Do daily news abstracts help nowcasting Swiss GDP growth?*

Marc Burri†

February 28, 2023

Abstract

This paper evaluates whether daily news abstracts help nowcasting Swiss GDP growth. I collect publicly available titles and lead texts from three Swiss newspapers and calculate text-based indicators covering various aspects of the economy. A composite indicator calculated from these measures is highly correlated with macroeconomic data and survey indicators of Swiss economic activity. In a pseudo out-of-sample forecasting exercise, the indicator outperforms a well-known Swiss business cycle indicator for nowcasting Swiss GDP growth if one month of information is available. The results show that short news texts are particularly useful for nowcasting during non-crisis periods. Furthermore, the daily availability of the data makes it possible to identify turning points at an early stage.

Keywords: Switzerland, mixed-frequency data, composite leading indicator, news sentiment, uncertainty, natural language processing

*I thank David Ardia, Daniel Kaufmann, and Philipp Wegmüller for helpful discussions and participants at 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, is more rapidly available than survey or hard data traditionally used by policymakers in their decision-making process and for forecasting economic variables. While traditional indicators have been reliable in the past, especially during non-crisis periods, the current global situation requires a more agile approach to help make far-reaching decisions in rapidly changing situations. Newspaper articles reporting on various areas of the economy are available daily. Decision-makers who read one or more of these articles could use the information to make decisions but must be aware that the authors may be mistaken. Nonetheless, the collective intelligence of many authors from multiple newspapers can be used to approximate reality more accurately. Obtaining all of these articles, however, is associated with considerable costs, such as subscribing to multiple newspapers. Additionally, it is impossible to read through all of these articles every day.

To address these issues, I propose a simple solution: using web scraping techniques to download publicly available abstracts and excerpts of newspaper articles from online archives and news websites, and applying text mining to extract relevant information. This method can save time and costs associated with manually reading multiple articles and subscribing to newspapers. Additionally, by using text mining techniques, it is possible to extract key information and identify patterns that may not be immediately obvious to the human eye. By using web-scraped and text-mined news articles, policymakers can stay informed of the latest developments and make more informed decisions. The question remains whether the information contained in these daily news abstracts can also be used to forecast economic variables.

This paper presents evidence that daily news abstracts contain valuable information for forecasting Swiss gross domestic product (GDP) growth. A context-based method is proposed to create various sentiment indicators, as well as indicators for uncertainty and recession. These indicators are combined to create a composite business cycle indicator, referred to as the "Short Economic News Indicator" (SENI). The results of an in-sample assessment indicate that the SENI is strongly correlated with various business cycle indicators. Furthermore, the analysis demonstrates that once one month of data is available in a given quarter, a model utilizing text data produces more accurate nowcasts than a model using a widely-used Swiss business cycle indicator. Additionally, the analysis shows that text data is particularly useful for forecasting during non-crisis periods. As the news abstracts are publicly available with a delay of no more than one day, they can be used to meet the high-frequency information demands that emerged during the Covid-19 crisis.

Textual data and news sentiment indicators are increasingly used to forecast economic activity. Studies such as Larsen and Thorsrud (2018), Shapiro et al. (2020) and Thorsrud (2020) have applied text mining techniques to news sources to obtain leading indicators and daily measures of the business cycle. Buckman et al. (2020) revealed that news sentiment indicators are a reliable and prompt source of information on the economy during the Covid-19 pandemic, even outpacing quickly accessible survey data. Furthermore, Ardia et al. (2019) showed that utilizing news sentiment can aid in forecasting U.S. industrial production growth, and Ellingsen et al. (2021) confirms that news data contains information not captured by traditional economic indicators. Finally, that short news texts also have predictive power has been demonstrated by Li et al. (2022) and Y. Bai et al. (2022), who found that news headlines can be used to predict asset prices.

Several related studies have been conducted using sentiment analysis for hand-selected economic concepts and topics. For instance, Barbaglia et al. (2022) have proposed a fine-grained aspect-based sentiment analysis to evaluate the informational content of sentiment contained in news articles about the state of the economy. Additionally, Kalamara et al. (2022) have performed a study that extracts timely economic signals from newspaper text and demonstrates that it improves forecasts of macroeconomic variables such as GDP, inflation, and unemployment. Aruoba and Drechsel (2022) have also conducted a study that creates hand-selected economic concepts from documents prepared for Federal Open Market Committee meetings, and utilizes the sentiments associated with these concepts to create a novel measure of monetary policy shocks.

Using textual data to measure uncertainty is also increasingly common in the literature following the initial contribution by Baker et al. (2016). Larsen (2021) computes a measure of uncertainty for different aspects of the economy using machine learning techniques. Similarly, Cieslak et al. (2021) use the information content of Federal Open Market Committee (FOMC) meetings to construct text-based measures of the uncertainty faced by policymakers.

In Switzerland, there are two primary contributions to using text data as a proxy for economic activity.[1] Burri and Kaufmann (2020) developed a daily composite indicator utilizing publicly available financial market and news data. They collected short economic news from online newspaper archives and calculated a news sentiment indicator by counting positive and negative words. In contrast to this simplistic approach, in this paper I propose utilizing more rigorous techniques to extract relevant information from news data, in order to draw an even clearer picture of the economy. Becerra et al. (2020) develop sentiment indicators using internet search engine data. They propose a text-based algorithm to extract relevant economic keywords from Google Trends data and create an index that tracks the relative popularity of these keywords over time.

This paper is the first, to the best of my knowledge, to evaluate the daily informational content from short, publicly available newspaper extracts for nowcasting GDP growth in Switzerland.

The paper proceeds as follows. In the next section, I describe the text data. Section 3 presents the methodology used to extract information from the text data and to summarize it into a business cycle indicator. Moreover, it discusses the models used for the out-of-sample nowcasting exercise. Section 4 presents an extensive evaluation of the newly created business cycle indicator. The last section concludes.

2 Data

I use publicly and quickly (usually with a delay of one day) available news data to create sentiment and uncertainty indicators based on three Swiss newspapers. The newspaper data stem from the online archives of the *Tages-Anzeiger* (TA)[2], the *Neue Zürcher Zeitung* (NZZ)[3] and the *Finanz und Wirtschaft* (FUW)[4]. 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.[5] For this paper, I use data from January 1, 2000 to December 31, 2021.

[1] There are various other initiatives in Switzerland, 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 labor market data, debit and credit card data, traffic and mobility data, payments and cash withdrawals data (Brown & Fengler, 2020; Eckert et al., 2020; Eckert & Mikosch, 2020; Wegmüller et al., 2021).

[2] See https://www.tagesanzeiger.ch/zeitungsarchiv-930530868737.

[3] See https://zeitungsarchiv.nzz.ch/archive.

[4] See https://www.fuw.ch/archiv.

[5] Sometimes the *Tages-Anzeiger* updates its archive with a relevant delay or not at all. Therefore, I additionally use abstracts from the *Tages-Anzeiger* website: https://www.tagesanzeiger.ch/wirtschaft.

One of the main benefits of using news data is its immediate availability, as well as its longer time period of coverage compared to other high-frequency data sources. Additionally, the process of extracting information from the data is straightforward, and since the data is publicly accessible, it can be utilized by anyone. 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 abstracts often succinctly convey the main message of the article and tend to have less extraneous information, making signal extraction more precise. In the end it is an empirical question whether short texts are sufficient for forecasting economic variables which is also addressed in this paper.

I further reduce the signal-to-noise ratio by utilizing 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. To filter out the most relevant articles, I focus on those that include specific German keywords related to the economy such as *Wirtschaft*, *Konjunktur* and *Rezession* (which translate to economy, business cycle, and recession respectively).[6] 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 uncertainty for both Switzerland and foreign countries by using location-specific keywords.[7] I use a total of 3 different search queries for articles pertaining to the domestic economy and 24 for the foreign economy. For more information on the search queries used to filter out relevant articles, see Table 6 in the appendix.

Table 1 presents an overview of the scraped text data. A few things are worth mentioning. First, over all sources, roughly 900'000 abstracts, text passages, or titles were collected. It is important to note that these 900'000 texts are not unique articles, but rather the results of multiple search queries. Second, the average number of words

[6] 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).

[7] 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 switzerland, recession germany).

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 6 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.

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 abstracts. Third, the more liberal *NZZ* and *FUW* have a more positive sentiment towards the economy than the *TA*, which is known to be rather left-leaning liberal. Finally, it is likely that the lower sentiment found in the *Tages Anzeiger* webpage (TAW) is due to its different time coverage.

3 Methodology

In this section, the method of extracting information from the data, the creation of various sub-indices covering different areas of the economy and the process of aggregating it to create a business cycle indicator are described. Moreover, the models used for the out-of-sample nowcasting exercise are explained.

3.1 Creating text-based indicators

In order 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 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, first, I use a sentiment lexicon that is not stemmed and, second, I do not use a topic model to classify the texts.

The data collected is used to create a variety of text-based indicators that capture the different contexts (topics) of the economy in Switzerland and abroad. This is achieved by using different sets of keywords, denoted by $\mathcal{K}$, which define these contexts. A detailed list of the topic defining keywords is shown in Table 2. The indicators are then created using two different methods.[8] The first method is used to create indicators for recessions and economic uncertainty by counting the occurrence of keywords related to these topics and summing them up to a daily time series. The recession index, also known as the R-word index, was invented by The Economist (2011) in the early 90's. Iselin and Siliverstovs (2013) create a R-word index for Switzerland and find that is has predictive power to forecast Swiss GDP growth. Uncertainty indicators became popular after the seminal work of Baker et al. (2016). These indicators measure uncertainty in the economy and recession fears and are therefore negatively correlated with the business cycle.

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

Table 2 — Keywords for economic topics

Topic	Keywords	English	Method
Recession	rezession, krise	recession, crisis	Count
Uncertainty	unsicher, angst, befürcht, ungewiss, gefahr, risiko, risiken, gefähr	uncertain, fear, anxiety, danger, risk	Count
Labor market	arbeit, job, beschäftigung	labor, job, employment	KWIC
Monetary policy	libor, nationalbank, notenbank, zentralbank, snb, ezb, \bfed\b, zins, geldpolitik	libor, national bank, central bank, snb, ecb, \bfed\b, interest rate, monetary policy	KWIC
Financial market	stock, asset, anlage, aktionär, aktie, dividend, börse, finanz, \bsmi\b, dax, \bspi\b, nasdaq, msci	stock, asset, investment, share, dividend, financial,\bsmi\b, dax, \bspi\b, nasdaq, msci	KWIC
Forex	währung, forex, exchange, wechselkurs	currency, forex, exchange, exchange rate	KWIC
Operating results	ertrag, geschäftsjahr, gewinn, verlust, umsatz, umsätze, ergebnis, quartalszahl	income, fiscal year, profit, loss, sales, revenues, result, quarterly figure	KWIC
Tourism	hotel, tourist, tourism, flug, air, reise, übernachtungen	hotel, tourist, tourism, flight, air, travel, overnight stays	KWIC
Government	regierung, staat, minister, govern, \bbund\b, steuer, politik	government, state, minister, federal, tax, policy	KWIC
Industrial sector	industrie, produktion, handel, export, import	industry, production, trade, export, import	KWIC
Investment	invest	invest	KWIC
Economy	wirtschaft, konjunktur	economy, business cycle	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, utilizes word co-occurrences to calculate topic-specific sentiment indicators. This is done by creating new sets of documents by screening the texts for keywords in $\mathcal{K}$. Whenever a keyword is found, the keyword, along with the ten preceding and ten following words, is extracted into a new document.[9] A sentiment score is then calculated for each of these documents.

[9]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.

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

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

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). Articles are classified into different topics based on the words they contain. If two articles contain similar words, they will probably be assigned to the same topics. It is likely that an article discusses different topics, thus it is difficult to assign the article-level sentiment score to a topic. This approach is more specific in comparison because it only looks at the topic specifying word and calculates a sentiment score from the a few words surrounding it. In addition, the fact that the texts are very short and sometimes consist only of text passages makes it difficult to estimate a topic model (Yan et al., 2013). Finally, a potential data leakage can be avoided which would be problematic for the nowcasting exercise. 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.

Figure 1 — Document-level sentiment score

Abstract 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 abstract 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 same method is applied to calculate the sentiment score for other topics, such as the industry topic, which in this example is given by $S_{t,d,\text{industry}} = (1-2)/19 = -0.05$.

3.2 Extracting factors

The objective of this paper is to assess the informational content of daily news-based sentiment for nowcasting quarterly Swiss GDP growth. To accomplish this, the use of models that allow for the inclusion of time series of varying frequencies in the same regression without the need for transformation is necessary. **Mi**xed-**da**ta **s**ampling (MIDAS) and bridge models are commonly used in literature for this purpose. 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. 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 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). [10]

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.[11] I use the proposed information criterion BIC_3 which is recommended for $N > 18$. 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 pseudo-real-time forecast evaluation.[12] 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

[10] 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.

[11] 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.

[12] The evaluation is not purely real-time forecast evaluation due to the utilization of three forms of in-sample information. First, the SENI is constructed based on knowledge of the business cycle in the past, in particular, the Covid–19 Crisis. Second, the selection of the underlying indicators is based on their strong correlation with the business cycle. Third, the normalization of indicators in the factor model may result in revisions that aren't considered in the forecast evaluation. Arguably, using this in-sample information in the evaluation is reasonable if the goal is to demonstrate the future usefulness of the indicator, not its past performance.

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. Similarly 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 available 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 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 latter is because $T_x \geq T_y$. The sample spans from January 1, 2000, to December 31, 2021.

To determine the informational content of the SENI I forecast GDP growth using 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). The forecasts are compared to three benchmarks. First, I use an autoregressive model (AR) of order 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. 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. To compute the forecast errors, I use the last available release of quarterly GDP from December 2021.

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))

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

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 $L^{1/60} x_{t_d} = x_{t_d-1/60}$. It is worth noting that I determine the effect of the daily indicator $x_{t_d+T_x-T_y}$ on y_{t_q+H+1} by estimating a regression coefficient β_p for every low-frequency lag included.[13] 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.[14]

For the polynomial specification, I use an exponential Almon lag of order two. This is extensively discussed in Ghysels et al. (2007) and has the following form[15]

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

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

[13] I estimate a model with only one regression coefficient for all included low-frequency lags as well. This, however, deteriorates forecasting performance. Results can be requested from the author.

[14] I also estimate a model with different polynomial specifications for every included low-frequency lag. However, this led to converging issues in the NLS estimation for some periods, and hence, this deteriorates forecasting performance. Results can be requested from the author.

[15] For robustness I also use a Legendre polynomial proposed by Babii et al. (2021). The results, however, are less promising. More details can be found in section A.1 in the Appendix.

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

Another common approach in the literature is the use of bridge equations that 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

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

is the time aggregated high-frequency variable. The aggregation function depends on the nature of the indicator, here it's a simple, equal weighted 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 first need to forecast the indicator variable to the end of the desired quarters. Several methods are possible. I use an AR(p) model 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).

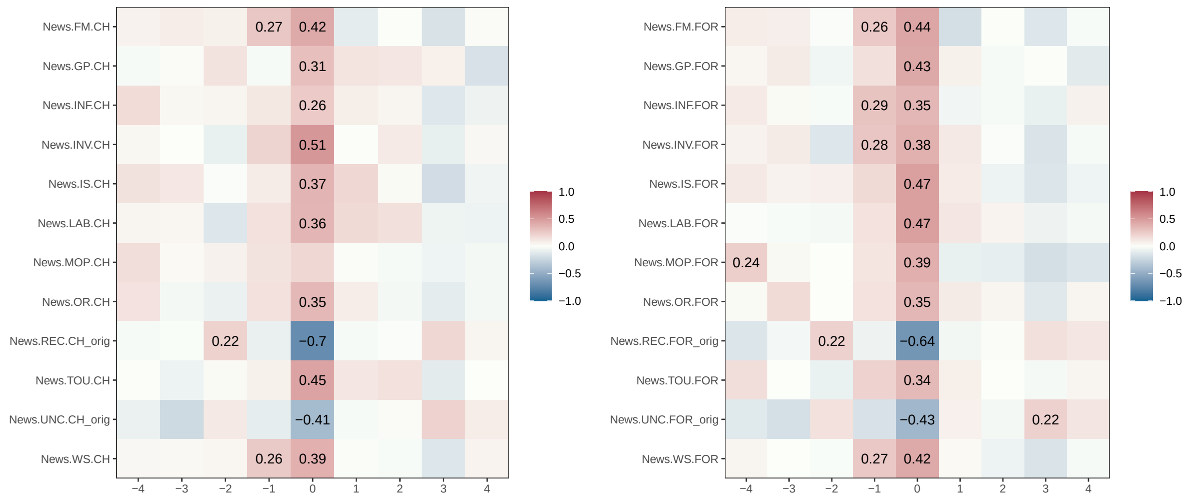
3.3.3 Iterative MIDAS

The iterative MIDAS (MIDAS-IT) model was introduced by Schumacher (2016). It's an intermediate model between bridge and MIDAS. In principle, it's a bridge model where 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.

4 Evaluation of the SENI

4.1 Descriptive analysis

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



Notes: Cross-correlation between news-based sentiment indicators and real Swiss GDP growth. I aggregate all data to quarterly frequency. Only statistically significant correlations are labeled. 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.

All the news-based indicators are fed into the factor model. Most of these indicators are substantially correlated with GDP growth (after pre-whitening with an AR(p) model (see Neusser, 2016, Ch. 12.1)). Figure 2 provides an overview of the cross-correlations. The correlation coefficients range from 0.26 to 0.7. The indicators for abroad tend to be slightly more correlated. In the end, nine indicators for the Swiss economy and eleven for abroad are included in the model. The news indicators are rather volatile (see Figures 6 and 7 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. I tried different time windows and found a good compromise with the ten-day window.

Panel (a) of Figure 3 displays the SENI together with actual GDP growth, revealing the indicator closely follows economic crises. It anticipates the downturn during the Global Financial Crisis, responds to the removal of the minimum exchange rate, and the euro area debt crisis. 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 than other news-based sentiment indicators, making it difficult to visually assess the state of the economy, but it is more accurate for nowcasting as daily values are less distorted. The indicator seems to indicate recovery from a crisis only with a delay.

4.2 In-sample evaluation

The SENI is correlated with many key macroeconomic variables (See Figure 9 in the Appendix) as it is not optimized to track any particular measure of economic activity in its current form. I, therefore, focus on the in-sample information content of the SENI, highlighting that it is available earlier than most other leading indicators.

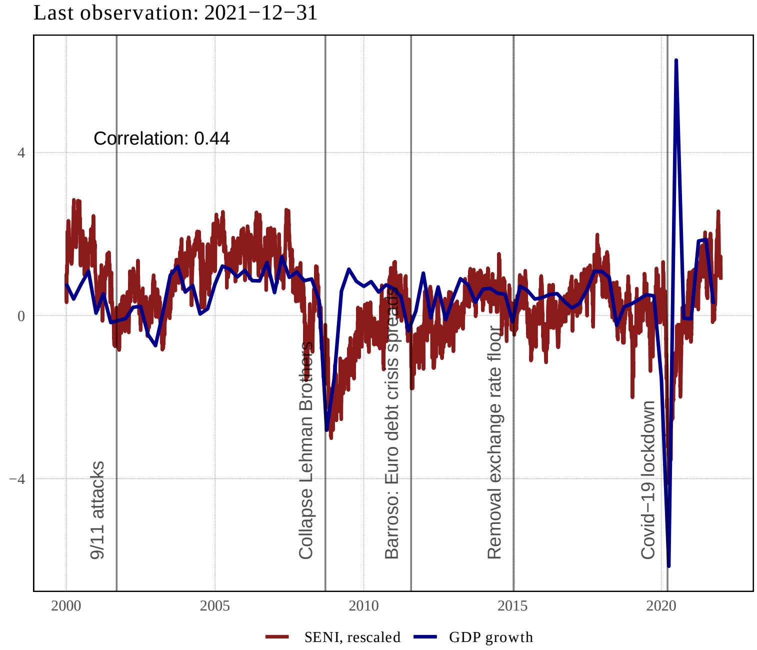
To compare the in-sample information content of the SENI to other leading indicators, I perform a cross-correlation test (see Neusser, 2016, Ch. 12.1).[16] Figure 4 shows a substantial correlation between the SENI and many prominent leading indicators. There is a coincident or a leading relationship with consumer confidence.[17] There is a leading, coincident, and lagging relationship with the KOF Economic Barometer and SECO's Swiss Economic Confidence (SEC). A coincident and a lagging correlation exists with the Organisation for Economic Co-operation and Development composite leading indicator (OECD CLI), trendEcon's perceived economic situation, and the SNB's Business Cycle Index (BCI). However, the OECD CLI is a smoothed indicator and is

[16]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 Table 5 in the Appendix); finally, some are based on lagged data (see, e.g., OECD, 2010).

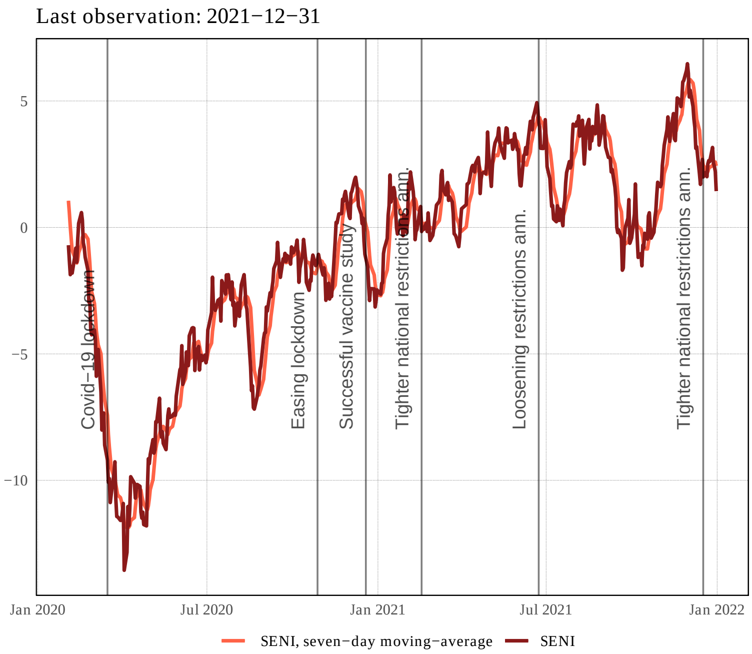
[17]All data sources are given in Table 5 in the Appendix.

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.

subject to substantial revisions, trendEcon's perceived economic situation starts only in 2006 and the BCI is published with a relevant delay. Moreover, the lagging correlation can be explained by the moving average. These findings can be confirmed by granger causality tests (Granger, 1969). The SENI granger causes all of the selected indicators. Whereas, it is granger caused by the OECD CLI, the BCI, and trendEcon's perceived economic situation. 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.

Table 3 — Pseudo real-time evaluation of real GDP growth: 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
a) Benchmark: AR(1) model – Hypothesis: Model > Benchmark												
Bridge	0.67	0.86**	1.02	1.11	0.85*	0.88	1.01	1.03	0.83***	0.87**	1.02	0.97
Midas	0.7	0.89*	0.99	1.01	0.78**	0.92	1.07	1	0.75**	0.93	1.08	0.93
Midas-IT	0.72	0.87**	0.97	1.08	0.81**	0.84*	0.99	0.98	0.72***	0.82***	0.99	0.91
b) Benchmark: Barometer bridge – Hypothesis: Model > Benchmark												
Bridge	1.08	0.96	0.96	1.1	0.86*	0.9	1	1.04	0.74**	0.82*	0.95	1
Midas	1.14	0.99	0.94	1	0.79**	0.95	1.06	1.02	0.67**	0.88	1.01	0.95
Midas-IT	1.17	0.96	0.92	1.07	0.82*	0.86*	0.98	0.99	0.64***	0.77**	0.92	0.93
c) Benchmark: First release – Hypothesis: Model < Benchmark												
Bridge	2.04*				1.17				1.17**			
Midas	2.15*				1.06				1.03			
Midas-IT	2.22*				1.11				0.98			

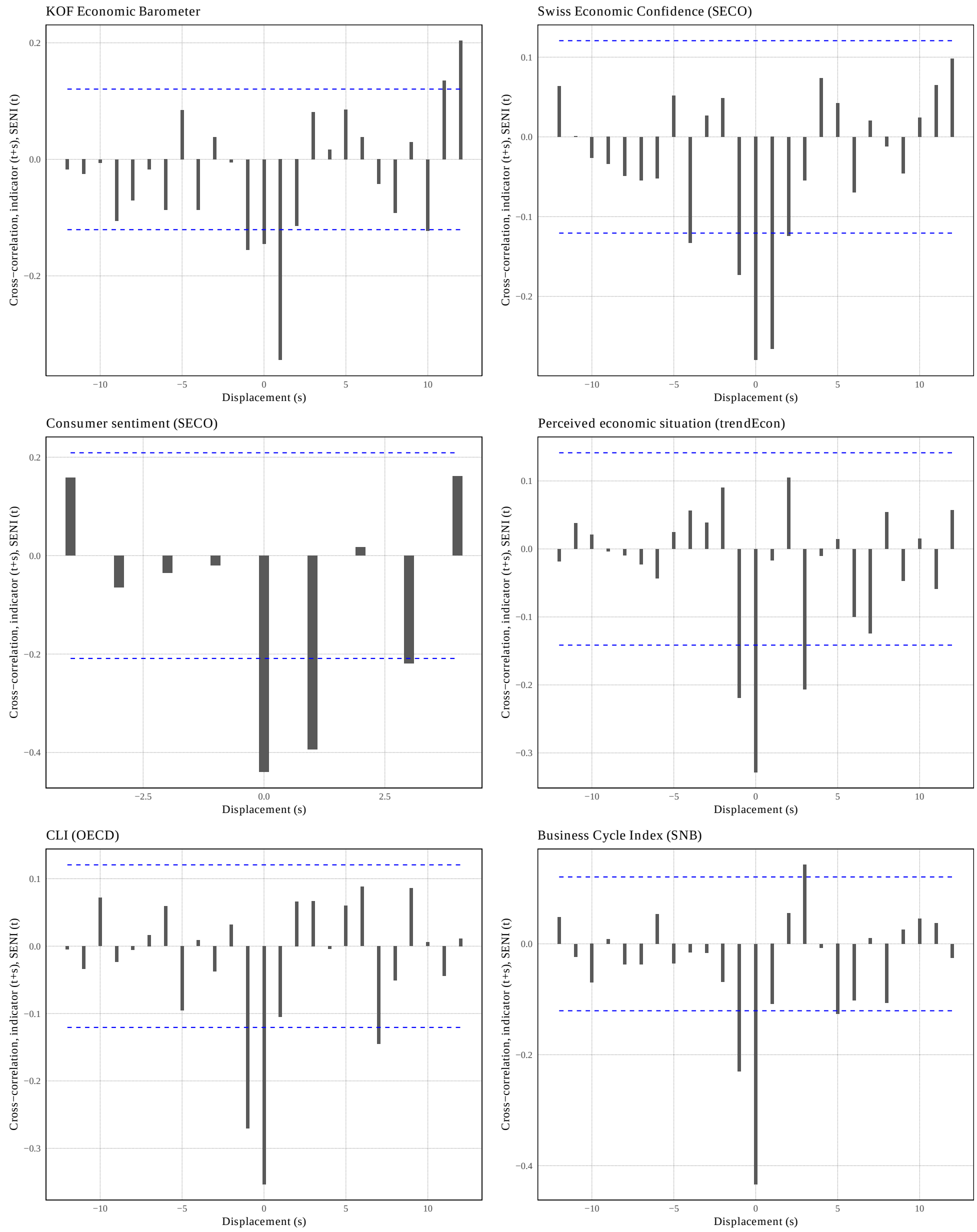
Notes: Relative Root-mean-squared errors (RMSE) for forecasts with selected forecast horizons. A lower RMSE implies higher predictive accuracy compared to the benchmark. Horizons from 0 to 59 (60 - 119) denote forecasts for the current (next) quarter. The full sample spans from January 1st 2002 to December 31st 2021. The sample exluding the Covid-19 crisis excludes quarters 2-4 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: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

4.3 Out-of-sample evaluation

Table 3 shows relative root-mean-squared errors (RMSE) of the pseudo-real-time out-of-sample nowcasting exercise.[18] Because the results for the full sample from 1st January 2002 to 31st December 2021 are largely driven by the Covid–19 crisis in 2020,

[18]For comparison, the results using the Legendre polynomial are shown in Table 4 in the Appendix.

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.

I additionally do a subsample analysis. That is, I exclude the last three quarters of 2020. And as a robustness check, I additionally exclude the Great Financial Crisis from 2008 to 2009. The used models do not perform significantly better compared to the AR(1) model in forecasting the next quarter's GDP growth which corresponds to horizons of 60 days and more (see Table 3, panel (a)). This holds true for all three (sub)samples. However, considering current quarter GDP growth nowcasts (horizons of 0 to 59 days) the models exhibit mostly significantly lower RMSE. The nowcast for the full sample when full information is available (horizon of 0) is much smaller than the benchmark but is not statistically significant. This is because of very large forecasting errors in the Covid-19 crisis, especially for the benchmark AR(1) model, which are distorting the test-statistics.

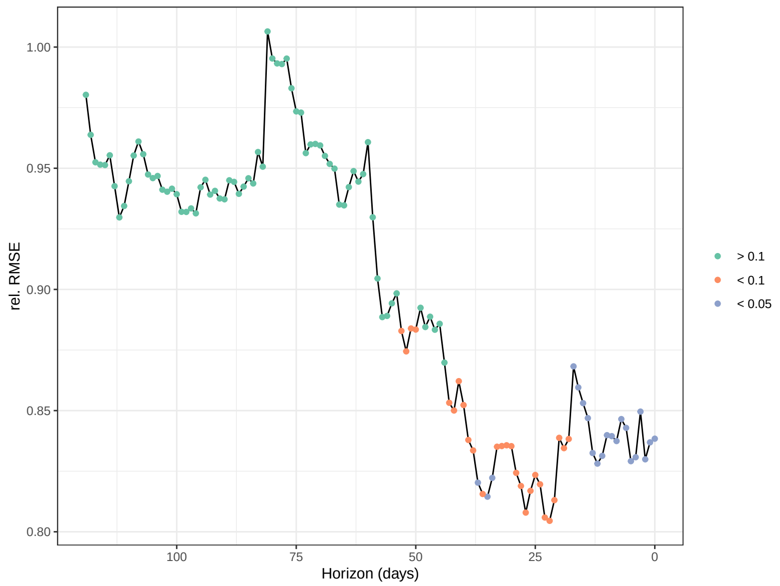
A similar picture emerges using a bridge equation model with the KOF Economic Barometer as the benchmark. Table 3, panel (b) shows that for the next quarter in general as well as over the full sample, the SENI does not outperform the KOF Barometer. However, looking at the two subsamples, the SENI is significantly more accurate for shorter horizons. This suggests that the KOF Barometer did well during crises, especially the Covid-19 crisis, however, in normal times the SENI performs much better. This suggests, that the advantage of the news indicator, in addition to its rapid availability, lies primarily in its accuracy in non-crisis periods.

Panel (c) shows that the RMSE of the SENI is higher than the RMSE of the first official GDP release, but the difference is not statistically significant in either sub-samples. The key advantage of the SENI is that it provides the full quarter's forecast about 2 months sooner than the first GDP release. Noteworthy is that the MIDAS-IT model exhibits lower RMSE than the first release in the sample excluding all crisis periods. The current vintage of GDP, which is used to compute the forecast errors, will likely be revised in the future. One of the reasons is that future vintages will include annual GDP estimates by the SFSO, which are based on comprehensive firm surveys. Therefore, I restrict the sample in panel (c) to years where the GDP figures already include these annual figures (to date up to 2020).

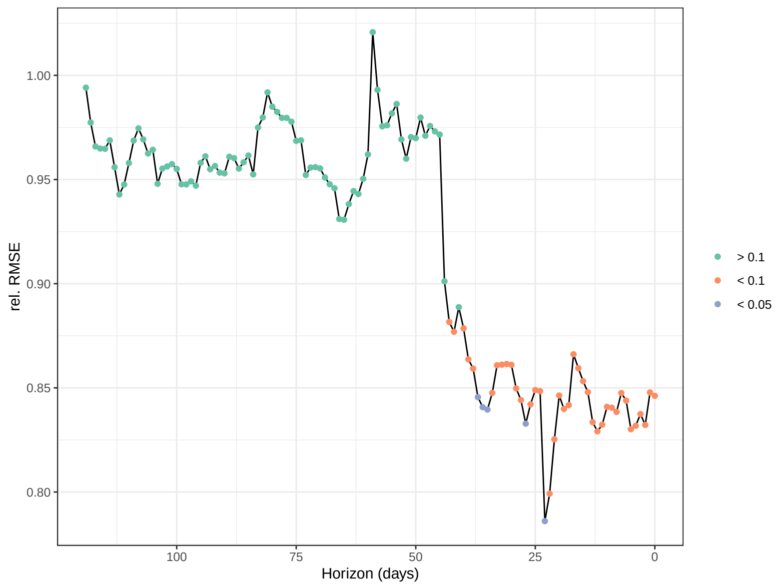
Is one of the models preferred over the others? In Figure 11 of the Appendix, I show the absolute RMSE of the out-of-sample nowcasting exercise. It becomes obvious that the RMSE of the MIDAS-IT model is below the RMSE of the other two models for most nowcasting horizons. Moreover, the bridge model mostly exhibits lower RMSE than the MIDAS model. These findings highlight two implications for the use of mixed frequency methods with daily data for nowcasting. First, the direct forecasting approach of the MIDAS model is less promising than an iterated approach. Second, it is advisable to use a nonlinear polynomial specification even if there are more parameters to estimate. Although, the nowcasts tend to be smoother when using a linear aggregation scheme.

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 forecasts with forecast horizons from 119 to 0 days. Periods of the Covid–19 crisis are excluded. 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. 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: ● $p < 0.1$, ● $p < 0.05$

We have seen that the SENI performs better for shorter horizons than a benchmark AR(1) and a bridge model with the KOF Barometer. The question is at what horizon is this difference getting significant? Figure 5 shows the evolution of the relative RMSE of the MIDAS-IT model vis-à-vis these two benchmark models for the sample excluding the Covid-19 crisis. Additionally, p-values for Diebold-Mariano-West (DMW) tests for the null hypothesis of equal predictive accuracy against the alternative that the MIDAS-IT model is more accurate are shown. Turning to panel (a), showing the relative RMSE against the AR(1) model, four points stand out. First, for next quarter GDP forecasts, that is horizons from 119 to 60 days, the MIDAS-IT model has a slightly lower RMSE, however, not statistically significant. Second, starting from a time horizon of 43 days, that is, roughly from the first month of the current quarter, MIDAS-IT is significantly more accurate than the AR(1) model in nowcasting current quarter GDP growth. Third, every time a new GDP vintage is released, which is usually around horizons of 20 and 80 days, respectively, the AR(1) model gets more accurate relative to the MIDAS-IT model. This is manifested in the spikes of the relative RMSE around these horizons. Finally, the rather large drop of the relative RMSE at a horizon of 59 days can be explained by the fact that I estimate two different models for forecasting current and next quarter GDP – and obviously, the MIDAS-IT model gets much more precise from this switch compared to the AR(1) model. Panel (b) shows similar behavior of the relative RMSE against the KOF Barometer bridge model. The SENI starts outperforming the KOF Barometer at a horizon of 40 days. It is surprising that an indicator based on short news abstracts is able to outperform a well-known and widely referenced economic indicator. A possible explanation could be that I take into account the publication lags of the KOF Barometer in the nowcasting exercise.

5 Concluding remarks

In this paper, I examine the informational content of daily news abstracts for nowcasting Swiss GDP growth. Using NLP methods, I create various sentiment and uncertainty indicators covering many aspects of the economy. Since the models that are able to quantify the daily informational content cannot deal with too many explanatory variables, I use factor analysis to extract a common component of the text-based indicators. The first principal component can be interpreted as a business cycle indicator as the creation of the sentiment and uncertainty indicators is based on economic reasoning. An evaluation of the indicator shows that it is not only correlated with other business cycle indicators but also accurately tracks Swiss GDP growth. Major strengths of the use of publicly available news abstracts are the low cost and that they can be updated with a delay of one day at most.

The lessons learned from a real-time out-of-sample nowcasting exercise using mixed frequency methods are twofold. First, already after the first month of the current quarter, a model with text data produces more accurate nowcasts than a model with a well-known economic indicator. Second, sub-sample analysis has shown that text data provide added value specifically in non-crisis periods. In contrast, they do not work that well during the recovery phase of crises. All in all, the use of textual data yields an accurate nowcast of Swiss GDP growth and is well suited for tracking Swiss economic activity at a high frequency.

However, there is still scope for enhancement. I identify five potential areas for future exploration: First, incorporating news sentiment indicators from additional sources such as French and Italian-speaking Swiss newspapers. Second, exploring the potential benefits of using full texts of articles instead of just publicly available abstracts. Third, customizing the lexicon specifically for economic news (see e.g. Shapiro et al., 2020). Fourth, assessing the predictive power of multiple factors and other macroeconomic data. Finally, I could investigate whether all the individual indicators in a model would improve the nowcast. Babii et al. (2021) have developed a sparse-group LASSO-MIDAS model which could be interesting in this context. All of this will likely further improve our understanding of using text data for assessing the health of the Swiss economy at 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. (Cit. on p. 12)

Almon, S. (1965). The Distributed Lag Between Capital Appropriations and Expenditures. *Econometrica*, *33*(1), 178. https://doi.org/10.2307/1911894 (cit. on p. 30)

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 (cit. on pp. 3, 9)

Aruoba, S. B., & Drechsel, T. (2022). Identifying Monetary Policy Shocks: A Natural Language Approach, 40 (cit. on p. 3).

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 (cit. on pp. 13, 24, 30)

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 (cit. on pp. 12, 14)

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 (cit. on p. 11)

Bai, J., & Ng, S. (2013). Principal components estimation and identification of static factors [Publisher: Elsevier]. *Journal of Econometrics*, *176*(1), 18–29. https://doi.org/10.1016/j.jeconom.2013.03.007 (cit. on p. 11)

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 (cit. on p. 3)

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 (cit. on pp. 3, 7)

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 (cit. on p. 3)

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 (cit. on p. 5)

Becerra, A., Eichenauer, V. Z., Indergand, R., Legge, S., Martinez, I., Mühlebach, N., Oguz, F., Sax, C., Schuepbach, K., & Thöni, S. (2020). trendEcon [Type: Website]. https://www.trendecon.org (cit. on pp. 4, 5)

Brown, M., & Fengler, M. (2020). MonitoringConsumptionSwitzerland [Type: Website]. https://public.tableau.com/profile/monitoringconsumptionswitzerland (cit. on p. 4)

Buckman, S. R., Shapiro, A. H., Sudhof, M., & Wilson, D. J. (2020). News sentiment in the time of COVID-19. *FRBSF Economic Letter*, *2020*(08), 1–05. https://www.frbsf.org/economic-research/publications/economic-letter/2020/april/news-sentiment-time-of-covid-19 (cit. on p. 3)

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 (cit. on p. 4)

Cieslak, A., Hansen, S., McMahon, M., & Xiao, S. (2021). Policymakers' Uncertainty. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3936999 (cit. on p. 3)

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 (cit. on pp. 18, 22, 31, 39)

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 (cit. on p. 14)

Eckert, F., Kronenberg, P., Mikosch, H., & Neuwirth, S. (2020). Tracking Economic Activity With Alternative High-Frequency Data [Artwork Size: 38 p. Medium: application/pdf Publisher: ETH Zurich], 38 p. https://doi.org/10.3929/ETHZ-B-000458723 (cit. on p. 4)

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 (cit. on p. 4)

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 (cit. on pp. 3, 9)

Feinerer, I., & Hornik, K. (2019). *Tm: Text mining package* (manual). https://CRAN.R-project.org/package=tm. (Cit. on p. 6)

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. https://doi.org/10.1111/rssa.12043 (cit. on p. 30)

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 (cit. on p. 14)

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2004). *The MIDAS touch: Mixed data sampling regression models* (CIRANO working papers). CIRANO. (Cit. on p. 12).

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 (cit. on pp. 12, 13)

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods [Publisher: [Wiley, Econometric Society]]. *Econometrica : journal of the Econometric Society*, *37*(3), 424–438. Retrieved May 31, 2022, from http://www.jstor.org/stable/1912791 (cit. on p. 18)

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 (cit. on p. 9)

Indergand, R., & Leist, S. (2014). A real-time data set for Switzerland. *Swiss Journal of Economics and Statistics*, *150*(4), 331–352. https://doi.org/10.1007/BF03399410 (cit. on p. 11)

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 (cit. on p. 7)

Kalamara, E., Turrell, A., Redl, C., Kapetanios, G., & Kapadia, S. (2022). Making text count: Economic forecasting using newspaper text*. *Journal of Applied Econometrics*, jae.2907. https://doi.org/10.1002/jae.2907 (cit. on p. 3)

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 (cit. on p. 12)

Larsen, V. H., & Thorsrud, L. A. (2018). Business Cycle Narratives. *CESifo Working Paper*, (7468) (cit. on p. 3).

Larsen, V. H. (2021). Components of Uncertainty, 20 (cit. on p. 3).

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 (cit. on p. 3)
Neusser, K. (2016). *Time Series Econometrics*. Springer International Publishing. https://doi.org/10.1007/978-3-319-32862-1. (Cit. on pp. 15, 16, 19)
OECD. (2010). Review of the CLI for 8 countries. *OECD Composite Indicators*. https://www.oecd.org/fr/sdd/indicateurs-avances/44556466.pdf (cit. on p. 16)
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 (cit. on pp. 6, 9)
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: STRUCTURAL TOPIC MODELS FOR SURVEY RESPONSES. *American Journal of Political Science*, *58*(4), 1064–1082. https://doi.org/10.1111/ajps.12103 (cit. on p. 9)
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 (cit. on pp. 12, 14)
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 (cit. on pp. 3, 9, 15, 24)
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 (cit. on p. 11)
The Economist. (2011). Up means down. The R-word index. The economist's gauge of gloom. http://www.economist.com/node/21529079 (cit. on p. 7)
Thorsrud, L. A. (2020). Words are the new numbers: A newsy coincident index of the business cycle [Publisher: Taylor & Francis]. *Journal of Business & Economic Statistics*, *38*(2), 393–409. https://doi.org/10.1080/07350015.2018.1506344 (cit. on pp. 3, 9, 15)
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 (cit. on p. 4)
West, K. (1996). Asymptotic inference about predictive ability. *Econometrica*, *64*(5), 1067–84. https://doi.org/10.2307/2171956 (cit. on pp. 18, 22, 31, 39)

Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013). *A Biterm Topic Model for Short Texts*. (Cit. on p. 9).

A Appendix

A.1 Legendre polynomial

Following Babii et al. (2021) I use a Legendre and an Almon polynomial of order two as a robustness check. The Legendre polynomial is an orthogonalization of the Almon polynomial (Almon, 1965). Both polynomials led to identical results. However, Babii et al. (2021) argue that it's preferable to use orthogonal polynomials in practice due to reduced multicollinearity and better numerical properties. The Almon polynomial is given by

$$b(k, \theta) = \sum_{j=0}^{2} \theta_j k^j = \theta_0 + \theta_1 k + \theta_2 k^2$$

As Babii et al. (2021) find, these polynomial specifications improve the performance substantially. Furthermore, the optimization problem becomes convex and can therefore be estimated with ordinary least squares (OLS)(See the Online Appendix of Babii et al., 2021, for more details). Normally, OLS can only be used for unrestricted MIDAS models as in Foroni et al. (2015).

Table 4 — Relative RMSE and DMW tests with legendre polynomial

	Full sample				Excluding Covid crisis				Excluding all crisis periods			
Horizon	0	39	79	119	0	39	79	119	0	39	79	119
a) Benchmark: AR(1) model – Hypothesis: Model > Benchmark												
Bridge	0.67	0.86**	1.02	1.11	0.85*	0.88	1.01	1.03	0.83***	0.87**	1.02	0.97
Midas	0.69	0.88	1.05	1.09	0.83*	0.95	1.05	1.05	0.8**	0.95	1.05	1.03
Midas-IT	0.69	0.9	1.01	1.11	0.83*	0.89	1.1	1.08	0.8**	0.88**	1.13	1.04
b) Benchmark: Barometer bridge – Hypothesis: Model > Benchmark												
Bridge	1.08	0.96	0.96	1.1	0.86*	0.9	1	1.04	0.74**	0.82*	0.95	1
Midas	1.11	0.98	1	1.08	0.85*	0.98	1.04	1.06	0.72**	0.9	0.98	1.05
Midas-IT	1.11	0.99	0.96	1.11	0.84*	0.92	1.08	1.09	0.71**	0.83*	1.05	1.06
c) Benchmark: First release – Hypothesis: Model < Benchmark												
Bridge	2.04*				1.17				1.17**			
Midas	2.11*				1.14				1.12*			
Midas-IT	2.11*				1.14				1.11*			

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 2-4 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: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

A.2 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 or uncertainty 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 or uncertainty topic (per article multiple phrases can match).

  b. Calculate daily recession/uncertainty 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.

Figure 6 — Daily news based indicators, part 1

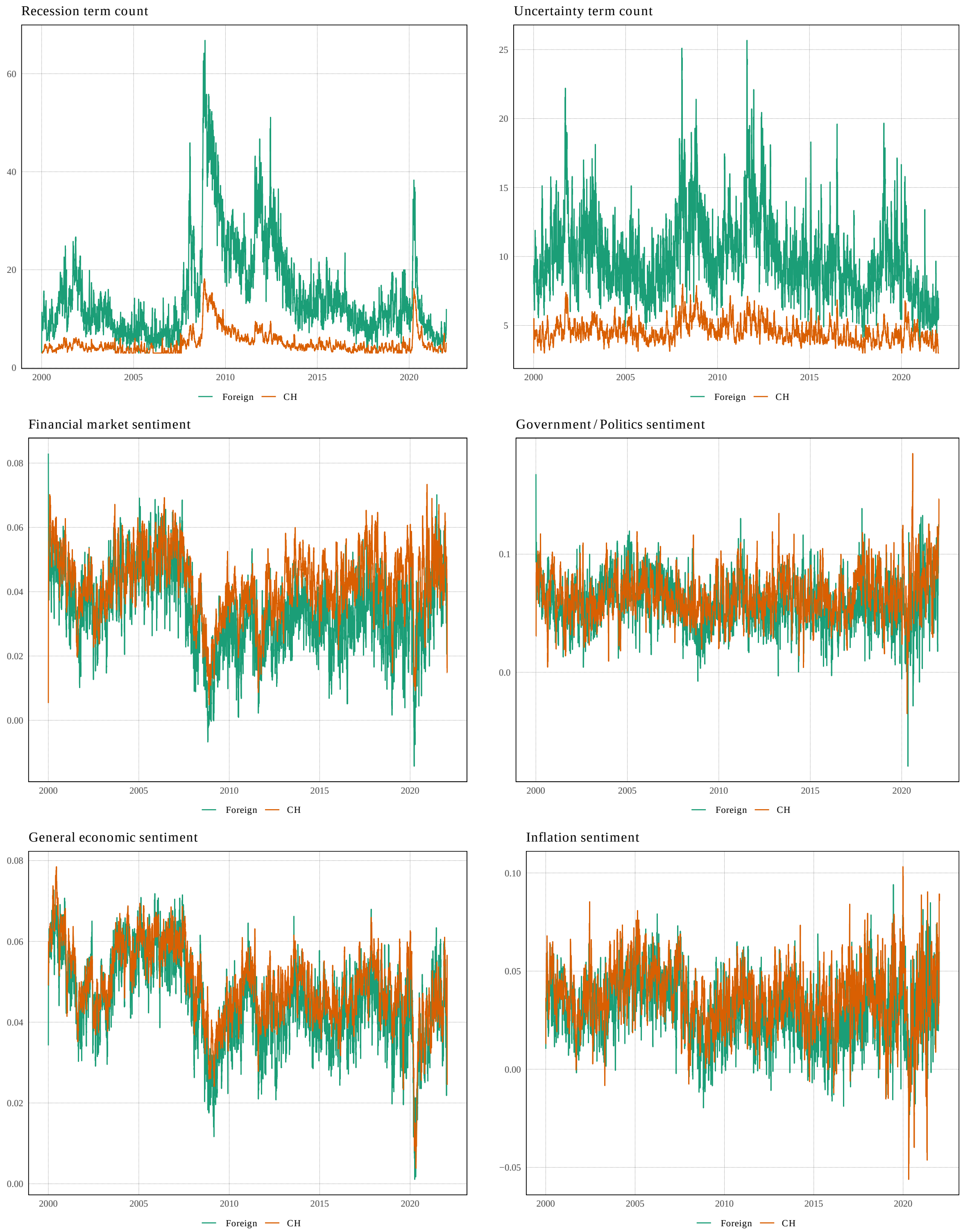


Figure 7 — Daily news based indicators, part 2

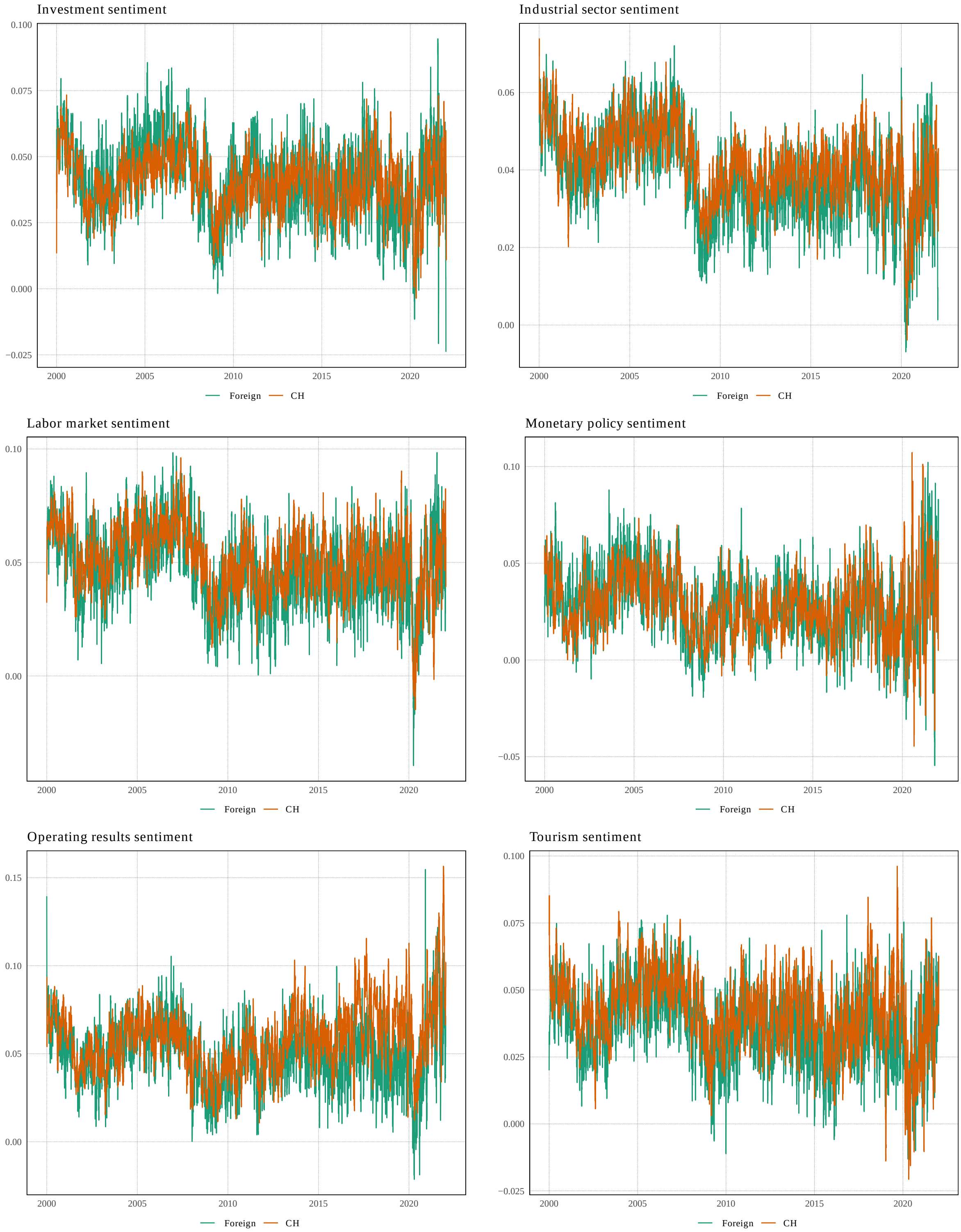
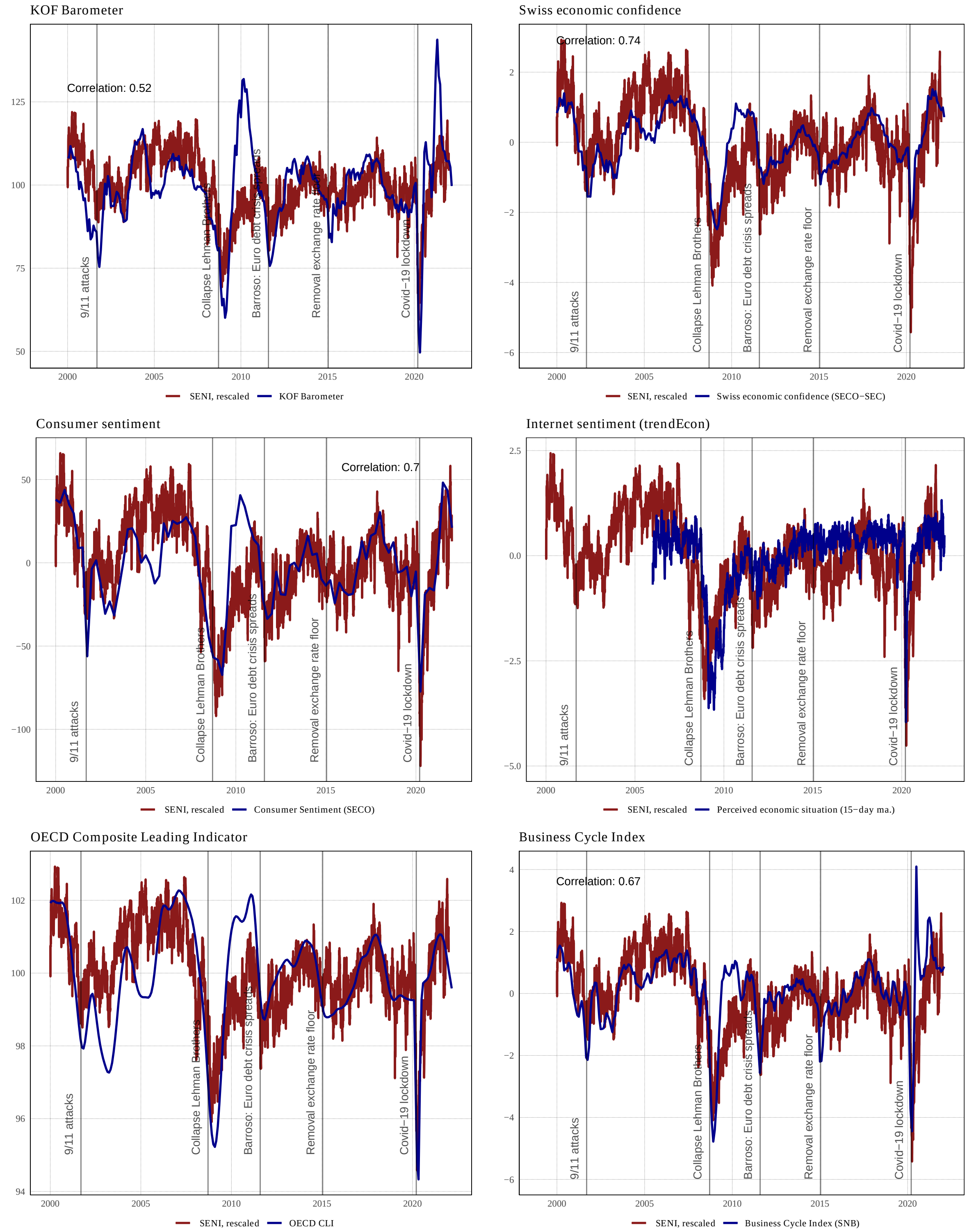


Figure 8 — Comparison with other indicators



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

Table 5 — 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

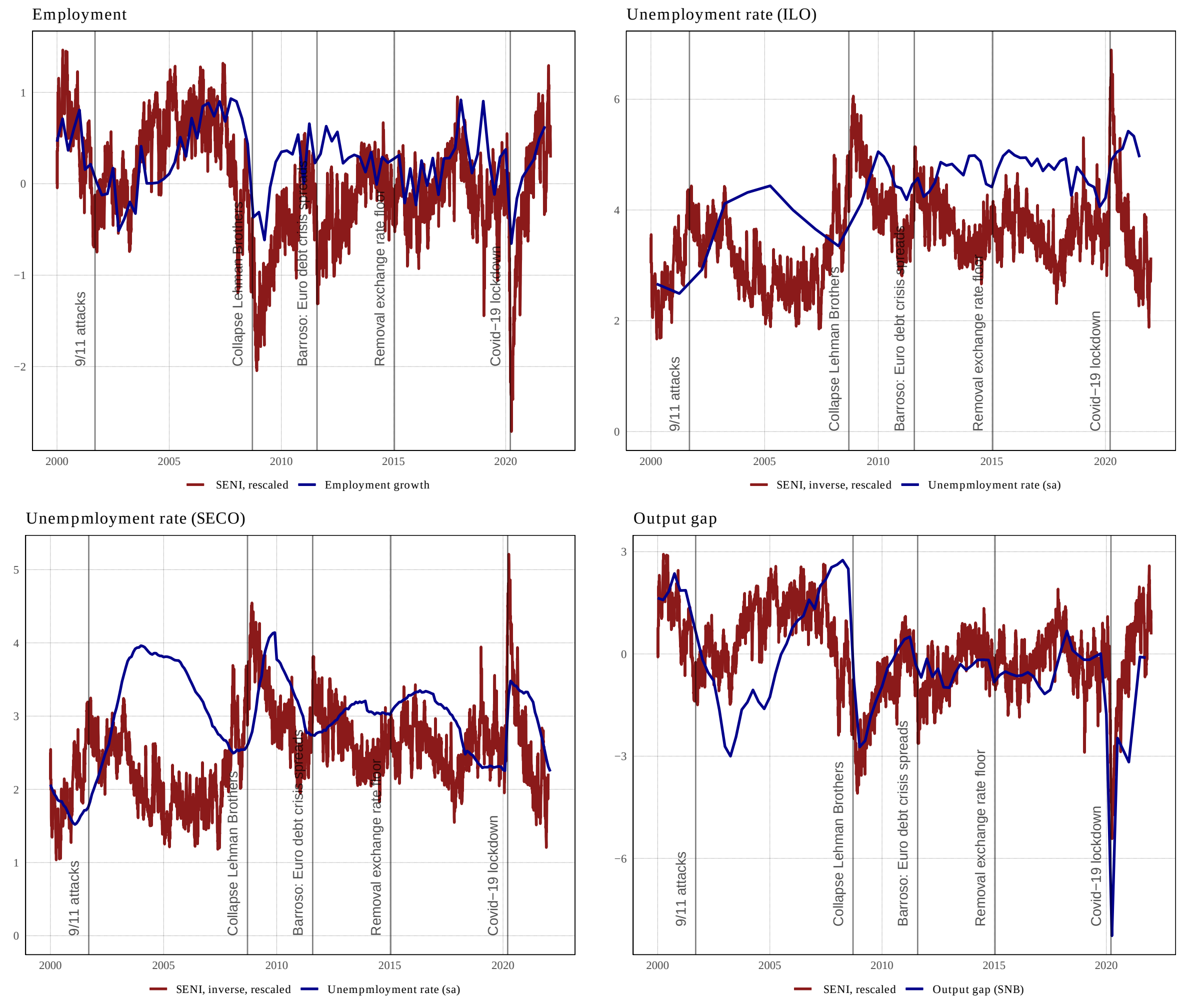
Notes: Publication lags between the last day of the variable frequency (i.e. last day of the quarter or last day of the month) and the publication date of a recent release. Therefore, all publication lags are approximate and may change over time.

Table 6 — 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.

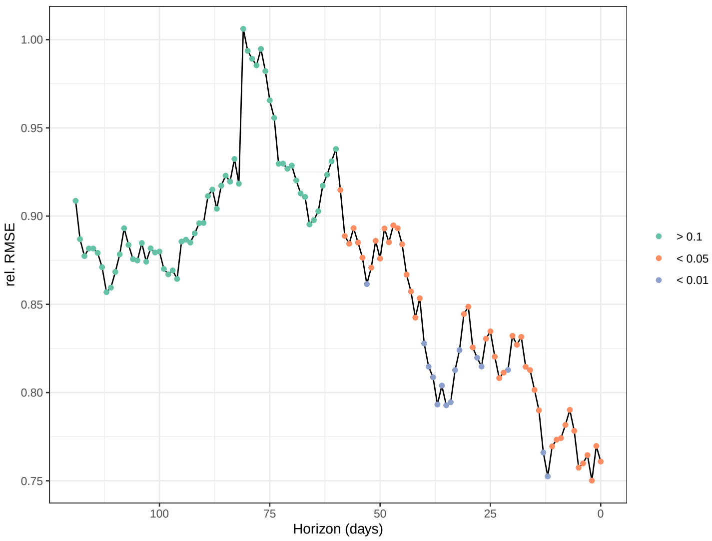
Figure 9 — Comparison with other macroeconomic data



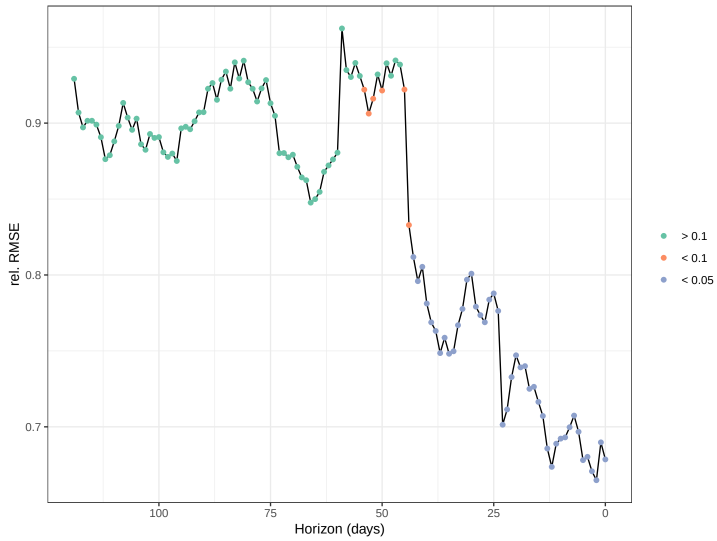
Notes: SENI rescaled such that it roughly matches the mean and volatility of the other data series.

Figure 10 — Evolution of relative RMSE

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

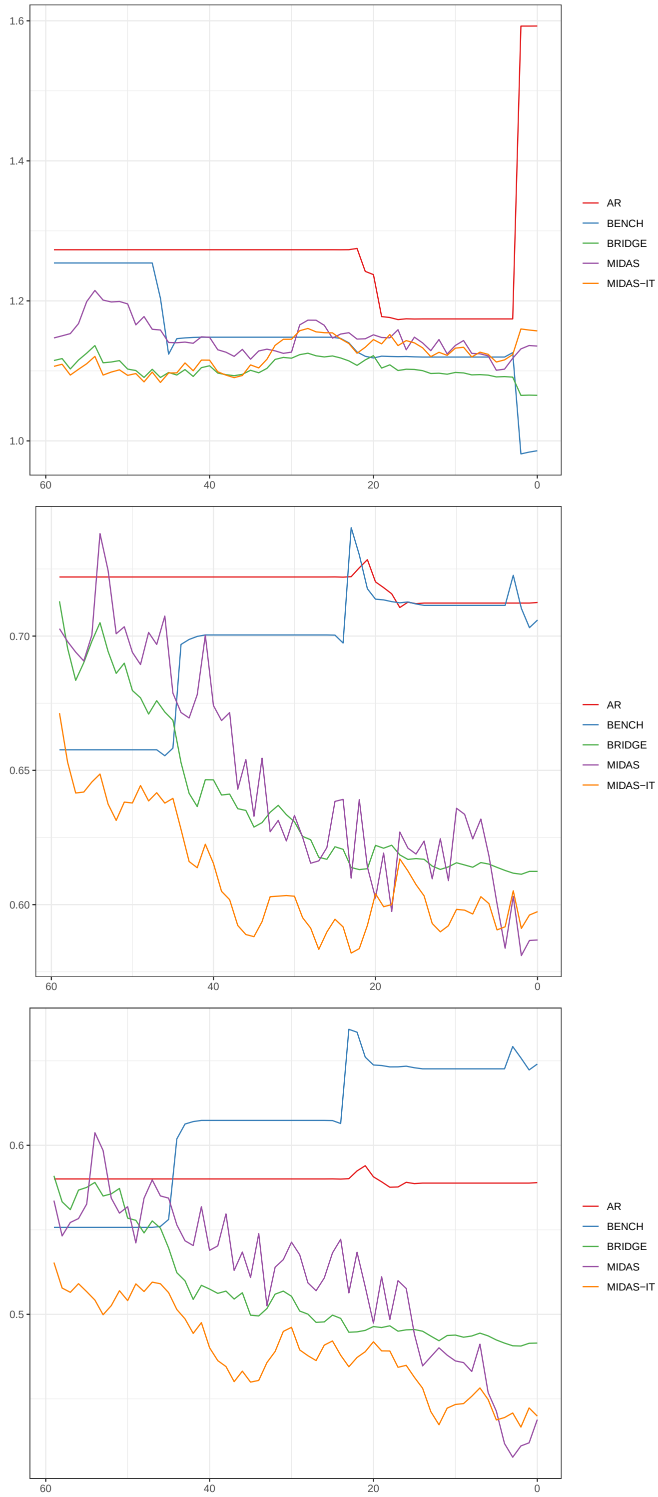


(b) MIDAS-IT vs. KOF Barometer bridge



Notes: Relative Root-mean-squared errors (RMSE) for forecasts with forecast horizons from 119 to 0 days. Periods of the Covid–19 crisis as well as of the GFC are excluded. 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. 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: ● $p < 0.1$, ● $p < 0.05$

Figure 11 — 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.