Economic forecasting with German newspaper articles

Tino Berger[a], Simon Wintter[b,c]

[a] *University of Göttingen, tino.berger@wiwi.uni-goettingen.de*
[b] *University of Göttingen, simon.wintter@wiwi.uni-goettingen.de*
[c] *Corresponding author*

Abstract

We introduce a new leading indicator for the German business cycle based on the content of newspaper articles from the *Süddeutsche Zeitung*. We use the rapidly evolving technique of Natural Language Processing (NLP) to transform the content of daily newspaper articles between 1992 and 2021 into topic time series using an LDA model. These topic time series reflect broad sectors of the German economy since 1992, in particular the precisely timed signals of the Great Financial Crisis and the Corona outbreak in Germany. We use the Newspaper Indicator in a Probit model to demonstrate that our data can be considered as a new leading indicator for predicting recession periods in Germany. Moreover, we show in an out-of-sample forecast experiment that our Newspaper Indicator has a predictive power for the German business cycle across all tested 29 target variables that is as strong as established survey indicators. Industrial Production, the Stock Market Index DAX and the Consumer Price Index for Germany can even be predicted out-of-sample more accurately with our indicator than with survey indices of the Ifo Institute and the OECD.

Keywords: Business cycle, Recession forecasting, Macroeconomic forecasting, Textual data, Natural language processing

Acknowledgements: The authors would like to thank Sebastian Hienzsch, Tore Dubbert and Christian Ochsner for insightful discussions and constructive feedback during the preparation of this paper.

Statements: The raw newspaper metadata that support the findings of this study are available from Süddeutsche Zeitung. Restrictions apply to the availability of these data, which were used under license for this study. The data of the constructed newspaper time series are available on request from the corresponding author.

1. Introduction

Developments in natural language processing (NLP) make it possible to process unstructured information from 'digital text' (Gentzkow et al., 2019)—such as newspapers—as a substitute for survey indicators for economic forecasts. Every day, thousands of people form their expectations about future economic developments based on information they read in newspapers. These expectations are incorporated monthly into survey indicators, which are an important source of data for forecasts (Lehmann & Reif, 2021). Research suggests that newspapers and other media sources should be increasingly considered for economic forecasting due to their information content (Barbaglia et al., 2022).

Newspapers are evaluated in the empirical literature as an important source of data for macroeconomic time series modelling (Alsem et al., 2008; Blood & Phillips, 1995; Goidel & Langley, 1995). The information conveyed in newspaper articles is primarily referred to as news sentiment or simply textual information. News sentiment is the tone in which certain topics are presented in newspaper articles. In the past, the assessment of these sentiments was done manually by researchers. Since newspaper articles and other sources of digital texts are very large data sets, reproducibility for further work is difficult, if not impossible, without the immense workload of a team of researchers.

The contribution of this paper is to add to the growing literature on business cycle forecasting using newspaper data. The general predictive power of newspaper data to forecast business cycles and macroeconomic variables is given (Barbaglia et al., 2022; Iselin & Siliverstovs, 2016; L. A. Thorsrud, 2020). To our knowledge, we are the first to evaluate the entire economic section of the *Süddeutsche Zeitung* from 1992 to 2021 and use it to forecast reference series of the German business cycle as well as German recession periods. Furthermore, we compare the forecasting performance of our compiled newspaper data with surveys and indices from the Organisation for Economic Co-operation and Development (OECD) and the German Ifo Institute.

As a starting point of our methodology, we process 346,756 articles from the economic section of the *Süddeutsche Zeitung* into numerical time series. The *Süddeutsche Zeitung* is the second largest newspaper in Germany in terms of print circulation. The newspaper articles are cleaned and sorted into topics using an Latent Dirichlet Allocation (LDA) model. The LDA model calculates topic frequencies, which are translated into time series based on the publication dates of the articles.

Our method continues in exploiting the fact that one of our topic time series contains information on the beginning of global shocks (2007, 2009, 2020). In the following, we refer to this time series as our Newspaper Indicator and use it to predict recession phases in Germany with probit models. We also determine the predictive performance of our Newspaper Indicator in a pseudo out-of-sample forecasting experiment. For this purpose, we generate forecasts of business cycle reference series. To evaluate this, we extend these forecasts by a term representing our Newspaper Indicator or one of the survey variables of the OECD or the Ifo Institute. We compare these very similar models - which differ only in one term - by their root mean square error (RMSE) over the out-of-sample period.

Our results for predicting recession probabilities suggest that our Newspaper Indicator performs better than the tested indices of the Ifo Institute and the OECD. Our indicator ranks third among the 14 variables tested. Based on our score measure, our indicator performs almost as well as the Stock Market Index DAX and Total Share Prices. We therefore conclude that our indicator can be seen as a robust new leading indicator for German recession phases. This conclusion is supported by our out-of-sample forecasting experiment. We find that our indicator performs strongly in forecasting of key German economic indicators. The Stock Market Index DAX, Industrial Production and the Consumer Price Index are predicted more accurately with our indicator than with the benchmark indicators of the Ifo Institute and the OECD.

Section 2 continues with a literature review. This is followed by our methodology in section 3. It first describes how the textual data of the newspaper articles are transformed into numerical, topic time series. It further explains the setup of the forecast experiment and the Probit model for recession prediction. Section 4 summarises the data framework consisting of newspaper and macro variables for our study. The results are described in section 5. The paper ends with a conclusion in section 6 and the Appendix.

2. Literature Review

The forecasting literature on the German business cycle can be divided into the prediction of turning points (Bandholz & Funke, 2003; Carstensen et al., 2020; Dreger & Schumacher, 2005; Fritsche & Kouzine, 2002; Fritsche & Kuzin, 2005; Ivanova et al., 2000; Sensier et al., 2004), the prediction of binary recession periods (Döpke, 1999; Döpke et al., 2017; Fornari & Lemke, 2010; Nyberg, 2010) as well as the forecasting of reference series (Askitas & Zimmermann, 2013; Hamburg et al., 2008; Kholodilin & Siliverstovs, 2006). While the methodology for turning points aims to predict the probability of regime changes (e.g. phases of upswing to phases of downswing), the methodology for recession periods predicts the probability of whether a given period falls into a recession phase. In both methods, the aim is to find leading indicators that predict recessions with a lead.

The literature suggests that the Ifo indices (Bandholz & Funke, 2003; Fritsche & Kouzine, 2002) and Order Inflows (Fritsche & Kouzine, 2002) are suitable leading indicators for turning points of the German economy. The interest rate spread (Döpke et al., 2017; Fritsche & Kuzin, 2005; Nyberg, 2010, 2014) and the slope of the US and German yield curves (Estrella et al., 2003; Nyberg, 2010; Pažickỳ, 2021) are considered important leading indicators for predicting recession phases in Germany. Pažickỳ (2021) notes, however, that the importance of the US yield curve for predicting recessions decreases after 1996, which, according to the author, indicates a desynchronisation of the US and German business cycles. Sensier et al. (2004) analyses the synchronisation of business cycles in four European countries between 1970 and 2001. The authors forecast business cycle expansions and contractions for Germany using domestic data and data from the other European countries as well as the USA. Their results suggest that there is an international synchronisation of the business cycles studied. For Germany, a composite US leading indicator and German short-term interest rates in particular have an important influence on the prediction of the business cycle.

We contribute to the growing literature that uses newspaper data as a source for economic forecasts and especially for forecasting the German economy. As far as we know, we are the first to analyse all newspaper articles published in the economic section of the *Süddeutsche Zeitung* between January 1992 and December 2021. We present a forecasting indicator that predicts the German Stock Market Index DAX, Industrial Production and the Consumer Price Index of Germany more accurately than the leading indicators mentioned in the previous paragraph. We thus provide a new source for predicting the German business cycle.

In understanding the general predictive power of newspaper data, we are in line with the literature on German and international newspaper data. Barbaglia et al. (2022) forecasts GDP in five European countries and finds that newspaper-based indicators are significant predictors. Ashwin et al. (2021) provides evidence that data from fifteen European newspapers improve forecasts of GDP in the euro area. Iselin and Siliverstovs (2016) finds that newspaper-based indices improve forecasts for the business cycle in Germany and Switzerland, while Shrub et al. (2022) finds that newspaper data from the *Handelsblatt* provide additional information for the nowcasting of German GDP. However, the authors point out that information collected on the basis of unsupervised text analysis tools still contains a considerable amount of noise and thus there is a need for research. A considerable body of work on international newspaper data (Larsen & Thorsrud, 2019; Larsen et al., 2021; L. Thorsrud, 2016; L. A. Thorsrud, 2020) uses the LDA topic modeling approach of Blei et al. (2003) to exploit newspaper data for macroeconomic forecasting. L. A. Thorsrud (2020) uses an LDA topic model to organize latent information from Norwegian newspapers from 1998 to 2016 into topic time series and use it for economic forecasting. The authors compare a principal component representation of their topics and a news sentiment index to output growth for Norway and find good results capturing both the 2001 recession and the 2008 financial crisis.

3. Methodology

3.1. Newspaper articles as time series

For econometric purposes, newspaper articles are inherently noisy and are in need of careful preprocessing. Common preprocessing steps for raw newspaper articles include removing stop words, stemming and tf-idf analysis (term frequency inverse document frequency). In tf-idf analysis, words that occur too frequently (more than 95%) or too rarely (less than 5%) are removed. In both cases, it is assumed that the words do not contribute to the information content of the articles.

Figure 1: Examples of word clouds. The word clouds show the most important words (in German as word stems) for each topic. The larger the font of a word, the more important the word is for its topic.

As a result of the preprocessing, a Bag-of-Words (BoW) corpus and a dictionary are created as input to the LDA model. The BoW corpus is a list of 2-tuples (word ID, word count) with the information about the number of each word ID in each article. An unsupervised LDA model introduced by Blei et al. (2003) is used with the package *gensim* (Řehřek, 2010) for the programming languages *Python* to sort the newspaper articles into topics. The LDA model calculates a fixed number of topics describing the newspaper articles. We select 60 topics based on the subjectively best result from several calculation runs.

The LDA model calculates which words are associated with the fixed number of topics. It also calculates how strongly each of these words is associated with a topic. These topic word weights represent the probability with which a word occurs in a topic. Topic names are chosen by the authors based on the most frequent words in each topic.

For a visual, economic interpretation of the topics, topic probabilities of each month are presented as time series (see Figure 2). The time series show how strongly a topic was represented in the newspaper articles compared to all other topics in each month. Figure 1 shows the most frequent words (in German as word stems) for the individual topics as word clouds. The naming of

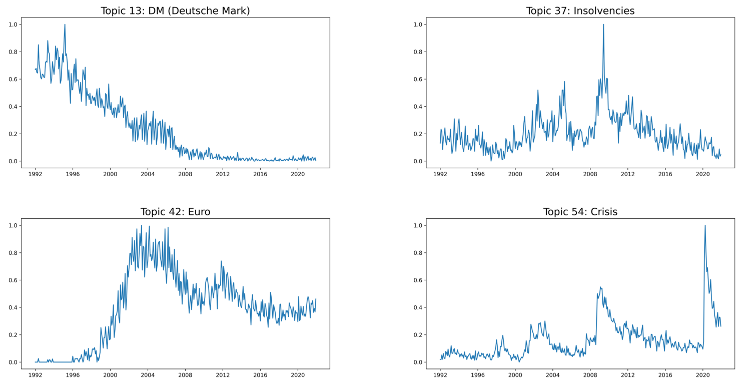

Figure 2: Topic frequencies. The time series show how strongly a topic is represented in the *Süddeutsche Zeitung* articles of a month compared to all other 59 topics.

the individual topics is based on the most frequent words. Shown are topic 13 about the German currency before the euro (DM stands for *Deutsche Mark*), topic 37 about insolvencies and restructurings, and topic 42 about the euro. For the topic 54, the most frequent word is 'crisis'; in addition, 'problem', 'consequence', 'bad' and also 'corona' are frequently mentioned. In Figure 2, the topics 13 and 42 reflect how the importance of the national currency, the DM, has been declining and the importance of the euro has been increasing since the decision to introduce the euro in the mid-1990s, peaking at the cash launch in 2001. Topic 37 shows the impact of the Great Financial Crisis after 2009 on corporate insolvencies in Germany.

In Figure 3 the topic 54 'crisis' closely mirrors the global shocks of 2007 (start of the Great Financial Crisis) and of 2008 (collapse of Lehman Brothers), as well as the Corona outbreak in Germany in 2020. Therefore, we test this topic time series to predict German recession periods and reference series of the German business cycle. In the following, we compare the crisis time series with established economic indicators and call it the Newspaper Indicator (see Table 1).

3.2. Forecasting recessions with newspaper data

The forecasting ability of our newspaper data is tested in a Probit model to forecast recession periods of the German economy. We perform in-sample estimations based on single-indicator models to determine whether our newspaper

Figure 3: Newspaper Indicator. We refer to the frequency of topic 54 as our Newspaper Indicator (blue line), which accurately reflects the timing of the beginnings of the Great Financial Crisis, the bankruptcy of Lehman Brothers and the Corona outbreak in Germany (red-dotted lines). The shaded areas denote the GCEE recession dates.

variable is able to capture information about recession periods in Germany. In addition, we forecast recession periods in a pseudo out-of-sample framework to see if our newspaper data serve as a leading indicator for the German economy.

For the definition of monthly recession periods we follow the German Council of Economic Experts (GCEE) with its *Business Cycle of the German Economy*. In its business cycle dating the GCEE follows the National Bureau of Economic Research (NBER), whose 'traditional definition of a recession is that it is a significant decline in economic activity that is spread across the economy and that lasts more than a few months.' The GCEE defines months as recessions that are in the phase between an cyclical peak and trough (Breuer et al., 2018). In a first step, the cyclical peaks and troughs are predetermined on the basis of German GDP. In a second step, a variety of monthly and quarterly macroeconomic indicators are used to finally determine the German business cycle.

Our single-indicator model (1) estimates the binary recession variable r_t for each variable x_t^i of Table 1 with a lead of $k = 1$. The recession variable r_t is zero for months of no recession and one if the economy was in a recession in a corresponding month t. This model specification is used to select the candidate variables to be tested further as leading indicator:

$$r_t = \beta_0 + \beta_1 x_{t-k}^i + \epsilon_t \tag{1}$$

To assess the leading indicator property of the selected variables, we conduct pseudo out-of-sample forecasts for one month ahead again with model (1). We forecast with an expanding estimation period a total of 276 months in 1-month forecast steps. We compute the model with a lead of $k = 1$, $k = 3$ and $k = 6$.

Moreover, the term x_t^i is replaced by a set of variables I_{t-k} to test the combined predictive performance of different variables. This multi-indicator model is defined as:

$$r_t = \beta_0 + \beta_2 I_{t-k} + \epsilon_t \tag{2}$$

The performance of both model specifications is measured by calculating the percentage at which the months with and without recession are correctly forecasted.[1] [2] In addition, the out-of-sample performance of the models is measured by the squared probability score (QPS). The value ranges from zero to two, with a lower score indicating a more accurate forecast. P denotes at time t the probability of a recession R for time $t + k$. Following (Döpke, 1999), we understand the QPS as a robust analogue of the root mean square error (RMSE):

$$QPS = \frac{1}{T} \sum_{t=1}^{T} 2(P_{t+k,t} - R_{t+k})^2$$

3.3. Forecasting experiment

In addition to predicting recession periods, we use our Newspaper Indicator to predict business cycle relevant variables. In this way, we further determine the predictive power of our Newspaper Indicator for the German business cycle. To do so, we conduct a pseudo out-of-sample forecasting experiment. We first follow Carriero et al. (2019) and assume that economic time series y_t can be adequately h-step predicted by autoregressive processes with one lag, AR(1), if their error term follows a white noise process (WN):

$$y_{t+h} = \alpha + \beta y_{t-1} + \epsilon_t \tag{3}$$

$$\epsilon_t \sim WN(0, \sigma^2)$$

Model 3 is used to predict all variables y_{t+h} from Table 1 (except the *Survey and News* variables) with h=1. We calculate the RMSE for the pseudo out-of-sample forecast of 288 months between the years 1998 and 2021. Thus we

[1]Percentage of recession months correctly predicted = sum of predicted recession periods in months with recession / sum of binary recession months

[2]Percentage of **non** recession months correctly predicted = 1 - (sum of predicted recession periods in months with **no** recession / sum of binary **non** recession months)

investigate two theses. First, we assume that all AR(1) forecasts with Model 3 can be improved by including our Newspaper Indicator. Accordingly, we test the forecasting performance of Model 3 against a model with an additional term, x_t, for our indicator:

$$y_{t+h} = \alpha + \beta y_{t-1} + \gamma x_t + \epsilon_t \tag{4}$$

Second, it is tested whether the forecasts with our Newspaper Indicator—as just described—perform as well or better than with benchmark variables from the Ifo Institute and the OECD as additional term x_t. Using Model 4, we can test this thesis by directly comparing the forecasting performances.

For both model specifications 3 and 4, the economic target variable y_{t+h} is forecasted h-steps ahead to $t + h$. At the time of the forecast, t, the current value of y_{t+h} is assumed to be unknown whereas the value of x_t is known. In other words, y_{t+1} is forecasted with a lagged value of y and the current value of x_t.

To measure the forecast performances, we follow Kalamara et al. (2022) and calculate a forecast score. We calculate this score as the RMSEs of the models that include the news and survey data related to the RMSE of the AR(1) model:

$$score = \frac{RMSE_{news}}{RMSE_{AR(1)}}$$

The forecast score is to be interpreted in two ways. First, a value less than one means that the added variable improves the prediction relative to the prediction without the added variable. The reverse is true for a value greater than one. Second, models with survey and news variables can be directly compared. Models with lower scores indicate a more accurate prediction.

In other words, we find evidence of predictive power in our newspaper data when the score for the model with Newspaper Indicator is less than zero or less than the other scores in the same target peer group.

4. Data

4.1. Newspaper data

We process XML files containing all articles and associated metadata published in the *Süddeutsche Zeitung* between January 1992 and December 2021.

In total, the XML files contain 2,058,795 unique articles of which we use 346,756 from the economic section for our study.

In Germany, the *Süddeutsche Zeitung* is the second-largest newspaper with around 280,000 copies printed in the second quarter of 2023. Most copies come from *BILD Zeitung* with more than 1.1 million copies, while the *Frankfurter Allgemeine Zeitung* is the third-largest with around 186,000 copies.

4.2. Economic data

As shown in the previous chapter, our newspaper data vividly captures information about economic activity and shocks. In particular, the global shocks of 2009 (Lehman Brothers collapse) and 2020 (Corona outbreak) are featured by one of the LDA topics. Our aim is therefore to model the information contained in our newspaper data with a set of economic data to statistically prove the information content of our newspaper data.

The compilation of the data framework in which we analyse our newspaper data follows (Berger & Ochsner, 2022), which uses a selection of German and international indicators grouped into economic blocks to measure the German output gap. In addition, we analyse our newspaper data together with a series of potential leading indicators for forecasting recession periods in Germany (Döpke, 1999). The transformations listed in Table 1 are performed to ensure stationarity of each time series.

5. Results

5.1. Probit results

In- and out-of-sample forecasts of recession probabilities are performed for Germany for the period from January 1993 to December 2021.[3] The in-sample estimations are performed over the whole period of observations for each variable from the List 1. The results suggest that 14 variables contain recession-relevant information (see Appendix A). These 14 variables, together with the survey data, are the ones whose recession month are correctly predicted by at least 30% in the individual estimates.

[3]The definition of monthly recessions is adopted from the Business Cycle of the German Economy (GCEE). See Section 3.2.

Table 1: Variable blocks, abbreviations, data transformations and data sources.

Variable	Abbreviation	Transformation	Source
External Relations			
Economic Policy Uncertainty Index for Europe	unce	12-months Rolling Demean	FRED
Interest Rate Spread USA	spus	Natural Logarithm	FRED
Terms of Trade Imports	glob	First Difference	Bundesbank
Finance			
Interbank Rate	inte	First Difference	FRED
Interest Rate Spread Germany	spge	Natural Logarithm	FRED
Stock Market Index DAX	daxm	12-months Rolling Demean	Yahoo Finance
Total acquisition of domestic and foreign investments	inve	12-months Rolling Demean	Bundesbank
Total Share Prices for All Shares for Germany	tosh	12-months Rolling Demean	FRED
Foreign Trade			
Capital Account Balance	kaps	First Difference Natural Logarithm	Bundesbank
Capital Account Balance Direct Investments	kapd	First Difference Natural Logarithm	Bundesbank
Capital account Balance: Portfolio Investment	kapw	First Difference Natural Logarithm	Bundesbank
Current account balance ('Leistungsbilanzsaldo')	leis	12-months Difference	Bundesbank
Current Account Balance Services	leid	12-months Difference	Bundesbank
Current Account Balance Trade in Goods	leiw	12-months Difference	Bundesbank
Current Account Exports	expo	12-months Rolling Demean	FRED
Current Account Imports	impo	12-months Rolling Demean	FRED
Labour			
Job Vacancies	jobs	First Difference Natural Logarithm	FRED
Unemployment	unem	First Difference Natural Logarithm	FRED
Unemployment Females	unfe	First Difference Natural Logarithm	FRED
Unemployment Males	unma	First Difference Natural Logarithm	FRED
Working Population	empl	12-months Rolling Demean	Bundesbank
Prices			
Consumer Price Index (CPI)	cpi	12-months Difference	Bundesbank
Consumer Price Index Energy	cpie	First Difference Natural Logarithm	FRED
Core Inflation (no energy, no food)	core	First Difference Natural Logarithm	FRED
Producer Prices of Industrial Products	prpr	First Difference	Bundesbank
Production			
Construction	cons	First Difference Natural Logarithm	FRED
Consumer Goods Durable	dura	12-months Rolling Demean	FRED
Consumer Goods Non Durable	ndur	12-months Rolling Demean	FRED
Energy Production	ener	12-months Rolling Demean	FRED
Industrial Production	indu	12-months Rolling Demean	Bundesbank
Sales			
Order Inflow Construction	orco	First Difference Natural Logarithm	Bundesbank
Order Inflow Industry	orin	12-months Difference	Bundesbank
Turnover Industry	rein	12-months Difference	Bundesbank
Survey and News			
Ifo Business Climate Index	ifoc	12-months Rolling Demean	Ifo Institute
Ifo Business Expectation Index	ifoe	12-months Rolling Demean	Ifo Institute
Ifo Business Situation Index	ifos	12-months Rolling Demean	Ifo Institute
Newspaper Indicator	news	12-months Rolling Demean	Süddeutsche
OECD Consumer Price Survey	oec1	12-months Rolling Demean	FRED
OECD Employment Survey	oec2	12-months Rolling Demean	FRED

Note: The variables refer to the German economy, except for the block *External Relations*.

Table 2: Results of out-of-sample recession forecasts with single-indicator model 1.

	k=1			k=3			k=6		
	QPS	rec.	no rec.	QPS	rec.	no rec.	QPS	rec.	no rec.
Stock Market Index DAX	0.16	73%	89%	0.17	73%	88%	0.23	54%	86%
Total Share Prices for All Shares for Germany	0.16	69%	89%	0.17	70%	88%	0.23	56%	86%
Newspaper Indicator	0.23	41%	85%	0.25	38%	85%	0.27	32%	85%
Ifo Business Climate Index	0.24	42%	85%	0.25	37%	85%	0.27	29%	85%
Ifo Business Situation Index	0.25	40%	85%	0.26	36%	86%	0.28	27%	85%
Ifo Business Expectation Index	0.25	32%	83%	0.26	31%	83%	0.28	25%	84%
Consumer Goods Durable	0.25	30%	85%	0.28	24%	84%	0.30	17%	84%
OECD Employment Survey	0.25	29%	87%	0.27	26%	88%	0.29	20%	87%
Industrial Production	0.26	34%	85%	0.29	26%	84%	0.31	16%	84%
Working Population	0.27	32%	87%	0.28	26%	85%	0.30	18%	84%
Interest Rate Spread Germany	0.28	31%	81%	0.30	20%	82%	*	*	*
OECD Consumer Price Survey	0.29	21%	81%	0.30	20%	82%	0.30	18%	83%

Note: 'GPS' stands for the squared probability score defined in section 3.2, 'rec.' for the percentage of correctly predicted recession months and 'no rec.' for the percentage of correctly predicted months without recession. * denotes forecasts where perfect separation is detected, so no results are available. k indicates the leads of 1, 3, and 6 months.

These 14 variables are used for single-indicator out-of-sample forecasts with model (1). For our first estimate, we use 71 months to forecast January 1999. Our last forecast is for December 2021. The out-of-sample performance is measured by the QPS score and the percentage of correctly estimated recession and non-recession months. For all three leads k, the results suggest that the Stock Market Index DAX performs best as a leading indicator for recession periods defined by the GCEE.

The forecasts for the individual indicators are listed in Table 2 in order of best to worst performance based on the QPS score. With a QPS value of 0.23, our Newspaper Indicator is only just behind the DAX and Total Share Prices in forecasting one month ahead (h=1). Compared to our indicator, the QPS score is worse for all three Ifo indices and both OECD surveys tested. This can be interpreted to mean that our indicator predicts somewhat more accurately.

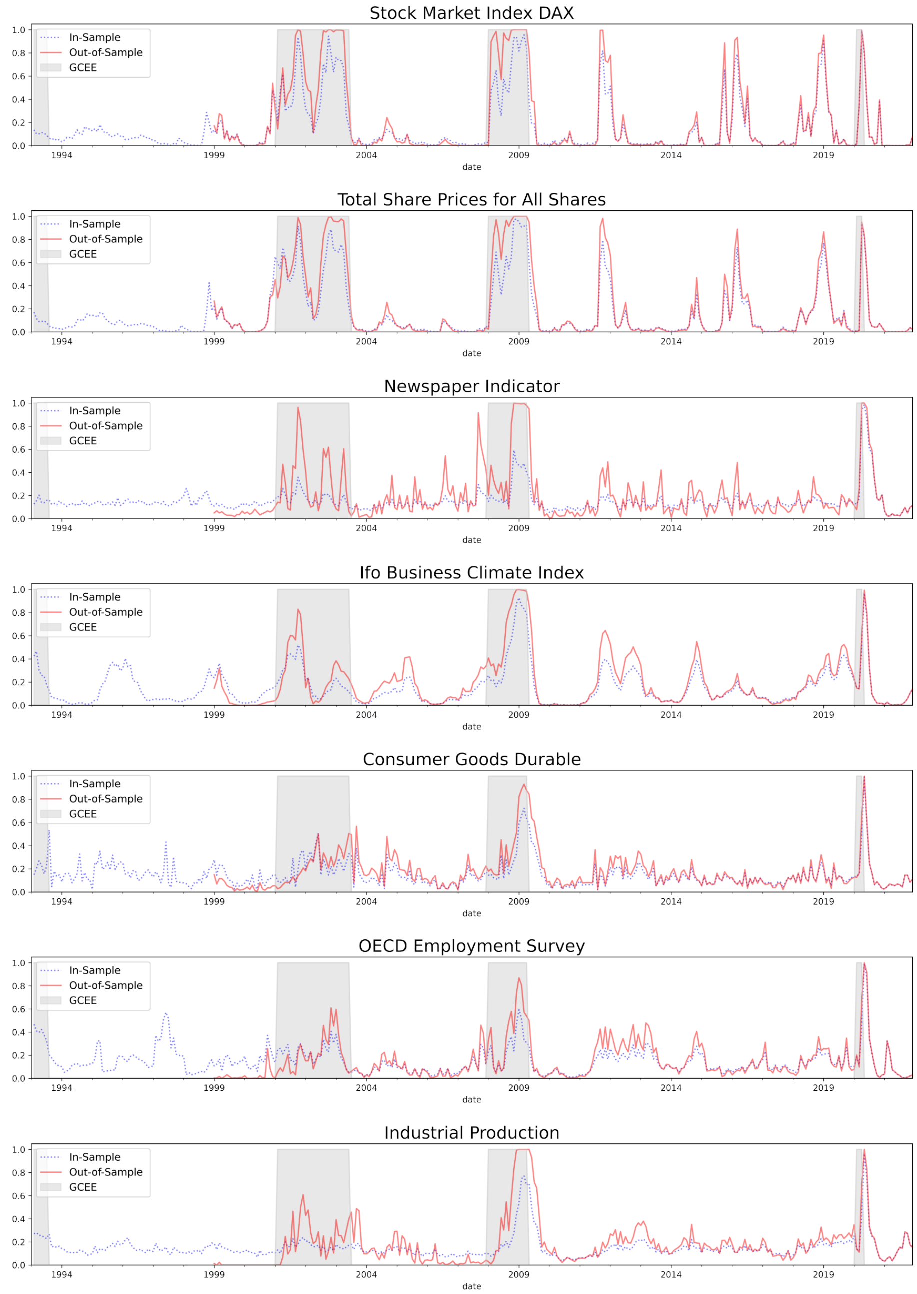

Figure 4: Plots of the Probit in-sample and out-of-sample single indicator recession forecasts. See Table 2 for performance results. The shaded areas denote GCEE recession dates.

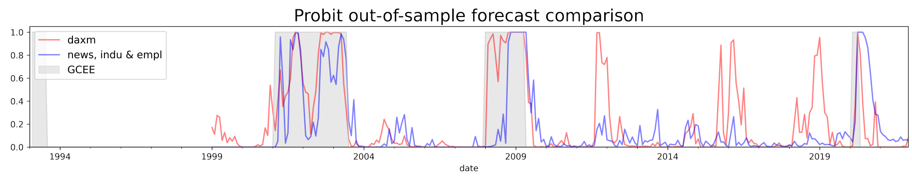

Figure 5: Recession forecasts of the single-indicator and multi-indicator models. The red line denotes the single-indicator forecast using DAX stock market data, which has high volatility between 2009 and 2020. The blue line denotes the multi-indicator forecast using Newspaper Indicator, Industrial Production and Working Population data, which reduces volatility between 2009 and 2020 while predicting the recession phases well. The shaded areas denote GCEE recession dates.

This finding can be confirmed when looking at the additional values of correctly predicted months with and without recession. At 41 percent, more than one in three recession months can be predicted as such by our Newspaper Indicator. Among the Ifo indices, only the climate index, at 42 percent, predicts recessions somewhat better, while the other two Ifo indices do worse. For the OECD surveys, the percentage of correctly predicted recessions is 29 and 21 percent. The best prediction performance for recession months is again achieved with the Stock Market Index DAX with 73 percent correctly predicted months. Months without recession are correctly predicted by our indicator and two Ifo indices at 85 percent, while our comparative indicators from the OECD are 87 and 81 percent. These findings can be confirmed by the results of the 3- and 6-month ahead forecasts (h=3 and h=6).

Using the Stock Market Index DAX, 89 per cent of all non-recession months can be correctly predicted; however, a visual inspection of this result in Figure 4 shows that both the DAX and Total Share Prices exhibit significant volatility in non-recession months. Despite the low QPS value and the high percentages of correctly predicted months in each case, visual inspection of the forecasts for the DAX and Total Share Prices therefore raises questions. The forecasts of both of the variables show three swings between 2008 and 2020, which indicate recession phases but are not defined by the GCEE. Our Newspaper Indicator, Consumer Goods Durable, Working Population and Industrial Production do not show these swings, while the Ifo Business Climate Index slightly hints at the swings.

Consequently, Consumer Goods Durable, Working Population and Industrial Production are tested in the multi-indicator model 2 alongside our Newspaper Indicator. Doing so, we aim to combine the forecast accuracy and reduce the forecast volatility achieved with the stock market data (DAX and Total Share Prices). Table B.4 in the Appendix shows the regression results and forecast performances of the four variables used in the multi-indicator model differently. The out-of-sample forecasting result for the combination with Newspaper Indicator, Industrial Production and Working Population (see Table B.4, Reg. 5) shows an equally good performance as the single-indicator result with the stock market data (see Table 2). However, as it can be seen in Figure 5 using the variable combination Reg. 5 reduces the forecast volatility in phases of no recession.

To conclude this section, we make two points. First, our Newspaper Indicator can predict recession phases in Germany better than the benchmark indicators of the Ifo Institute and the OECD. We demonstrate this using the QPS-score and the visual interpretation using Figure 4. Second, the Stock Market Index DAX and Total Share Prices predict recession phases in Germany better than our indicator based on our measures. However, both stock market variables indicate recession phases that are not defined as such. A combination of variables containing our Newspaper Indicator lead to equally good out-of-sample performance with low volatility.

5.2. Forecast experiment results

Our forecasting experiment aims to compare the forecasting performance of models with and without our newspaper indicator on business cycle reference series. The forecast period is 288 months, which we predict in 1-month steps. The forecast period is chosen to cover the three recession phases defined by the GCEE between 2000 and 2021. We evaluate the forecasting performance of the models based on their forecast scores (see Figure 6 and 8).

To achieve the results of Figure 6 and 8, we test if the model 4 with our newspaper variable outperforms the same model with a established survey variable (OECD surveys and Ifo indices) as well as the AR(1) benchmark model 3 without an additional variable. We thus follow (Kalamara et al., 2022), who

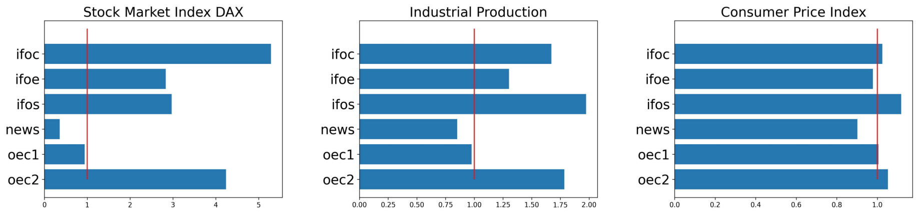

Figure 6: Forecast scores (blue bars) for those business cycle reference series that are more accurately predicted by the newspaper data (news) compared to the survey data of the OECD and the Ifo Institute. Note: Blue bars smaller than a value of 1 indicate that the forecast performance of the indicator variable (y-axis) is better than the benchmark of the AR(1) model (red, vertical line).

use a AR(1) model as a benchmark to test the forecast performance of their text-based variables. According to (Carriero et al., 2019), AR(1) models provide robust predictive performance for a wide range of economic variables and are therefore seen suitable as a benchmark for our experiment.

With the Stock Market Index DAX, Industrial Production and the Consumer Price Index, our Newspaper Indicator succeeds in predicting macroeconomic variables important for the business cycle better than the comparison variables from the OECD and the Ifo Institute can. Measured by our forecast score based on the RMSE, Figure 6 shows that our indicator predicts the three reference series both more accurately than the comparison variables and more accurately than the pure AR(1) process. Figure 7 shows the importance of the three target variables in determining at least the last two recession phases (2008, 2020), as well as the accuracy with which the true path of the reference series can be forecasted using our newspaper indicator. As discussed in Section 5.1, forecasts using stock market data exhibit high volatility between the 2008 and 2020 recession periods.

Table 3 presents the differences in forecast accuracy (as measured by the forecast scores in Figure 6) using the Diebold-Mariano test. The test is whether the forecast scores of the Newspaper Indicator differ significantly from those of the OECD Consumer Price Survey and the Ifo Business Expectation Index. For the prediction of the Stock Market Index DAX and Industrial Production, the results of the Diebold-Mariano test indicate that the predictions with the Newspaper Indicator and the OECD Survey do not differ significantly, but the

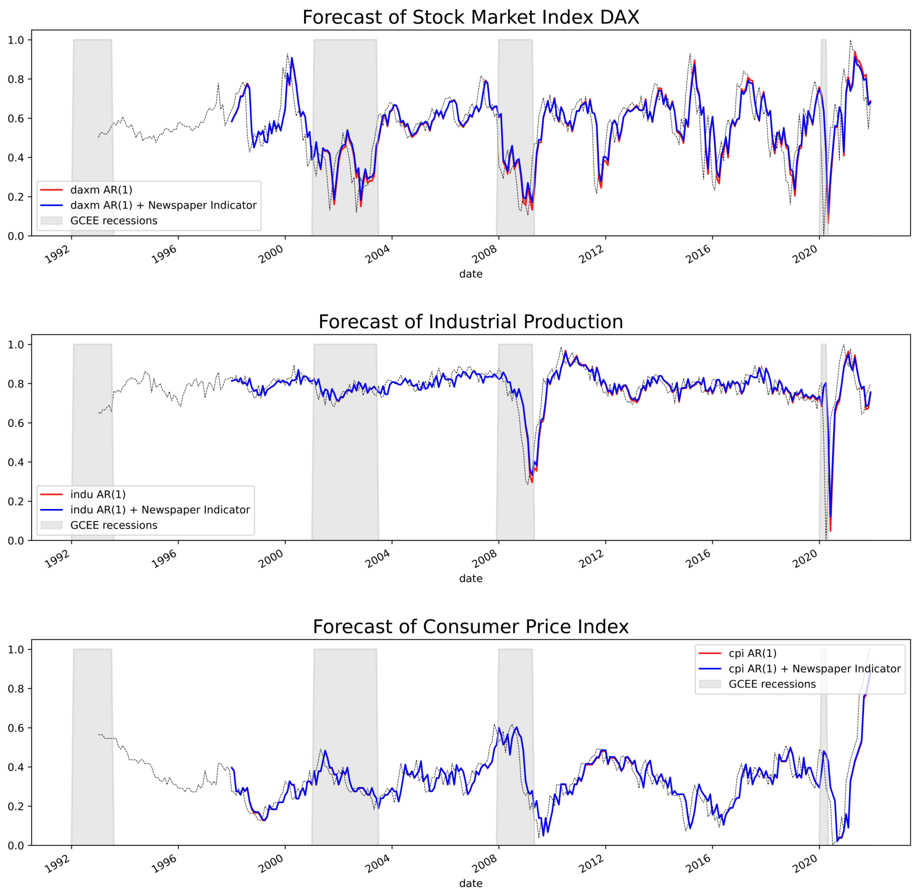

Figure 7: The blue and red lines denote the 1-month pseudo out-of-sample forecasts of the Stock Market Index DAX (daxm), Industrial Production (indu) and the Consumer Price Index (cpi) with and without the Newspaper Indicator. The shaded areas denote the GCEE recession dates. Note: The forecasts with Newspaper Indicator (blue lines) correspond more closely to the actual data than the pure AR(1) forecast (red lines). See forecast scores smaller than 1 in Figure 6.

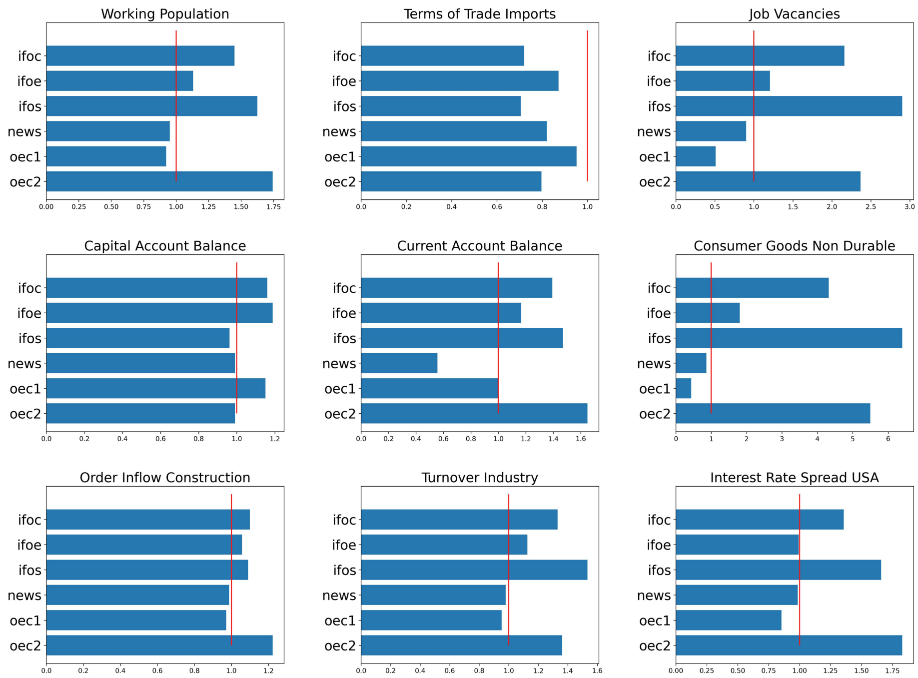

Figure 8: Forecast scores (blue bars) for those business cycle reference series that are more accurately predicted by the Newspaper Indicator (news) compared to the benchmark forecast of the AR(1) model (red, vertical line). Note: Blue bars smaller than a value of 1 indicate that the forecast performance of the indicator variable (y-axis) is better than the benchmark of the AR(1) model.

Table 3: Diebold-Mariano test results

	Newspaper Indicator against	
	OECD Consumer Price Survey	Ifo Business Expectation Index
Stock Market Index DAX	0.368 (0.999)	-7.773 (0.000)
Industrial Production	1.689 (0.954)	-1.736 (0.041)
Consumer Price Index	-17.259 (0.000)	-1.365 (0.087)

Note: Shown are the results of the Diebold-Mariano test (Mariano & Diebold, 1995) with the modification proposed by Harvey et al. (1997). Compared are the forecasts of the Stock Market Index DAX, Industrial Production and Consumer Price Index predicted with the model 4 including the Newspaper Indicator against the forecasts of the model 4 including the OECD Consumer Price Survey and the Ifo Business Expectation Index, respectively. Reported are the test statistics based on the mean squared errors and the p-values in brackets for the null hypothesis that there is no difference in the accuracy of the forecasts with and without the Newspaper Indicator.

predictions with the Newspaper Indicator and the Ifo Business Climate Index do (*p-values* < 0.05). The results are reversed for the prediction of the Consumer Price Index. The results indicate that the predictions with Newspaper Indicator and those with Ifo Index do not differ significantly, while the predictions with Newspaper Indicator and OECD Survey differ (*p-value* $= 0.00$).

Figure 8 shows the forecast scores for the business cycle reference series that are more accurately predicted when the Newspaper Indicator is included in the forecast model compared to the pure AR(1) process (*forecast score* < 1). We see these results as further evidence that our newspaper data is a suitable predictive indicator along established survey indicators for a wide range of economic variables.

As a conclusion of this section, we find that our Newspaper Indicator helps to predict important reference series of the business cycle more accurately than with its AR(1) processes. Furthermore, we show that the prediction of the Stock Market Index DAX, Industrial Production and the Consumer Price Index using our Newspaper Indicator is significantly more accurate than the prediction of established indicators of the Ifo Institute and the OECD.

6. Conclusion

To our knowledge, we are the first to evaluate the entire business section of the *Süddeutsche Zeitung* from 1992 to 2021 for forecasting reference series of the German business cycle as well as German recession periods. Our results are in line with the literature that newspaper articles are a suitable variable for business cycle and macroeconomic forecasting in general. Moreover, our Newspaper Indicator predicts recession periods in Germany and important reference series of the German business cycle more accurately than the benchmark indicators of the Ifo Institute and the OECD examined.

References

Alsem, K. J., Brakman, S., Hoogduin, L., & Kuper, G. (2008). The impact of newspapers on consumer confidence: Does spin bias exist? *Applied Economics*, *40*(5), 531–539.

Ashwin, J., Kalamara, E., & Saiz, L. (2021). *Nowcasting euro area gdp with news sentiment: A tale of two crises* (tech. rep.). ECB Working Paper.

Askitas, N., & Zimmermann, K. F. (2013). Nowcasting business cycles using toll data. *Journal of Forecasting*, *32*(4), 299–306.

Bandholz, H., & Funke, M. (2003). In search of leading indicators of economic activity in germany. *Journal of Forecasting*, *22*(4), 277–297.

Barbaglia, L., Consoli, S., & Manzan, S. (2022). Forecasting gdp in europe with textual data. *Available at SSRN 3898680*.

Berger, T., & Ochsner, C. (2022). *Robust real-time estimates of the german output gap based on a multivariate trend-cycle decomposition* (tech. rep.). Bundesbank Discussion Paper.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*(Jan), 993–1022.

Blood, D. J., & Phillips, P. C. (1995). Recession headline news, consumer sentiment, the state of the economy and presidential popularity: A time series analysis 1989–1993. *International Journal of Public Opinion Research*, *7*(1), 2–22.

Breuer, S., Elstner, S., Kirsch, F., & Wieland, V. (2018). Datierung der deutschen konjunkturzyklen–die methode des sachverständigenrates. *Arbeitspapier*, *13*, 2018.

Carriero, A., Galvao, A. B., & Kapetanios, G. (2019). A comprehensive evaluation of macroeconomic forecasting methods. *International Journal of Forecasting*, *35*(4), 1226–1239.

Carstensen, K., Heinrich, M., Reif, M., & Wolters, M. H. (2020). Predicting ordinary and severe recessions with a three-state markov-switching dynamic factor model: An application to the german business cycle. *International Journal of Forecasting*, *36*(3), 829–850.

Döpke, J. (1999). *Predicting germany's recessions with leading indicators: Evidence from probit models* (tech. rep.). Kiel Working Paper.

Döpke, J., Fritsche, U., & Pierdzioch, C. (2017). Predicting recessions with boosted regression trees. *International Journal of Forecasting*, *33*(4), 745–759.

Dreger, C., & Schumacher, C. (2005). Out-of-sample performance of leading indicators for the german business cycle: Single vs. combined forecasts. *Journal of Business Cycle Measurement and Analysis*, *2005*(1), 71–87.

Estrella, A., Rodrigues, A. P., & Schich, S. (2003). How stable is the predictive power of the yield curve? evidence from germany and the united states. *Review of Economics and Statistics*, *85*(3), 629–644.

Fornari, F., & Lemke, W. (2010). *Predicting recession probabilities with financial variables over multiple horizons* (tech. rep.). ECB Working Paper.

Fritsche, U., & Kouzine, V. (2002). *Do leading indicators help to predict business cycle turning points in germany?* (Tech. rep.). DIW Discussion Papers.

Fritsche, U., & Kuzin, V. (2005). Prediction of business cycle turning points in germany/prognose konjunktureller wendepunkte in deutschland. *Jahrbücher für Nationalökonomie und Statistik*, *225*(1), 22–43.

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as data. *Journal of Economic Literature*, *57*(3), 535–574.

Goidel, R. K., & Langley, R. E. (1995). Media coverage of the economy and aggregate economic evaluations: Uncovering evidence of indirect media effects. *Political Research Quarterly*, *48*(2), 313–328.

Hamburg, B., Hoffmann, M., & Keller, J. (2008). Consumption, wealth and business cycles in germany. *Empirical Economics*, *34*, 451–476.

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of forecasting*, *13*(2), 281–291.

Iselin, D., & Siliverstovs, B. (2016). Using newspapers for tracking the business cycle: A comparative study for germany and switzerland. *Applied Economics*, *48*(12), 1103–1118.

Ivanova, D., Lahiri, K., & Seitz, F. (2000). Interest rate spreads as predictors of german inflation and business cycles. *International Journal of Forecasting*, *16*(1), 39–58.

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.

Kholodilin, K. A., & Siliverstovs, B. (2006). On the forecasting properties of the alternative leading indicators for the german gdp: Recent evidence. *Jahrbücher für Nationalökonomie und Statistik*, *226*(3), 234–259.

Larsen, V. H., & Thorsrud, L. A. (2019). The value of news for economic developments. *Journal of Econometrics*, *210*(1), 203–218.

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

Lehmann, R., & Reif, M. (2021). Predicting the german economy: Headline survey indices under test. *Journal of Business Cycle Research*, *17*(2), 215–232.

Mariano, R. S., & Diebold, F. X. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*, 253.

Nyberg, H. (2010). Dynamic probit models and financial variables in recession forecasting. *Journal of Forecasting*, *29*(1-2), 215–230.

Nyberg, H. (2014). A bivariate autoregressive probit model: Business cycle linkages and transmission of recession probabilities. *Macroeconomic Dynamics*, *18*(4), 838–862.

Pažickỳ, M. (2021). Predicting recessions in germany using the german and the us yield curve. *Journal of Business Cycle Research*, *17*(3), 263–291.

Řehřek, R. (2010). Software framework for topic modelling with large corpora. *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, 45.

Sensier, M., Artis, M., Osborn, D. R., & Birchenhall, C. (2004). Domestic and international influences on business cycle regimes in europe. *International Journal of Forecasting*, *20*(2), 343–357.

Shrub, Y., Rieger, J., Müller, H., & Jentsch, C. (2022). *Text data rule-don't they? a study on the (additional) information of handelsblatt data for nowcasting german gdp in comparison to established economic indicators*. Ruhr Economic Papers.

Thorsrud, L. (2016). *Nowcasting using news topics big data versus big bank* (tech. rep.). Norges Bank Working Paper.

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.

Appendix A. Probit in-sample estimations

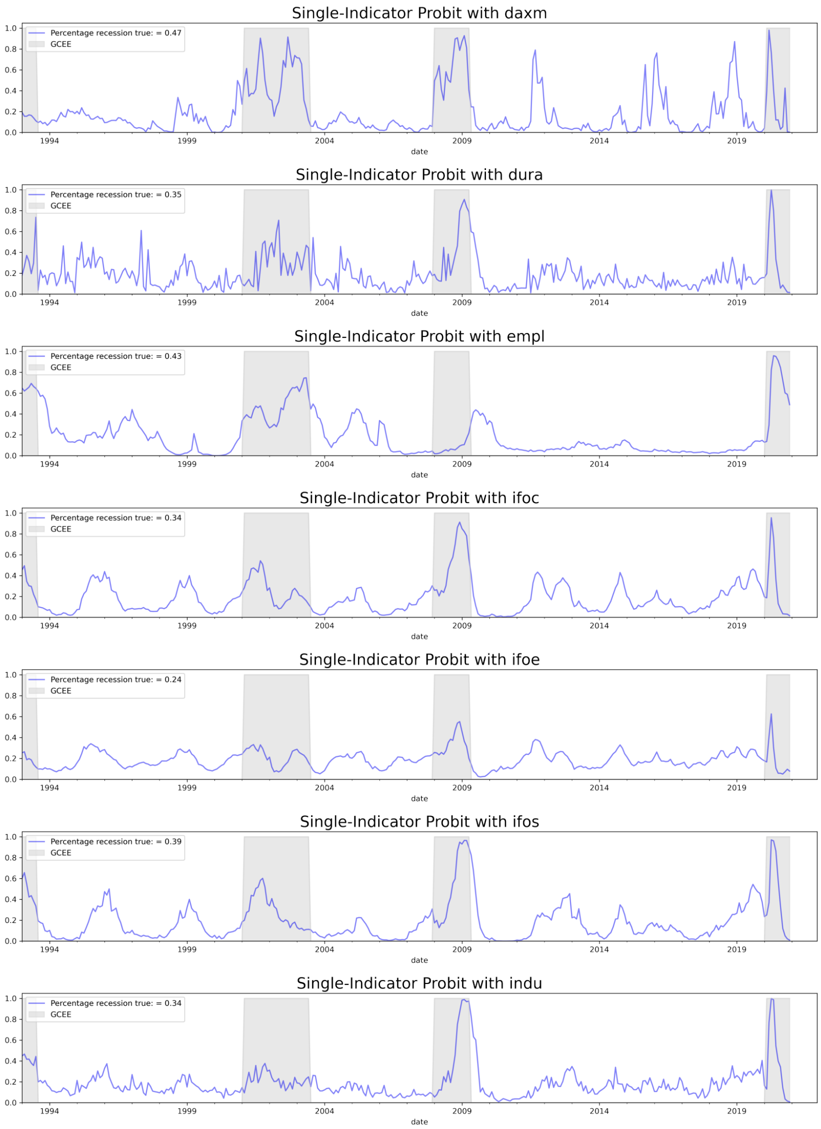

Figure A.9: Single-indicator model estimates (Model 1, Section 3.2) with at least 30% correct estimates of recession periods. The shaded areas denote GCEE recession dates.

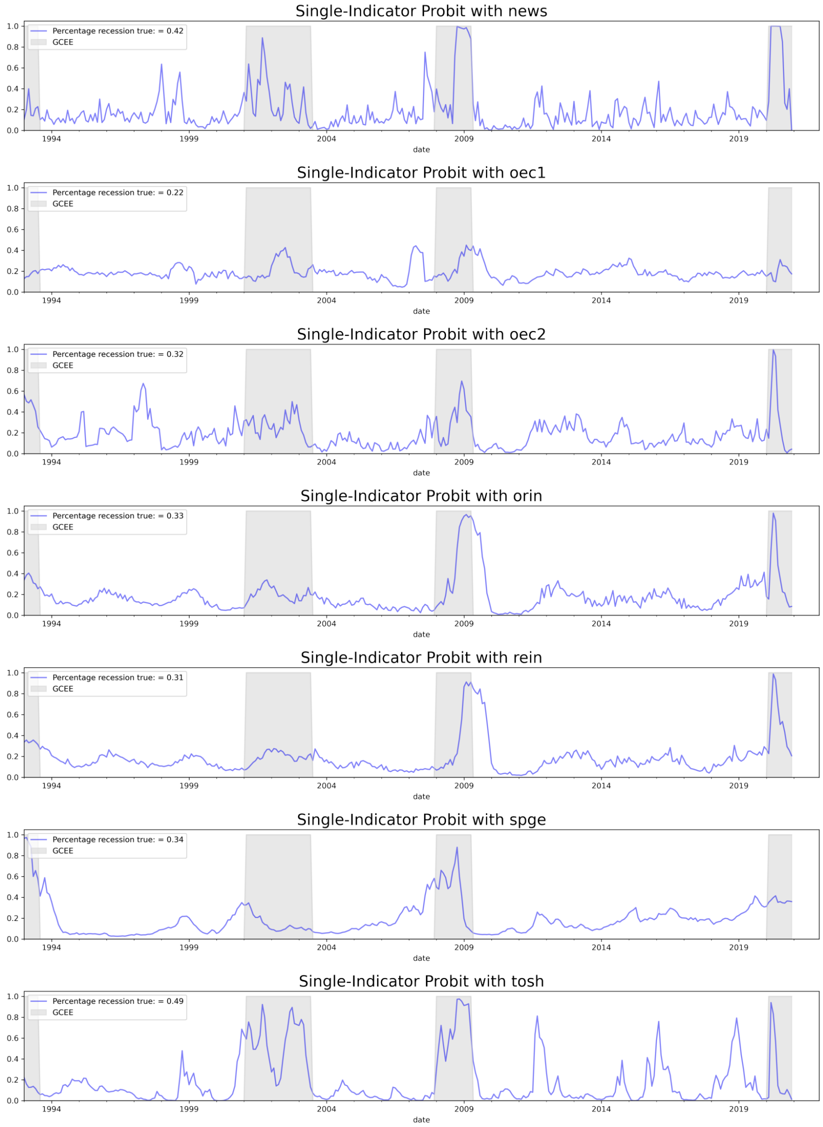

Figure A.10: Single-indicator model estimates (Model 1, Section 3.2) with at least 30% correct estimates of recession periods. The shaded areas denote GCEE recession dates.

Appendix B. Probit regression & forecast results

Table B.4: Probit Multi-Indicator Model

	In-sample regressions					
	Reg. 1	Reg. 2	Reg. 3	Reg. 4	Reg. 5	Reg. 6
Constant	1.13	-0.22	-0.47	0.48	0.59	2.86
	(0.21)	(0.84)	(0.25)	(0.67)	(0.59)	(0.00)
Newspaper Indicator	1.83	2.66	3.03	2.50	2.10	
	(0.08)	(0.03)	(0.00)	(0.05)	(0.08)	
Consumer Goods Durable	-3.38			-4.29		-3.45
	(0.00)			(0.00)		(0.00)
Industrial Production		-1.88		1.56	-1.15	
		(0.08)		(0.31)	(0.30)	
Working Population			-2.35		-2.28	-2.26
			(0.00)		(0.00)	(0.00)
Pseudo R2	0.13	0.10	0.15	0.14	0.16	0.18
	Out-of-sample forecast performances (t=1)					
QPS-Score	0.21	0.23	0.20	0.22	0.19	0.24
% Recession correct	0.45	0.42	0.53	0.42	0.52	0.39
% No recession correct	0.88	0.88	0.92	0.90	0.92	0.89

Note: For in-sample results, p-values in brackets. See Model 2, Section 3.2.