the financial markets.

Figure 3 — A short economic news indicator for the Swiss economy

(a) Correlation with real GDP growth



(b) Evolution during the Covid–19 crisis



Notes: Panel (a) compares the SENI (rescaled) to quarterly GDP growth. Panel (b) panel gives daily values of the SENI along with important policy decisions.

state of the economy, however this makes the SENI more timely and as shown in the next sections, more accurate for nowcasting.

4.2 In-sample evaluation

Because the SENI is a combination of sentiment indicators covering several economic topics, it is correlated with many key macroeconomic variables (See Figure 11 in the Appendix). Since it is not optimized to track any particular measure of economic activity in its current form, I evaluate the in-sample information content of the SENI.

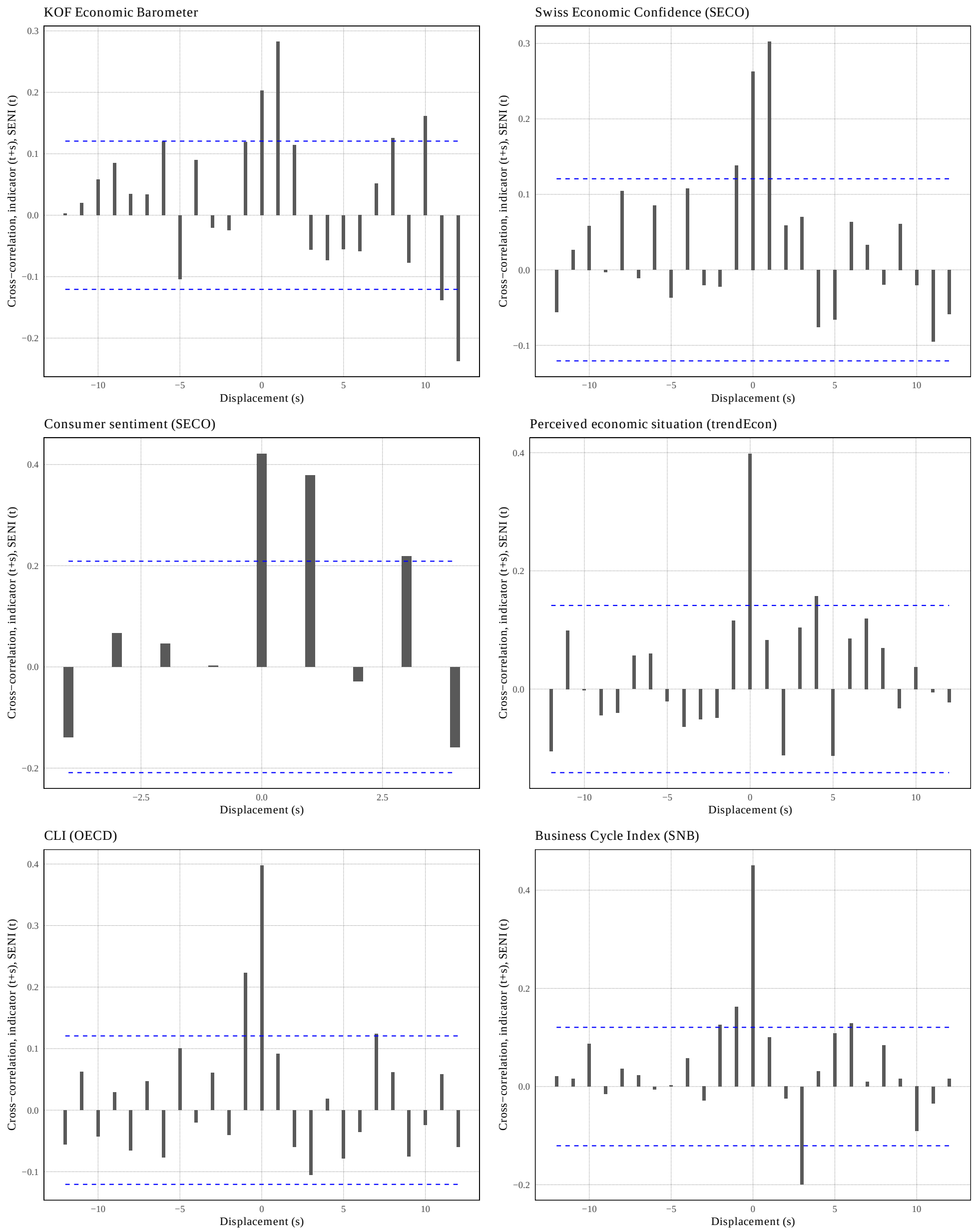
To compare the in-sample information content of the SENI to other leading indicators, I perform a cross-correlation test (see Neusser, 2016, Ch. 12.1).[23] Figure 4 shows a substantial correlation between the SENI and many prominent leading indicators. There is a coincident or a leading relationship with consumer confidence and trendEcon's perceived economic situation (Becerra et al., 2020).[24] There is a leading, coincident, and lagging relationship with the KOF Economic Barometer and SECO's Swiss Economic Confidence (SEC, Wegmüller & Glocker, 2019).[25] A coincident and a lagging correlation exists with the Organisation for Economic Co-operation and Development composite leading indicator (OECD CLI, OECD, 2010) and the SNB's Business Cycle Index (BCI, Galli, 2018). However, the OECD CLI is a smoothed indicator and is subject to substantial revisions and the BCI is published with a relevant delay. The in-sample analysis shows that the SENI provides information comparable to other indicators, with the added benefit of being more quickly accessible or covering a longer time frame.

[23] It is noteworthy that other indicators are estimated or smoothed such that they undergo substantial revisions over time. Moreover, some of the indicators are published with significant delays (see, e.g., Galli, 2018). Finally, some are based on lagged data (see, e.g., OECD, 2010). See Table 4 in the Appendix for details.

[24] All data sources are given in Table 4 in the Appendix.

[25] Both are composite indicators. The KOF barometer combines various economic indicators, including surveys of businesses, to assess the overall economic outlook. The SEC contains 30 domestic survey indicators with favourable leading properties.

Figure 4 — Cross-correlation with other indicators



Notes: Cross correlation between the SENI and other prominent leading and sentiment indicators. I aggregate all data either to quarterly frequency (consumer sentiment) or monthly frequency (remaining indicators). The dashed lines give 95% confidence intervals. A bar outside of the interval suggests a statistically significant correlation between the indicators at a lead/lag of s. Before computing the cross-correlation the series have been pre-whitened with an AR(p) model (see Neusser, 2016, Ch. 12.1). The lag order has been determined using the Bayesian Information Criterion. The only exception is the OECD CLI for which an AR(4) model is used.

4.3 Out-of-sample evaluation

Table 3 presents the relative root-mean-squared errors (RMSE) of the pseudo-real-time out-of-sample nowcasting exercise. To gain a deeper understanding of the results, I conduct a sub-sample analysis by excluding quarters two and three of 2020. Further, I exclude the Great Financial Crisis from 2008 to 2009 as a robustness check. The models used in the analysis do not significantly outperform the AR(1) model in forecasting GDP growth for horizons of 60 days and beyond (Table 3, panel (a)). This observation holds true across all sub-samples. However, when examining nowcasts for the current quarter's GDP growth (horizons of 0 to 59 days), the models consistently exhibit significantly lower RMSE. The full-sample nowcast (horizon of 0) is slightly lower than the benchmark, but this difference is not statistically significant due to large forecasting errors during the Covid-19 crisis.

Table 3 — Real-time evaluation: Relative RMSE and DMW tests

	Full sample				**Excluding Covid crisis**				**Excluding all crisis periods**			
Horizon	0	39	79	119	0	39	79	119	0	39	79	119
Hypothesis: RMSE Model < RMSE AR(1) model												
Bridge	0.67	0.86**	1.01	1.1	0.87*	0.81*	1.02	1.03	0.86**	0.78*	1.01	0.97
Midas	0.69	0.92	0.99	1.08	0.78**	0.92	1.11	1.02	0.74***	0.9	1.11	0.98
Midas-IT	0.72	0.86*	0.98	1.07	0.8**	0.78**	0.99	0.96	0.77***	0.73**	0.98	0.88
Hypothesis: RMSE Model < RMSE Barometer bridge												
Bridge	1.07	0.96	0.96	1.1	0.87*	0.9	1	1.05	0.76**	0.82**	0.94	0.99
Midas	1.12	1.01	0.94	1.07	0.78**	1.01	1.09	1.03	0.66***	0.95	1.03	1
Midas-IT	1.17	0.96*	0.93	1.06	0.8**	0.86**	0.98	0.98	0.68***	0.77**	0.91	0.9
Hypothesis: RMSE Model > RMSE First Release												
Bridge	2.03*				1.2				1.22**			
Midas	2.13*				1.09				1.08			
Midas-IT	2.22*				1.08				1.05			

Notes: Relative Root-mean-squared errors (RMSE) for forecasts with selected forecast horizons. A lower RMSE implies higher predictive accuracy compared to the benchmark. Horizons from 0 to 59 (60 - 119) denote forecasts for the current (next) quarter. The full sample spans from January 1st 2002 to December 31st 2021. The sample exluding the Covid-19 crisis excludes quarters 2-3 of 2020. The sample excluding all crisis periods additionally excludes the GFC from 2008 -2009. I use three benchmarks. First, I use an AR(1) model (panel a). Second, I use a bridge model with the KOF Economic Barometer (panel b). Third, I use the first quarterly release of the corresponding quarter (panel c). Note, I restrict the sample in panel (c) to years where the GDP figures already include official annual figures by the SFSO (up to 2020). The Diebold-Mariano-West (DMW) test provides a p-value for the null hypothesis of equal predictive accuracy against the alternative written in the row header (Diebold & Mariano, 2002; West, 1996). I assume a quadratic loss function. Significance levels are given by: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Similar patterns are observed when employing a bridge equation model with the KOF

Economic Barometer as the benchmark. In panel (b) of Table 3, it is evident that the SENI does not exhibit superior performance compared to the KOF Barometer in forecasting GDP growth for the next quarter. However, focusing on the current quarter nowcast for both sub-samples, it becomes apparent that the SENI significantly outperforms the Barometer bridge model. The relative RMSEs for the sample excluding the Covid-19 crisis are slightly higher than those for the sample also excluding the GFC.

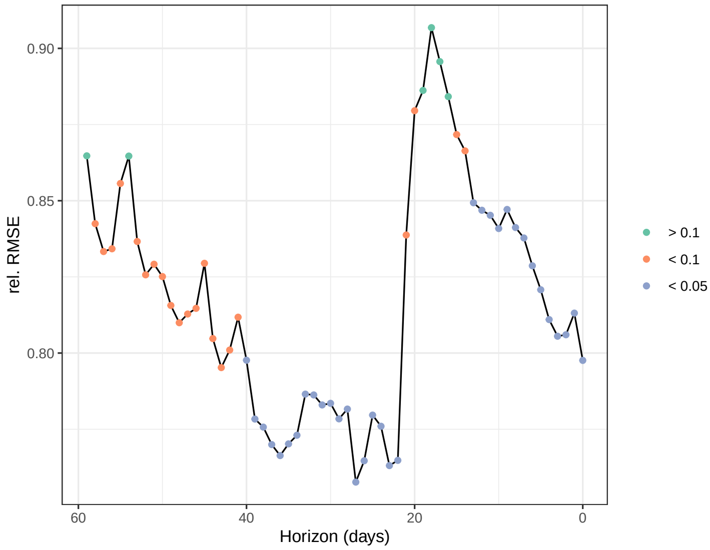
Moving to panel (c), the RMSE of the SENI is higher than that of the first official GDP release, but the difference is mostly not statistically significant in either sub-sample. The key advantage of the SENI is its ability to provide a full quarter's forecast approximately two months earlier than the first GDP release. Note that as the current vintage of GDP is subject to future revisions, particularly with the inclusion of annual GDP estimates based on comprehensive firm surveys by the SFSO, I restrict the sample in panel (c) to years where the GDP figures already incorporate these annual figures (up to 2020).

The results show that the MIDAS-IT model outperforms the other two models across most nowcasting horizons, exhibiting lower RMSE values.[26] The bridge model and the MIDAS model perform comparably well. These findings have important implications for utilizing mixed-frequency methods with daily data for nowcasting. Firstly, the bridge model's direct forecasting approach shows more promise compared to the iterative approach of the MIDAS model. Secondly, the employment of a nonlinear polynomial specification, despite requiring estimation of more parameters, proves to be beneficial.

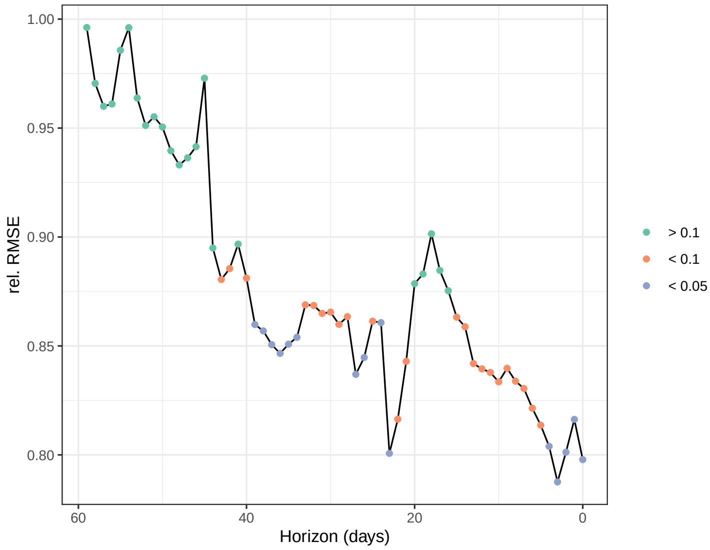
[26]Figure 13 in the Appendix presents the absolute RMSE of the out-of-sample nowcasting exercise

Figure 5 — Evolution of relative RMSE

(a) MIDAS-IT vs. AR(1)



(b) MIDAS-IT vs. KOF Barometer bridge



Notes: Relative Root-mean-squared errors (RMSE) for current quarter nowcasts with horizons from 59 to 0 days. Periods of the Covid–19 crisis are excluded. A lower relative RMSE implies higher predictive accuracy compared to the benchmark. I use two benchmarks. First, I use an AR(1) model (panel a). Second, I use a bridge model with the KOF Economic Barometer (panel b). The Diebold-Mariano-West (DMW) test provides a p-value for the null hypothesis of equal predictive accuracy against the alternative that the MIDAS-IT model is more accurate (Diebold & Mariano, 2002; West, 1996). I assume a quadratic loss function. Significance levels are given by: ● $p > 0.1$, ● $p < 0.1$, ● $p < 0.05$, ● $p < 0.01$

The performance of the SENI compared to the benchmark AR(1) model and the bridge model with the KOF Barometer demonstrates its superiority for shorter horizons. The question arises: at what horizon does this difference become significant? Figure 5 illustrates the evolution of the relative RMSE of the MIDAS-IT model compared to these benchmark models, focusing on the sample excluding the Covid-19 crisis. The figure also presents p-values for Diebold-Mariano-West (DMW) tests, assessing the null hypothesis of equal predictive accuracy against the alternative hypothesis that the MIDAS-IT model is more accurate.

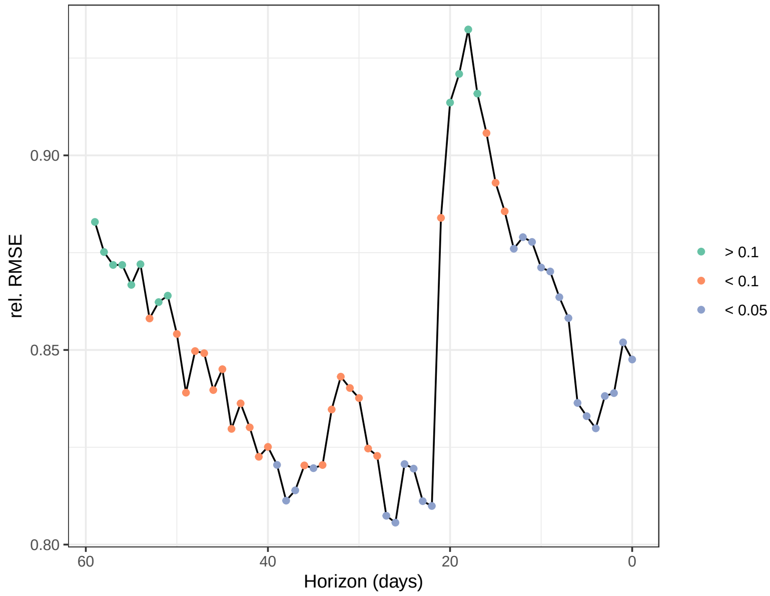
Examining panel (a) of Figure 5, which displays the relative RMSE against the AR(1) model, three key observations emerge. First, starting from a horizon of 53 days, the MIDAS-IT model significantly outperforms the AR(1) model in nowcasting the current quarter's GDP growth. Thirdly, as indicated by spikes in the relative RMSE at horizons of around 20 days, the AR(1) model demonstrates increased accuracy relative to the MIDAS-IT model when new GDP vintages get released. Finally, with exception of this peak, the more data points get available over the quarter, the more accurate the nowcasts get. Panel (b) exhibits similar behavior of the relative RMSE against the KOF Barometer bridge model. The SENI starts surpassing the KOF Barometer at a horizon of 38 days, approximately after the first month of the current quarter has passed.

A further innovation of this study lies in the integration of recession indicators with sentiment indicators. This combination has important implications for nowcasting performance. Figure 6 displays the relative RMSE values obtained solely from sentiment indicators. It becomes apparent that the inclusion of recession indicators greatly enhances the performance. Therefore, based on the findings of this research, it is recommended that text-based nowcasting endeavors take into account both sentiment indicators and recession indicators for more accurate and robust predictions of GDP growth.

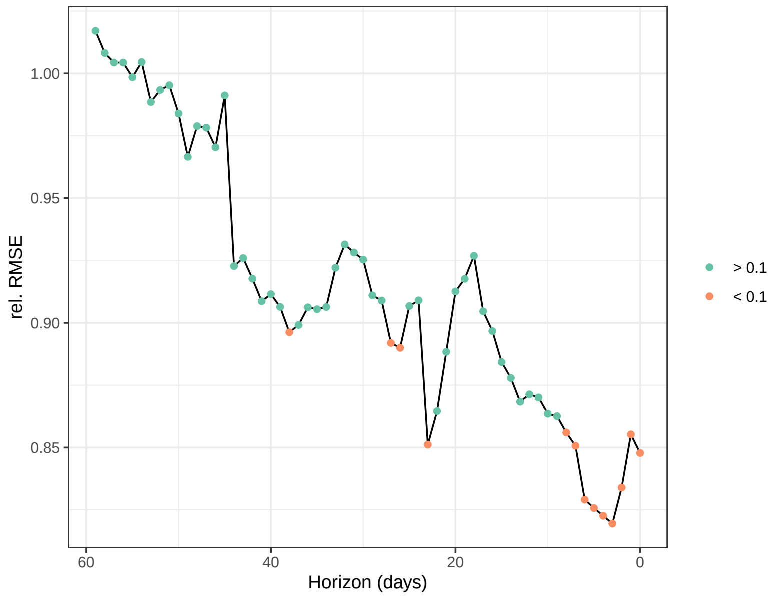
Having established the usefulness of text data for nowcasting, an important question arises: when does text data exert the most significant influence on forecast accuracy? To explore this, Figure 7 provides a comprehensive breakdown of squared error differences between the Midas.IT model (horizon 0) and the benchmark models. The lines in the graph represent the squared error differences in nowcasts compared to the benchmark model, with values below zero indicating superior performance of the text-based model. The analysis reveals that the majority of performance improvement relative to the benchmarks is concentrated during the financial crisis. Hence, forecast enhancements

Figure 6 — Evolution of relative RMSE - excl. recession indices

(a) MIDAS-IT vs. AR(1)



(b) MIDAS-IT vs. KOF Barometer bridge



Notes: Relative Root-mean-squared errors (RMSE) for current quarter nowcasts with horizons from 59 to 0 days. Periods of the Covid–19 crisis are excluded. A lower RMSE implies higher predictive accuracy compared to the benchmark. I use two benchmarks. First, I use an AR(1) model (panel a). Second, I use a bridge model with the KOF Economic Barometer (panel b). The Diebold-Mariano-West (DMW) test provides a p-value for the null hypothesis of equal predictive accuracy against the alternative that the MIDAS-IT model is more accurate (Diebold & Mariano, 2002; West, 1996). I assume a quadratic loss function. Significance levels are given by: ● $p > 0.1$, ● $p < 0.1$, ● $p < 0.05$, ● $p < 0.01$

are often observed during turning points. These findings are consistent with recent literature on the subject (see Barbaglia et al., 2022; Ellingsen et al., 2021; Kalamara et al., 2022). However, it is important to acknowledge that the text-based model exhibits certain weaknesses in nowcasting recovery periods. The Barometer bridge model, for example, produced more accurate nowcasts during the recovery from the GFC and the rebound of the Covid–19 crisis. Nonetheless, the overall strong performance of news-based predictions cannot be solely attributed to recession periods.

5 Concluding remarks

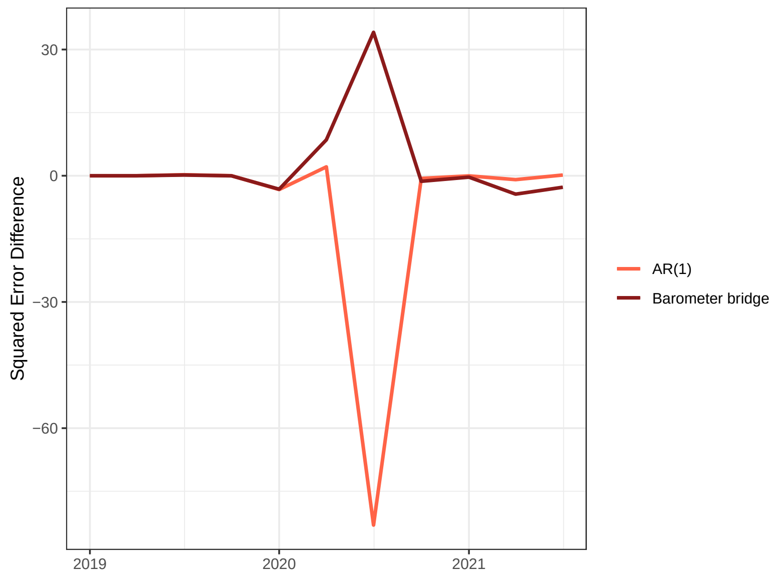
This paper explores the potential of utilizing news data, specifically daily lead texts, for nowcasting real Swiss GDP growth. The use of news data offers several advantages, including real-time insights, that are particularly valuable in dynamic situations such as the Covid-19 pandemic. To overcome the challenges associated with acquiring and reviewing multiple articles daily, the paper proposes to employ web-scraping techniques to retrieve publicly available news lead texts and to leverage text mining methods to extract relevant information. The public availability of lead texts with a maximum delay of one day makes them a valuable resource for fulfilling quick high-frequency information demands.

The study presents compelling evidence supporting the effectiveness of daily news lead texts for nowcasting Swiss GDP growth. By creating sentiment indicators for various hand-selected economic concepts and combining them with recession indicators, the paper develops a daily composite business cycle indicator. In-sample assessments demonstrate a strong correlation between this indicator and several existing business cycle indicators. Furthermore, the analysis reveals that once one month of data becomes available within a given quarter, a model using the text-based indicator outperforms a widely-used Swiss business cycle indicator in terms of nowcasting accuracy. Inline with recent literature on the subject, the analysis demonstrates that compared to the benchmark models, the text-based model significantly improves performance, particularly during turning points.

Figure 7 — Squared error differences: Midas.IT - Benchmark

(a) 2004 - 2019

(b) 2019 - 2022



Notes: Squared error differences between Midas.IT model for horizon 0 and two benchmarks models. The sample in panel (a) spans 2004 - 2019 and (b) 2019 - 2022.

The proposed approach for constructing sentiment and recession indicators presents several advantages. First, it is particularly suitable for analyzing short text data, which has proven to be challenging for topic modeling algorithms (Yan et al., 2013). Second, unlike topic modelling methods, the approach avoids data leakage since it does not incorporate future information to create the indicators. Finally, the approach is simple and based on economic reasoning, making it intuitive and easy to comprehend.

However, there are still areas with untapped potential for improvement. I have identified three specific areas that deserve attention in future research. First, customizing the lexicon specifically for economic news, as exemplified by the work of Shapiro et al. (2020), has the potential to augment the accuracy of sentiment analysis in capturing economic trends. Second, the evaluation of the predictive power derived from multiple factors and for other macroeconomic data would bolster the robustness of the forecasting model. Finally, it would be valuable to investigate the potential gains obtained by incorporating all individual indicators in a model, with the sparse-group LASSO-MIDAS model proposed by Babii et al. (2021) offering a promising avenue for exploration in this context. By delving into these areas of research, we can further deepen our understanding of utilizing text data to assess the state of the economy at a high frequency.

References

Abberger, K., Graff, M., Siliverstovs, B., & Sturm, J.-E. (2014). *The KOF economic barometer, version 2014. a composite leading indicator for the swiss business cycle* (KOF working papers No. 353). KOF Swiss Economic Institute, ETH Zurich. https://doi.org/10.3929/ethz-a-010102658

Ardia, D., Bluteau, K., & Boudt, K. (2019). Questioning the news about economic growth: Sparse forecasting using thousands of news-based sentiment values. *International Journal of Forecasting*, *35*(4), 1370–1386. https://doi.org/10.1016/j.ijforecast.2018.10.010

Babii, A., Ghysels, E., & Striaukas, J. (2021). Machine learning time series regressions with an application to nowcasting. *Journal of Business & Economic Statistics*, 1–23. https://doi.org/10.1080/07350015.2021.1899933

Baffigi, A., Golinelli, R., & Parigi, G. (2004). Bridge models to forecast the euro area GDP. *International Journal of Forecasting*, *20*(3), 447–460. https://doi.org/10.1016/S0169-2070(03)00067-0

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

Bai, J., & Ng, S. (2013). Principal components estimation and identification of static factors. *Journal of Econometrics*, *176*(1), 18–29. https://doi.org/10.1016/j.jeconom.2013.03.007

Bai, Y., Li, X., Yu, H., & Jia, S. (2022). Crude oil price forecasting incorporating news text. *International Journal of Forecasting*, *38*(1), 367–383. https://doi.org/10.1016/j.ijforecast.2021.06.006

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, *131*(4), 1593–1636. https://doi.org/10.1093/qje/qjw024

Bańbura, M., Giannone, D., Modugno, M., & Reichlin, L. (2013). Now-casting and the real-time data flow. In *Handbook of economic forecasting* (pp. 195–237). https://doi.org/10.1016/B978-0-444-53683-9.00004-9

Bańbura, M., Giannone, D., & Reichlin, L. (2012, September 18). Nowcasting. In M. P. Clements & D. F. Hendry (Eds.), *The oxford handbook of economic forecasting* (1st ed., pp. 193–224). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195398649.013.0008

Barbaglia, L., Consoli, S., & Manzan, S. (2022). Forecasting with economic news. *Journal of Business & Economic Statistics*, 1–12. https://doi.org/10.1080/07350015.2022.2060988

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323–370. https://doi.org/10.1037/1089-2680.5.4.323

Becerra, A., Eichenauer, V. Z., Indergand, R., Legge, S., Martinez, I., Mühlebach, N., Oguz, F., Sax, C., Schuepbach, K., & Thöni, S. (2020). trendEcon. https://www.trendecon.org

Brown, M., & Fengler, M. (2020). MonitoringConsumptionSwitzerland. https://public.tableau.com/profile/monitoringconsumptionswitzerland

Burri, M., & Kaufmann, D. (2020). A daily fever curve for the swiss economy. *Swiss Journal of Economics and Statistics*, *156*(1), 6. https://doi.org/10.1186/s41937-020-00051-z

Bybee, L., Kelly, B., Manela, A., & Xiu, D. (2021, October). *Business news and business cycles* (w29344). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w29344

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, *20*(1), 134–144. https://doi.org/10.1198/073500102753410444

Diron, M. (2008). Short-term forecasts of euro area real GDP growth: An assessment of real-time performance based on vintage data. *Journal of Forecasting*, *27*(5), 371–390. https://doi.org/10.1002/for.1067

Eckert, F., Kronenberg, P., Mikosch, H., & Neuwirth, S. (2020). Tracking economic activity with alternative high-frequency data. https://doi.org/10.3929/ETHZ-B-000458723

Eckert, F., & Mikosch, H. (2020). Mobility and sales activity during the corona crisis: Daily indicators for switzerland. *Swiss Journal of Economics and Statistics*, *156*(1), 9. https://doi.org/10.1186/s41937-020-00055-9

Ellingsen, J., Larsen, V. H., & Thorsrud, L. A. (2021). News media versus FRED-MD for macroeconomic forecasting. *Journal of Applied Econometrics*, jae.2859. https://doi.org/10.1002/jae.2859

Feinerer, I., & Hornik, K. (2019). *Tm: Text mining package* (manual). https://CRAN.R-project.org/package=tm

Foroni, C., & Marcellino, M. G. (2013). A survey of econometric methods for mixed-frequency data. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2268912

Galli, A. (2018). Which indicators matter? analyzing the swiss business cycle using a large-scale mixed-frequency dynamic factor model. *Journal of Business Cycle Research*, *14*(2), 179–218. https://doi.org/10.1007/s41549-018-0030-4

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2004). *The MIDAS touch: Mixed data sampling regression models* (CIRANO working papers).

Ghysels, E., Sinko, A., & Valkanov, R. (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews*, *26*(1), 53–90. https://doi.org/10.1080/07474930600972467

Hansen, S., McMahon, M., & Prat, A. (2018). Transparency and deliberation within the FOMC: A computational linguistics approach*. *The Quarterly Journal of Economics*, *133*(2), 801–870. https://doi.org/10.1093/qje/qjx045

Indergand, R., & Leist, S. (2014). A real-time data set for switzerland. *Swiss Journal of Economics and Statistics*, *150*, 331–352. https://doi.org/10.1007/BF03399410

Iselin, D., & Siliverstovs, B. (2013). The r-word index for switzerland. *Applied Economics Letters*, *20*(11), 1032–1035. https://doi.org/10.1080/13504851.2013.772290

Kalamara, E., Turrell, A., Redl, C., Kapetanios, G., & Kapadia, S. (2022). Making text count: Economic forecasting using newspaper text*. *Journal of Applied Econometrics*. https://doi.org/10.1002/jae.2907

Kuzin, V., Marcellino, M., & Schumacher, C. (2011). MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area. *International Journal of Forecasting*, *27*(2), 529–542. https://doi.org/10.1016/j.ijforecast.2010.02.006

Larsen, V. H., & Thorsrud, L. A. (2018). Business cycle narratives. *CESifo Working Paper*, (7468).

Larsen, V. H. (2021). Components of uncertainty. *International Economic Review*, *62*(2), 769–788. https://doi.org/10.1111/iere.12499

Lewis, D. J., Mertens, K., Stock, J. H., & Trivedi, M. (2022). Measuring real activity using a weekly economic index. *Journal of Applied Econometrics*, *37*(4), 667–687. https://doi.org/10.1002/jae.2873

Li, J., Li, G., Liu, M., Zhu, X., & Wei, L. (2022). A novel text-based framework for forecasting agricultural futures using massive online news headlines. *International Journal of Forecasting*, *38*(1), 35–50. https://doi.org/10.1016/j.ijforecast.2020.02.002

Luhn, H. P. (1960). Key word-in-context index for technical literature (kwic index). *American Documentation*, *11*(4), 288–295. https://doi.org/10.1002/asi.5090110403
Neusser, K. (2016). *Time series econometrics*. Springer International Publishing. https://doi.org/10.1007/978-3-319-32862-1
OECD. (2010). Review of the CLI for 8 countries. *OECD Composite Indicators*. https://www.oecd.org/fr/sdd/indicateurs-avances/44556466.pdf
Remus, R., Quasthoff, U., & Heyer, G. (2010). SentiWS - a publicly available german-language resource for sentiment analysis. *Proceedings of the seventh international conference on language resources and evaluation (LREC'10)*. http://www.lrec-conf.org/proceedings/lrec2010/pdf/490%3Csub%3EP%3C/sub%3Eaper.pdf
Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, *58*(4), 1064–1082. https://doi.org/10.1111/ajps.12103
Schumacher, C. (2016). A comparison of MIDAS and bridge equations. *International Journal of Forecasting*, *32*(2), 257–270. https://doi.org/10.1016/j.ijforecast.2015.07.004
Shapiro, A. H., Sudhof, M., & Wilson, D. J. (2020). Measuring news sentiment. *Journal of Econometrics*, S0304407620303535. https://doi.org/10.1016/j.jeconom.2020.07.053
Shiller, R. J. (2019). *Narrative economics: How stories go viral & drive major economic events*. Princeton University Press.
Stock, J. H., & Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, *20*(2), 147–162. https://doi.org/10.1198/073500102317351921
The Economist. (2011). Up means down. the economist's gauge of gloom. http://www.economist.com/node/21529079
Thorsrud, L. A. (2020). Words are the new numbers: A newsy coincident index of the business cycle. *Journal of Business & Economic Statistics*, *38*(2), 393–409. https://doi.org/10.1080/07350015.2018.1506344
Wegmüller, P., & Glocker, C. (2019). 30 indikatoren auf einen schlag. *Die Volkswirtschaft*, 19–22. https://dievolkswirtschaft.ch/de/2019/10/30-indikatoren-auf-einen-schlag/
Wegmüller, P., Glocker, C., & Guggia, V. (2021). Weekly economic activity: Measurement and informational content. *International Journal of Forecasting*, S0169207021001734. https://doi.org/10.1016/j.ijforecast.2021.10.010

West, K. (1996). Asymptotic inference about predictive ability. *Econometrica*, *64*(5), 1067–84. https://doi.org/10.2307/2171956

Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013). A biterm topic model for short texts. *Proceedings of the 22nd international conference on World Wide Web*, 1445–1456. https://doi.org/10.1145/2488388.2488514

A Appendix

A.1 Supplementary material

Algorithm 1: Keyword in context for economic sentiment analysis

1. Define sets of keywords $\mathcal{K}$ describing the j topics.
2. Define context window size ws.
3. **for** *each set of keywords* $\mathcal{K}_j$ *in* $\mathcal{K}$ **do**

 if $\mathcal{K}_j$ *is recession topic* **then**

 a. **for** *each article a in each location i* **do**

 i. Identify phrase $p = w_{t,a,i,n,j} \in \mathcal{K}_j$ that matches recession topic (per article multiple phrases can match).

 b. Calculate daily recession indicators, $S_{t,i,j}$, about the domestic and foreign economy by simply counting the matched phrases

$$S_{t,i,j} = P_{t,i,j}$$

 else

 a. **for** *each article a in each location i* **do**

 i. Identify phrase $p = w_{t,a,i,n,j} \in \mathcal{K}_j$ that matches topic j (per article multiple phrases can match).

 ii. Keep phrase p including ws terms before and after. Let $w_{t,p,i,j} = (w_{t,p,i,j,1}, w_{t,p,i,j,2}, ...w_{t,p,i,j,N_{t,p,i,j}})$ be the list of terms around phrase p. $N_{t,p,i,j}$, the total number of words is at most $2 * ws + 1$.

 iii. Count the number of positive, negative and the total number of words: $\sum_n \mathbb{1}(w_{t,p,i,j,n} \in \mathcal{P})$, $\sum_n \mathbb{1}(w_{t,p,i,j,n} \in \mathcal{N})$ and $N_{t,p,i,j}$

 b. Calculate sentiment per matched phrase as

$$S_{t,p,i,j} = \frac{\sum_n \mathbb{1}(w_{t,p,i,j,n} \in \mathcal{P}) - \sum_n \mathbb{1}(w_{t,p,i,j,n} \in \mathcal{N})}{N_{t,p,i,j}}$$

 c. Finally, daily news sentiment indicators,$S_{t,i,j}$, about the domestic and foreign economy for topic j are given by a simple average

$$S_{t,i,j} = \frac{1}{P_{t,i,j}} \sum_{p=1}^{P_{t,i,j}} S_{t,p,i,j}$$

 where $P_{t,i,j}$ is the number of matched phrases.

Table 4 — Macroeconomic data and leading indicators

	Type	Publication	Frequency	Source	Comments
GDP	Target	+9 weeks	Quarter	SECO	First publication subject to further revisions
Employment	Target	+9 weeks	Quarter	SFSO	
Registered unemployment	Target	+1 week	Month	SECO	
ILO unemployment	Target	+6 weeks	Month	SFSO	
Output gap	Target	> +4 months	Quarter	SNB	
SNB Business Cycle Index	Indicator	> +2 months	Month	SNB	
Internet search sentiment	Indicator	+1 day	Day	trendEcon	Indicator based on internet search engine
KOF Economic Barometer	Indicator	+0 days	Month	KOF	Some underlying data probably missing at the end of the sample
Consumer sentiment	Indicator	+4 weeks	Quarter	SECO	Survey during first month of quarter. Indicator published at beginning of second month
OECD CLI	Indicator	> +1 week	Month	OECD	Many underlying data are lagged two months

Notes: Publication lags between the last day of the variable frequency (i.e. last day of the quarter or last day of the month) and the publication date of a recent release. Therefore, all publication lags are approximate and may change over time. Table is from Burri and Kaufmann (2020).

Table 5 — Queries underlying news indicators

	URL	Keywords
		Domestic news
FuW	fuw.ch/unternehmen/ fuw.ch/makro/	I use all articles listed in *Makro* and *Unternehmen* and select those containing the word *schweiz** in either lead text, tag or cateogry.
NZZ	zeitungsarchiv.nzz.ch	[*konjunktur** OR *wirtschaft** OR *rezession**] AND *schweiz**
TA	tagesanzeiger.ch/zeitungsarchiv-930530868737	[*konjunktur* OR *wirtschaft* OR *rezession*] AND *schweiz*
TA Web	tagesanzeiger.ch	[*konjunktur* OR *wirtschaft* OR *rezession*] AND *schweiz*
		Foreign news
FuW	fuw.ch/unternehmen/ fuw.ch/makro/	I use all articles listed in *Makro* and *Unternehmen* and select those containing [*ausland* OR *eu* OR *euro** OR *deutsch** OR *us** OR *amerika**] in either lead text, tag or cateogry.
NZZ	zeitungsarchiv.nzz.ch	[*konjunktur** OR *wirtschaft** OR *rezession**] AND [*ausland* OR *eu* OR *euro** OR *deutsch** OR *us** OR *amerika**]
TA	tagesanzeiger.ch/zeitungsarchiv-930530868737	[*konjunktur* OR *wirtschaft* OR *rezession*] AND [*ausland* OR *eu* OR *euro* OR *europa* OR *deutschland* OR *us* OR *usa* OR *amerika*]
TA Web	tagesanzeiger.ch	[*konjunktur* OR *wirtschaft* OR *rezession*] AND [*ausland* OR *eu* OR *euro* OR *europa* OR *deutschland* OR *us* OR *usa* OR *amerika*]

Notes: Since the *Finanz und Wirtschaft* is a business newspaper, I do not restrict the search with keywords related to the economy. The asterisk (*) represents a wildcard search operator. E.g. the query *schweiz** matches also *schweizerische*. Wildcards are allowed only in the NZZ archive. Table is from Burri and Kaufmann (2020).

Figure 8 — Daily news based recession indicators

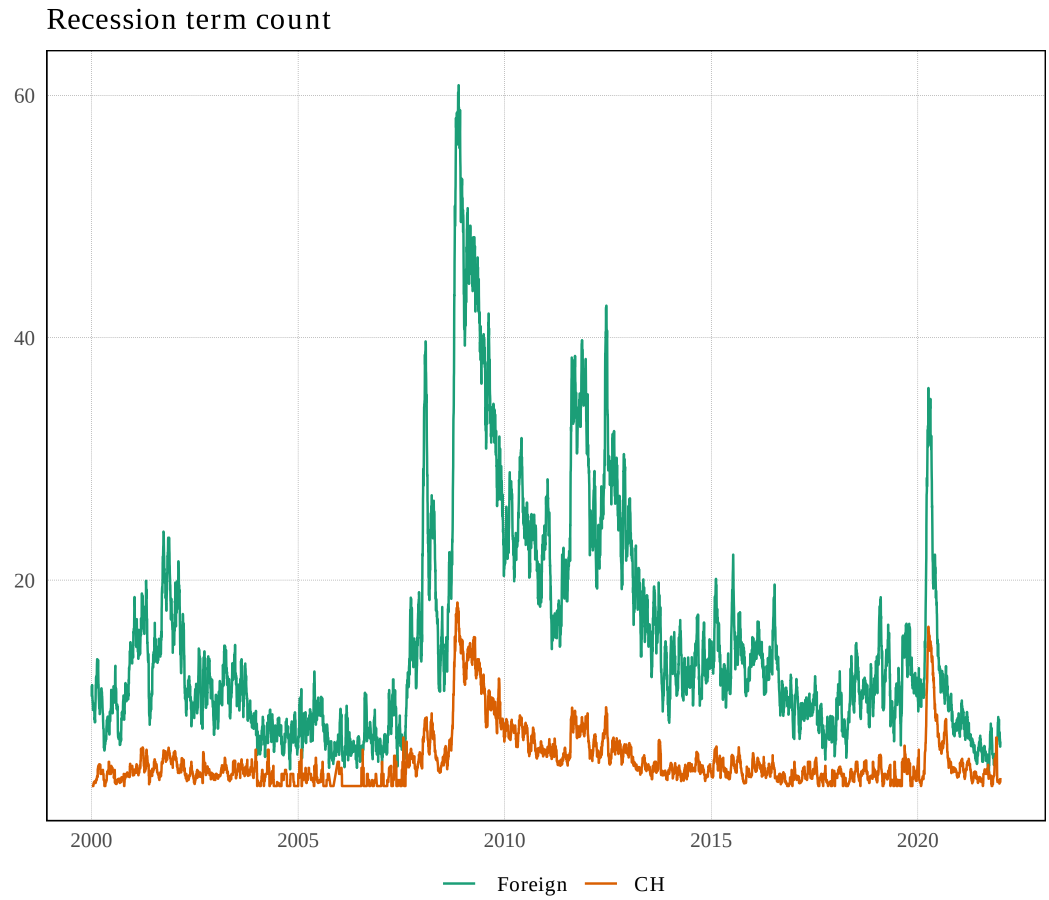


Figure 9 — Daily news based sentiment indicators

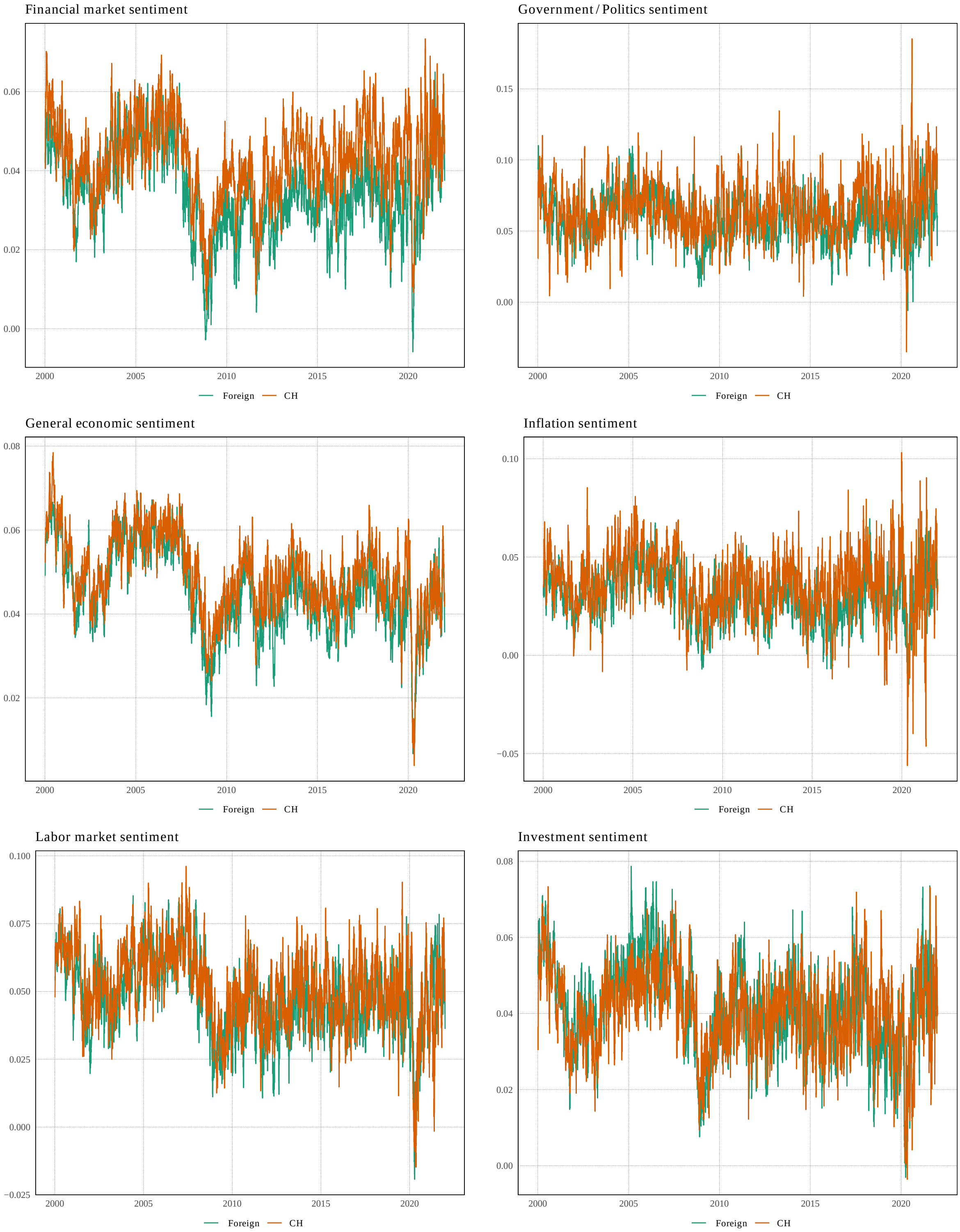
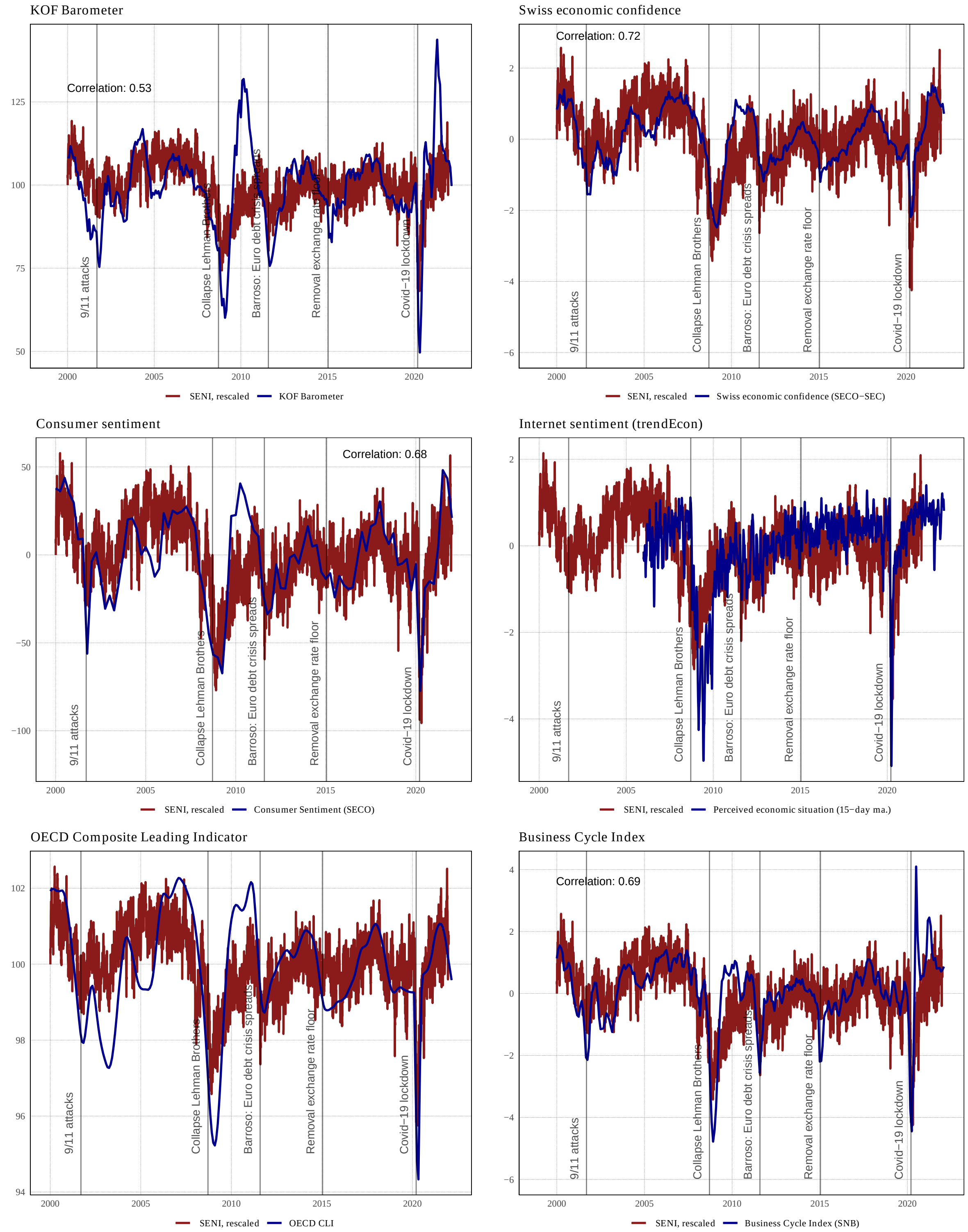
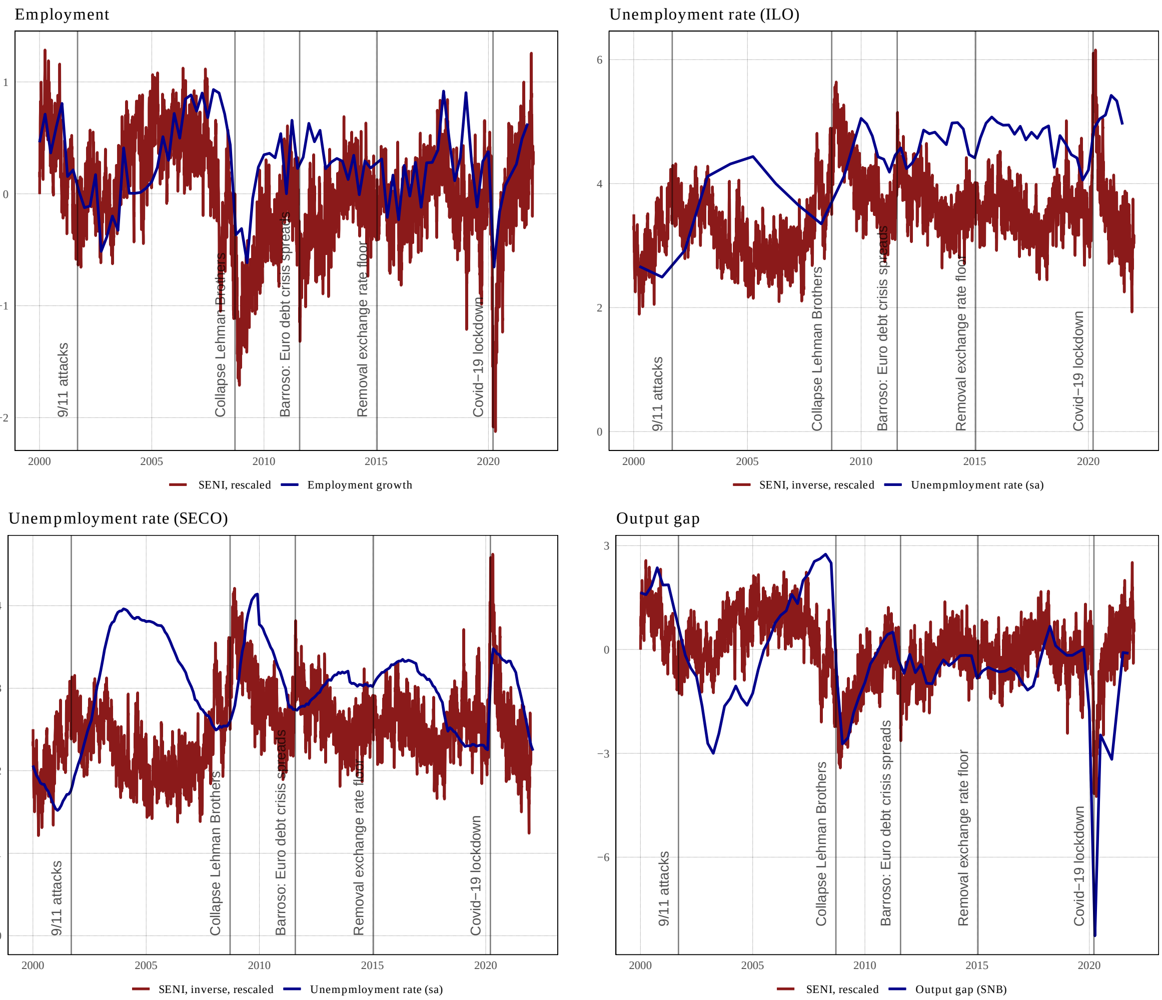


Figure 10 — Comparison with other indicators



Notes: SENI rescaled such that it roughly matches the mean and volatility of the other data series.

Figure 11 — Comparison with other macroeconomic data



Notes: SENI rescaled such that it roughly matches the mean and volatility of the other data series.

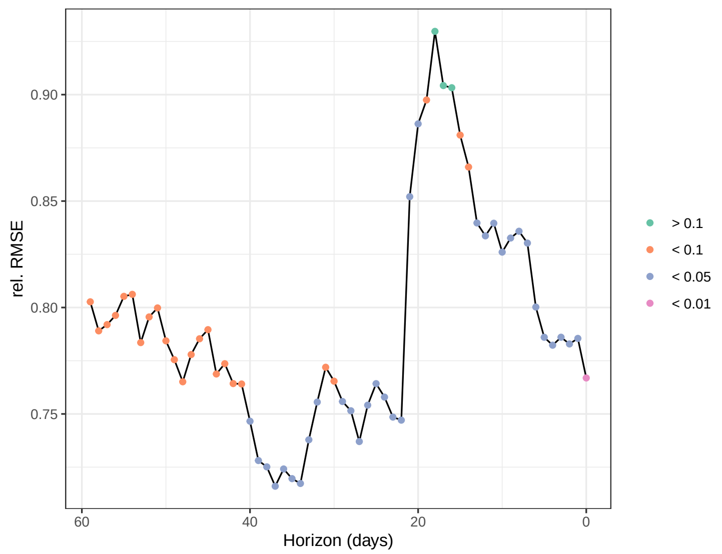
Table 6 — Real-time evaluation with legendre polynomial: Relative RMSE and DMW tests

	Full sample				Excluding Covid crisis				Excluding all crisis periods			
Horizon	0	39	79	119	0	39	79	119	0	39	79	119
Hypothesis: RMSE Model < RMSE AR(1) model												
Bridge	0.67	0.86**	1.01	1.1	0.87*	0.81*	1.02	1.03	0.86**	0.78*	1.01	0.97
Midas	0.69	0.89	1.06	1.09	0.87	0.89	1.06	1.03	0.87*	0.85	1.05	1
Midas-IT	0.69	0.9	1.02	1.1	0.87	0.83*	1.11	1.05	0.87*	0.8*	1.15	0.98
Hypothesis: RMSE Model < RMSE Barometer bridge												
Bridge	1.07	0.96	0.96	1.1	0.87*	0.9	1	1.05	0.76**	0.82**	0.94	0.99
Midas	1.12	0.99	1	1.08	0.87	0.99	1.05	1.05	0.77**	0.91	0.98	1.02
Midas-IT	1.12	1	0.97	1.1	0.87*	0.92	1.09	1.06	0.77**	0.84*	1.07	1
Hypothesis: RMSE Model > RMSE First Release												
Bridge	2.03*				1.2				1.22**			
Midas	2.12*				1.19				1.22**			
Midas-IT	2.11*				1.19				1.21**			

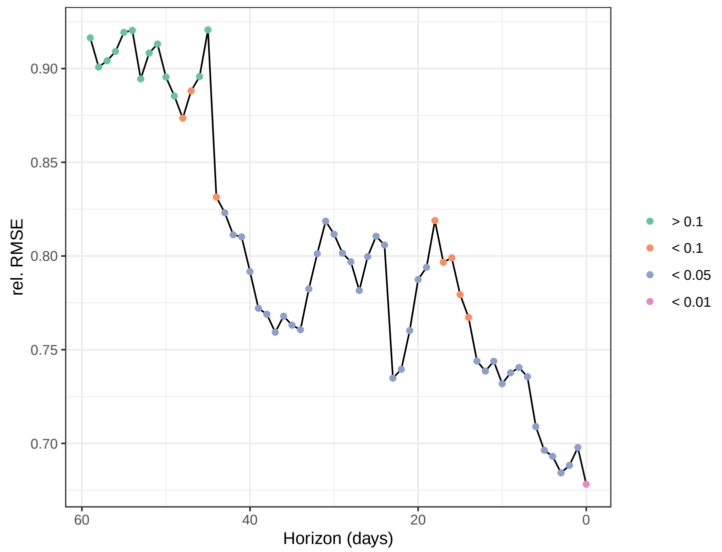
Notes: Relative Root-mean-squared errors (RMSE) for forecasts with selected forecast horizons. A lower RMSE implies higher predictive accuracy compared to the benchmark. Horizons from 0 to 59 (60 - 119) denote forecasts for the current (next) quarter. The full sample spans from January 1st 2002 to December 31st 2021. The sample exluding the Covid-19 crisis excludes quarters 2-3 of 2020. The sample excluding all crisis periods additionally excludes the GFC from 2008 -2009. I use three benchmarks. First, I use an AR(1) model (panel a). Second, I use a bridge model with the KOF Economic Barometer (panel b). Third, I use the first quarterly release of the corresponding quarter (panel c). Note, I restrict the sample in panel (c) to years where the GDP figures already include official annual figures by the SFSO (up to 2020). The Diebold-Mariano-West (DMW) test provides a p-value for the null hypothesis of equal predictive accuracy against the alternative written in the row header (Diebold & Mariano, 2002; West, 1996). I assume a quadratic loss function. Significance levels are given by: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Figure 12 — Evolution of relative RMSE on sample excluding crisis periods

(a) MIDAS-IT vs. AR(1)

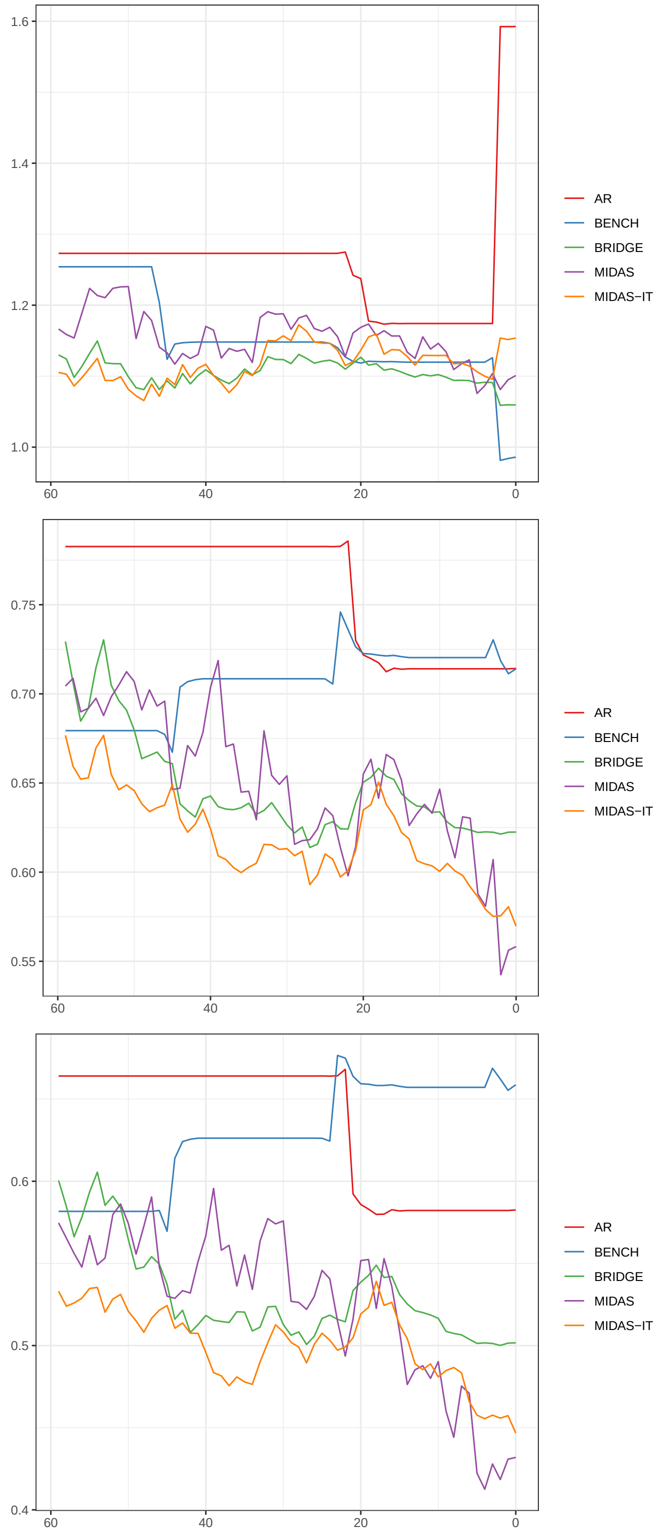


(b) MIDAS-IT vs. KOF Barometer bridge



Notes: Relative Root-mean-squared errors (RMSE) for current quarter nowcasts with horizons from 59 to 0 days. Periods of the Covid–19 crisis are excluded. A lower RMSE implies higher predictive accuracy compared to the benchmark. I use two benchmarks. First, I use an AR(1) model (panel a). Second, I use a bridge model with the KOF Economic Barometer (panel b). The Diebold-Mariano-West (DMW) test provides a p-value for the null hypothesis of equal predictive accuracy against the alternative that the MIDAS-IT model is more accurate (Diebold & Mariano, 2002; West, 1996). I assume a quadratic loss function. Significance levels are given by: ● $p > 0.1$, ● $p < 0.1$, ● $p < 0.05$, ● $p < 0.01$

Figure 13 — Pseudo real-time evaluation of real GDP growth: Absolute RMSE



Notes: RMSE for the current quarter nowcast. From top to bottom: Full sample, sample excluding Covid–19 crisis, sample excluding Covid–19 and GFC.