

Discussion Papers

Deutsches Institut für Wirtschaftsforschung

2017

Reading between the Lines: Using Media to Improve German Inflation Forecasts

Benjamin Beckers, Konstantin A. Kholodilin and Dirk Ulbricht

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

IMPRESSUM

© DIW Berlin, 2017

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>

Reading Between the Lines: Using Media to Improve German Inflation Forecasts

Benjamin Beckers^a, Konstantin A. Kholodilin^b, Dirk Ulbricht^c

April 28, 2017

Abstract

In this paper, we examine the predictive ability of automatic and expert-rated media sentiment indicators for German inflation. We find that sentiment indicators are competitive in providing inflation forecasts against a large set of common macroeconomic and financial predictors. Sophisticated linguistic sentiment algorithms and business cycle news rated by experts perform best and are superior to simple word-count indicators and autoregressive forecasts.

Keywords: inflation prediction, media sentiment indicators, news reports, real-time forecasting.

JEL Classification: C53; E31; E37.

^aDIW Berlin, Macroeconomics, Mohrenstraße 58, 10117 Berlin, Germany, bbeckers@diw.de.

^bDIW Berlin, Macroeconomics, Mohrenstraße 58, 10117 Berlin, Germany, kkholodilin@diw.de; NRU HSE, Kantemirovskaya ul, 3, korp.1, lit. A, 194100, St. Petersburg, Russia.

^cInstitut für Finanzdienstleistungen, Rödingsmarkt 31/33, 20459, Hamburg, Germany, dirk.ulbricht@iff-hamburg.de.

1 Introduction

Today, most major central banks aim to stabilize inflation at a low but positive rate over the medium-term. In order to meet these targets, monetary policy makers consequently require accurate forecasts of the future path of inflation. Likewise, other economic agents, such as households, firms, and financial institutions, also share great interest in reliable inflation forecasts either for gauging future policy rate decisions by the central bank, or for their price setting, wage negotiations, nominal asset holdings and investment decisions. These needs have stimulated extensive and ongoing research aiming to obtain reliable real-time forecasts of future inflation. In this paper, we develop a novel set of indicators that capture *sentiment* in TV and print media reporting on inflation, the business cycle, and on policy news, and assess their potential for real-time inflation prediction.

So far, the literature has long focused on indicators that can be grouped in two categories. First, real economic variables such as the unemployment rate, vacancies or output measures are employed, which is motivated by the insights of the Phillips-curve literature (see, e.g., Phillips, 1958, Gordon, 1981). Second, the literature has employed monetary or financial variables such as policy rates, the term spread, stock prices or exchange rates, arguing that asset prices are inherently forward-looking and contain information about market expectations on future real economic variables. The predictive ability of these indicators, however, is mixed at best: While real variables, depending on their specification and the sample, are frequently found to improve inflation forecasts relative to simple autoregressive models, this generally does not hold for financial variables (see Stock and Watson, 1999, 2003, 2007; Ang et al., 2007).

In the past decade, the literature has, therefore, extended the set of predictors by survey-based indicators that measure agents' expectations about future inflation directly instead of relying on information inferred from asset prices (see, e.g., Thomas, 1999 and Mehra, 2002). As argued above, these expectations formed by economic agents may, in effect, determine future inflation through their impact on today's price setting or wage negotiations. Expectation measures from surveys are found to indeed outperform univariate model forecasts (Clements, 2015), or forecasts building on real and financial variables (Ang et al., 2007).

Yet, there may be additional room to improve upon the forecast performance of survey measures, in particular since the predictive ability of survey forecasts has started to deteriorate around 2002 (Trehan, 2015). While surveys are timely available and can measure expectations of either broad or specific groups of agents (e.g., consumers, financial market participants, business managers), it is not obvious that these expectations are updated regularly and reflect all newest information all the time. Mankiw and Reis (2001) and Sims (2003) argue that, if economic agents find it too costly to constantly gather and process all new information, they occasionally do not update their expectations. Due to this rational

inattention, expectations are sticky. Supporting this proposition, Carroll (2003) presents microfoundations for these sticky expectations, proposing that households mostly lack the skills to form rational expectations and do not process economic data themselves. Instead, they form their expectations based on the views of professional forecasters as reported by media, which in turn may be rational. However, as households do not pay attention to macroeconomic news all the time, expectations adapt only gradually. Along those lines, Carroll (2003) empirically shows that there is a gap between inflation expectations of experts and households, which diminishes the more the media reports on inflation. Yet, more news reporting may not only provide better information to households, but also bias their expectations, as shown by Lamla and Maag (2012) and Lamla and Lein (2014). Here, the authors consider how the *tone*, or *sentiment*, of media coverage on inflation influences households' expectation formation, and show that these expectations may, in turn, be biased if reports exaggerate certain developments or are badly toned. In either case, it is important to note that news reporting likely influences households' expectations, which, in turn, affects their decision-making such as price and wage negotiations or saving decisions, and, hence, future inflation.

Therefore, employing media reports for forecasting purposes is particularly appealing. In the best of all cases, survey results about households' expectations can be anticipated and traditional forecast models can be improved upon. For this purpose, news can be analyzed by humans following a code book to rate them according to their content. However, information content in the media can also be quantified employing automated models. In contrast to human-rated media indicators, these automated models present an inexpensive, replicable, and objective alternative. One example for such a simple but very effective automated approach is the *R*-word index counting the frequency of the word "recession" in the Wall Street Journal and the New York Times (Doms and Morin, 2004). Such content-based indicators have previously been used for forecasting macroeconomic outcomes such as output growth and recession events, however, the evidence for their predictive value is mixed at best (Bandholz and Funke, 2003; Grossarth-Maticek and Mayr, 2008; Iselin and Siliverstovs, 2013).

Recently however, more refined computer linguistic tools attempting to detect the *tone*, or *sentiment*, of news reports have been developed and applied to economic questions (e.g. Uhl, 2011). As argued above, it may be the tone and not only the content of news reporting that influences agents' expectations decisively. Therefore, evaluating the content of news by a mere word count of the frequency of topics covered, may ignore an important dimension of media coverage. Hence, using such novel indicators for media sentiment may provide useful information for predicting key economic variables. Yet, despite their potential, such media sentiment indicators have been employed for forecasting only few macroeconomic variables

(Ulbricht et al., 2016, Uhl, 2011, 2014) outside of financial markets (e.g. Tetlock, 2007) and have, to the best of our knowledge, not been used to forecast inflation. This paper aims to fill this gap by comparing the usefulness of media *sentiment* indicators – both automatic ones and indicators based on human expert ratings – to that of a comprehensive set of standard leading indicators.

The comparison is conducted in form of an experiment of forecasting the German inflation rate in real-time. First, we construct an automatic sentiment indicator building on general economic or inflation news in the national weekly German newspaper *Die Zeit*. To the best of our knowledge, this is the first time this indicator is employed in the literature to make inflation forecasts. Second, we compare its forecasting performance to that of a simple word-count index and indicators obtained from human analysts’ ratings of reporting in general TV and print media. Finally, we compare the predictive ability of these media (sentiment) indicators to the forecast performance of standard real macroeconomic and financial variables, as well as of indicators based on the standard business or consumer surveys for Germany.

We find that our sophisticated automatic sentiment indicator based on economic reporting as well as sentiment ratings by human experts of business cycle news reports provide considerable gains to autoregressive forecasts of inflation for up to one year and are consistently among the best predictors of the entire set of predictors across all forecast horizons. Furthermore, for each horizon, there is at least one news sentiment indicator that outperforms the best performing survey indicator.

The remainder of this paper is structured as follows. The next section introduces the media sentiment indicators and alternative predictors. Section 3 describes the forecast experiment. Section 4 presents the results, and the last section concludes.

2 Media sentiment indicators and inflation predictors

As outlined, we are not only interested in the amount of media reporting on inflation in Germany, but also in the *sentiment* of news reporting as this may affect expectation formation by households. To measure this, we first make use of an automatic sentiment index extracted using a lexicon-based algorithm.¹ This index has the advantage that it can be readily applied to any available database of newspaper articles and other reports or transcripts, and therefore provides a low-cost and objective alternative predictor for inflation. However, the accuracy

¹This indicator is related to the broad natural language processing (NLP) literature, which is becoming increasingly popular for textual analysis. In addition to the sentiment indicators employed by Tetlock (2007), Ulbricht et al. (2016), and Uhl (2011, 2014), several more developed indicators are available to assess not sentiment but context of (newspaper) text (see for example Gentzkow and Shapiro, 2010). We only focus on sentiment analysis here as we restrict our analysis to news items on inflation only. Further, for a full context analysis, it is necessary to develop a lexicon database linked to context, which is not readily available for the German language.

of such a computer-based algorithm cannot be assessed in a straightforward way and may be prone to error. Therefore, in Section 2.2 we describe expert-rated sentiment indices, which we use as an alternative for inflation forecasting. Finally, this section briefly describes the set of rival predictors as employed in the previous literature.

2.1 Automatic sentiment indices

For the automatic sentiment indices, we downloaded 15,447 articles of the economics section of the weekly newspaper *Die Zeit* published between 2000:M1 and 2014:M3 from its publicly available online archive. Among all weekly newspapers, *Die Zeit* is the most popular one with an average readership of 2.34 million purchases per issue in 2016 (Institut für Demoskopie Allensbach, 2016).² Of the articles we downloaded, we identified 958 articles concerned with inflation, that is, articles that contain the word “*inflation*” or its synonyms. However, we do not distinguish between articles referring exclusively to inflation in Germany, or to inflation both in Germany and in other countries or in the EU. Since our sample starts after the gradual introduction of the Euro beginning in 1999 and the implementation of a monetary union with a common monetary policy, news about inflation in the Euro area are also likely to influence the decision making process of German households and firms.³

We then extract the sentiment of these news articles using a lexicon-based algorithm in the line of Taboada et al. (2011) using the *qdap* package (see Rinker, 2013) in the open software project *R* (see R Core Team, 2013). To each article i in month t a sentiment value, $ArticleValue_{i,t}$, is attributed, which is defined as:

$$ArticleValue_{i,t} = \frac{\sum_{j=1}^{N_{i,t}} (WordValue_{j,i,t} \times n_{j,i,t})}{N_{i,t}}, \quad (1)$$

where $WordValue_{j,i,t}$ is the sentiment value attributed to each word j of article i , (with frequency $n_{j,i,t}$) where $j = 1, \dots, N_{i,t}$ and $N_{i,t}$ is the number of unique words of the respective article. The value is based on the lexicon *SentimentWortschatz* (see Remus et al., 2010), which is a publicly available German-language resource for sentiment analysis. It contains positive and negative sentiment bearing words weighted within the interval of $[-1, 1]$ plus their part of speech tag (adjective, adverbs, nouns, and verbs), and if applicable, their inflections. As an example, the word “*harmonisch*” (harmonious) and its inflections (“*harmonische*”, . . . , “*harmonischst*”) are assigned the positive value +0.5243. The German word

²The main competitors are *Welt am Sonntag* and *Frankfurter Allgemeine Sonntagszeitung* with on average 1.5 and 1.1 million readers (Institut für Demoskopie Allensbach, 2016). The online archives of these newspapers are not available to us. To compare the relevance of *Die Zeit* to other news sources, see section 2.2.

³Specifically, news about inflation in the Euro area likely triggers a response by the ECB, and further affects real effective exchange rates between Germany and its trading partners countries. Hence, German households and firms have an incentive to incorporate these news to form their expectations about inflation in Germany.

for crisis (“*Krise*”) on the other hand, is associated with a negative value of $-.3631$.⁴ The version employed here (SentimentWortschatz v1.8b) contains 1,650 negative and 1,818 positive baseforms, which, together with their inflections, sum up to 16,406 positive and 16,328 negative vocables.

SentimentWortschatz accounts for the fact that some words have different parts of speech and can, thus, have different word values. However, tagging of parts of speech automatically is very time consuming and may still be inaccurate. As an alternative, we take advantage of the fact that, in German, nouns always begin with upper-case letters to easily identify them in a text. In contrast to sentiment analysis for English, we do not convert upper-case letters to lower-case letters. This leads to a small error, as upper-case at the beginning of a sentence will not be correctly identified. However, this leaves only a small fraction of about 100 (lower-case) words to which no unique value can be attributed. Thus, they are dropped from the lexicon prior to the analysis. Then, the value found in the lexicon is refined in the following way. If negators like “*nicht*” (no/not) are found in the proximity (four words before and two words after the word j) and if the number of negators is odd the polarity of a word is inverted. For example, “*good*” is a word with a positive *WordValue* while “*not good*” has a negative *WordValue*.

By this procedure, two indices are computed as monthly averages of $ArticleValue_{i,t}$. First, we obtain the index *Zeit*, based on all articles downloaded from the economic section for the respective month. Second, we refine the search and obtain the index *Zeit Inflation*, based only on economic articles concerned with inflation.

Finally, we construct a third automatic indicator that does not measure the sentiment of news reports but only the amount of media reporting on inflation in line with the propositions outlined by Carroll (2003). For this, we construct a simple word-count index (*Genios I-word-count*) of “Inflation” and its synonyms⁵ using the online news archive Genios covering all 181 major German newspapers. For this, we obtain all occurrences of “inflation” in all articles across our sample and aggregate them to the monthly frequency.

2.2 Human sentiment indices

To compare the forecasting performance of our automatic sentiment indicators to indices based on expert ratings instead of computer algorithms, we also investigate the predictive ability of sentiment indices obtained from Media Tenor International (MTI). MTI employs

⁴Further documentation on the determination of the weights and other examples can be found in Remus et al. (2010).

⁵Namely: price increase (Preisanstieg, steigende Preise, Preissteigerung), inflation (Teuerung, Teuerungsrate, Verteuerung, Inflationsrate, inflationär), inflation risk (Inflationsrisiko, Inflationsangst), combating inflation (Inflationsbekämpfung), inflationary pressure (Inflationsdruck), price level (Preisniveau), price stability (Preisstabilität), purchase power (Kaufkraft), loss in purchase power (Kaufkraftverlust), and Stagflation.

human analysts to rate news in TV and print media, where the following subset of indicators is available according to the news' time reference (*present or future*), their polarity (positive, negative, and neutral) and to their economic category (*monetary indicators, inflation, taxation, the business cycle, labor markets, and the government budget*). The media sources included are the five most watched TV news programs (ARD Tagesschau, ARD Tagesthemmen, ZDF heute, ZDF heute journal, RTL Aktuell), two popular weekly magazines (Spiegel, Focus), and the most-sold daily newspaper (Bild).⁶

Hence, we use a total of 51,304 economic news items as the subset from all news items from 2000:M1 to 2014:M3 to compute 15 indices in total. The first nine of these 15 indices are based on all news items (*All*), items referring to the *Present* or *Future*, items related to the exchange rate, monetary policy, and interest rates (*Monetary*), to *Inflation*, *Taxation*, the business cycle (*Cycle*), the labor market (*Labor*), or to the government budget (*Budget*). Each news item published at time τ in month t is then assigned an aggregate positive (+), negative (−) or neutral (0) rating. The final indices B_t are then computed as the differences between the percentage share of the positively rated and the share of negatively rated news items within a category and a month:

$$B_t = \frac{A_{\tau,t}^+ - A_{\tau,t}^-}{A_{\tau,t}^+ + A_{\tau,t}^- + A_{\tau,t}^0} \times 100, \quad (2)$$

where $A_{\tau,t}^\bullet$ is the number of positive (+), negative (−) or neutral (0) media reports about events happening at time τ published in month t . The index varies between -100 (all reports are negatively rated) and 100 (all reports are positively rated).

The second subset of the 15 indicators is the same as employed by Lamla and Maag (2012) and is based on subcategories of *Inflation*. As there are comparatively few items (1195) available, they are mere word-counts, therefore resembling the *Genios I-word-count*. *Volume* is the number of all items relating to inflation, *Falling*, *Rising*, and *Notrend* comprise the items referring to falling and rising inflation, and to items where no trend is indicated, respectively. *Goodrising*, *Badrising*, and *Otherrising* refer to items about rising inflation that have been rated positively, negatively, or that are neutral. Table 1 presents descriptive statistics of automatic and human sentiment indicators.

[Table 1 about here.]

⁶The ARD Tagesschau is a 15-minute, 8pm daily newscast with 8.96 million viewers on average in 2014. ZDF heute and RTL aktuell are the competitors airing daily at 7pm or 6.45pm with 3.8 and 3.3 million viewers on average, respectively. ZDF heute journal and ARD Tagesthemmen are 30 minute news magazines broadcasted later in the evening at 9.45pm and 10.30pm with 3.9 and 2.5 million viewers, respectively (Media Perspektiven, 2016). Der Spiegel and Focus are national weekly magazines with an average readership of 5.9 and 3.6 million per issue in 2016. Finally, Bild is the largest daily (tabloid) newspaper with an average readership of 8.5 million in 2016 (Institut für Demoskopie Allensbach, 2016).

2.3 Non-media indicators

As a reference to compare the predictive ability of the media indicators to other commonly used predictors, we include the comprehensive set of early indicators of Drechsel and Scheufele (2012). This includes real economic variables comprising measures of output, the labor market and prices, as well as financial variables and a large set of survey data. Since the major advantage of media data is that they are unrevised and timely, we use a real-time version of the data. This means that we use the data only as they would have been available at the moment of time when forecasts are made. For an overview of all variables, their publication lag and the transformations see Table 2.

[Table 2 about here.]

3 Real-time forecast experiment and evaluation

In this section, we describe in detail our model specifications and estimation procedure for our real-time out-of-sample forecasting exercise. In a second step, the forecast evaluation is explained.

3.1 Model specification

We replicate the forecast experiment of Drechsel and Scheufele (2012). Therefore, we first assess the forecast performance of a univariate autoregressive model (AR) of inflation and then augment this AR by one individual media or non-media indicator from Table 2. With 18 media and 77 non-media indicators plus all their transformations indicated in Table 2, we therefore arrive at $N = 179$ unique bivariate forecast models (plus the AR, denoted by $m = 0$).⁷ Each individual model $m = 0, 1, \dots, N$ is defined as:

$$y_{t+h} = \alpha + \sum_{p=l_0}^{P+l_0} \beta_p \tilde{y}_{t-p} + \sum_{q=l_i}^{Q+l_i} \gamma_{i,q} x_{i,t-q} + u_{t+h} \quad (3)$$

where α , β_p , and $\gamma_{i,q}$ are parameters to be estimated, y_t are h -period inflation rates of the consumer price index (CPI) in time period t ($t = 1, \dots, T$) and \tilde{y}_t are monthly inflation rates. This definition of the autoregressive process with monthly inflation rates as predictors is in line with the literature of Rossi and Sekhposyan (2010). Our benchmark model $m = 0$ is the AR model for which all $\gamma_{i,q}$ are zero. $x_{i,t}$ is the time t exogenous indicator variable i ($i = 1, \dots, N$), and u_t is a disturbance term. The different number of minimum lags l_j for

⁷For example, we use the level of the money market rate for one bivariate prediction model, and the difference of the money market rate for another prediction model, as specified in Table 2.

each regressor, with $j = m = 0, \dots, N$, reflects the varying degree of data availability. For example, if a predictor, such as vacancies, is available with a lag of one month, then $l_j = 1$. The maximum lag orders are identical for all regressors and set to 12. However, the lag orders employed in the forecasting exercise are then determined for each model individually using the Bayesian Information Criterion. This is carried out sequentially. First, P is determined for the autoregressive forecast at each forecast margin $\tau = T_E, T_E + 1, \dots, T - h$, where T_E is the first forecast margin at the end of the training sample. The lag order Q for each model $m = 1, \dots, N$ is then determined conditional on the fixed lag length P . As such, the lag order of the autoregressive component is unchanged.

The whole sample stretches from 2000:M1 to 2014:M3. We perform an out-of-sample forecast experiment, for which we continuously re-estimate the autoregressive model of inflation and 179 candidate predictor models using data for the last 60 months (rolling window approach). For each of the forecast horizons (one, three, six, and twelve months) we estimate separate models. The first forecast is made from the end of 2005:M12 (defined as T_E), using non-revised data that have been available until then. For the next iteration, we then employ data as of 2006:M1, and so on.

3.2 Forecast evaluation

We evaluate our large set of forecast models according to their forecast accuracy. For this, we compare their relative root mean square forecast errors (RMSFE) and their overall forecast ranks. We then assess the significance of differences of forecasts by standard tests and compute the model confidence set of Hansen et al. (2011).

The forecast errors are computed using time series of inflation as of 2015:M4. For each model m and each horizon h , the RMSFE can be computed as:

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

where $\hat{y}_{m,t}$ is the forecast made by model m for period t . Then, Theil's U is constructed to compare the forecast performance of model m for horizon h to that of the benchmark AR-model, as:

$$Theil's U_{m,h} = \frac{RMSFE_{m,h}}{RMSFE_{0,h}}, \quad m = 1, \dots, N \quad (5)$$

where model $m = 0$ is the AR model without additional predictors.

In order to test the hypothesis of a candidate model performing equally good as the AR we employ the test proposed by Clark and West (2007) for nested models as applicable

here. This test compares the mean square forecast errors of two nested model at a time and is applicable to real-time data. This allows testing the accuracy of a candidate model to a predetermined benchmark, but, hence, cannot indicate the best models overall as it is sensitive to the choice of the benchmark. Therefore, in order to make the comparison free of the choice of the benchmark, we lastly identify the five percent model confidence set (MCS, Hansen et al. (2011)). The MCS iteratively eliminates models with significantly poorer predictive power from the large set of $N + 1$ models until a subset of models is identified, that cannot be distinguished as statistically inferior among each other at the given confidence level.

4 Results

Table 3 shows the results based on the forecasts for all periods from 2005:M1 to 2013:M4 for the one, three, six, and twelve months forecast horizons. The number of models out of the total of 180 models that are included in the MCS for each horizon is given in the second row. The RMSFE of the AR forecast is given in the third row. The results for the media models are presented in the lower part of the table. It shows Theil’s U , such that a value below one indicates a model with a lower RMSFE than the AR process, and vice versa. Values smaller than one are printed in bold and one or two stars indicate that the hypothesis of equal forecast performance of the AR and the respective model is rejected at the five or one percent significance level. The forecast rank is given in parentheses. For the ease of presentation, only the results of media models included in the model confidence set are reported. As a reference, the results for the best non-media models for each forecast horizon are shown in the upper part of the table.

[Table 3 about here.]

Depending on the forecast horizons, about 50 – 60% of all 180 models can be excluded from the MCS (row 2). For the one month forecasts the MTI index based on news items related to the business cycle (*MTI Cycle*) ranks first overall having a relative RMSFE to the AR of 0.96 and the null hypothesis of equal predictive performance is rejected at the five percent level. For the three and six months horizon forecasts, the automatic sentiment indicator based on all articles of the economic section in “*Die Zeit*” is the overall best model with an improvement to the AR forecast of eleven and six percent, respectively. However, the null of equal predictive forecast accuracy is rejected at the five percent level only in case of the six months horizon forecasts. For the twelve months horizon forecasts, the best media model is again *MTI Cycle*, ranking second overall. Its relative RMSFE of 0.83 is very close

to the best non-media model, the unemployment rate, and the null of equal forecast accuracy compared to the AR model is rejected at the five percent level. From this, it follows that, for all four horizons, there is at least one media sentiment indicator obtained from economic or business cycle news that outperforms other, standard survey indicators.

In contrast to the sentiment indicators obtained from general economic articles or news reporting of the business cycle, the automatic sentiment indicator based on *Zeit* articles concerned only with inflation performs rather poorly and can be excluded from the MCS for most forecast horizons. Similarly, also the *Genios I-word-count* index for a large set of newspapers (37th at best for the six months horizon) and the MTI count indices on news items related to inflation (bottom six rows, 39th at best for *MTI Notrend* and the six month horizon) are providing comparatively inaccurate forecasts.

Finally, mixed evidence can be found for the MTI *Inflation* index, which predicts well for the short horizons of one and three months, ranking second and eleventh, but performs relatively poorly for longer horizons of one year, ranking 49th. Furthermore, the MTI sentiment indicator on items referring to the *future* performs better as an inflation predictor than the MTI *Present* indicator, yet only outperforms the AR benchmark for the three month horizon. MTI sentiment indicators on taxation, the labor market, or fiscal news can in general not be excluded from the MCS but are never among the best 30 predictors with the exception of MTI *Taxation* for the one year horizon.

Nonetheless, there is strong evidence that sentiment indicators attempting to capture the tone of media reporting on business cycle or general economic news can provide valuable information for real-time inflation prediction. These indicators not only outperform simple AR forecasts but are also competitive, in particular, against forecasts obtained from other survey information. This is especially relevant in view of the recent findings of Ang et al. (2007) who demonstrate that standard survey-based forecasts for inflation are superior to forecasts obtained from real or financial variables.

Similarly, we thus also find that our sentiment indicators *Zeit* and *MTI Cycle* frequently outperform financial indicators. Since financial indicators should capture the forecasts made by highly informed financial analysts and market participants, this result of better predictions by our sentiment indicators is surprising, but in line with e.g. Stock and Watson (1999, 2003). Two reasons may account for this: First, financial variables such as interest rates and spreads do not only capture inflation expectations, but also expectations about the future real rate and risk premia. In particular, the expectations hypothesis of the term structure of interest rates relates the forward rate to market expectations of future inflation and the future real rate only under the assumption of risk-neutrality (Stock and Watson, 2003). With time-varying risk premia during the financial crisis, this link likely does not hold, which impairs the predictive content of the forward rate. Second, the superior predictive content of survey

and our sentiment indicators may point to a causal link between the formation of household expectations based on news sentiment and future inflation as argued by Carroll (2003). If financial analysts underestimate the importance of news on consumers, and particularly of the *sentiment* of news coverage, they may weigh available information incorrectly, which could explain the superior predictive ability of our indicators.

5 Conclusion

This paper analyzes the usefulness of automatic and expert-rated media sentiment indices for the prediction of German inflation and compares them to a comprehensive set of leading indicators identified in the previous literature. In particular, it evaluates the real-time predictive ability of 180 indicators, including real macroeconomic and financial variables, consumer and business survey-based indicators, and media sentiment indices. The latter are based either on the human made ratings or on the automatic algorithm developed by the authors of the paper.

We demonstrate that the tone of news reports on the business cycle rated by humans and that of articles from the economic section of a German quality weekly newspaper *Die Zeit* extracted by a sophisticated automatic computer algorithm are among the best predictors for inflation forecasts up to one year. Thus, we extend the previous literature by a novel set of valuable leading indicators that are available in real time. While traditional predictors remain helpful, it can be recommended to use media sentiment indicators as a supplementary real-time information in inflation forecasts.

References

- Ang, A., G. Bekaert, and M. Wei (2007). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics* 54(4), 1163 – 1212.
- Bandholz, H. and M. Funke (2003). In search of leading indicators of economic activity in germany. *Journal of Forecasting* 22(4), 277–297.
- Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. *The Quarterly Journal of Economics* 118(1), pp. 269–298.
- Clark, T. E. and K. D. West (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138(1), 291–311.
- Clements, M. P. (2015). Are professional macroeconomic forecasters able to do better than forecasting trends? *Journal of Money, Credit and Banking* 47(2-3), 349–382.
- 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.).
- Drechsel, K. and R. Scheufele (2012). The performance of short-term forecasts of the German economy before and during the 2008/2009 recession. *International Journal of Forecasting* 28(2), 428–445.
- Gentzkow, M. and J. M. Shapiro (2010). What drives media slant? evidence from u.s. daily newspapers. *Econometrica* 78(1), 35–71.
- Gordon, R. J. (1981). Inflation, flexible exchange rates, and the natural rate of unemployment. Working Paper 708, National Bureau of Economic Research.
- Grossarth-Maticek, J. and J. Mayr (2008). Medienberichte als Konjunkturindikator. *Ifo-Schnelldienst* 61, 17–29.
- Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica* 79(2), 453–497.
- Institut für Demoskopie Allensbach (2016). Allensbacher Markt- und Werbeträgeranalyse 2016: Printmedien. published online: <http://www.ifd-allensbach.de/awa/medien/printmedien.html>. Last access: 23 February 2017.
- Iselin, D. and B. Siliverstovs (2013). The R-word index for Switzerland. *Applied Economics Letters* 20(11), 1032–1035.
- Lamla, M. J. and S. M. Lein (2014). The role of media for consumers’ inflation expectation formation. *Journal of Economic Behavior & Organization*.

- Lamla, M. J. and T. Maag (2012). The role of media for inflation forecast disagreement of households and professional forecasters. *Journal of Money, Credit and Banking* 2012(7), 1325–1350.
- Mankiw, N. G. and R. Reis (2001). Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve. Technical report, National Bureau of Economic Research.
- Media Perspektiven (2016). Basisdaten 2015. Technical report, Vertrieb Media Perspektiven.
- Mehra, Y. P. (2002). Survey measures of expected inflation: revisiting the issues of predictive content and rationality. *Economic Quarterly* (Sum), 17–36.
- Phillips, A. W. (1958). The relation between unemployment and the rate of change of money wage rates in the united kingdom, 1861-1957. *Economica* 25(100), 283–299.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Remus, R., U. Quasthoff, and G. Heyer (2010). SentiWS – A Publicly Available German-language Resource for Sentiment Analysis. In *LREC*.
- Rossi, B. and T. Sekhposyan (2010). Have economic models’ forecasting performance for US output growth and inflation changed over time, and when? *International Journal of Forecasting* 26(4), 808–835.
- Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics* 50(3), 665–690.
- Stock, J. H. and M. W. Watson (1999). Forecasting inflation. *Journal of Monetary Economics* 44(2), 293–335.
- Stock, J. H. and M. W. Watson (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature* 41(3), 788–829.
- Stock, J. H. and M. W. Watson (2007). Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking* 39, 3–33.
- Taboada, M., J. Brooke, M. Tofiloski, K. Voll, and M. Stede (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics* 37(2), 267–307.
- Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance* 62(3), 1139–1168.
- Thomas, L. B. (1999). Survey measures of expected U.S. inflation. *Journal of Economic Perspectives* 13(4), 125–144.

- Trehan, B. (2015). Survey measures of expected inflation and the inflation process. *Journal of Money, Credit and Banking* 47(1), 207–222.
- Uhl, M. W. (2011). Explaining US consumer behavior with news sentiment. *ACM Transactions on Management Information Systems (TMIS)* 2(2), 9.
- Uhl, M. W. (2014). Reuters sentiment and stock returns. *Journal of Behavioral Finance* 15(4), 287–298.
- Ulbricht, D., K. A. Kholodilin, and T. Thomas (2016). Do Media Data Help to Predict German Industrial Production? *Journal of Forecasting*. first published: 3 November 2016.

Appendix

Table 1: Descriptives: Media data

Statistic	Mean	St. Dev.	Min	Max
Automatic media indicators				
Zeit	-0.003	0.001	-0.005	-0.001
Zeit inflation	-0.0004	0.0002	-0.001	0.00002
Genios I-Word count	0.00467	0.00105	0.00299	0.00991
Human media indicators				
MTI Notrend	3.144	2.596	0	16
MTI Rising	3.169	4.335	0	24
MTI Falling	0.869	1.563	0	7
MTI Volume	7.181	5.437	0	28
MTI Badrising	2.300	3.736	0	21
MTI Otherrising	0.831	1.280	0	8
MTI All	-20.851	19.075	-66.815	27.660
MTI Present	-24.916	22.272	-72.811	26.131
MTI Future	-14.413	18.940	-61.290	57.143
MTI Monetary	-22.327	39.257	-100.000	83.333
MTI Inflation	-22.060	46.810	-100.000	100.000
MTI Taxation	-26.671	15.614	-80.000	7.759
MTI Cycle	-1.570	47.908	-84.211	100.000
MTI Labor	-23.349	22.005	-74.054	40.000
MTI Budget	-29.068	33.478	-89.474	58.621

Table 2: Data: definitions, transformations, and sources

Block	Name	Label	L	D	Dln	D2ln	Lag	Source
Dependent variable	Inflation	CPI	-	-	1	-	-	Buba RTDB
Financial	Money market rate (monthly average)	IS-M	1	1	-	-	0	Buba
	Discount rate/short term repo rate (monthly average)	IS-D	1	1	-	-	0	Buba
	3m-money market rate (monthly average)	IS-3M	1	1	-	-	0	Buba
	Yields on debt securities outstanding (maturity 3-5 years)	IL-3	1	1	-	-	0	Buba
	Yields on debt securities outstanding (maturity 5-8 years)	IL-5	1	1	-	-	0	Buba
	Long term government bond yield-9-10 years	IL-10	1	1	-	-	0	Buba
	Term spread (10 years - money market rate)	SPR-10Y-M	1	-	-	-	0	Buba
	Term spread (10 years - discount rate)	SPR-10Y-D	1	-	-	-	0	Buba
	Term spread (10 years - 3 months-money market rate)	SPR-10Y-3M	1	-	-	-	0	Buba
	Term spread (discount rate - money market rate)	SPR-1D-M	1	-	-	-	0	Buba
	Corporate bond-government bonds	SPR-C-G	1	-	-	-	0	Buba
	Nominal effective exchange rate	EX	-	-	1	-	1	Buba
	Real effective exchange rate	EXR	-	-	1	-	1	Buba
	DAX	DAX	-	-	1	-	0	Buba
	DAX volatility new	VOLA1	1	1	-	-	0	Buba
	DAX volatility old	VOLA2	1	1	-	-	0	Buba
	Hwwa index of world market prices of raw materials	HWWA	-	-	1	1	1	Datastream
	HWWA index, real	HWWAR	-	-	1	1	-	Datastream
	HWWA index ,energy	HWWA-E	-	-	1	1	1	Buba
	HWWA index, energy real	HWWA-ER	-	-	1	1	-	Buba
	HWWA index ,excl. energy	HWWA-EX	-	-	1	1	1	Buba
	HWWA index, excl. energy real	HWWA-EXR	-	-	1	1	-	Buba
	Oil prices (euros per barrel)	OIL	-	-	1	1	0	ECB
	Oil prices (euros per barrel), real	OILR	-	-	1	1	-	ECB
Surveys	Ifo index climate	IFO-C	1	1	-	-	0	ifo
	Ifo expectations climate	IFO-EXP	1	1	-	-	0	ifo
	Ifo index manufacturing	IFOM-C	1	1	-	-	0	ifo
	Ifo expectations manufacturing	IFOM-EXP	1	1	-	-	0	ifo
	Ifo index capital goods	IFOMI-C	1	1	-	-	0	ifo
	Ifo expecations capital goods	IFOMI-EXP	1	1	-	-	0	ifo
	Ifo index intermediate goods	IFOMV-C	1	1	-	-	0	ifo
	Ifo expectations intermediate goods	IFOMV-EXP	1	1	-	-	0	ifo
	Ifo index wholesale	IFOWH-C	1	1	-	-	0	ifo

Table 2: Data: definitions, transformations, and sources (continued)

Block	Name	Label	L	D	Dln	D2ln	Lag	Source
	Ifo expectations wholesale	IFOWH-EXP	1	1	-	-	0	ifo
	GfK consumer climate survey - business cycle expectations	GfK-EXP	1	1	-	-	0	Datastream
	ZEW economic sentiment	ZEW	1	1	-	-	0	Datastream
	Assessment of order-book levels	ECBS2	1	1	-	-	0	EC
	Assessment of export order-book levels	ECBS3	1	1	-	-	0	EC
	Assessment of stocks of finished products	ECBS4	1	1	-	-	0	EC
	Production expectations for the months ahead	ECBS5	1	1	-	-	0	EC
	Selling price expectations for the months ahead	ECBS6	1	1	-	-	0	EC
	Employment expectations for the months ahead	ECBS7	1	1	-	-	0	EC
	Industrial confidence indicator (40%)	ESI-INDU	1	1	-	-	0	EC
	Services confidence indicator (30%)	ESI-SERV	1	1	-	-	0	EC
	Consumer confidence indicator (20%)	ESI-C	1	1	-	-	0	EC
	Retail trade confidence indicator (5%)	ESI-TRADE	1	1	-	-	0	EC
	Construction confidence indicator (5%)	ESI-CTR	1	1	-	-	0	EC
	Economic sentiment indicator (average)	ESI	1	1	-	-	0	EC
	Confidence Indicator (Q2 + Q4 - Q7 + Q11) / 4	ECCS99	1	1	-	-	0	EC
	Financial situation over last 12 months	ECCS1	1	1	-	-	0	EC
	Financial situation over next 12 months	ECCS2	1	1	-	-	0	EC
	General economic situation over last 12 months	ECCS3	1	1	-	-	0	EC
	General economic situation over next 12 months	ECCS4	1	1	-	-	0	EC
	Price trends over last 12 months	ECCS5	1	1	-	-	0	EC
	Price trends over next 12 months	ECCS6	1	1	-	-	0	EC
	Unemployment expectations over next 12 months	ECCS7	1	1	-	-	0	EC
	Major purchases at present	ECCS8	1	1	-	-	0	EC
	Major purchases over next 12 months	ECCS9	1	1	-	-	0	EC
	Savings at present	ECCS10	1	1	-	-	0	EC
	Savings over next 12 months	ECCS11	1	1	-	-	0	EC
	Statement on financial situation of household	ECCS12	1	1	-	-	0	EC
Prices and wages	Consumer price index	CPI	-	-	1	1	-	Buba RTDB
	Core consumer price index	CPI-EX	-	-	1	1	-	Buba RTDB
	Negotiated wage and salary level	TARIF	-	-	1	1	-	Buba RTDB
Real economy	Industrial production	IP	-	-	1	-	-	Buba RTDB
	Intermediate goods production	IP-VORL	-	-	1	-	-	Buba RTDB
	Manufacturing orders - consumer goods	ORD-C	-	-	1	-	-	Buba RTDB
	Manufacturing orders - capital goods	ORD-I	-	-	1	-	-	Buba RTDB

Table 2: Data: definitions, transformations, and sources (continued)

Block	Name	Label	L	D	Dln	D2ln	Lag	Source
Composite indicators	Employed persons (work-place concept)	EW	-	-	1	-	-	Buba RTDB
	1+unemployment (% civilian labour)	ALQ	-	1	-	-	1	Datastream
	Vacancies	VAC	-	-	1	-	1	Datastream
	Capacity utilisation	CAPA	1	1	-	-	0	Datastream
	Hours worked	WHOUR	1	1	-	-	0	Buba RTDB
	FAZ indicator	FAZ	-	-	1	-	1	Datastream
	Composite leading indicator (amplitude restored)	OECDL1	1	1	-	-	2	OECD
	Composite leading indicator (trend restored)	OECDL2	-	1	-	-	2	OECD
	Composite leading indicator (normalized)	OECDL3	1	1	-	-	2	OECD
Automatic media indicators	Sentiment, all articles	Zeit	1	1	-	-	0	Zeit
	Sentiment, inflation articles	Zeit inflation	1	1	-	-	0	Zeit
	Number of articles mentioning inflation	Genios	1	1	-	-	0	Genios
		I-word						
Human media indicators	Sentiment, all observations	All	1	1	-	-	0	MTI
	Sentiment, all observations related to the future	Future	1	1	-	-	0	MTI
	Sentiment, all observations related to the present	Present	1	1	-	-	0	MTI
	Sentiment, all observations related to monetary issues	Monetary	1	1	-	-	0	MTI
	Sentiment, all observations related to Inflation	Inflation	1	1	-	-	0	MTI
	Sentiment, all observations related to taxation	Taxation	1	1	-	-	0	MTI
	Sentiment, all observations related to the business cycle	Cycle	1	1	-	-	0	MTI
	Sentiment, all observations related to the labor market	Labor	1	1	-	-	0	MTI
	Sentiment, all observations related to government budget	Budget	1	1	-	-	0	MTI
	Number of observations, inflation	Volume	1	1	-	-	0	MTI
	Number of observations, falling inflation	Falling	1	1	-	-	0	MTI
	Number of observations, rising inflation	Rising	1	1	-	-	0	MTI
	Number of observations, inflation without trend	Notrend	1	1	-	-	0	MTI
	Number of observations, rising inflation, positive	Goodrising	1	1	-	-	0	MTI
	Number of observations, rising inflation, negative	Badrising	1	1	-	-	0	MTI
Number of observations, rising inflation, neutral	Otherrising	1	1	-	-	0	MTI	

Note: The different transformations of the raw data are either level (L), differences (D), differences of natural logarithms (Dln) or second differences of natural logarithms (D2ln). The publication lag (Lag) ranges from 0 to 3 months. The sources are Datastream, Deutsche Bundesbank (Buba), Deutsche Bundesbank Realtime Database (Buba RTDB), European Commission (EC), European Central Bank (ECB), Ifo institute for

Table 3: Theil's U (and overall rank in parentheses) for all models included in MCS, all horizons

Forecast horizon	1	3	6	12
Number of models in MCS	87	77	92	68
<i>AR-benchmark model</i>				
RMSFE	1.45 (13)	0.86 (13)	0.62 (6)	0.68 (14)
<i>Best non-media models</i>				
Survey on financial situation (DECCS1)	0.98** (3)			
Yields on debt securities (IL.5)		0.94** (3)		
Hours worked (DWHOUR)			0.97* (2)	
Unemployment rate (DALQ)				0.82** (1)
<i>Media models</i>				
Zeit	1.00 (17)	0.89 (1)	0.94* (1)	0.92** (5)
Zeit inflation		1.29 (125)		
Genios I-word-count		1.09 (65)	1.13 (37)	1.30 (88)
MTI All	1.04 (81)	1.31 (131)		1.44 (106)
MTI Present				
MTI Future	1.03* (70)	0.97** (6)	1.19 (54)	1.06** (25)
MTI Monetary	1.01** (28)	0.97* (5)	1.22 (62)	1.13 (51)
MTI Inflation	0.97 (2)	1.00 (11)	1.07 (20)	1.12 (49)
MTI Taxation	1.02** (50)	1.13** (82)	1.20* (55)	0.93* (7)
MTI Cycle	0.96* (1)	0.99 (8)	1.17 (50)	0.83* (2)
MTI Labor	1.03** (72)	1.14* (89)	1.31 (85)	1.10** (39)
MTI Budget	1.03 (57)	1.16* (92)		1.89 (143)
MTI Notrend	1.02 (55)	1.08 (60)	1.14 (39)	1.13 (52)
MTI Rising				
MTI Falling	1.04 (75)	1.11 (77)		1.19 (67)
MTI Volume				
MTI Badrising				1.54 (118)
MTI Otherrising	1.04 (84)			1.23 (74)

The second row shows the number of models out of the total of 180 models that are included in the model confidence set (MCS). The third row shows the RMSFE of the AR forecast and its rank in parenthesis. Rows four to seven show the Theil's U and overall rank of the best non-media indicators. The bottom rows show the Theil's U and overall rank for the media indicators. Only the results of media models included in the model confidence set are reported for ease of presentation. Values smaller than one are printed in bold and one or two stars indicate that the hypothesis of equal forecast performance of the AR and the respective model is rejected at the five or one percent significance level.