

1193

Discussion  
Papers

# Do Media Data Help to Predict German Industrial Production?

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

#### IMPRESSUM

© DIW Berlin, 2014

DIW Berlin  
German Institute for Economic Research  
Mohrenstr. 58  
10117 Berlin

Tel. +49 (30) 897 89-0  
Fax +49 (30) 897 89-200  
<http://www.diw.de>

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:  
<http://www.diw.de/discussionpapers>

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:  
<http://ideas.repec.org/s/diw/diwwpp.html>  
<http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html>

# Do media data help to predict German industrial production?

Konstantin A. Kholodilin\*      Tobias Thomas<sup>†</sup>      Dirk Ulbricht<sup>‡</sup>

June 25, 2014

## Abstract

Expectations form the basis of economic decisions of market participants in an uncertain world. Sentiment indicators reflect those expectations and thus have a proven track record for predicting economic variables. However, respondents of surveys perceive the world to a large extent with the help of media. So far, mainly very crude media information, such as word-count indices, has been used in the prediction of macroeconomic and financial variables. In this paper, we employ a rich data set provided by Media Tenor International, based on the sentiment analysis of opinion-leading media in Germany from 2001 to 2014, whose results are transformed into several monthly indices. German industrial production is predicted in a real-time out-of-sample forecasting experiment using more than 17,000 models formed of all possible combinations with a maximum of 3 out of 48 macroeconomic, survey, and media indicators. It is demonstrated that media data are indispensable when it comes to the prediction of German industrial production both for individual models and as a part of combined forecasts. They increase reliability by improving accuracy and reducing instability of the forecasts, particularly during the recent global financial crisis.

**Keywords:** forecast combination, media data, German industrial production, reliability index, R-word.

**JEL classification:** C10; C52; C53; E32.

---

\*Research associate, DIW Berlin, Mohrenstraße 58, 10117 Berlin, Germany, e-mail: [kkholodilin@diw.de](mailto:kkholodilin@diw.de).

<sup>†</sup>Head of research, Media Tenor International, Alte Jonastraße 48, 8640, Switzerland, and research affiliate, Düsseldorf Institute for Competition Economics (DICE), Heinrich-Heine-Universität, Universitätsstraße 1, 40225 Düsseldorf, e-mail: [t.thomas@mediatenor.com](mailto:t.thomas@mediatenor.com).

<sup>‡</sup>Research associate, DIW Berlin, Mohrenstraße 58, 10117 Berlin, Germany, e-mail: [dulbricht@diw.de](mailto:dulbricht@diw.de).

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Empirical approach and data</b>	<b>3</b>
2.1	Empirical approach . . . . .	3
2.2	Data . . . . .	5
<b>3</b>	<b>Forecast evaluation</b>	<b>8</b>
3.1	Measures for comparing performance . . . . .	8
3.2	Performance of individual models . . . . .	9
3.3	Performance of forecast combinations . . . . .	12
<b>4</b>	<b>Conclusion</b>	<b>14</b>
	<b>References</b>	<b>15</b>
	<b>Appendix</b>	<b>18</b>

## List of Tables

1	Macroeconomic indicators: definitions and descriptive statistics . . . . .	18
2	Analyzed Media Set . . . . .	19
3	Sentiment indicators: definitions and descriptive statistics . . . . .	19
4	Media indicators: definitions and descriptive statistics . . . . .	20
5	Best models: July 2001 to April 2014 . . . . .	21
6	Best models: May 2008 to January 2009, recession period . . . . .	22
7	Combinations: July 2001 to April 2014, sorted by coefficient of reliability . . . . .	23
8	Combinations: May 2008 to January 2009, recession period, sorted by coefficient of reliability . . . . .	24

## List of Figures

1	Precision and stability over all periods: individual models versus combinations . . . . .	25
2	Precision and stability during recession: individual models versus combinations . . . . .	26

# 1 Introduction

Typically, the data on gross domestic product (GDP) are available on a quarterly basis. In addition, they are published half a quarter after the end of the reference quarter. Therefore, in order to get a quick insight into the current economic situation, a monthly series of industrial production is used. It is thus a central monthly indicator for the business activity. This is especially the case for Germany. Although the share of industrial production has been shrinking over the past decades, it remains at high levels compared to other OECD and especially other EU member countries<sup>1</sup>. Moreover, the secular decline of the share is expected to reverse in the coming years. Among others, the European Commission stresses the importance of the contribution of industrial competitiveness to the overall competitiveness performance of the EU and aspires raising the contribution of industry to GDP to as much as 20% by 2020 [Commission \(2014\)](#). Moreover, it substantially contributes to the business cycle dynamics.

Consequently, there have been many attempts to improve forecast accuracy of this variable <sup>2</sup>. Most of these studies employ hard economic indicators such as interest rates, manufacturing orders, etc. There have also been several studies using soft data such as business surveys like the ifo or ZEW indicator (see, for example, [Abberger and Wohlrabe, 2006](#) or [Hüfner and Schröder, 2002](#)). It was demonstrated that due to their forward-looking nature they are well-suited for forecasting industrial production. The underlying idea of this approach is to employ a measure of the intentions or the expectations of the managers or analysts, respectively. The main advantages of these indicators are: their high frequency, timeliness, and that they are almost not subject to revisions in comparison to many statistical indicators.

While in classical economics the homo oeconomicus is omniscient and decides independently, and decisions lead to efficient outcomes at the market level, [Keynes \(1937\)](#) underlined the role of uncertainty concerning decisions and behavior and the related (suboptimal) outcomes on the macro level just as [von Hayek \(1992\)](#) pointed to the pretense of knowledge. Similarly, [Simon \(1957\)](#) as well as [Kahneman and Tversky \(1979\)](#) have shown that in reality human behavior clearly deviates from the behavior predicted by standard economic models. Due to their limited information processing capacity, individuals use subjective models for the perception of

---

<sup>1</sup>According to the OECD Factbook 2011: Economic, Environmental and Social Statistics, in 2010, the percentage of total value added in industry (including energy) was 24% in Germany, 19% in the EU, and 21% in the OECD countries.

<sup>2</sup>See, for example, [Kholodilin and Siliverstovs, 2006](#).

reality. If these models are shared because of common cultural background and experience, in accordance with [Denzau and North \(1994\)](#) one can speak of shared mental models. In media societies, media reporting does form relevant parts of those shared mental models not only because investors, consumers, politicians, and voters receive lots of information via the media, but because additional information perceived directly is interpreted on the basis of the frame determined by the media reporting. Therefore, what is on the agenda (“agenda setting”) and what is not (“agenda cutting”) becomes highly relevant, as well as the way in which these things are described in the media, such as with a positive, negative or neutral tone. Individuals decide and behave at least in parts on what information they perceive through the media. This is also important in the context of business surveys, as respondents interpret their own economic situation and build their expectations within the frame set by the media.

A growing literature employs media data to explain economic sentiment. For instance, [Goidel and Langley \(1995\)](#) as well as [Doms and Morin \(2004\)](#) show an impact of media reporting on consumer climate. For [Nadeau et al. \(2000\)](#) and [Soroka \(2006\)](#) the assessment of the state of the economy depends at least in parts on the media reporting.

The literature can be split into two main classes. The first one simply counts the number of times a single word or a group of words, which can be associated with a certain event, occur in the media. The second strand of literature captures content expressed in the media:

The most popular word-count indicator has been introduced by the weekly journal *The Economist*. It counts how many articles in the *Washington Post* and the *New York Times* use the word “recession” in a quarter. For Germany this concept has been adapted by the HypoVereinsbank basing the count on articles in the *Frankfurter Allgemeine Zeitung*, *Handelsblatt*, and *WirtschaftsWoche* and was published until 2002. More recently it has been revived by [Grossarth-Maticek and Mayr \(2008\)](#) who, however, do not consider the Great Recession period. [Doms and Morin \(2004\)](#) count the number of articles in 30 American newspapers that contain 9 keywords or expressions in the title or the first section of the article and use this statistic to forecast US private consumption. [Ammann et al. \(2011\)](#) compute the number of mentionings of a lexicon of 236 words in the online archive of the *Handelsblatt* with the aim of predicting yields of the German stock market index DAX. [Bordino et al. \(2011\)](#) use the number of queries of listed companies in the US search engine Yahoo! as a predictor for stock market

volumes. Using the number of queries in the alternative search engine Google, [Kholodilin et al. \(2010\)](#) try to improve forecasts of US private consumption.

Media indicators that are based on the sentiment expressed in media reports use both automated methods and human analyst to evaluate the news. [Bollen et al. \(2011\)](#) employ the software OpinionFinder to analyze Twitter tweeds with the aim of forecasting stock prices. To the same end, [Tetlock \(2007\)](#) evaluates the sentiment of articles of the Wall Street Journal. [Uhl \(2010, 2011\)](#) uses sentiment data of newspaper and TV-news provided by Media Tenor International to forecast US private consumption. [Iselin and Siliverstovs \(2013\)](#) use the R-word index to forecast the growth rates of real GDP in Germany and Switzerland.

The first and until now the only attempt to use media indices to predict German industrial production was undertaken by [Grossarth-Maticek and Mayr \(2008\)](#). They contrast the R-word index for Germany and a Media Tenor International index to predict growth rates of industrial production and recession probabilities. Our approach differs from their approach in several respects. First, we consider a more recent period including the Great Recession. Second, we examine all possible combinations of a much wider set of indicators. Third, we evaluate the usefulness of media indicators in forecast combinations. Fourth, unlike [Grossarth-Maticek and Mayr \(2008\)](#) who use a single aggregate Media Tenor International business conditions index, we employ 18 more indicators that differ both in their time perspective (present, future, and climate) and their underlying topic (fiscal policy, foreign exchange, labour market, etc.). Fifth, we employ monthly instead of quarterly data. Sixth, we develop and apply a novel measure of reliability to assess the forecasts. Seventh, we employ real-time series of the dependent variable.

This paper is structured as follows. The second section presents the empirical approach and the data used in the analysis. In section three the forecasts are evaluated. The fourth section concludes.

## 2 Empirical approach and data

### 2.1 Empirical approach

Many existing studies concentrate on the comparison of single models including one different alternative indicator at a time in a horse race with respect to average forecast accuracy of a benchmark model such as the simple

autoregressive model (AR). However, as [Stock and Watson \(2004\)](#) have demonstrated, single models are prone to structural breaks and tend to be less reliable when compared to combinations of many different forecasts. To address this issue, we suggest a novel approach. In a first step, we estimate the models including all possible combinations of indicators varying from one to a given maximum number of exogenous variables.

In a second step, we construct combined forecasts as weighted averages of the individual models predictions.

The individual models are defined as

$$y_t = \alpha + \sum_{p=l_0}^{P+l_0} \beta_p y_{t-p} + \sum_{i=1}^K \sum_{p=l_i}^{P+l_i} \gamma_{i,p} x_{i,t-p} + u_t, \quad (1)$$

where  $\alpha$ ,  $\beta_p$ , and  $\gamma_{i,p}$  are parameters to estimate,  $y_t$  are year-on-year growth rates of industrial production in time period  $t$  ( $t = 1, \dots, T$ ),  $x_{i,t}$  is an indicator variable  $i$  in time period  $t$ , and  $u_t$  is a disturbance term. The total number of indicator variables is  $N$ . Each individual model can contain a subset of  $K$  indicators. We let  $K$  vary between 0 and 3. The different number of minimum lags  $l_q$  for each regressor, with  $q = 0, \dots, N$ , used reflects the varying degree of data availability. For example, as the dependent variable is published with a lag of 2 months,  $l_0 = 3$ . The number is dictated both by data limitations (the sample is relatively short) and computational intensity. The number of parameters of an individual model ranges from 2 to 50.

The total number of individual models can be computed as  $M = \frac{N!}{K! \times (N-K)!} + 1$ . Given that we employ 48 regressors and up to three regressors in one model, the maximum number of models is  $M = 17,296$ . In fact, the number of individual models in our case is slightly smaller, since we excluded some combinations of regressors due to their extremely high mutual correlation (with a correlation coefficient more than or equal to 0.95). Likewise, a model containing short-term, long-term interest rates, and the spread between them was dropped to avoid multicollinearity. In the end, we are left with 17,135 individual models. With 4 regressors the number of models attains 194,580, whereas with 5 regressors it would reach 1,712,304. Our computational capacities preclude the estimation of so many models.

The lag order,  $P$ , is identical for all regressors and is determined using the Bayesian Information Criterion (BIC) with a maximum of 12.

In the simplest case, when  $N = 0$  the model boils down to an autoregressive process, which we employ as a benchmark model.

The whole sample stretches from January 2001 to April 2014. The data set is unbalanced: some series start in March 2001. On the other hand, the publication delays are different, so the data are characterized by a ragged edge. In order to address this problem, the series are shifted forward correspondingly.

We perform an out-of-sample forecast experiment. The first estimation sub-sample,  $T_E$ , ends in June 2004. The first forecast is performed for July 2004. The estimation and forecasting are implemented in a recursive way. The forecast horizon is  $h = 1$  month. Thus, the number of forecasts for each model is 112.

All the computations in this paper have been carried out using the codes written by the authors in the statistical programming language *R* (see [R Core Team, 2013](#)).

## 2.2 Data

The dependent variable is the monthly series of real-time German industrial production, taken from the Deutsche Bundesbank database (see [Table 1](#)).

The set of regressors includes 15 macroeconomic indicators, 11 purely business survey data and two composite indicators<sup>3</sup>, and 19 media indicators. [Tables 1, 3, and 4](#) list the variables, their sources, and report some descriptive statistics.

In this paper, two types of media indicators are considered: word-count indices and sentiment-analysis indices.

The word-count indices are the simplest form of the media sentiment indicators. The idea is simple: one counts the occurrences of a word or group of words, whose polarity can be determined more or less unambiguously, in several media. One example of such index is the famous recession index, or R-word index, of *The Economist*. It counts the number of articles in the *Washington Post* and the *New York Times* using the word “recession” in a quarter. In Germany, similar indicator had been developed at the HypoVereinsbank. However, several years ago the publication of the index was given up. Therefore, we had to reconstruct it. For this purpose we computed the number of articles published in the most influential German general and economic newspapers (*Frankfurter Allgemeine Zeitung*, *Handelsblatt*, and *Süddeutsche Zeitung*) and in one business journal (*WirtschaftsWoche*) containing the word “Rezession”. The counts for *Frankfurter Allgemeine*

---

<sup>3</sup>The two OECD composite leading indicators for Germany are based on several components such as macroeconomic variables (new orders, spread, etc.) and ifo business survey indicator.

Zeitung were obtained using the online archive search of the newspaper<sup>4</sup>. To calculate the number of articles in Handelsblatt and WirtschaftsWoche we used their joint article database<sup>5</sup>. Finally, for Süddeutsche Zeitung the word occurrences were recovered from the database Genios.<sup>6</sup>

The simple R-word index was constructed in a two-step procedure: First, the “Rezession” word occurrences were aggregated to the monthly frequency by computing the monthly means. Secondly, the monthly series were added up across the four media. However, since our sample includes both general and specialized media, we have to account for their different exposure to the word “recession”: the relative frequency of the word varies from 0.4% in Süddeutsche Zeitung to 2.4% in WirtschaftsWoche. Hence, the simple adding of the medium-specific averages could introduce a bias. In order to address the problem we computed a scaled R-word index by dividing the number of monthly occurrences of the word “Rezession” by those of the word “der” for each medium. The latter word was chosen as a proxy for the overall text size, given that it is the most frequent word in German language.

A more sophisticated way to analyze media is the method of content analysis. Content analysis “is a research technique for the objective, systematic, and quantitative description of the manifest content of communication” (Berelson (1952), 18). There exist many different types of content analysis, going beyond simple frequency counts such as complicated assessments of arguments and media frames. Our contribution is based on the analysis of the content of opinion-leading media in Germany, including five TV news programs, two weekly magazines, and one daily tabloid newspaper by the Swiss-based media analysis institute Media Tenor International. News items only referring to the state of the economy in the media set were analyzed over the period from January 1, 2001 through March 31, 2014. Hence, the data set analyzed can be seen as a subset of a much bigger data set including news items on all possible protagonists, such as persons (politicians, entrepreneurs, managers, celebrities, etc.) and institutions (political parties, companies, football clubs, etc.). Each of these news items was analyzed with regard to the topic mentioned (unemployment, inflation, etc.), the region of reference (for example, Germany, EU, USA, UK, BRIC, worldwide), the time reference (such as past, present, and future), the source of information (journalist, politician, expert, etc.), as well as with regard to the tone of the information (negative, positive or neutral).<sup>7</sup> Overall 80,675 news items about the state of the economy have been analyzed.

---

<sup>4</sup>[www.faz.de](http://www.faz.de)

<sup>5</sup>[www.wirtschaftspresse.biz](http://www.wirtschaftspresse.biz)

<sup>6</sup>[www.genios.de](http://www.genios.de)

<sup>7</sup>Media Tenor International employs professional coders to carry out media-analysis. Only coders that achieved a minimum

For a detailed description of the analyzed media set see Table 2.

Table [Analyzed Media Set] about here

Based on the rating we computed Media Tenor International indices (MT) as the differences between the percentage share of the positive ratings and the that of the negative ratings:

$$B_{i,j,t} = 100 \times \frac{A_{i,j,t}^+ - A_{i,j,t}^-}{A_{i,j,t}^+ + A_{i,j,t}^- + A_{i,j,t}^0} \quad (2)$$

where  $A_{it}^+$  is the number of positive ratings of medium reports about events happening in the time  $i$  in the country  $j$ , published in the period  $t$ ,  $A_{i,j,t}^-$  is the number of negative ratings, and  $A_{i,j,t}^0$  is the number of neutral ratings. The index varies between  $-100$  (all reports are negatively rated) and  $100$  (all reports are positively rated).

In this study, we construct four overall indices: media sentiments regarding all countries in the present, media sentiments concerning all countries in the future, media sentiments regarding only Germany in the present, and media sentiments concerning only Germany in the future. In addition, we compute similar indices for 5 most frequent economic topics (budget, currency, labour market, business cycle, and taxation, see Table 4).

Moreover, the indices of the present and the future sentiment are employed to construct a so-called **media climate index**:

$$MCI = \sqrt{(MS_{it}^{present} + 100)(MS_{it}^{future} + 100)} \quad (3)$$

where  $MS_{it}^{present}$  is the present sentiment index and  $MS_{it}^{future}$  is the future sentiment index. By construction, the MCI can take values between 0 indicating extremely bad media climate and 200 pointing to an excellent media climate.

---

reliability of 0.85 are cleared for coding. That means that the coding of these coders deviate at most by 0.15 from the trainers' master-versions. The reliability of the coding is checked on an ongoing basis, both with quarterly standard tests and random spot checks. For each month and coder, three analyzed reports are selected randomly and checked. Coders scoring lower than 0.80 are removed from the coding process. In none of the months the mean deviation among all coders was above 0.15.

### 3 Forecast evaluation

#### 3.1 Measures for comparing performance

Typically, the usefulness of a forecasting model is evaluated based on its precision. Here, the precision of the models over all periods is measured by the Root Mean Squared Forecast Error (RMSFE) and the Theil's U. The RMSFE is calculated as

$$RMSFE = \sqrt{\sum_{t=T_E+1}^T (\hat{y}_{i,t} - y_t)^2}, \quad (4)$$

where  $\hat{y}_{m,t}$  is the forecast made by model  $m$  ( $m = 1, \dots, M$ ) for period  $t$ ,  $t = T_E + 1, \dots, T$ , where  $T_E$  is the first estimation subsample and  $y_t$  is the realized value. Here, the Theil's U is constructed such that it compares the forecast performance of model  $m$  to that of the benchmark AR-model. It is computed as ratio of the RMSFE of model  $m$  and the RMSFE of the autoregressive model

$$TheilsU_m = \frac{RMSFE_m}{RMSFE_{AR}} \quad (5)$$

The RMSFE and Theil's U are average measures over all periods and therefore do not reflect the instability of performance of individual models over time. In fact, the rank of a model according to its accuracy can fluctuate enormously: being the best model in some periods, in others it can rank the worst. Surely, huge instability is not a desirable property of a forecasting model. In order to take this into account we need a new forecast performance measure. Firstly, let us define a single-period rank of model  $m$  in period  $t$  as  $\rho_{m,t} = rank(RMSFE_{m,t})$ . Thus, the model with the lowest  $RMSFE$  in period  $t$  obtains the rank of 1. Secondly, with an eye to the construction of our third measure below, we want the rank to be independent of the number of models and negatively correlated to the  $RMSFE$ . Therefore, we compute the transformed rank by calculating the percentage of all models outperformed by model  $m$  in period  $t$

$$\tilde{\rho}_{m,t} = \left(1 - \frac{\rho_{m,t}}{M}\right) \times 100. \quad (6)$$

Thirdly, we compute the average transformed rank for each model  $m$  over all periods

$$PercOut_m = \frac{\sum_{t=T_E+1}^T \tilde{\rho}_{m,t}}{T - T_E}. \quad (7)$$

This measure can be interpreted as the average percentage of models outperformed by model  $m$  over time. It can vary between 0 and 100 per cent. The larger its value, the better the precision of the respective model. It can be considered as a complement to RMSFE, although it can be expected that both are highly correlated.

Fourthly, the instability of model  $m$  in each period can be measured as the standard deviation of its respective transformed rank over time,  $\sigma_{\tilde{\rho}} = sd(\tilde{\rho}_{m,t})$ . The larger its value, the more instable the forecasting performance of a model over time. It is the average percentage point dispersion of  $\tilde{\rho}_{m,t}$  around its mean,  $PercOut_m$ . Of two models with the same  $PercOut_m$  we would prefer the one with the lower  $\sigma_{\tilde{\rho}}$ .

Finally, we construct a measure of reliability, which takes into account both precision and stability. A reliable model is the model with a high precision and a low instability. Thus, we define the measure  $R_m$  as

$$R_m = \frac{PercOut_m}{\sigma_{\tilde{\rho}}}. \quad (8)$$

$R_m$  is an increasing function of the average relative precision with respect to the alternative models and a decreasing function of its instability. In fact, it is an inverted coefficient of variation. It is analogous to the Sharpe ratio in finance.

### 3.2 Performance of individual models

Table [[Best models: July 2001 to April 2014](#), 5] about here

Table 5 compares the performance of the five best individual models, which is, without considering combined forecasts, over all forecasting periods. Our analysis provide here allows for the comparison of media and non-media indicators. However, due to the differences in the underlying media set, R-word and Media Tenor International based indicators are not comparable to each other.<sup>8</sup> Columns I to III show the RMSFE and Theil's U as well as the ranking of each model according to the two measures, columns IV and V present the mean percentage of models outperformed by the respective model and the corresponding rank. Columns VI

---

<sup>8</sup>For a description of the underlying media sets see section 2.2.

and VII show the standard deviation of the percentage of models outperformed by the respective model and its ranking. Columns VIII and IX report the coefficient of reliability and the corresponding rank, and column X and XI present its best and worst rank over all periods. Lines 1 through 5 show the five best models with respect to RMSFE and Theil's U, lines 6 through 10 the five best models with respect to the mean percentage of models outperformed, lines 11 through 15 the five best models according to stability, and lines 16 through 20 the five best models with respect to reliability.

According to RMSFE, standard deviation of rank, and coefficient of reliability, models using media data clearly outperform models without media data. In particular, according to RMSFE, standard deviation of rank, and coefficient of reliability, as well as the second best model with respect to the number of models outperformed on average employs MT.currency, cli.ger, and manuf.order. Its Root Mean Forecast Error is 43% lower than that of the benchmark AR model giving a Theil's U of 57. On average it outperforms 69.8% of the alternative individual models. The standard deviation of its rank is 21.84, however, it oscillates between a minimum of 99 and a maximum of 15859. This wide range of ranks is also observed for the other high performing models. Its coefficient of reliability is 319.6. About 50% of German exports are directed to countries outside the Euro area. For some important sectors, like machinery, investment goods, and cars more than 50% of the overall production are exported. Thus, media information on currency issues such as provided by MT.currency is crucial for predicting industrial production.<sup>9</sup>

Apart from MT.currency, two more Media Tenor International based indicators, namely MT.taxation with its particular information on tax issues and MT.de, which consists all economy-related topics with an effect on the German economy, form part of models ranking among the five best models in each category. The model employing MT.taxation, cli.ger, and manuf.order ranks fifth with respect to the mean percentage of models outperformed and 31st in terms of the standard deviation of this measure. In terms of the reliability indicator, it ranks 6th. The model consisting of MT.de, esi.ger, and dax is the third best with respect to stability. However, it only ranks 5081st according to RMSFE and 6254th according to the mean percentage of models outperformed. Most of the media reports on taxes relate to taxes in general. Taxes increase budget constraints, and so negatively affect demand for industrial production, for both companies and households. Hence, news related to tax changes influence sales expectations, too.

---

<sup>9</sup>Source: German Federal Statistical Office.

Both R-word indicators form part of the the five best models. Together with R1 and `manuf.order`, `rword` ranks fourth with respect to the standard deviation of the percentage of models outperformed and fifth according to reliability. There is a striking contrast of its performance according to its mean forecast error and its mean percentage of outperforming alternative models, ranking 820th regarding the former and 17th with respect to the latter.

Table [\[Best models: May 2008 to January 2009, recession period, 6\]](#) about here

The period under consideration includes a so called “Great Recession”. To show which models are especially good at predicting it, we separately analyze the recession period, which starts in May 2008 and ends in January 2009. It is based on the ECRI<sup>10</sup> classical business cycle chronology.

Table 6 shows the corresponding results for the recession period only. When compared to the outcome over all periods, the correlation of accuracy and reliability is lower for the recession period.

Again according to RMSFE, standard deviation of rank, and coefficient of reliability, models using media data clearly outperform models without media data. In particular, according to RMSFE all five best models contain media based indices such as MT indices as well as R-word. The number one model with respect to RMSFE consists of both R-word and `MT.taxation`. Nevertheless, none of the models appears among the five best models in more than one category. The improvement of accuracy when compared to the AR is higher than when looking at the best models over all periods, the Theil’s U giving values between 43.4 and 43.7. However, no model containing a media indicator ranges among the five best models, according to the mean percentage of outperforming alternative models. The model containing `MT.de.labor`, `trade.bal`, and `imp.pr` is the third best one according to the standard deviation of ranks. However, the most stable models are performing poorly with respect to accuracy. Its rank with respect to RMSFE is 17,134 and with respect to the mean percentage of models outperformed it is 17,135. Looking at reliability, four media indicators appear in the best models. The model containing `MT.de.cycle`, `dax`, and `usd` is the best model with a value of 10.84, followed by the model containing `rword`, `cli.eur`, and `cons.conf` has a value of 9.14, the model consisting of `MT.all`, `zew`, and `esi.ger` ranks fourth with a value of 8.52, and `MT.de` in combination with `zew` and `esi.ger` ranks fifth with a value of 8.50. The standard deviations of ranks and their ranges are smaller when compared to the respective values

---

<sup>10</sup>Economic Cycle Research Institute, <https://www.businesscycle.com>.

of all periods. The standard deviation of the best model is 5.5, its minimum rank is 5729th and its maximum rank is 8283th. With the exception of the second best model containing rword, which ranks 75th with respect to RMSFE and 21th according to the mean percentage of models outperformed, the relative accuracy of the best models is much lower when compared to the results over all periods. For RMSFE and the percentage of models outperformed the best rank of the remaining 4 models is 4810 with a Theil's U of 50.9, respectively 2936, with a value of 83.7, both for the model containing R6, gfk, and dax. During the recession period, the media information that directly addresses the overall situation of the economy, or those that reject it representing the sentiment on all sectors such as MTL.all for all economies and MTL.de for the German economy are best suited to predict industrial production.

### 3.3 Performance of forecast combinations

Table [[Combinations: July 2001 to April 2014, sorted by coefficient of reliability, 7](#)] about here

As shown, individual models are very instable over time. To illustrate the relationship of precision and stability, we draw bivariate highest density regions plots<sup>11</sup> in Figures 1 and 2 for the whole period and the recession period, respectively. Each point in these graphs represents a single model. The horizontal axis shows the percentage of models outperformed by the respective model. The vertical axis depicts the standard deviation of the percentage of models outperformed by the respective model. The light-gray and yellow areas are bivariate high density regions that cover 50 and 95% of the distribution, respectively. As Figure 1 shows, higher precision of individual models is weakly positively correlated with stability. The best models are those with a high precision and a high stability.

Following the literature, we try to improve upon this by evaluating the usefulness of media data in forecast combinations of individual models. Indeed, as can be seen in Figures 1 and 2, the combination models improve the relationship between precision and stability. They allow reducing substantially the instability without incurring large losses in precision. Figure 1 presenting the results for all periods shows the dominance of the combined forecasts with respect to stability. At the same time, they are among the most accurate forecasting models. Figure 2 reports the results for the recession period. As the averages are based on fewer observations

---

<sup>11</sup>See [Hyndman, 1996](#).

and due to the higher uncertainty over economic downturns, the bivariate highest density distribution is much more spread in both dimensions. There is a group of models characterized by low accuracy and high stability. At the same time there is a group of models having a higher accuracy but a very large instability. The advantages of combinations are smaller but still pronounced. Higher precision of individual models can only be obtained at the cost of increased instability.

Table 7 contrasts the results of the combined forecast of all models not containing media indices (benchmark, in italics) with combinations that employ both all non-media and one media index at a time to see, whether this improves the performance. The table is based on the results for all periods and the models are sorted by the coefficient of reliability. The standard deviation of the combined forecasts is markedly lower than that of the best individual models. The worst ranks of the combinations are much lower, as well. This results in the combinations being the best models with respect to the coefficient of reliability.

As shown in the analysis of the individual models, concerning combinations again models using media data clearly outperform models without media data. According to the coefficient of reliability the best combination contains models including *rword\_sc*. Six more media indices improve the benchmark. In descending order with respect to reliability these are *MT.taxation*, *MT.de.taxation*, *rword*, *MT.de*, *MT.de.cycle*, and *MT.currency*. However, difference in the ranks in between the media-based models can be determined by the differences in the media set used.

The addition of models containing the remaining media indices leads to a deterioration with respect to reliability when compared to the benchmark model. In general, the performance of the combined forecasts does not display substantial differences.

Table [[Combinations: May 2008 to January 2009, recession period, sorted by coefficient of reliability, 8](#)] about

here

Table 8 reports the forecast performance of combined models for the recession period from May 2008 to January 2009 only. Again, models using media data clearly outperform models without media data. The combinations of models including R-word indicators rank first and second according to the reliability measure followed by eleven other combinations including media indices, such as *MT.cycle*, *MT.future*, *MT.climate*, and *MT.all*. These combinations are superior in terms of precision to the remaining combinations, in particular,

to the combination of all non-media models, which rank 14th. The media combination has on average a 4% smaller forecast error than the combination of all non-media models. In comparison to the benchmark model, the autoregressive model, the improvement is about 20%. The mean percentage of models outperformed by the best media combination is by 3 percentage points higher than the combination of all non-media models. The standard deviation of 7.28 is about 1.5 percentage points lower. Thus, the media models are both more accurate and less instable resulting in a higher reliability measure. Interestingly, MT.all that is based on all news ranks relatively high. This means, that the overall media sentiment can be useful in predicting recessions.

## 4 Conclusion

In this paper, we analyzed the usefulness of media indicators for the prediction of monthly series of German industrial production growth. We used two types of media indicators: a simple word-count index of the word “recession” and several Media Tenor International indices which are based on a more sophisticated method using human analysts of reports in German opinion-leading media. The forecast performance was evaluated through forecast experiment covering the period from July 2004 to April 2014. In addition, we consider the period of the Great Recession using the business cycle chronology of ECRI, to see whether the media indices improve recession forecasts. More than 17,000 individual models representing all possible combinations with a maximum of 3 out of 48 macroeconomic, survey, and media indicators were employed.

The forecasting performance was evaluated using four different criteria. First, we use two measures of forecast accuracy, namely the Root Mean Squared Forecast Error and the mean percentage of outperformed alternative models each period. Then, as a measure of stability we employ the standard deviation of the percentage of outperformed alternative models each period. Finally, we introduce and apply our own measure of reliability, which aggregates the information on accuracy and stability.

The results clearly show, models using media data outperform models without media data. This is case according to both individual models as well as combinations of the individual models.

Individual models using media data are among the best models with respect to accuracy and stability over the whole sample period. For the overall sample, the Media Tenor International index based on news related to foreign exchange market stands on top of the rankings in terms of all four criteria considered. This might

be due to the strong export-orientation of German industrial production. For the recession period, the models including the R-word indices focusing on recessions by construction are particularly useful.

Combinations of the individual models improve the stability of the forecasts and lead to highly accurate models at the same time. We tested if augmenting the combination of models not making use of media data with models making use of one additional media index improves forecasts. Over the complete sample and the recession period, some of the media augmented combinations lead to an improvement of forecast reliability. In addition, media sentiment on the overall situation implicitly rejecting information on the business cycle improves forecast combinations for the recession period.

Under the common heading of media data two very different groups of indicators have been employed. The main differences is in the technique used to extract information from the media. Media Tenor International extracts the overall sentiments from the media items with the help of specialized analysts, while R-word simply counts occurrences of one word. However, the data sets employed here are not comparable: they are both non-overlapping and cover different segments of media. Although these differences do not preclude the comparison of their helpfulness in the prediction of industrial production, a deeper analysis is needed to understand the impact of these differences on the forecasting performance. This is left to future research.

Nevertheless, our analysis have clearly shown that when it comes to the forecast of industrial production models using media data clearly outperform models without media data.

## References

- Abberger, K. and K. Wohlrabe (2006). Einige Prognoseeigenschaften des ifo Geschäftsklimas — Ein Überblick über die neuere wissenschaftliche Literatur. *ifo Schnelldienst* 59(22), 19–26.
- Ammann, M., R. Frey, and M. Verhofen (2011). Do newspaper articles predict aggregate stock returns? Working paper, University of St. Gallen - Swiss Institute of Banking and Finance.
- Berelson, B. (1952). Content analysis in communication research.
- Bollen, J., H. Mao, and X. Zeng (2011). Twitter mood predicts the stock market. *Journal of Computational Science* 2, 1–8.

- Bordino, I., S. Battiston, G. Caldarelli, M. Cristelli, A. Ukkonen, and I. Weber (2011). Web search queries can predict stock market volumes. *CoRR abs/1110.4784*.
- Commission, E. (2014). For a european industrial renaissance. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions COM 14/2*.
- Denzau, A. T. and D. C. North (1994). Shared mental models: ideologies and institutions. *Kyklos* 47(1), 3–31.
- Doms, M. and N. Morin (2004). Consumer sentiment, the economy, and the news media. Finance and Economics Discussion Series 2004-51, Board of Governors of the Federal Reserve System (U.S.).
- Goidel, R. K. and R. E. Langley (1995). Media coverage of the economy and aggregate economic evaluations: Uncovering evidence of indirect media effects. *Political Research Quarterly* 48, 313–328.
- Grossarth-Maticek, J. and J. Mayr (2008). Medienberichte als Konjunkturindikator. *Ifo-Schnelldienst* 61, 17–29.
- Hüfner, F. and M. Schröder (2002). Prognosegehalt von ifo-Geschäftserwartungen und ZEW-Konjunkturerwartungen: Ein ökonometrischer Vergleich. *Jahrbücher für Nationalökonomie und Statistik* 222(3), 316–336.
- Hyndman, R. J. (1996). Computing and graphing highest density regions. *The American Statistician* 50(2), 120–126.
- Iselin, D. and B. Siliverstovs (2013). Using newspapers for tracking the business cycle: A comparative study for Germany and Switzerland. KOF Working papers 13-337, KOF Swiss Economic Institute, ETH Zurich.
- Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47(2), 263–91.
- Keynes, J. M. (1937). The general theory of employment. *The Quarterly Journal of Economics* 51(2), 209–223.
- Kholodilin, K. A., M. Podstawski, and B. Siliverstovs (2010). Do google searches help in nowcasting private consumption?: A real-time evidence for the US. Discussion Papers of DIW Berlin 997, DIW Berlin, German Institute for Economic Research.

- Kholodilin, K. A. and B. Siliverstovs (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.
- Nadeau, R., R. G. Niemi, and T. Amato (2000). Elite economic forecasts, economic news, mass economic expectations, and voting intentions in Great Britain. *European Journal of Political Research* 38, 135–170.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Simon, H. A. (1957). Models of man; social and rational.
- Soroka, S. N. (2006). Good news and bad news: Asymmetric responses to economic information. *The Journal of Politics* 68, 372–385.
- Stock, J. H. and M. W. Watson (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting* 23(6), 405–430.
- Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance* 62(3), 1139–1168.
- Uhl, M. W. (2010). Explaining US consumer behavior and expectations with news sentiment. KOF Working papers 10-263, KOF Swiss Economic Institute, ETH Zurich.
- Uhl, M. W. (2011). Nowcasting private consumption with TV sentiment. KOF Working papers 11-293, KOF Swiss Economic Institute, ETH Zurich.
- von Hayek, F. A. (1992). The pretence of knowledge. *Economic Sciences, 1969-1980: The Sveriges Riksbank (Bank of Sweden) Prize in Economic Sciences in Memory of Alfred Nobel* 1, 179.

## Appendix

Table 1: Macroeconomic indicators: definitions and descriptive statistics

Indicator	Description	Source	Transformation	Mean	Standard deviation
dax	DAX, German stock market index	Deutsche Börse AG	year-on-year change rates	5.05	25.52
eur.stox	Eurostoxx 50, European stock market index	STOXX Ltd.	year-on-year change rates	-1.54	21.44
long.rate	long-term government bond yields, 9-10 years	Datastream		3.43	1.08
short.rate	short-term euro repo rate	European Central Bank		2.19	1.26
oil	crude Brent oil in US dollar per barrel	Datastream	year-on-year change rates	15.57	33.32
manuf.order	manufacturing orders	German Federal Statistical Office	year-on-year change rates	2.44	11.35
usd	US dollar – euro exchange rate	Datastream	year-on-year change rates	3.29	9.90
ex	German exports of goods and services	Deutsche Bundesbank	year-on-year change rates	5.26	9.57
im	German imports goods and services	Deutsche Bundesbank	year-on-year change rates	4.60	10.72
trade.bal	German trade balance	Deutsche Bundesbank		13.13	2.82
ex.pr	German export price inflation	Deutsche Bundesbank	year-on-year change rates	0.90	1.60
im.pr	German import price inflation	Deutsche Bundesbank	year-on-year change rates	1.12	4.67
tot	terms of trade	Deutsche Bundesbank	year-on-year change rates	-0.07	3.25
infl	consumer price inflation	German Federal Statistical Office		1.62	0.68
spread	long-term minus and short-term rates	authors' calculation		1.22	0.71
ip	industrial production	Deutsche Bundesbank <sup>a</sup>	year-on-year change rates	1.82	7.31

<sup>a</sup>[http://www.bundesbank.de/Navigation/DE/Statistiken/Suche\\_Statistik/Echtzeitdaten/statistiksuche\\_rtd\\_node.html](http://www.bundesbank.de/Navigation/DE/Statistiken/Suche_Statistik/Echtzeitdaten/statistiksuche_rtd_node.html)

Table 2: Analyzed Media Set

TV-Program / Newspaper	Number of news items analysed
TV-newscasts:	
ARD Tagesschau	11,472
ARD Tagesthemen	14,933
ZDF heute	10,158
ZDF heute journal	15,415
RTL Aktuell	6,167
Weekly magazines:	
Spiegel	4,833
Focus	7,111
Daily newspaper:	
Bild	10,586
<b>Total</b>	<b>80,675</b>

Table 3: Sentiment indicators: definitions and descriptive statistics

Indicator	Description	Source	Transformation	Mean	Standard deviation
R1	business climate, levels	ifo Institute for Economic Research		102.36	7.63
R2	business situation, levels	ifo Institute for Economic Research		104.50	10.87
R3	business expectations, levels	ifo Institute for Economic Research		100.44	6.34
R4	business climate, balances	ifo Institute for Economic Research		-2.33	14.74
R5	business situation, balances	ifo Institute for Economic Research		-1.66	20.64
R6	business expectations, balances	ifo Institute for Economic Research		-2.61	12.44
zew	ZEW indicator of economic sentiment	Centre for European Economic Research		17.40	35.05
esi.eu	economic sentiment indicator, European Union	European Commission		99.15	9.43
esi.ger	economic sentiment indicator, Germany	European Commission		98.18	9.60
cli.eur	composite leading indicator, Euro area (18 countries)	OECD		99.95	1.19
cli.ger	composite leading indicator, Germany	OECD		99.98	1.48
cons.conf	confidence indicator	European Commission		-7.94	9.66
gfk	GfK consumer index	Society for Consumer Research		4.86	3.88

Table 4: Media indicators: definitions and descriptive statistics

Indicator	Description	Source	Transformation	Mean	Standard deviation
MT.all	all countries, assessment of current situation and expectation	Media Tenor International		-29.89	16.04
MT.future	all countries, expectation	Media Tenor International		-20.13	18.91
MT.present	all countries, assessment of current situation	Media Tenor International		-35.14	16.96
MT.climate	all countries, Media Climate Index, see equation 3	Media Tenor International		71.58	16.26
MT.de	Germany, assessment of current situation and expectation	Media Tenor International		-20.85	19.07
MT.de.future	Germany, expectation	Media Tenor International		-14.41	18.94
MT.de.present	Germany, assessment of current situation	Media Tenor International		-24.92	22.27
MT.de.climate	Germany, Media Climate Index, see equation 3	Media Tenor International		79.57	18.62
MT.budget	all countries, assessment of current situation and expectation, government budget	Media Tenor International		-42.83	26.92
MT.currency	all countries, assessment of current situation and expectation, currency related issues	Media Tenor International		-26.77	34.24
MT.labor	all countries, assessment of current situation and expectation, labor market related issues	Media Tenor International		-28.06	19.74
MT.cycle	all countries, assessment of current situation and expectation, business cycle related issue	Media Tenor International		-22.37	35.43
MT.taxation	all countries, assessment of current situation and expectation, taxation related issues	Media Tenor International		-26.33	15.17
MT.de.budget	Germany, assessment of current situation and expectation, government budget	Media Tenor International		-29.07	33.48
MT.de.labor	Germany, assessment of current situation and expectation, labor market related issues	Media Tenor International		-23.35	22.00
MT.de.cycle	Germany, assessment of current situation and expectation, business cycle related issue	Media Tenor International		-1.57	47.91
MT.de.taxation	Germany, assessment of current situation and expectation, taxation related issues	Media Tenor International		-26.67	15.61
rword_sc	recession word indicator scaled by the overall number of words	authors' calculation		9.37	8.00
rword	recession word indicator	authors' calculation		13.93	12.46

Table 5: Best models: July 2001 to April 2014

criterion	Line #	Variables included in the model	RMSFE			Mean % of models outperformed each period ( <i>PercOut</i> )		Standard deviation of rank each period		Coefficient of reliability ( <i>CoeffRel</i> )		Best Rank	Worst Rank
			I Value	II Theil's U	III Rank	IV Value	V Rank	VI Value	VII Rank	VIII Value	IX Rank		
RMSFE	1	MT.currency, cli.ger, manuf.order	3.76	56.67	1	69.80	2	21.84	1	3.20	1	99	15859
	2	cli.ger, dax, manuf.order	3.79	57.13	2	69.16	4	22.29	5	3.10	4	51	15831
	3	rword.sc, esi.eu, manuf.order	3.90	58.78	3	67.50	13	25.38	1282	2.66	92	8	16505
	4	cli.ger, manuf.order, im	3.91	58.93	4	70.08	1	22.56	15	3.11	3	133	16435
	5	R5, manuf.order, infl	3.91	58.94	5	66.29	58	26.50	3461	2.50	394	103	16635
Mean % of models outperformed each period ( <i>PercOut</i> )	6	cli.ger, manuf.order, im	3.91	58.93	4	70.08	1	22.56	15	3.11	3	133	16435
	7	MT.currency, cli.ger, manuf.order	3.76	56.67	1	69.80	2	21.84	1	3.20	1	99	15859
	8	cli.ger, manuf.order, ex	4.28	64.65	387	69.67	3	22.33	7	3.12	2	22	15900
	9	cli.ger, dax, manuf.order	3.79	57.13	2	69.16	4	22.29	5	3.10	4	51	15831
	10	MT.taxation, cli.ger, manuf.order	4.77	71.93	1087	69.05	5	22.98	31	3.00	6	20	16511
Standard deviation	11	MT.currency, cli.ger, manuf.order	3.76	56.67	1	69.80	2	21.84	1	3.20	1	99	15859
	12	rword.sc, cli.ger, tot	4.52	68.22	762	62.72	692	21.91	2	2.86	17	66	15805
	13	MT.de, esi.ger, dax	5.84	88.07	5081	52.07	6254	22.25	3	2.34	1124	431	16049
	14	rword, R1, manuf.order	4.56	68.77	820	67.27	17	22.28	4	3.02	5	135	17076
	15	cli.ger, dax, manuf.order	3.79	57.13	2	69.16	4	22.29	5	3.10	4	51	15831
Coefficient of reliability ( <i>CoeffRel</i> )	16	MT.currency, cli.ger, manuf.order	3.76	56.67	1	69.80	2	21.84	1	3.20	1	99	15859
	17	cli.ger, manuf.order, ex	4.28	64.65	387	69.67	3	22.33	7	3.12	2	22	15900
	18	cli.ger, manuf.order, im	3.91	58.93	4	70.08	1	22.56	15	3.11	3	133	16435
	19	cli.ger, dax, manuf.order	3.79	57.13	2	69.16	4	22.29	5	3.10	4	51	15831
	20	rword, R1, manuf.order	4.56	68.77	820	67.27	17	22.28	4	3.02	5	135	17076

Table 6: Best models: May 2008 to January 2009, recession period

criterion	Line #	Variables included in the model	RMSFE			Mean % of models outperformed each period ( <i>PercOut</i> )		Standard deviation of rank each period		Coefficient of reliability ( <i>CoeffRel</i> )		Best Rank	Worst Rank
			I Value	II Theil's U	III Rank	IV Value	V Rank	VI Value	VII Rank	VIII Value	IX Rank		
RMSFE	1	ip, rword, MT.taxation, esi.eu	2.09	43.44	<b>1</b>	84.13	13	30.94	16351	2.72	5102	48	16682
	2	ip, rword, MT.de.taxation, esi.eu	2.10	43.59	<b>2</b>	83.79	19	30.86	16318	2.71	5134	40	16696
	3	ip, rword, esi.eu, usd	2.10	43.60	<b>3</b>	84.73	9	30.85	16313	2.75	4914	51	16592
	4	ip, rword, MT.future, esi.eu	2.11	43.72	<b>4</b>	83.46	35	31.62	16554	2.64	5749	127	16937
	5	ip, rword, esi.ger, cli.eur	2.11	43.72	<b>5</b>	83.93	16	20.55	6428	4.08	803	1	11659
Mean % of models outperformed each period ( <i>PercOut</i> )	6	ip, manuf.order, im.pr, infl	2.29	47.60	33	87.28	<b>1</b>	15.58	2500	5.60	133	26	7625
	7	ip, manuf.order, ex, im.pr	2.48	51.38	81	86.69	<b>2</b>	14.75	2012	5.88	86	4	5935
	8	ip, manuf.order, ex, infl	2.35	48.76	41	86.66	<b>3</b>	21.54	7397	4.02	852	140	11434
	9	ip, manuf.order, tot, infl	2.36	48.89	44	86.52	<b>4</b>	20.35	6219	4.25	657	26	9653
	10	ip, manuf.order, ex, tot	2.46	51.05	76	85.55	<b>5</b>	16.76	3247	5.11	251	1	7014
Standard deviation	11	ip, R2, usd, tot	6.57	136.45	17086	2.20	17130	2.06	<b>1</b>	1.06	15451	16094	17134
	12	ip, R5, usd, tot	6.57	136.44	17085	2.21	17129	2.07	<b>2</b>	1.06	15450	16092	17135
	13	ip, MT.de.labor, trade.bal, im.pr	6.91	143.50	17134	1.75	17135	2.14	<b>3</b>	0.81	16517	16027	17130
	14	ip, R2, gfk, oil	6.51	135.07	17055	2.34	17123	2.40	<b>4</b>	0.97	15854	16165	17106
	15	ip, R5, gfk, oil	6.51	135.06	17054	2.34	17122	2.41	<b>5</b>	0.97	15871	16145	17107
Coefficient of reliability ( <i>CoeffRel</i> )	16	ip, MT.de.cycle, dax, usd	3.94	81.83	7126	59.65	6072	5.50	81	10.84	<b>1</b>	5729	8283
	17	ip, rword, cli.eur, cons.conf	2.45	50.94	75	83.74	21	9.16	312	9.14	<b>2</b>	61	5597
	18	ip, R6, gfk, dax	3.60	74.77	4810	70.11	2936	7.78	201	9.02	<b>3</b>	3031	7230
	19	ip, MT.all, zew, esi.ger	4.11	85.41	8234	52.80	8412	6.20	101	8.52	<b>4</b>	5821	9417
	20	ip, MT.de, zew, esi.ger	4.15	86.05	8412	52.11	8685	6.13	98	8.50	<b>5</b>	5960	9415

Table 7: Combinations: July 2001 to April 2014, sorted by coefficient of reliability

Variables included in the model	Line #	RMSE		Mean % of models outperformed each period ( <i>PercOut</i> )		Standard deviation of rank each period		Coefficient of reliability ( <i>CoeRel</i> )		Best Rank	Worst Rank	
		I Value	II Theil's U	III Rank	IV Value	V Rank	VI Value	VII Rank	VIII Value			IX Rank
Combination including rword_sc	1	4.90	73.98	1268	65.81	92	14.66	4	4.49	1	75	11603
Combination including MT.taxation	2	4.94	74.52	1329	65.38	128	14.65	3	4.46	2	81	11778
Combination including MT.de.taxation	3	4.94	74.53	1330	65.29	143	14.64	2	4.46	3	7	11906
Combination including rword	4	4.89	73.82	1247	66.19	64	14.93	10	4.43	4	17	11496
Combination including MT.de	5	5.14	77.56	1680	63.10	599	14.30	1	4.41	5	272	10945
Combination including MT.de.cycle	6	4.93	74.42	1320	65.31	139	14.82	5	4.41	6	226	11470
Combination including MT.currency	7	4.95	74.62	1349	65.19	157	14.84	6	4.39	7	12	11743
Combination of non-media data	8	4.91	74.12	1291	65.44	124	14.93	11	4.38	8	240	11788
Combination including MT.cycle	9	4.93	74.42	1321	65.48	116	14.95	13	4.38	9	1	11479
Combination including MT.de.future	10	4.94	74.57	1335	65.24	151	14.92	8	4.37	10	66	11728
Combination including MT.future	11	4.94	74.55	1332	65.22	153	14.93	12	4.37	11	341	11732
Combination including MT.de.budget	12	4.95	74.63	1350	65.07	172	14.92	9	4.36	12	22	11559
Combination including MT.all	13	4.94	74.56	1334	65.13	168	14.95	14	4.36	13	29	11605
Combination including MT.labor	14	4.95	74.67	1356	64.97	183	14.91	7	4.36	14	24	11664
Combination including MT.climate	15	4.94	74.54	1331	65.13	167	14.97	15	4.35	15	108	11702
Combination including MT.de.present	16	4.95	74.64	1351	65.14	164	15.00	17	4.34	16	130	11737
Combination including MT.de.climate	17	4.94	74.58	1337	65.21	154	15.03	20	4.34	17	114	11681
Combination including MT.present	18	4.94	74.59	1344	65.14	166	15.02	19	4.34	18	90	11708
Combination including MT.de.labor	19	4.95	74.66	1354	65.02	179	15.00	16	4.34	19	21	11735
Combination including MT.budget	20	4.95	74.69	1358	64.96	184	15.01	18	4.33	20	16	11642

Table 8: Combinations: May 2008 to January 2009, recession period, sorted by coefficient of reliability

Variables included in the model	Line #	RMSE			Mean % of models outperformed each period ( <i>PercOut</i> )		Standard deviation of rank each period		Coefficient of reliability ( <i>CoeffRel</i> )		Best Rank	Worst Rank
		I Value	II Theil's U	III Rank	IV Value	V Rank	VI Value	VII Rank	VIII Value	IX Rank		
Combination including rword	1	3.88	80.63	6688	57.36	6709	7.28	175	7.87	8	5619	9144
Combination including rword_sc	2	3.98	82.65	7368	55.85	7204	7.11	158	7.86	9	5978	9643
Combination including MT.cycle	3	4.06	84.34	7890	54.46	7744	8.23	233	6.62	32	5086	10143
Combination including MT.future	4	4.08	84.63	7990	54.06	7901	8.22	231	6.57	34	5071	10182
Combination including MT.climate	5	4.08	84.62	7986	54.09	7890	8.36	237	6.47	37	4994	10210
Combination including MT.all	6	4.08	84.59	7980	54.19	7853	8.39	242	6.46	38	4955	10208
Combination including MT.de.future	7	4.08	84.68	8013	53.95	7959	8.37	238	6.45	39	5011	10204
Combination including MT.de.climate	8	4.08	84.70	8020	53.94	7963	8.50	256	6.34	44	4926	10223
Combination including MT.present	9	4.08	84.65	8001	54.12	7880	8.54	260	6.33	46	4899	10240
Combination including MT.taxation	10	4.08	84.66	8004	53.89	7985	8.61	267	6.26	51	4778	10161
Combination including MT.de.taxation	11	4.08	84.62	7987	53.88	7989	8.65	270	6.23	56	4765	10173
Combination including MT.de.budget	12	4.09	84.86	8072	53.76	8029	8.69	275	6.18	61	4844	10308
Combination including MT.currency	13	4.09	84.96	8105	53.45	8170	8.66	271	6.17	63	4990	10337
<i>Combination of non-media data</i>	14	<i>4.06</i>	<i>84.40</i>	<i>7920</i>	<i>54.32</i>	<i>7800</i>	<i>8.82</i>	<i>287</i>	<i>6.16</i>	<i>66</i>	<i>4823</i>	<i>10297</i>
Combination including MT.de.present	15	4.08	84.80	8055	53.98	7944	8.77	283	6.16	68	4805	10294
Combination including MT.labor	16	4.10	85.19	8191	53.28	8238	8.68	273	6.14	71	4767	10147
Combination including MT.budget	17	4.09	85.01	8117	53.51	8141	8.82	288	6.07	82	4784	10300
Combination including MT.de.cycle	18	4.07	84.51	7952	54.21	7848	8.97	304	6.04	84	4748	10246
Combination including MT.de	19	4.18	86.89	8729	51.32	8981	8.50	253	6.04	85	4794	9966
Combination including MT.de.labor	20	4.10	85.21	8195	53.20	8266	8.98	306	5.92	96	4677	10214

Figure 1: Precision and stability over all periods: individual models versus combinations

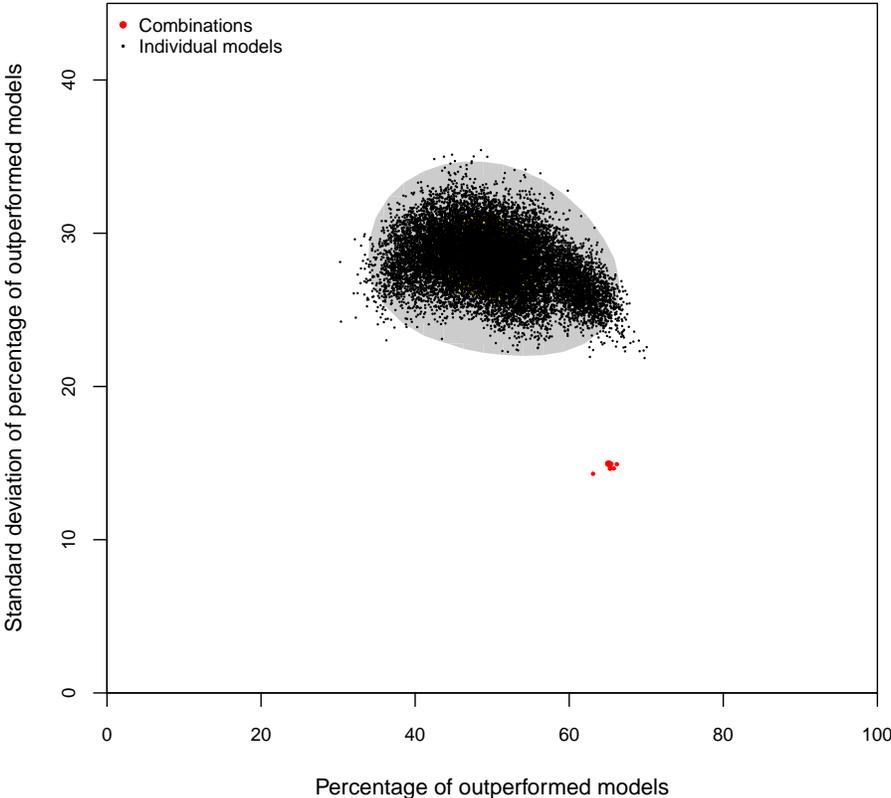


Figure 2: Precision and stability during recession: individual models versus combinations

