A daily fever curve for the Swiss economy*

Marc Burri† Daniel Kaufmann‡ §¶

November 19, 2024

Abstract

Because macroeconomic data is published with a substantial delay, assessing the health of the economy in real-time is challenging. This chapter develops a daily business cycle indicator for the Swiss economy using publicly available daily financial market and news data. The indicator can be computed with a delay of one day. Moreover, it is highly correlated with macroeconomic data and survey indicators of Swiss economic activity. Therefore, it provides timely and reliable warning signals if the health of the economy takes a turn for the worse.

Keywords: Composite leading indicator, high-frequency data, financial market data, news data, natural language processing, news sentiment, Switzerland

*We thank David Ardia, Ronald Indergand, Alexander Rathke, Maximilian Schröder, Jonas Striaukas, Jan-Egbert Sturm, Leif Anders Thorsrud, Mark W. Watson, Philipp Wegmüller, and Vaidotas Zemlys-Balevičius for helpful comments and discussions and participants at the IRENE PhD meeting, the IAAE Annual Conference in Oslo, the EEA-ESEM Conference in Barcelona, and the Money-Macro-Finance Conference in Warsaw for valuable comments. This chapter brought forward publication of a research paper co-authored by Daniel Kaufmann (Burri & Kaufmann, 2020) and a working paper (Burri, 2023).

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

‡University of Neuchâtel, Institute of Economic Research, Rue A.-L. Breguet 2, CH-2000 Neuchâtel, daniel.kaufmann@unine.ch

§KOF Swiss Economic Institute

¶Corresponding author

1 Introduction

Because macroeconomic data is published with a substantial delay, assessing the health of the economy during the rapidly evolving coronavirus disease of 2019 (Covid–19) crisis was challenging. Usually, policymakers and researchers make predictions and take decisions based on early information from surveys, financial markets, and rapidly available statistics (see, e.g., Abberger et al., 2014; Galli, 2018; Kaufmann & Scheufele, 2017; OECD, 2010; Stuart, 2020; Wegmüller & Glocker, 2019, for Swiss applications). These indicators and forecasts are published with a delay of at least one to two months.[1] Although this approach works well in stable periods, the Covid–19 crisis has shown that these indicators are not sufficient to make important decisions in a rapidly changing situation. Consequently, there is a pressing need for high-frequency information to accurately evaluate the current economic conditions and assess the impact of dynamic factors such as health restrictions and economic stimulus measures.

We introduce a novel approach to assess the well-being of the Swiss economy, proposing a daily fever curve (f–curve) based on publicly available financial market and news data. To construct the f–curve, we utilize risk premia on corporate bonds, term spreads, and stock market volatility indices from 2000 onwards. Additionally, we gather short economic news lead texts from online newspaper archives. Using these news articles, we generate sentiment indicators for various hand-selected economic concepts (topics), such as the labor market or inflation.[2] Subsequently, we use factor analysis to estimate a composite indicator that can be interpreted as a fever curve. Analogous to monitoring a patient's tempereature, an upward trend in the fever curve serves as a reliable and timely warning signal indicating a deterioration in the economy's health.

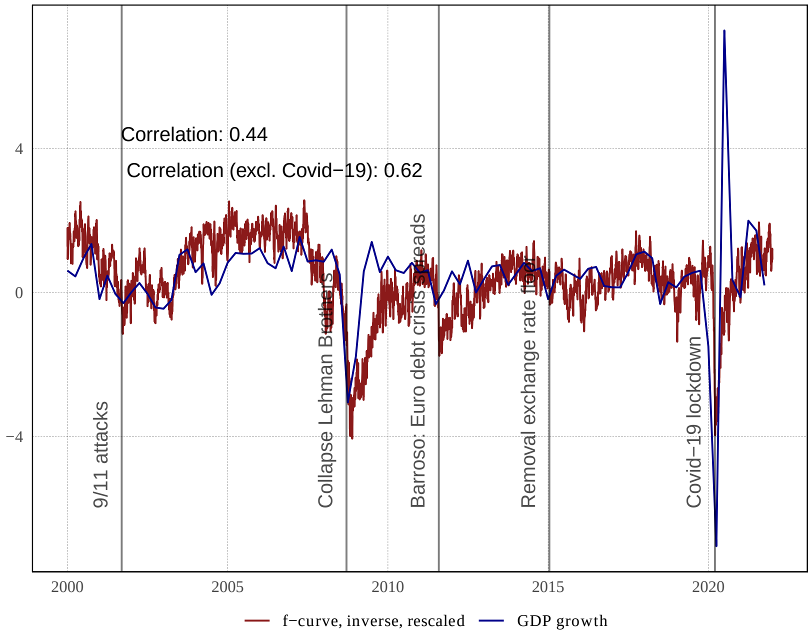
Panel (a) of Figure 1 shows the f–curve (on an inverted scale) jointly with real gross domestic product (GDP) growth: the indicator closely tracks economic crises. It presages the downturn during the Global Financial Crisis, responds to the removal of the minimum exchange rate in 2015, and to the euro area debt crisis. The f–curve also responds strongly to the Covid–19 crisis (see panel (b)). The indicator starts to rise in late February. By then, it became evident that the Covid–19 crisis would hit most European countries; in Switzerland, the first public events were canceled. It reaches a peak shortly after the lockdown. Afterwards, the fever curve gradually declines with

[1] See table 1 in the Appendix for publication lags of some important macroeconomic data and leading indicators.

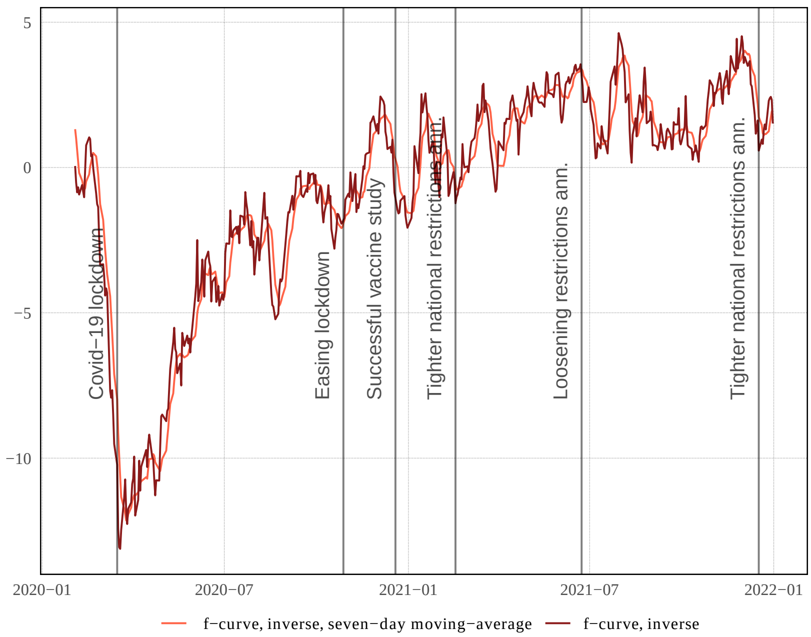
[2] In this context, "hand-selected" implies that we carefully choose keywords to define each topic by reviewing various lead texts.

Figure 1 — A fever curve for the Swiss economy

(a) Correlation with real GDP growth



(b) Evolution during the Covid–19 crisis



Notes: Panel (a) compares the *f*–curve (inverted and rescaled) to quarterly GDP growth. To calculate correlations, the *f*–curve is aggregated to quarterly. Periods of Covid–19 are defined as quarters 2 and 3 of 2020. Panel (b) panel gives daily values of the fever curve along with important policy decisions.

news about economic stimulus packages and the gradual loosening of the lockdown. The peak during the Covid–19 crisis is comparable to the Global Financial Crisis. But the speed of the downturn is considerably higher. In addition, the crisis is less persistent. By the end of July 2020, the f–curve improved to 1/4 of its peak value during the lockdown. The indicator also reflects the health situation very well. It improves continuously as measures are relaxed one after the other and starts to rise just before tougher restrictions are introduced.

To evaluate the informational content of the f–curve more formally, we conduct an in-sample assessment, as well as 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).[3] In-sample, the indicator has a coincident or leading relationship with many business cycle indicators. In addition, it Granger causes many of them. Therefore, it allows us 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 f–curve are more accurate than those using existing business cycle indicators. Therefore, the f–curve provides accurate and timely information about the state of the economy. We then investigate the role of the state of information and the predictive performance over the business cycle. The f–curve provides more accurate forecasts once one month of daily information of the current quarter is available. Considering that most monthly survey data for the current month are released at the start of the following month, the f–curve contains valuable information even after accounting for realistic publication lags. Furthermore, the findings emphasize that nowcasting performance improvements mostly occur during business cycle turning points, underscoring the f–curve's effectiveness in detecting recessions in real-time..

Various initiatives in Switzerland and abroad exist to satisfy the demand for reliable high-frequency information that emerged during the Covid–19 crisis. Becerra et al. (2020) and Eichenauer et al. (2021) develop sentiment indicators using internet search engine data. Brown and Fengler (2020) provide information on Swiss consumption behavior based on debit and credit card payment data. Eckert et al. (2020) develop a daily mobility index using data on traffic, payments, and cash withdrawals. 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., 2023). Moreover, Shapiro et al. (2022) create a daily news sentiment indicator that leads U.S.

[3] 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.

traditional consumer sentiment based on surveys.[4] News texts are also increasingly used for forecasting economic variables.[5] Ardia et al. (2019) show that news sentiments improve forecasts of U.S. industrial production growth. Ellingsen et al. (2022) find that news data are particularly informative for forecasting consumption developments. Moreover, Barbaglia et al. (2023) and Kalamara et al. (2022) demonstrate that it improves forecasts of macroeconomic variables such as GDP, inflation, and unemployment.

This chapter contributes to this literature in several ways. First, it investigates the information content of the daily financial market and news data for measuring business cycle fluctuations. So far, most studies in this field evaluated the information content of monthly indicators. Although daily business cycle indicators exist, there is no analysis of how many daily observations are needed to forecast GDP growth accurately. Second, unlike previous studies that relied on expensive or confidential data sources, we use publicly available financial market and news data, making it more accessible and cost-effective. Third, an innovation of this study lies in integrating financial market and text data. While studies evaluating the predictive value for both separately exist, their combined value has not been examined before. Moreover, the methodology to create news-based indicators for hand-selected economic concepts and the aggregation process is simple and robust to data leakage.[6]

The chapter proceeds as follows. In the next section, we describe the data and explain how they relate to the business cycle. Section 3 presents the methodology for creating text-based indicators and aggregating the data into a composite business cycle indicator. Moreover, it explains the mixed frequency methods for evaluating the indicators' out-of-sample performance. Section 4 provides a descriptive analysis and evaluates the f–curve in- and out-of-sample. The last section concludes.

2 Data

When selecting variables for the indicator, special attention is given to various properties. On the one hand, the data selection process is based on economic theory and intuition. On the other hand, daily data going back to at least the year 2000 is used, and the

[4] Research on the use of textual data for measuring and forecasting economic activity has already been conducted before the Covid–19 pandemic, but usually at monthly frequency (See Bybee et al., 2023; Larsen & Thorsrud, 2018; Shiller, 2019; Thorsrud, 2020, for a few examples).

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

[6] Data leakage is a situation where future information is accidentally incorporated into the model. Topic models (see, e.g. Thorsrud, 2020) are particularly prone to data leakage because they use the entire text to classify articles into topics.

time series must be freely and quickly available. The selected variables can be updated with a delay of one day, provided that the data providers' websites are accessible. The following description outlines the two data types underlying the f–curve.

2.1 Financial market data

We use publicly available bond yields underlying the SIX Swiss Bond Indices® (SBI) (SIX, 2020a). These data are accessible and updated daily, albeit with a one-day delay. To account for the fact that many bond yields only began around 2007, the series is extended by closely matching government and corporate bond yields from the Swiss National Bank. The detailed information can be found in Figure 2 in the Appendix.[7]

Various spreads are then calculated, which are expected to correlate with economic activity. These include the government bond term spread (8Y - 2Y), the interest rate differential relative to the euro area (1Y), and the risk premia of short- and long-term corporate debt. In addition to interest rate spreads for Switzerland, the risk premia of foreign companies that issue short- and long-term debt in Swiss francs are also computed. Term spreads for the U.S. and the euro area are included as well. For the latter, short-term interest rates in euro (European Central Bank, 2020) and long-term yields of German government debt (Deutsche Bundesbank, 2020) are utilized.

Furthermore, two implied volatility measures of the Swiss and U.S. stock markets are incorporated. The Swiss data is sourced from SIX (2020b) and is published with a one-day delay, while the U.S. data is obtained from the Chicago Board Options Exchange (2020).

These financial market data should be related to the Swiss business cycle. Stuart (2020) shows that the term spread exhibits a lead on the Swiss business cycle.[8] Kaufmann (2020) argues that a narrowing of the interest rate differential appreciates the Swiss franc and thereby dampens economic activity. Finally, risk premia are correlated with the default risk of companies, which should increase during economic crises. Finally, recent research documents increased uncertainty during economic downturns (Baker et al., 2016; Scotti, 2016). There are various ways to measure uncertainty (see, e.g. Dibiasi & Iselin, 2016). A measure of stock market volatility is preferred because the aim is to exploit quickly and freely available financial market data.

[7] Data from the Swiss National Bank are published with a longer delay. Therefore, these bond yields cannot be used to track the economy on a daily basis.

[8] Therefore, all term spreads are moved forward by half a year.

2.2 News lead texts

We complement the financial market data with publicly and quickly (usually with a delay of one day) 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),[9] the *Neue Zürcher Zeitung* (NZZ)[10], and the *Finanz und Wirtschaft* (FUW).[11] These three newspapers are among the most relevant German-language newspapers reporting on economic affairs in Switzerland and abroad. Their archives cover the period from 2000 until today.[12] For this chapter, we use data from January 1, 2000 to December 31, 2021.

One of the main benefits of using news data is its immediate availability and longer time coverage compared to other high-frequency data sources. Additionally, extracting information from the data is straightforward, and since the data is publicly accessible, it can be used by anyone. While the public availability of news data is a positive aspect, it should be noted that only the titles, lead texts, or specific passages of articles are typically available instead of full articles. However, this should not be considered a significant drawback. These lead texts often succinctly convey the article's main message and tend to have less extraneous information, making signal extraction more precise.

We further reduce the signal-to-noise ratio by utilizing solely texts about the economy. Articles about subjects unrelated 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, we 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).[13] As a small open economy, Switzerland is greatly affected by economic developments in other countries. To account

[9]See tagesanzeiger.ch/zeitungsarchiv-930530868737.

[10]See zeitungsarchiv.nzz.ch/archive.

[11]See fuw.ch/archiv.

[12]Sometimes the *Tages-Anzeiger* updates its archive with a relevant delay or not at all. Therefore, we additionally use lead texts from the *Tages-Anzeiger* website: tagesanzeiger.ch/wirtschaft.

[13]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 finding highlights that people's interest in the economy is not symmetrical. It 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).

for this, we create indicators that measure sentiments and recession prevalence for both Switzerland and foreign countries by using location-specific keywords.[14] For more information on the search queries used to filter out relevant articles, see Table 3 in the Appendix.

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 3 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.

Table 1 presents an overview of the web-scraped text data. Roughly 900'000 lead texts, text passages, or titles were collected. It is important to note that these 900'000 texts stem not from unique articles but are the results of multiple search queries. 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. An interesting observation is that 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. The lower sentiment found in the *Tages Anzeiger* webpage (TAW) is likely due to its different time coverage.

3 Methodology

This section describes the method of extracting information from the textual data, creating various sub-indices covering different areas of the economy, and aggregating all indicators into a business cycle indicator. Moreover, the models used for the

[14]We 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).

out-of-sample nowcasting exercise are explained.

3.1 Creating text-based indicators

To convert the high-dimensional and unstructured newspaper texts into time series, they must be preprocessed (cleaned). We, therefore, filter out irrelevant information, as is standard in the natural language processing (NLP) literature. We remove Hyper Text Markup Language (HTML) tags, punctuation, numbers, and stopwords, which are words that are not informative, typically conjunctions such as "or" and "if". The stop words are provided by Feinerer and Hornik (2019). Finally, we 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 unnecessary because we use a sentiment lexicon that is not stemmed.

The collected data is used to create text-based indicators that capture the different contexts (topics) of the economy in Switzerland and abroad. We use different sets of keywords denoted by $\mathcal{K}$ to 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.[15]

The first method creates recession indicators by counting the occurrence of keywords related to the recession topic 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 an R-word index for Switzerland and find that it has predictive power to forecast Swiss GDP growth.[16] These indicators measure economic uncertainty and recession fears and are negatively correlated with the business cycle.

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

[16] The recession indicators are highly correlated with uncertainty indicators invented by Baker et al. (2016). An evaluation has shown that creating uncertainty indicators provides no added value.

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 **krise** which also matches *krisengeschwächt*). The symbol \b reverses the wildcard operator (i.e \bspi\b doesn't match *spillovers*).

The second approach, the keyword-in-context (KWIC) method, utilizes word co-occurrences to calculate topic-specific sentiment indicators (Luhn, 1960). For this, we create new documents by screening the texts for keywords in $\mathcal{K}$. Whenever a keyword appears in a text, the keyword, along with the ten preceding and ten following words, is extracted into a new document.[17] A sentiment score is then calculated for each of these documents. Thus, the sentiment score is local because it considers only the text related to a topic of interest. Denote by $\mathcal{P}$ and $\mathcal{N}$ the list of phrases identified as 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., 2022; 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, we drop the subscript i in what follows. The document-level news sentiment is hence given by

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

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

where $N_{t,d,j}$ is the number of terms in the document. Figure 2 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.

Figure 2 — 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 article lead text 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 $S_{t,d,\text{industry}} = (1-2)/19 = -0.05$.

Several studies have used a probabilistic topic model to classify articles into topics (see, e.g., Ellingsen et al., 2022; Hansen et al., 2018; Thorsrud, 2020). With 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, assigning the article-level sentiment score to a specific topic becomes challenging due to the possibility of articles discussing multiple topics. The KWIC 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 the 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).[18]

3.2 Estimation of indicator

This chapter aims to provide a daily indicator of the business cycle and evaluate the value added of the daily frequency for nowcasting quarterly Swiss GDP growth. To accomplish this, we use models that can handle time series of different frequencies within the same regression without needing to transform or aggregate them. MIDAS and bridge models are commonly used in the literature for this purpose. However, including many explanatory variables in a MIDAS model can lead to parameter proliferation.[19] Additionally, financial market data and news indicators are quite volatile and correlated. To effectively summarize the information content of the data and eliminate idiosyncratic noise while avoiding parameter proliferation, we estimate a factor model in static form:[20]

$$x_{i,t} = \lambda_i f_t + e_{i,t} \tag{2}$$

[18]Nevertheless, we 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, which 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 those of the KWIC method used here.

[19]Parameter proliferation in the context of MIDAS models refers to the issue that arises when too many parameters are included in the model.

[20]The news indicators are much more volatile than the financial market data (see Figures 1, 3 and 4 in the Appendix). We, 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., 2022; Thorsrud, 2020). The moving average time window choice is a trade-off between less volatility and more timeliness. We tried different time windows and found a good compromise with the ten-day window.

where $x_{i,t}$ denotes one of the $n = 24$ indicators $(i = 1, ..., n)$, λ_i is the factor loading, f_t are the common factors at time t, and e_t is the idiosyncratic component. The model comprises T daily observations $(t = 1, ..., T)$. The advantage of using a factor model is that it allows for summarizing the information in a large data matrix with a small number of common factors. Factors and loadings can be estimated through principal components assuming that the idiosyncratic components are only weakly serially and cross-sectionally correlated (Bai & Ng, 2013; Stock & Watson, 2002).[21]

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 Bai and Ng (2002) confirms that one factor is representing the data well enough.[22] We use the information criterion BIC_3, which is recommended for $n > 18$.

3.3 Out-of-sample evaluation models

How reliable is the f–curve and what is the informational content of the daily frequency? We perform a daily pseudo-real-time forecast evaluation using mixed frequency methods to answer these questions.

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. We use the real-time data set for quarterly GDP vintages by Indergand and Leist (2014), accounting for the ragged-edge structure due to the

[21] We exclude weekends and holidays. Then, we interpolate additional missing values using an EM algorithm (Stock & Watson, 2002) after standardizing the data to have zero mean and unit variance. We choose a relatively large number of factors for interpolating the data ($r = 4$). Finally, we use the first principal component of the interpolated data set. As a robustness test, we additionally estimate the factor model with missing values using least squares instead of the EM algorithm. The results are almost identical. Consider the model

$$x_{i,t} = \lambda_i f_t + e_{i,t}$$

and let $\delta_{i,t} = 1$ if $x_{i,t}$ is observed and zero otherwise. Then the principal components estimates of λ_i and f_t solve

$$\min_{\lambda_i, f_t} \sum_{t=1}^{T} \sum_{i=1}^{n} (x_{i,t} - \lambda_i f_t)^2 \delta_{i,t}$$

subject to the normalization

$$\sum_{i=1}^{n} \lambda_i' \lambda_i = I_r.$$

The problem can then be solved by using least squares iteratively.

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

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. We 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), we 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, we 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 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 of $T_x \geq T_y$. The sample spans from January 1, 2000, to December 31, 2021.

To determine the informational content of the f–curve we forecast GDP growth using three models that exploit the information in the high-frequency indicator and link it to the low-frequency GDP. First, we estimate a MIDAS model introduced by Ghysels et al. (2004) and Ghysels et al. (2007). Second, we employ bridge equations following Baffigi et al. (2004). Third, we consider an iterative MIDAS model, a mixture of both, as discussed by Schumacher (2016).

3.3.1 MIDAS model

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

$$y_{t_q+H+1} = \alpha + \sum_{p=0}^{P-1} \beta_p \sum_{k=0}^{K-1} b(k, \theta) L^{(pD+k)/D} x_{t_d+T_x-T_y} + \varepsilon_{t+H+1} \tag{3}$$

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. We 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 $L^{1/60} x_{t_d} = x_{t_d - 1/60}$.

We 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.[23] Because x_{t_d} is sampled at a much higher frequency than y_{t_q}, we potentially have to include many high-frequency lags to achieve adequate modeling. This can easily lead to overparameterization in the unrestricted linear case. We use a non-linear weighting scheme given by the polynomial $b(k, \theta)$ to avoid parameter proliferation. The same polynomial specification is applied to all low-frequency lags included in the model.[24]

We use an exponential Almon lag of order two for the polynomial specification. This polynomial is extensively discussed in Ghysels et al. (2007) and has the following form:[25]

$$b(k, \theta) = \frac{exp(\theta_1 k + \theta_2 k^2)}{\sum_{j=0}^{K} exp(\theta_1 j + \theta_2 j^2)}. \tag{4}$$

As 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 $P = 2$ we only have to estimate five parameters. The parameters are estimated by non-linear least squares (NLS). Since MIDAS models are a direct forecasting tool and depend on the forecast horizon H, we have to estimate a model for every H and re-estimate them whenever 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. We estimate a bridge model for a

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

[24] We also estimate a model with different polynomial specifications for every included low-frequency lag. However, this led to converging issues in the non-linear least squares (NLS) estimation for some periods, which deteriorates forecasting performance. Results can be requested from the author.

[25] For robustness, we also use a Legendre polynomial proposed by Babii et al. (2021). The results, however, are less promising and shown in Table 4 in the Appendix.

forecast horizon H of the following form:

$$y_{t_q+H+1} = \alpha + \sum_{p=0}^{P-1} \beta_p L^p x_{t_q+H+1} + \varepsilon_{t_q+H+1} \tag{5}$$

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

$$x_{t_q} = \sum_{k=0}^{K-1} \omega(k) L^{k/D} x_{t_d} = \sum_{k=0}^{K-1} \omega_k L^{k/D} x_{t_d} \tag{6}$$

is the time-aggregated high-frequency variable. The aggregation function, $\omega(k)$, depends on the nature of the indicator. Here it is a simple, equal-weighted average (i.e., $\omega_k = 1/D \quad \forall k$). The bridge equation in (5) can be estimated by OLS only on sample periods where all the high-frequency variables are available. We first need to forecast the high-frequency indicator to the end of the desired quarters to get a forecast of the low-frequency variable. To do so, we use an AR(p) model where the Bayesian Information Criterion (BIC) determines the lag order. These predictions are then aggregated according to equation (6) and plugged into the estimated equation (5).

3.3.3 Iterative MIDAS

The iterative MIDAS (MIDAS-IT) model was introduced by Schumacher (2016). It is an intermediate model between bridge and MIDAS. In principle, it is a bridge model where the aggregation function $\omega(k)$ is replaced with the restricted weighting polynomial $b(k, \theta)$. As for the bridge model, we use an AR(p) model to forecast the indicator variable until the end of the desired quarters. We use the same polynomial specification as for the MIDAS model. Using these three model types allows us to identify the advantages of selected aspects of MIDAS and bridge models.

3.3.4 Benchmarks

The forecasts are compared to three benchmarks. First, we use an autoregressive model (AR) of order one estimated on the corresponding real-time vintage for GDP growth. Second, using bridge equations, we 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, we use the release from March 2022. For the out-of-sample exercise, we assume that the Barometer 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, we 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, we consider the initial quarterly GDP release to be a forecast of the final GDP outcome. To compute the forecast errors, we use the release of quarterly GDP from December 2021.

4 Evaluation of the f–curve

This section first provides a descriptive analysis of the f–curve and its underlying indicators. It then demonstrates the in-sample information content of the f–curve, highlighting that it is available earlier than most 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 f–curve is correlated with many key macroeconomic variables (See Figure 6 in the Appendix).

4.1 Descriptive analysis

Most of the indicators underlying the f–curve are substantially correlated with GDP growth (after pre-whitening with an AR(p) model (see Neusser, 2016, Ch. 12.1)). Figure 3 provides an overview of the cross-correlations. The absolute correlation coefficients range from 0.26 to 0.75. The coefficients for financial market data (panel a) and the text-based recession indices (panel b) are negative and significant. The term spreads are somewhat less strongly correlated. Moreover, term spreads for Switzerland and the euro area lead GDP growth by about half a year. This confirms that uncertainty measures such as risk premia and volatility indicators are highly correlated with the business cycle. The news sentiment indices in panel (b) feature a strong positive correlation with GDP growth. This correlation indicates that journalists' discussion of the economy holds valuable information about economic activity.

The f–curve is determined by computing the first principal component using all these variables. It is, therefore, valuable to assess the individual contributions of each variable to the factor. These contributions are represented by the factor loadings, which indicate how well the factor explains each original variable. Variables with higher absolute loadings are more strongly associated with the factor. Table 3 presents the factor loadings for the input variables, indicating their relationship with the f–curve. Positive loadings indicate a positive (negative) association with the business cycle (f–curve), while negative loadings suggest a negative (positive) association. The factor loadings show that the Swiss business cycle is well represented by news sentiment on the general economy, the financial markets, investment, and the labor markets. Risk premia, volatility indices, and text-based recession indices also contribute substantially to the factor. By contrast, news sentiments on inflation and politics, as well as term spreads and the interest rates differential, have lower loadings and contribute only a little. Finally, foreign variables tend to have higher factor loadings than domestic variables, reflecting Switzerland's strong dependence on foreign countries.

Table 3 — Factor loadings

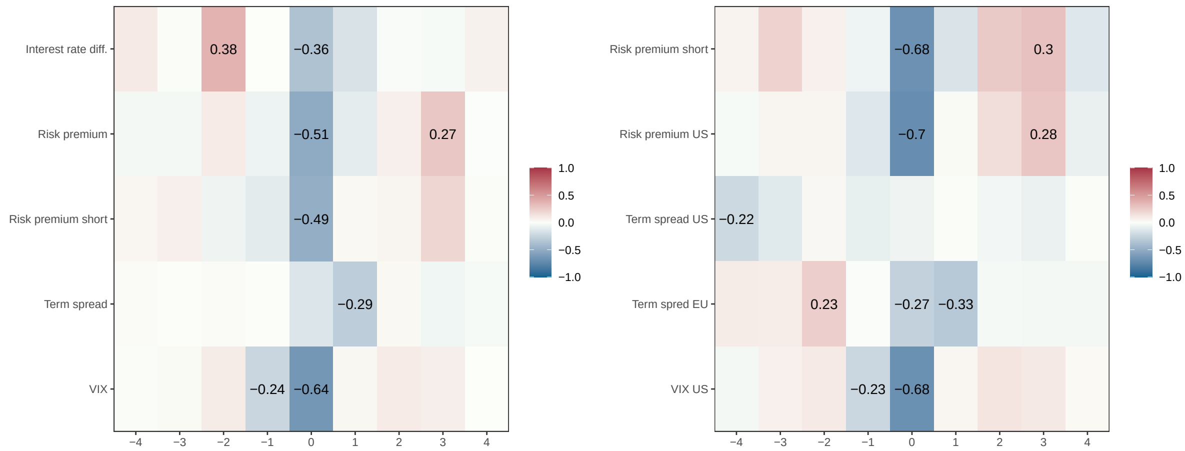
Variable	Country	Loading
News Economy	Foreign	0.29
Long term Risk Premium	Foreign	-0.27
News Financial Market	Foreign	0.27
News Economy	Switzerland	0.26
Long term Risk Premium	Switzerland	-0.25
News Recession	Foreign	-0.24
News Recession	Switzerland	-0.24
VIX	USA	-0.23
News Investment	Foreign	0.23
Short term Risk Premium	Switzerland	-0.23
News Labor Market	Foreign	0.23
VIX	Switzerland	-0.22
News Financial Market	Switzerland	0.21
Short term Risk Premium	Foreign	-0.20
News Labor Market	Switzerland	0.19
News Investment	Switzerland	0.18
News Inflation	Foreign	0.18
News Politics	Foreign	0.18
News Inflation	Switzerland	0.14
News Politics	Switzerland	0.07
Term Spread	USA	0.07
Interest rate differential	Switzerland	-0.07
Term Spread	Euro Area	-0.03
Term Spread	Switzerland	-0.02

Notes: Factor loadings of the first principal component (f–curve). Positive loadings indicate a negative association with the factor, while negative loadings suggest a positive association.

After updating the f–curve on daily basis over one year since May 2020, it is possible to judge the actual real-time performance of the indicator. Figure 4 provides preliminary

Figure 3 — Cross-correlations of data underlying indicator with real GDP growth

(a) Financial market data



(b) Text-based indicators

Notes: Cross-correlation between financial market data (Panel a) and news-based sentiment indicators (Panel b) underlying the f–curve and real Swiss GDP growth. On the left, indicators for domestic data are displayed, while indicators for foreign data are on the right. The sample ends in 2020 Q1. We aggregate all data to quarterly frequency. Only statistically significant correlations at displacement s given on the x-axis are labeled. A significant correlation at $s > 0$ means the series 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.

results on how many input variables of the f–curve are available. The indicator is revised because not all data series are available in real-time (ragged edge problem). It shows results over the first one and a half year of daily updating the indicator. On average, more than 12 out of 24 series are available with a delay of one day. After three days, almost all indicators are available. The main reason why the average lies below

24 is that the archive of the *NZZ* was not accessible between November 2020 and March 2021. This led to a rather large revision. Moreover, the *Tages-Anzeiger* has not been updated for a couple of months in 2020. For this reason, we augmented the indicator with information from this newspaper's online edition. Finally, on rare occasions, the websites of financial sources were not available. All revisions to the indicator are solely because specific sources are unavailable at the time of the update. The inputs of the f–curve are not revised. The availability of inputs within one to two days and relatively small revisions if a data source was inaccessible confirms the indicator's excellent suitability for real-time tracking of the Swiss economy.

4.2 In-sample analysis

Because the f–curve is a combination of sentiment indicators covering several economic topics and financial market indicators, it is correlated with many key macroeconomic variables (see Figure 6 in the Appendix). Since it is not optimized to track any particular measure of economic activity in its current form, we evaluate the in-sample information content of the f–curve.

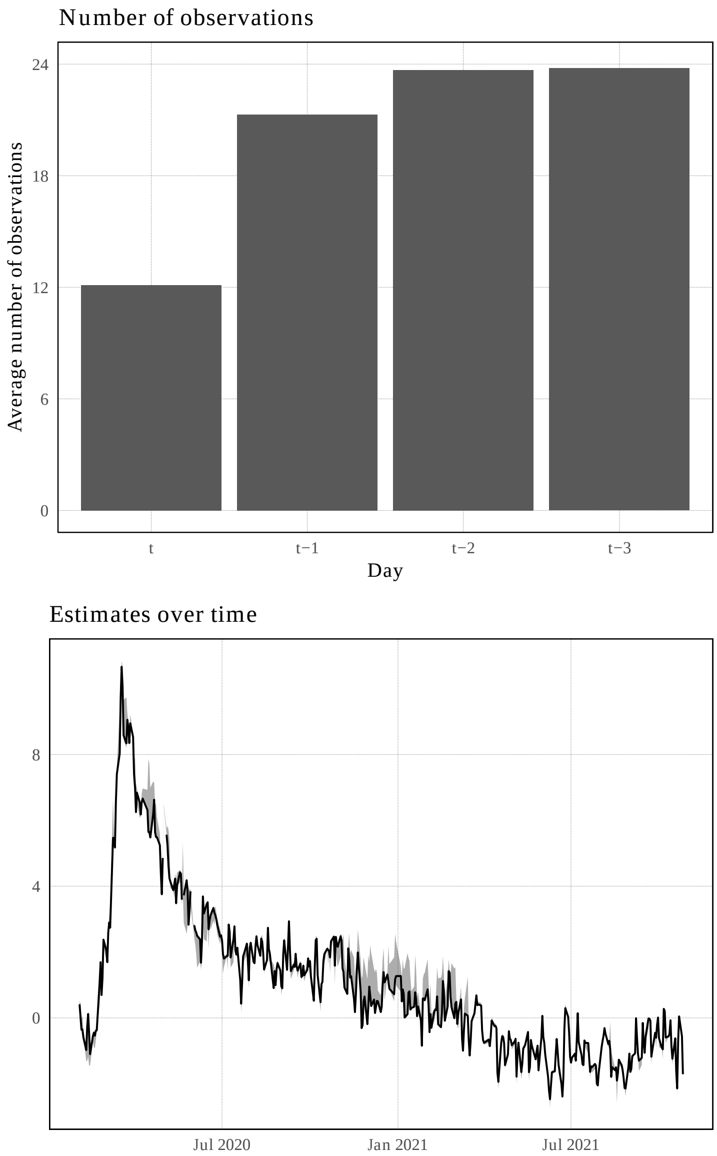
To compare the in-sample information content of the f–curve to other leading indicators, we perform a cross-correlation test (see Neusser, 2016, Ch. 12.1).[26] Figure 5 shows a substantial correlation between the f–curve and many prominent leading indicators. There is a coincident or a leading relationship with the KOF Economic Barometer, SECO's Swiss Economic Confidence (SEC), consumer confidence, and trendEcon's perceived economic situation.[27] There is a leading, a coincident, and a lagging relationship with the Organisation for Economic Co-operation and Development composite leading indicator (OECD CLI) and with the SNB's Business Cycle Index (BCI). However, the OECD CLI is a smoothed indicator subject to substantial revisions. TrendEcon's perceived economic situation starts only in 2006 and the BCI is published with a relevant delay.

Another way of assessing the in-sample information of the f–curve is to test whether it is Granger causing other indicators (Granger, 1969). The f–curve, f_t, is said to Granger cause another indicator, I_t, if it contains statistically significant information about the future values of the other indicator. Therefore, the f–curve should Granger cause the indicator. To test this condition, the following model is estimated

[26] 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 1 in the Appendix); finally, some are based on lagged data (see, e.g., OECD, 2010).

[27] All data sources are given in Table 1 in the Appendix.

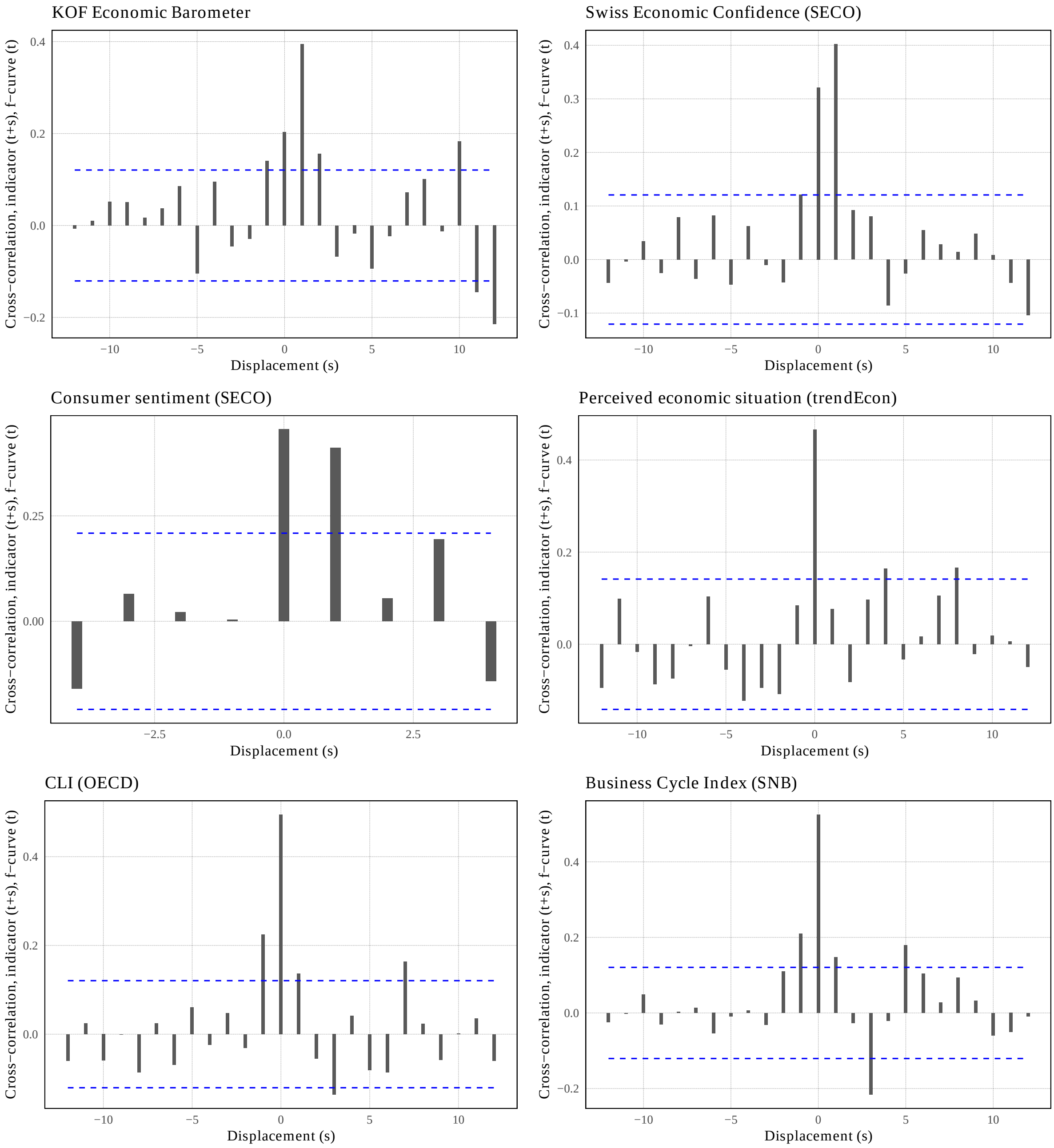
Figure 4 — Real-time results since initial version of the f–curve



Notes: Average number of variables available for calculation of the f–curve (top figure). The gray shaded area represents minimum and maximum estimates from 5 May 2020 to 27 October 2021 (bottom figure). Estimates over time are based on the published version (Burri & Kaufmann, 2020). See https://github.com/dankaufmann/f-curve for more information.

$$I_t = \sum_i \alpha_i f_{t-i} + \sum_i \beta_i I_{t-i} + \varepsilon_t. \tag{7}$$

Figure 5 — Cross-correlation with other indicators



Notes: Cross-correlation between the *f*–curve and other prominent leading and sentiment indicators. We 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.

The joint hypothesis that $\alpha_i = 0$ is then tested using a Wald test. The rejection of the hypothesis suggests that the f–curve Granger causes the indicator. The other indicator should not Granger cause the f–curve. To test this condition, the following model is estimated

$$f_t = \sum_i \alpha_i f_{t-i} + \sum_i \beta_i I_{t-i} + \varepsilon_t \tag{8}$$

and the joint hypothesis that $\beta_i = 0$ is tested. To satisfy the condition, this hypothesis should not be rejected.

Table 4 — Testing Granger causality properties of f–curve

Indicator	f–curve granger causes Ind.	Ind. doesn't granger cause f–curve
KOF Barometer	✓	✓
SECO SEC	✓	✓
Consumer Sentiment	✓	✓
Trendecon	✓	✓
OECD CLI	✓	✗
SNB BCI	✓	✗
SECO WEA	✓	✓

Notes: All tests are conducted on the 5% significance level.

If both conditions are satisfied, there is a one-directional Granger causality. As shown in Table 4, the f–curve Granger causes all indicators under consideration. However, the f–curve is Granger caused by the OECD CLI and the SNB BCI. The reason for this might be technical and due to smoothing or estimation procedures to calculate the indicators. The tests for Granger causality confirm the conclusions from the cross-correlation tests above.

Overall, these results suggest the f–curve provides sensible information comparable with other existing indicators. The key advantage of the f–curve is its prompt availability and availability over a longer period.

4.3 Out-of-sample evaluation

Table 5 presents the relative root-mean-squared errors (RMSE) of the pseudo-real-time out-of-sample nowcasting exercise.[28] To better understand the results, we conduct a sub-sample analysis by excluding quarters two and three of 2020. Further, we exclude

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

Table 5 — Real-time evaluation: Relative RMSE and DMW tests

	Full sample				Excluding Covid crisis				Excluding all crisis periods			
Horizon	0	40	80	120	0	40	80	120	0	40	80	120
a) Hypothesis: Model > AR(1) model												
Bridge	0.64	0.84**	1.02	1.1	0.82*	0.78*	1.03	1.03	0.83***	0.76*	1	0.93
Midas	0.73	0.88*	0.98	1.08	0.77**	0.85	1.07	1.01	0.73***	0.82	1.07	0.92
Midas-IT	0.73	0.88*	0.98	1.06	0.78**	0.81*	1.01	0.98	0.75***	0.74**	0.98	0.88
b) Hypothesis: Model > Barometer bridge												
Bridge	1.04	0.93*	0.97	1.1	0.82**	0.86*	1.02	1.04	0.73**	0.81**	0.93	0.95
Midas	1.17	0.98	0.93	1.07	0.77**	0.94	1.05	1.02	0.65***	0.87	0.99	0.94*
Midas-IT	1.17	0.97	0.93	1.06	0.78**	0.89*	1	0.99	0.66***	0.78**	0.91	0.9*
c) Hypothesis: Model < First Release												
Bridge	1.97*				1.13				1.17*			
Midas	2.23*				1.06				1.03			
Midas-IT	2.23*				1.06				1.04			

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 (where $A > B$ means A has lower RMSE than B) written in the row header (Diebold & Mariano, 2002; West, 1996). I assume a quadratic loss function. Significance levels are given by: *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$

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 5, 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.

Similar patterns are observed when employing a bridge equation model with the KOF Economic Barometer as the benchmark. Panel (b) of Table 5 shows that the f–curve 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, the f–curve 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 f–curve 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 f–curve 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, we 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.[29] Additionally, 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 than the MIDAS model's iterative approach. Secondly, the employment of a nonlinear polynomial specification, despite requiring the estimation of more parameters, proves to be beneficial.

The performance of the f–curve 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 6 illustrates the relative RMSE of the MIDAS-IT model compared to these benchmark models for the current quarter nowcast. To prevent significant nowcasting errors during the pandemic from distorting the analysis, the errors are calculated on the sample excluding the Covid-19 crisis.[30] 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.

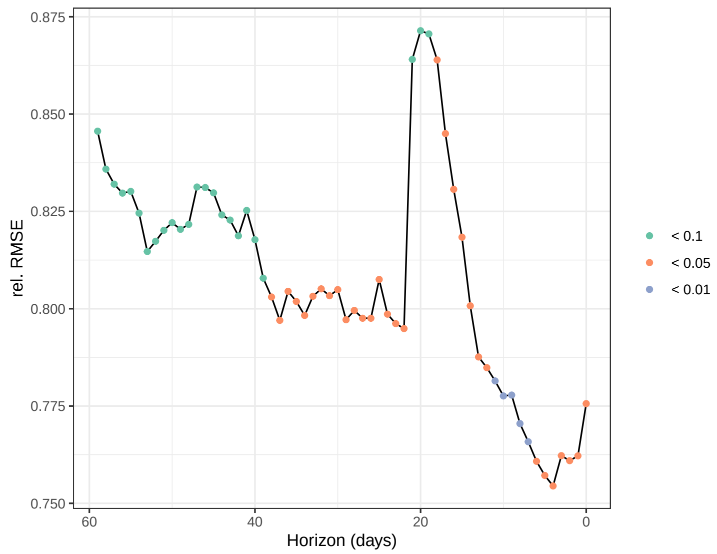
Examining panel (a) of Figure 6, which displays the relative RMSE against the AR(1) model, two key observations emerge. First, starting from a horizon of 40 days, the MIDAS-IT model significantly outperforms the AR(1) model in nowcasting the current quarter's GDP growth. Second, as indicated by spikes in the relative RMSE

[29]Figure 9 in the Appendix presents the absolute RMSE of the out-of-sample nowcasting exercise

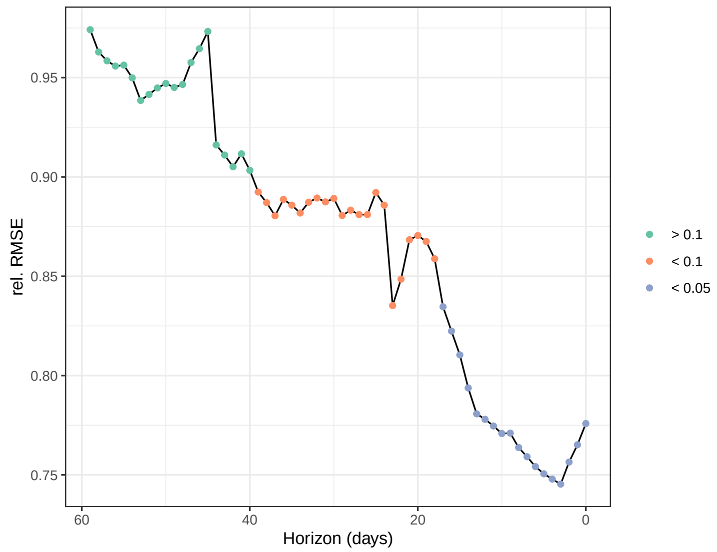
[30]For comparison Figure 7 in the Appendix shows the evolution of the relative RMSE for the sample excluding all crisis periods.

Figure 6 — Relative RMSE of daily current quarter nowcasts

(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. We use two benchmarks. First, we use an AR(1) model (panel a). Second, we 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). We assume a quadratic loss function. Significance levels are given by: ● $p > 0.1$, ● $p < 0.1$, ● $p < 0.05$, ● $p < 0.01$

at a horizon of around 20 days, the AR(1) model demonstrates increased accuracy relative to the MIDAS-IT model when new GDP vintages are released. Panel (b) shows a similar behavior of the relative RMSE against the KOF Barometer bridge model. The f–curve starts surpassing the KOF Barometer at a horizon of around 40 days, approximately after the first month of the current quarter has passed.

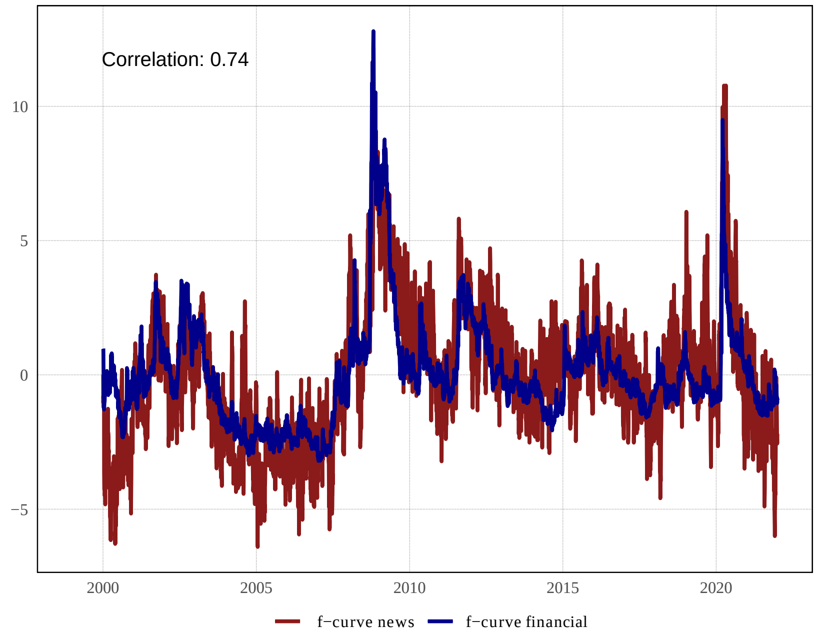
Are the financial market or news data more important for the forecasting performance of the f–curve? Figure 7 shows two indicators only calculated with financial market and news data, respectively. Although the indicators are highly positively correlated, there is one key difference. The news data are more volatile.[31] This suggests that financial market data provide a more accurate signal of the business cycle than the news data. Looking at relative RMSEs in tables 5 and 6 in the Appendix, this cannot be confirmed. For different models and different time horizons, sometimes the news-based f–curve and at other times the financial f–curve perform better. For larger time horizons, the financial f–curve outperforms the news-based f–curve. Hence, the more volatile news data does not worsen the f–curve because the factor model, including financial market data, removes the idiosyncratic fluctuations. The overall f–curve therefore benefits more from the financial market data for some nowcasting horizons and more from the news data for others.[32]

Having established the usefulness of daily data for nowcasting, an important question arises: when does this data exert the most significant influence on forecast accuracy? To explore this, Figure 8 in the Appendix 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 f–curve-based model. The analysis reveals that most performance improvements relative to the benchmarks are concentrated during the financial crisis. In a broader context, forecast enhancements are often observed during turning points. These findings are consistent with recent literature on the use of textual data for predicting economic activity (see Barbaglia et al., 2023; Ellingsen et al., 2022; Kalamara et al., 2022). However, it is essential to acknowledge that the f–curve-based model exhibits certain weaknesses

[31] This is also because we smooth the news indicators with a moving average of only ten days. Comparable studies smooth over a much longer period. For example, Thorsrud (2020) uses a moving average of 60 days. On the one hand, this reduces the volatility of the news sentiment. On the other hand, this renders the indicator less helpful in detecting rapid daily changes.

[32] An interesting avenue for future research would be to conduct forecast encompassing tests (Harvey, 1989).

Figure 7 — Comparison news and financial market data



Notes: The graph shows two indicators estimated only on financial market and news data, respectively.

in nowcasting recovery periods. The Barometer bridge model, for example, was more accurate during the recovery from the GFC and the Covid–19 crisis. Nonetheless, it is crucial to note that the overall strong performance of the f–curve cannot be solely attributed to recession periods.

5 Concluding remarks

We develop a daily indicator of Swiss economic activity. The indicator consists of publicly available financial market and news data. It has the interpretation of a fever curve: As for monitoring the temperature of a patient, an increase in the fever curve provides a reliable and timely warning signal if health takes a turn for the worse.

A major strength of the indicator is that it can be updated with a delay of only one day and it is hardly revised ex-post. An evaluation of the indicator shows that it correlates with other business cycle indicators and accurately tracks Swiss GDP growth. 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 indicator outperforms a widely-used Swiss business cycle indicator in nowcasting GDP growth. In line with recent literature on the subject, the analysis demonstrates that the indicator significantly improves nowcasting performance compared to the benchmark models, particularly during turning points. Therefore, the f–curve provides an accurate and flexible framework to track Swiss economic activity at high frequency.

There is still room for improvement. We see six promising avenues for future research. First, the news sentiment indicators could exploit other publicly available news sources, particularly newspapers from Switzerland's French- and Italian-speaking parts. Second, instead of publicly available news lead texts, full texts of these articles could be incorporated. Third, customizing the lexicon specifically for economic news, as exemplified by Shapiro et al. (2022), has the potential to augment the accuracy of sentiment analysis in capturing economic trends. Fourth, we could examine the predictive ability of multiple factors and for other macroeconomic data. Fifth, the information could be used to disaggregate quarterly GDP and industrial production into monthly or weekly series. Finally, it would be valuable to investigate the potential gains obtained by incorporating all individual indicators in a nowcasting model, with the sparse-group LASSO-MIDAS model proposed by Babii et al. (2021) offering a promising avenue for exploration in this context. Exploiting all this new information will likely further improve our understanding of 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*. 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. *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. (2023). Forecasting with economic news. *Journal of Business & Economic Statistics*, *41*(3), 708–719. 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. (2023). Do daily lead texts help nowcasting GDP growth? *IRENE Working Papers 23-02*.

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. T., Manela, A., & Xiu, D. (2023). Business news and business cycles. *Journal of Finance, Forthcoming*. http://dx.doi.org/10.2139/ssrn.3446225

Chicago Board Options Exchange. (2020). CBOE volatility index: VIX. https://fred.stlouisfed.org/series/VIXCLS

Deutsche Bundesbank. (2020). Zeitreihe BBK01.WT1010: Rendite der jeweils jüngsten bundesanleihe mit einer vereinbarten laufzeit von 10 jahren. https://www.bundesbank.de/dynamic/action/de/statistiken/zeitreihen-datenbanken/zeitreihen-datenbank/723452/723452?tsId=BBK01.WT1010

Dibiasi, A., & Iselin, D. (2016). *Measuring uncertainty* (KOF bulletin No. 101). KOF Swiss Economic Institute, ETH Zurich. https://ethz.ch/content/dam/ethz/special-interest/dual/kof-dam/documents/KOF_Bulletin/kof_bulletin_2016_11_en.pdf

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. *KOF Working Papers No. 488*. https://doi.org/10.3929/ETHZ-B-000458723

Eichenauer, V. Z., Indergand, R., Martínez, I. Z., & Sax, C. (2021). Obtaining consistent time series from google trends. *Economic Inquiry*, ecin.13049. https://doi.org/10.1111/ecin.13049

Ellingsen, J., Larsen, V. H., & Thorsrud, L. A. (2022). News media versus FRED-MD for macroeconomic forecasting. *Journal of Applied Econometrics*, *37*(1), 63–81. https://doi.org/10.1002/jae.2859

European Central Bank. (2020). Yield curve spot rate, 1-year maturity - government bond, nominal, all issuers whose rating is triple. https://sdw.ecb.europa.eu/browseExplanation.do?node=qview&SERIES_KEY=165.YC.B.U2.EUR.4F.G_N_A.SV_C_YM.SR_1Y

Feinerer, I., & Hornik, K. (2019). *Tm: Text mining package*. 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 2004s-20*.

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

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*(3), 424. https://doi.org/10.2307/1912791

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

Harvey, C. R. (1989). Forecasts of economic growth from the bond and stock markets. *Financial Analysts Journal*, *45*(5), 38–45. https://doi.org/10.2469/faj.v45.n5.38

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*, *37*(5), 896–919. https://doi.org/10.1002/jae.2907

Kaufmann, D. (2020, April). *Wie weiter mit der tiefzinspolitik? szenarien und alternativen* (IRENE policy reports No. 20-01). IRENE Institute of Economic Research. https://ideas.repec.org/p/irn/polrep/20-01.html

Kaufmann, D., & Scheufele, R. (2017). Business tendency surveys and macroeconomic fluctuations. *International Journal of Forecasting*, *33*(4), 878–893. https://doi.org/10.1016/j.ijforecast.2017

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 No. 7468*.

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. *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)*.

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

Scotti, C. (2016). Surprise and uncertainty indexes: Real-time aggregation of real-activity macro-surprises. *Journal of Monetary Economics*, *82*, 1–19. https://doi.org/10.1016/j.jmoneco.2016.06.002

Shapiro, A. H., Sudhof, M., & Wilson, D. J. (2022). Measuring news sentiment. *Journal of Econometrics*, *228*(2), 221–243. https://doi.org/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.

SIX. (2020a). SBI–swiss bond indices. https://www.six-group.com/exchanges/indices/data_centre/bonds/sbi_en.html

SIX. (2020b). VSMI–volatility index on the SMI. https://www.six-group.com/exchanges/indices/data_centre/strategy_indices/vsmi_en.html

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

Stuart, R. (2020). The term structure, leading indicators, and recessions: Evidence from switzerland, 1974–2017. *Swiss Journal of Economics and Statistics*, *156*(1), 2–17. https://doi.org/10.1186/s41937-019-0044-4

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. (2023). Weekly economic activity: Measurement and informational content. *International Journal of Forecasting*, *39*, 228–243. 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 Supplementary material

Table 1 — 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.

Figure 1 — Daily financial market indicators for *f*–curve

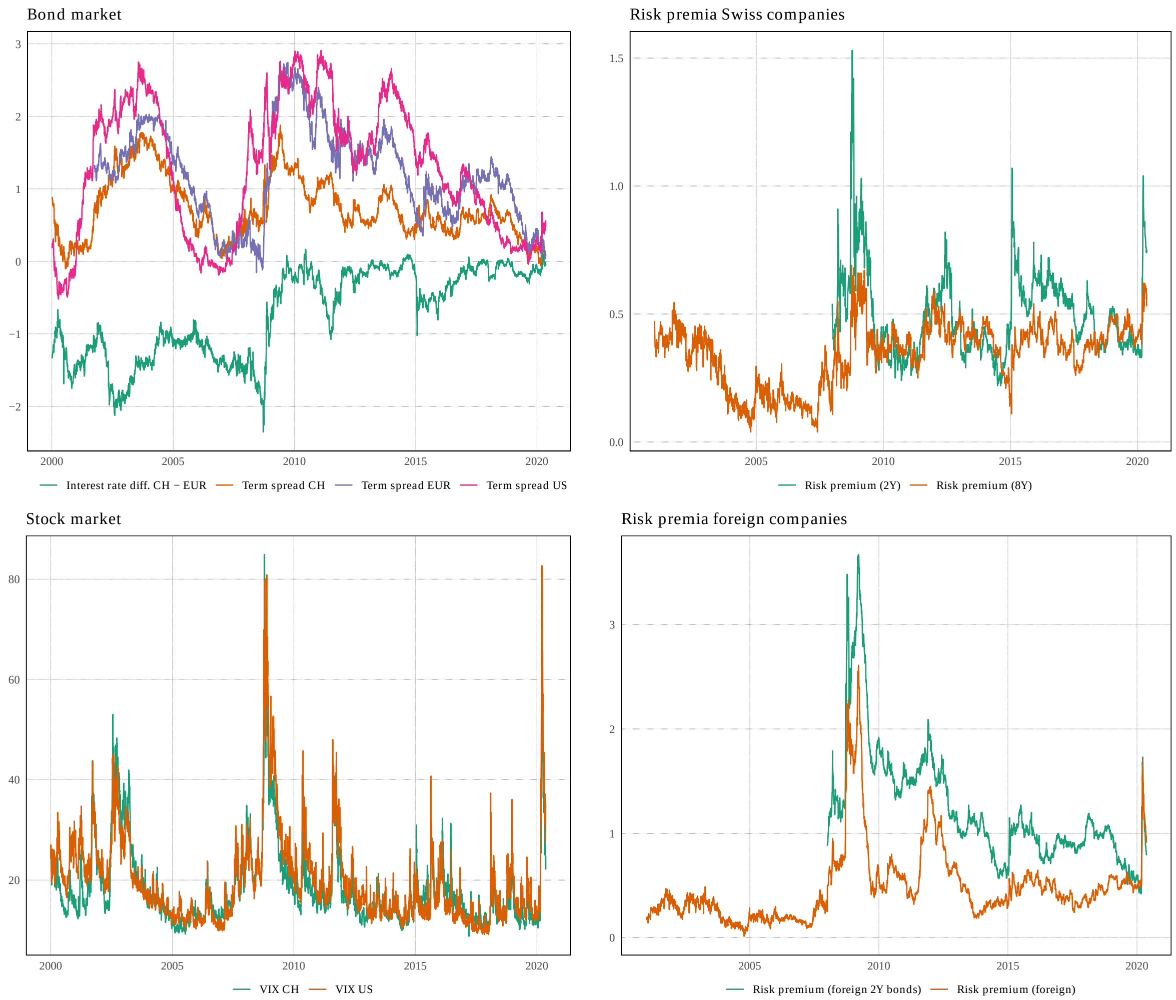


Table 2 — Data underlying f–curve

	Type	Publication	Frequency	Source	Comments
Term spread CH	f–curve	+1 day	Day	SIX, SNB	8Y – 2Y. SNB data used before SIX data available. Maturity of SIX data is approximate
Term spread USA	f–curve	+1 day	Day	Fed Board	10Y – 2Y
Term spread Europe	f–curve	+1 day	Day	Buba, ECB	10Y Germany – 1Y euro area. 1Y EUR Libor used before 2004
Risk premium CH	f–curve	+1 day	Day	SIX, SNB	8Y AAA-AA – 8Y government. SNB data for debt issues by banks used before SIX data available. Maturity of SIX data is approximate
Short-term risk premium CH	f–curve	+1 day	Day	SIX	1-3Y AAA-BBB – 1-3Y government. Start in 2008
Risk premium foreign	f–curve	+1 day	Day	SIX, SNB	8Y Foreign corp. – 8Y government. SNB data used before SIX data available. Average of various credit ratings
Short-term risk premium foreign	f–curve	+1 day	Day	SIX	1-3Y AAA-AA – 8Y government. Start in 2008
Stock market volatility CH	f–curve	+1 day	Day	SIX	
Stock market volatility USA	f–curve	+1 day	Day	CBOE	
Interest rate differential	f–curve	+1 day	Day	SIX, SNB, ECB	1-3 year government bonds CH – 1 year government bond yields euro area. 1Y EUR Libor used before 2004
Domestic news sentiment	f–curve	+1 day	Day	FuW, NZZ, TA	More details in Table 3
Foreign news sentiment	f–curve	+1 day	Day	FuW, NZZ, TA	More details in Table 3

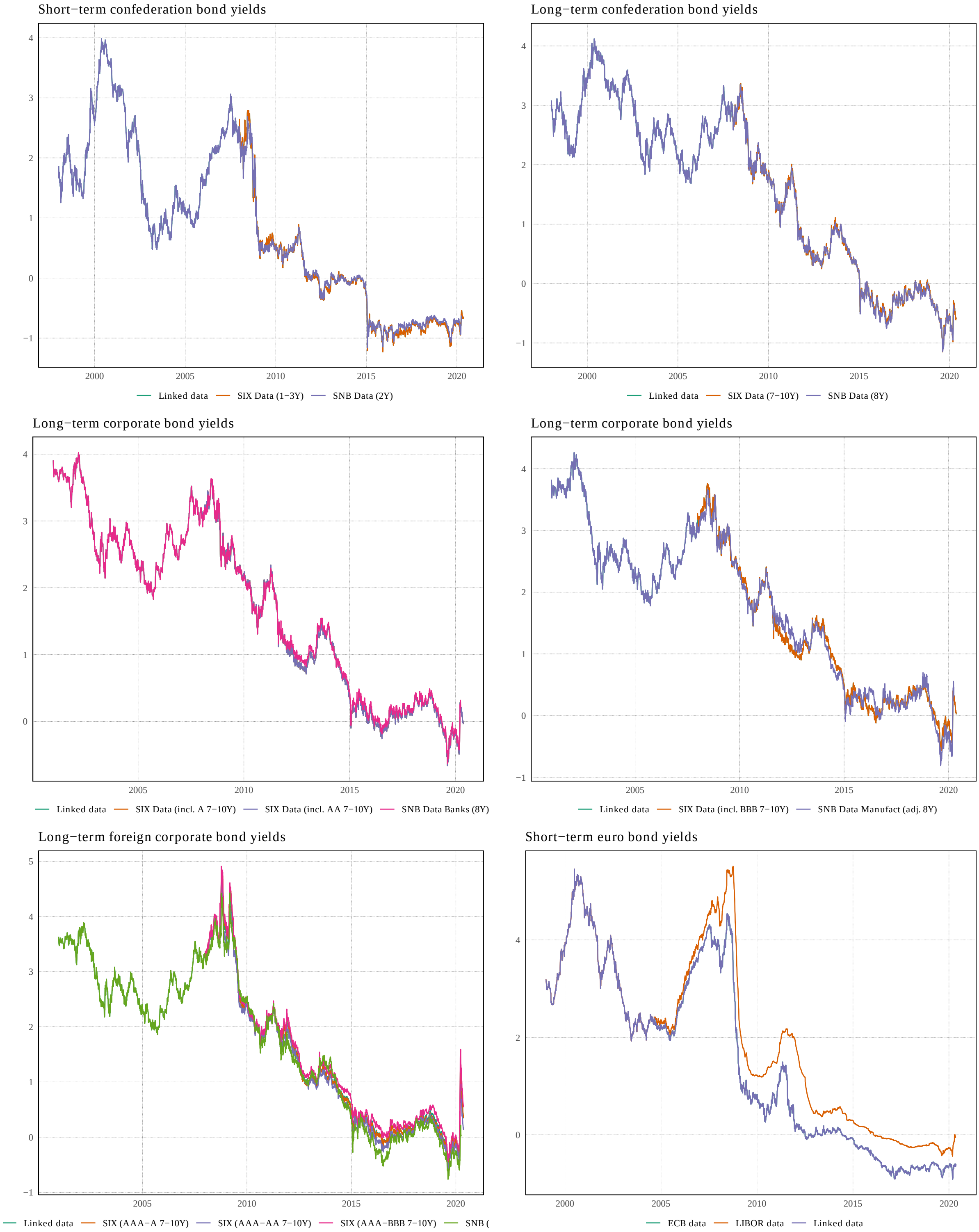
Notes: The SIX Swiss Exchange AG disclaimer applies to the SIX data: https://www.six-group.com/exchanges/download/market/data_services/six_disclaimer.pdf

Table 3 — Queries underlying news indicators

	URL	Keywords
	Domestic news sentiment	
FuW	fuw.ch/unternehmen/ fuw.ch/makro/	We use all articles listed in *Makro* and *Unternehmen* and select those containing the word *schweiz** in either lead text, tag or category.
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 sentiment	
FuW	fuw.ch/unternehmen/ fuw.ch/makro/	We 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, we 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 2 — Spliced data underlying f–curve



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 the 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.

Figure 3 — Daily news based recession indicators

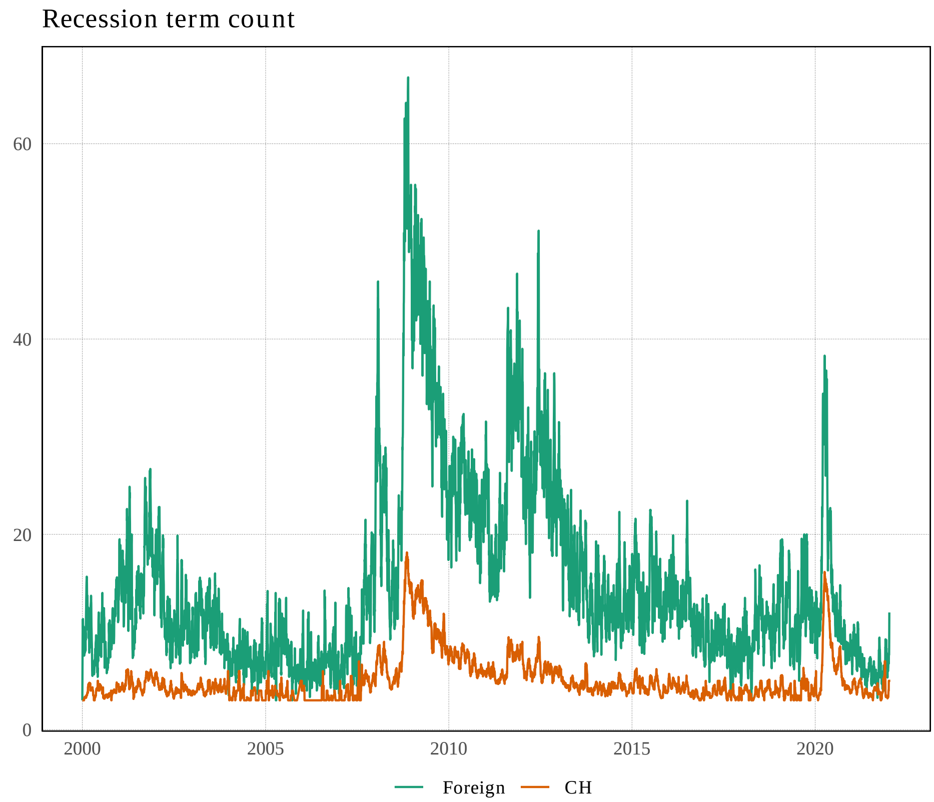


Figure 4 — Daily news based sentiment indicators

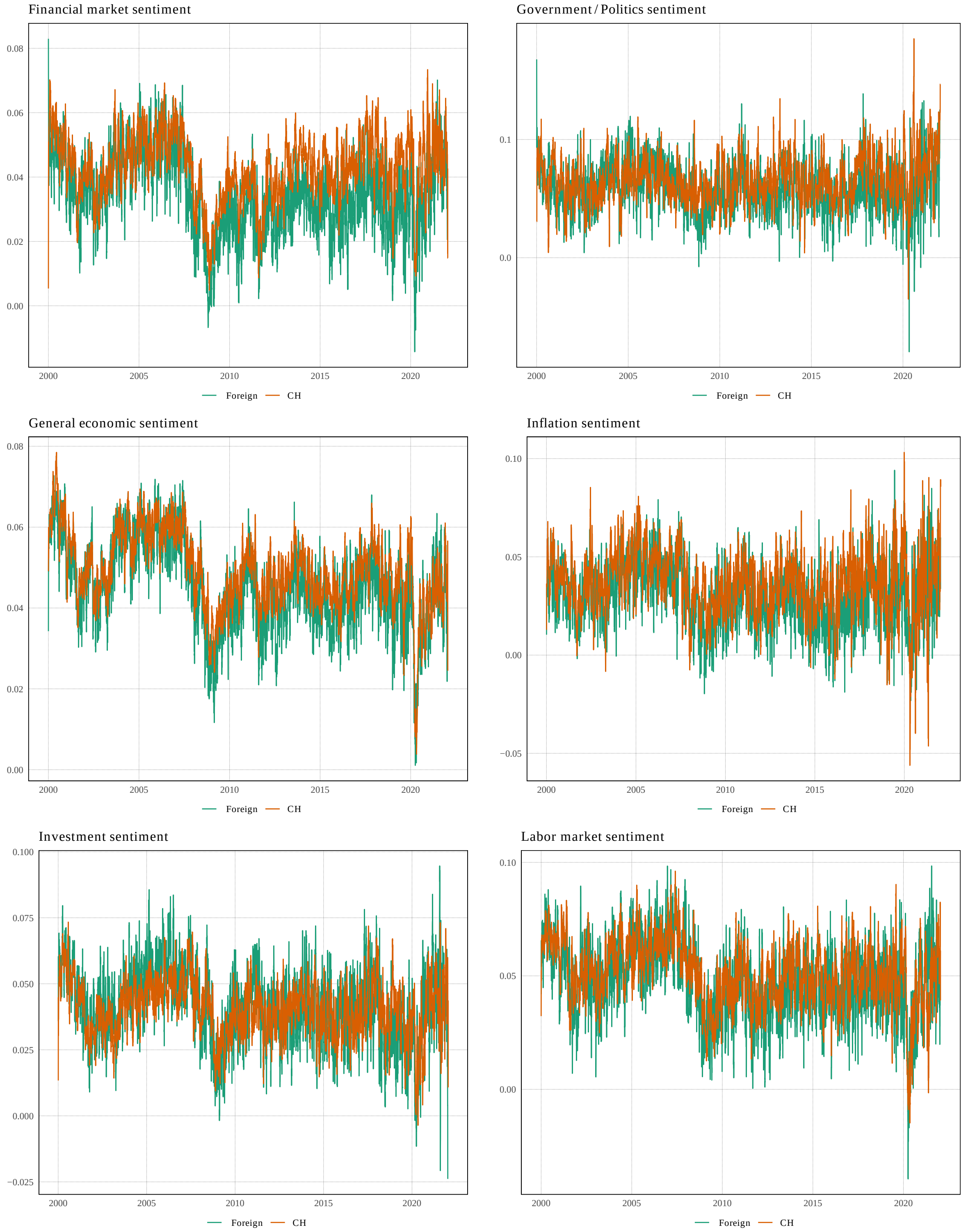
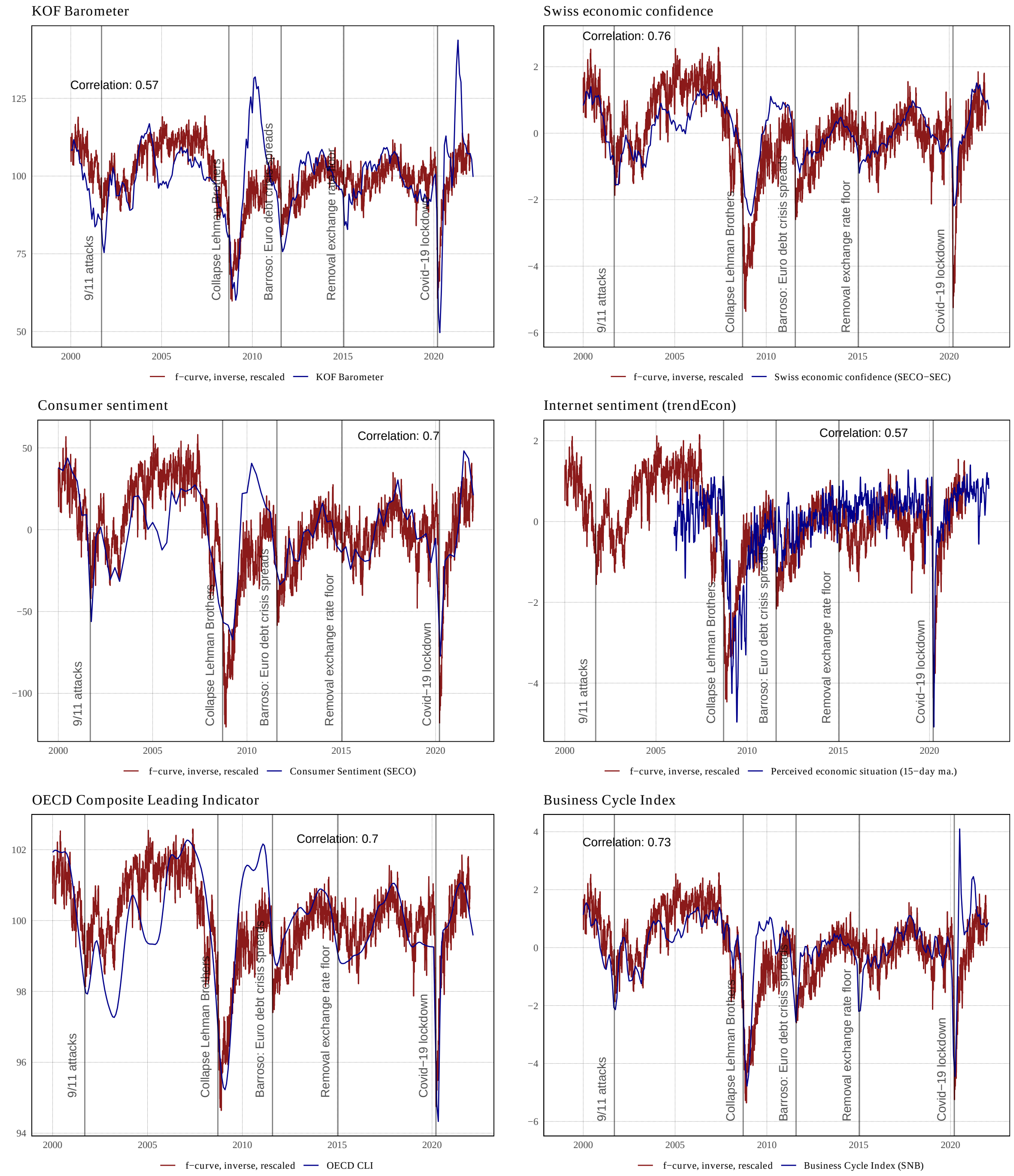
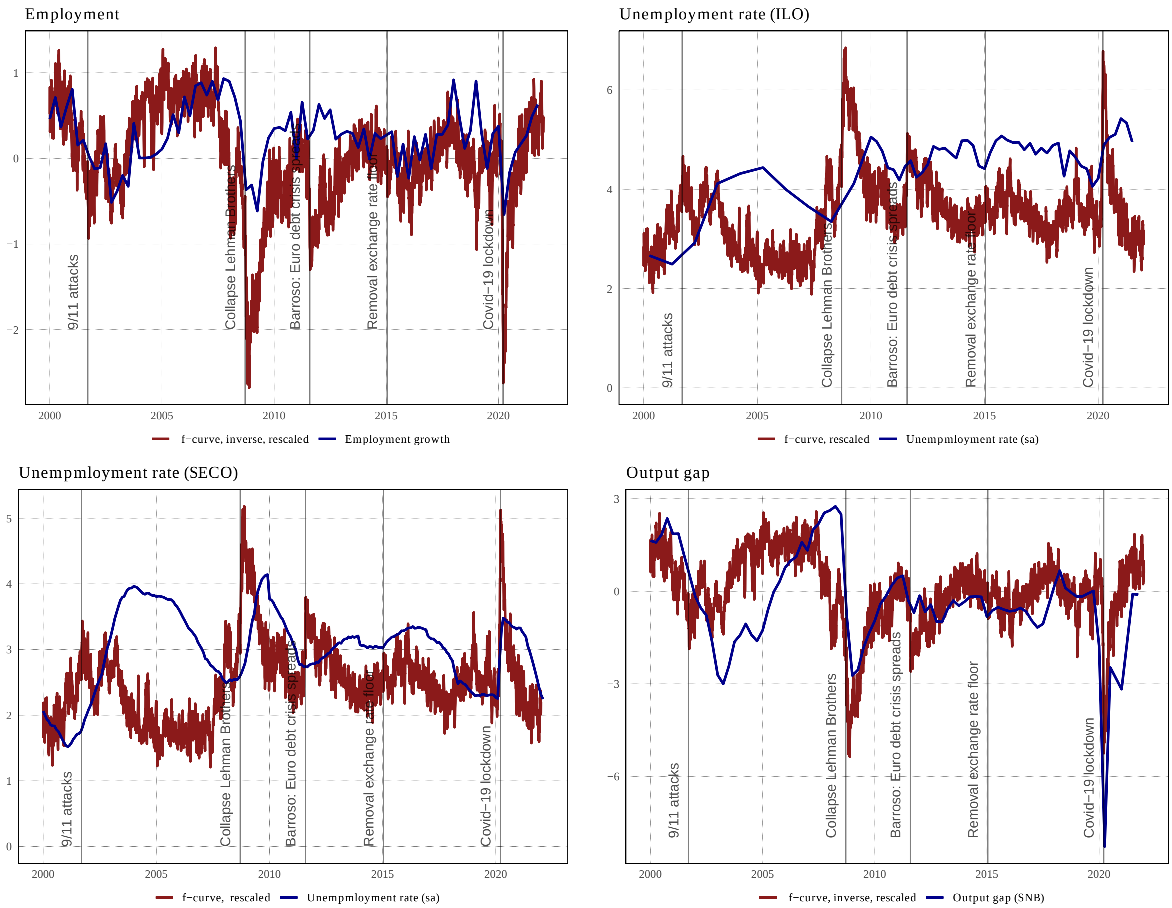


Figure 5 — Comparison with other indicators



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

Figure 6 — Comparison with other macroeconomic data



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

Table 4 — Relative RMSE and DMW tests using legendre polynomial

	Full sample				**Excluding Covid crisis**				**Excluding all crisis periods**			
Horizon	0	40	80	120	0	40	80	120	0	40	80	120
a) Hypothesis: Model > AR(1) model												
Bridge	0.64	0.84**	1.02	1.1	0.82*	0.78*	1.03	1.03	0.83***	0.76*	1	0.93
Midas	0.68	0.87*	1.06	1.09	0.83*	0.87	1.08	1.05	0.82**	0.83	1.04	0.96
Midas-IT	0.67	0.89*	1.02	1.11	0.82*	0.82*	1.09	1.05	0.82**	0.79*	1.09	0.97
b) Hypothesis: Model > Barometer bridge												
Bridge	1.04	0.93*	0.97	1.1	0.82**	0.86*	1.02	1.04	0.73**	0.81**	0.93	0.95
Midas	1.09	0.96	1	1.08	0.83*	0.96	1.06	1.06	0.73**	0.88	0.97	0.98
Midas-IT	1.09	0.98	0.97	1.1	0.82*	0.9	1.07	1.07	0.73**	0.84*	1.01	0.99
c) Hypothesis: Model < First Release												
Bridge	1.97*				1.13				1.17*			
Midas	2.06*				1.12				1.14*			
Midas-IT	2.05*				1.12				1.14*			

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 (where $A > B$ means A has lower RMSE than B) written in the row header (Diebold & Mariano, 2002; West, 1996). I assume a quadratic loss function. Significance levels are given by: *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$

Table 5 — Relative RMSE and DMW tests for financial f–curve

	Full sample				**Excluding Covid crisis**				**Excluding all crisis periods**			
Horizon	0	40	80	120	0	40	80	120	0	40	80	120
a) Hypothesis: Model > AR(1) model												
Bridge	0.67	0.85**	1.01	1.07	0.77**	0.73*	1.02	0.98	0.83***	0.77*	0.96	0.84**
Midas	0.7	0.87**	0.99	1.06	0.71**	0.76*	1.07	1	0.76***	0.79*	1.01	0.91*
Midas-IT	0.72	0.87**	0.98	1.05	0.74**	0.74**	1	0.95	0.78**	0.77*	0.95	0.83*
b) Hypothesis: Model > Barometer bridge												
Bridge	1.09	0.95	0.96	1.07	0.77**	0.81*	1.01	1	0.73**	0.82**	0.89	0.86**
Midas	1.13	0.96	0.94	1.06	0.71***	0.84*	1.06	1.01	0.67***	0.84*	0.94	0.93*
Midas-IT	1.16	0.96	0.93	1.04	0.74**	0.81*	0.98	0.97	0.69**	0.82**	0.88	0.85**
c) Hypothesis: Model < First Release												
Bridge	2.07*				1.06				1.18**			
Midas	2.14*				0.96				1.06			
Midas-IT	2.19*				1				1.08			

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 (where $A > B$ means A has lower RMSE than B) written in the row header (Diebold & Mariano, 2002; West, 1996). I assume a quadratic loss function. Significance levels are given by: *** $p < 0.01$, ** $p < 0.05$, *$p < 0.1$

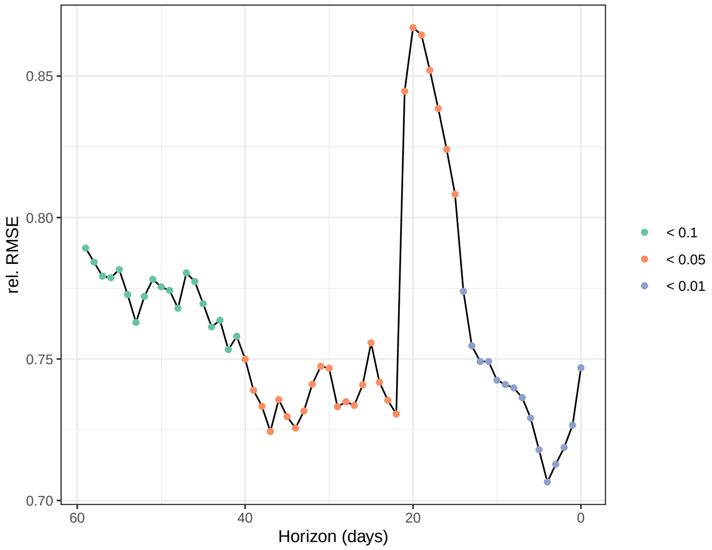
Table 6 — Relative RMSE and DMW tests for news-based f–curve

	Full sample				**Excluding Covid crisis**				**Excluding all crisis periods**			
Horizon	0	40	80	120	0	40	80	120	0	40	80	120
a) Hypothesis: Model > 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
b) Hypothesis: Model > 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
c) Hypothesis: Model < First Release												
Bridge	2.03*				1.2				1.22**			
Midas	2.13*				1.09				1.08			
Midas-IT	2.22*				1.08				1.05			

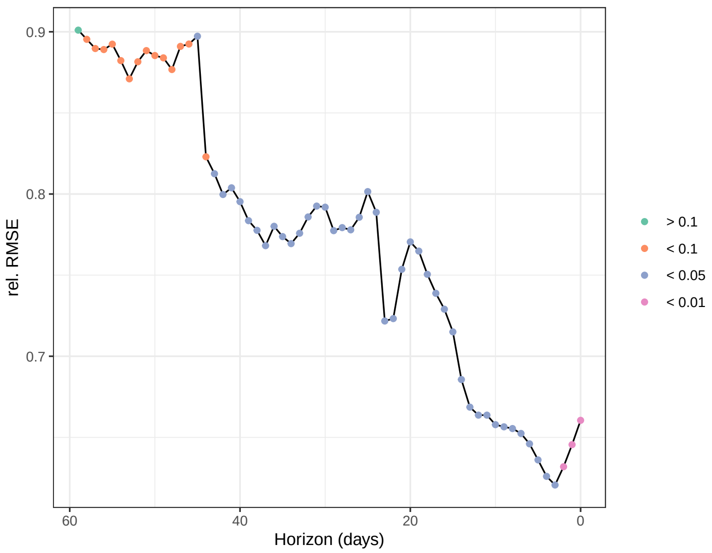
Notes: Relative Root-mean-squared errors (RMSE) for forecasts with selected forecast horizons. A lower RMSE implies higher predictive accuracy compared to the benchmark. Horizons from 0 to 59 (60 - 119) denote forecasts for the current (next) quarter. The full sample spans from January 1st 2002 to December 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 (where $A > B$ means A has lower RMSE than B) 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 7 — Relative RMSE of daily current quarter nowcasts on sample excluding crisis periods

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



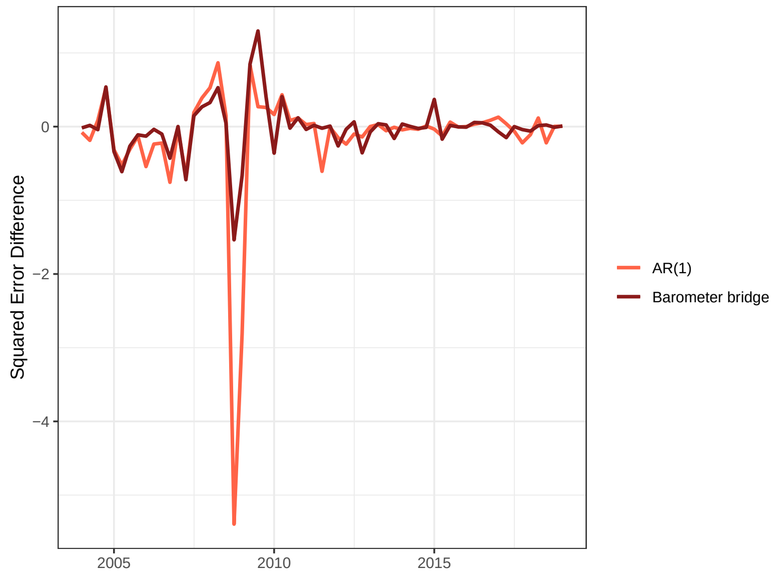
(b) MIDAS-IT vs. KOF Barometer bridge



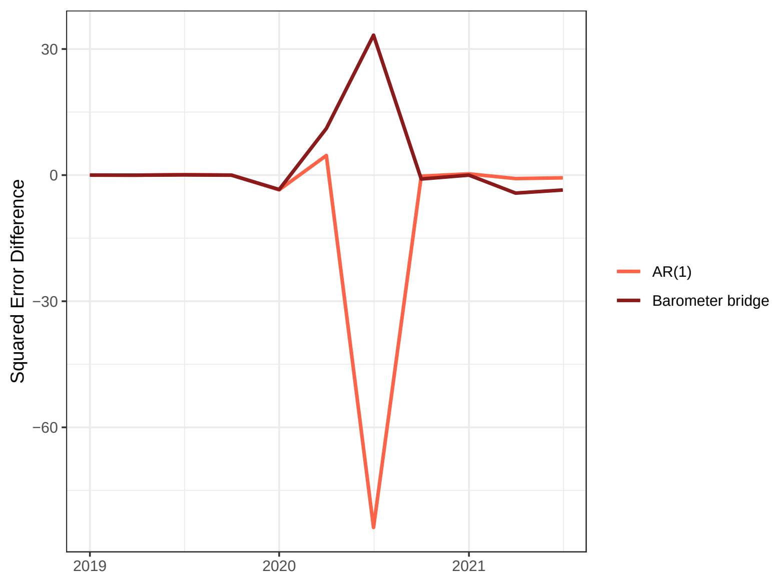
Notes: Relative Root-mean-squared errors (RMSE) for forecasts with forecast horizons from 119 to 0 days. Periods of the Covid–19 and the great financial crisis are excluded. A lower RMSE implies higher predictive accuracy compared to the benchmark. Horizons from 0 to 59 (60 - 119) denote forecasts for the current (next) quarter. We use two benchmarks. First, we use an AR(1) model (panel a). Second, we 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). We assume a quadratic loss function. Significance levels are given by: ● $p > 0.1$, ● $p < 0.1$, ● $p < 0.05$, ● $p < 0.01$

Figure 8 — Squared error differences: Midas.IT - Benchmark

(a) 2004 - 2019

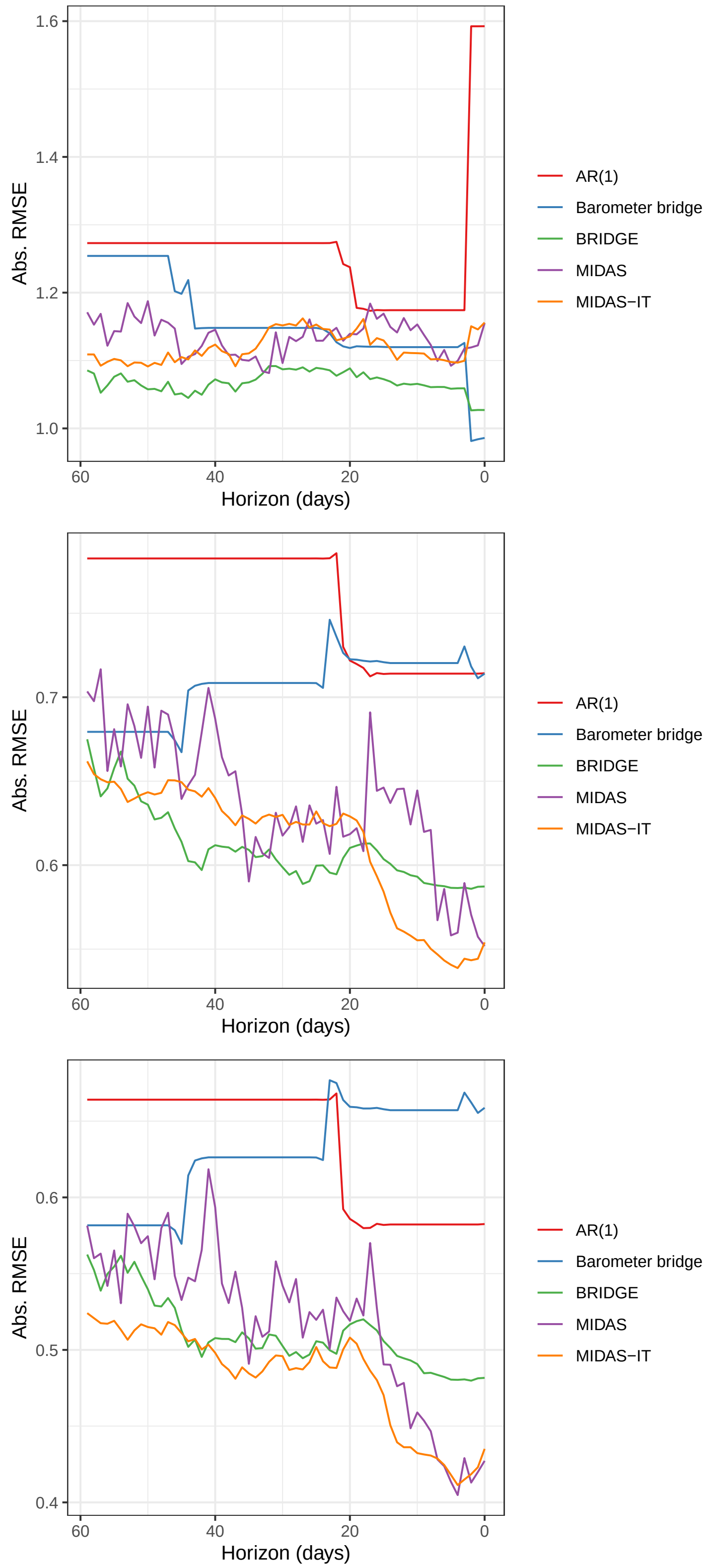


(b) 2019 - 2022



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

Figure 9 — Absolute RMSE of daily current quarter nowcasts



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