Using Newspapers for Tracking the Business Cycle

A comparative study for Germany and Switzerland

KOF

KOF Working Papers

Using Newspapers for Tracking the Business Cycle: A comparative study for Germany and Switzerland

David Iselin and Boriss Siliverstovs

No. 337
June 2013

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

KOF

ETH Zurich
KOF Swiss Economic Institute
WEH D 4
Weinbergstrasse 35
8092 Zurich
Switzerland

Phone +41 44 632 42 39
Fax +41 44 632 12 18
www.kof.ethz.ch
kof@kof.ethz.ch

Using Newspapers for Tracking the Business Cycle: A comparative study for Germany and Switzerland*

David Iselin
ETH Zurich
KOF Swiss Economic Institute
Weinbergstrasse 35
8092 Zurich, Switzerland
iselin@kof.ethz.ch

Boriss Siliverstovs
ETH Zurich
KOF Swiss Economic Institute
Weinbergstrasse 35
8092 Zurich, Switzerland
siliverstovs@kof.ethz.ch

June 14, 2013

Abstract

On the basis of keyword searches in newspaper articles several versions of the Recession-word Index (RWI) are constructed for Germany and Switzerland. We use these indices in order to track the business cycle dynamics in these two countries. Our main findings are the following. First, we show that augmenting benchmark autoregressive models with the RWI generally leads to improvement in accuracy of one-step ahead forecasts of GDP growth compared to those obtained by the benchmark model. Second, the accuracy of out-of-sample forecasts obtained with models augmented with the RWI is comparable to that of models augmented with established economic indicators in both countries, such as the Ifo Business Climate Index and the ZEW Indicator of Economic Sentiment for Germany, and the KOF Economic Barometer and the Purchasing Managers Index in manufacturing for Switzerland. Third, we show that the RWI-based forecasts are more accurate than the consensus forecasts (published by Consensus Economics Inc.) for Switzerland, whereas we reach the opposite conclusion for Germany. In fact, the accuracy of the consensus forecasts of GDP growth for Germany appears to be superior to that of any other indicator considered in our study. These results are robust to changes in estimation/forecast samples, the use of rolling vs expanding estimation windows, and the inclusion of a web-based recession indicator extracted from Google Trends into a set of the competing models.

Keywords: Nowcasting, Recession, R-word Index, Google Trends
JEL code: C22, C53

*We thank Dirk Heilmann from the *Handelsblatt* for providing us their data of the R-word index for Germany. We also thank the participants of the KOF Brown Bag Seminar at ETH Zurich (March 18, 2013) as well as two anonymous reviewers for useful suggestions. Earlier versions of the paper benefited from comments of the participants at the following conferences: 7th Annual International Symposium on Economic Theory, Policy and Applications in Athens (July 24, 2012) and CIRET Conference in Vienna (September 6, 2012). All computations were performed in Ox version 6.30 Doornik (2007).

1 Introduction

The use of media-based data for tracking current trends in real economy as well as in financial sector has gained popularity among both academics and practitioners in recent years. This surge of interest of using publicly available media data for inferring economic and financial conditions can be traced to the following reasons. First, attempts to establish a relationship between economic and financial processes on the one hand and their reporting in the public media on the other hand are based on the appealing idea that the intensity of media coverage of particular events generally reflects the scope of public sentiment or concern about these events. Second, the advances in information technology allows an immediate access and processing of the vast source of information provided by the various media, including the internet and digitalised newspaper archives, which in turn practically enables a real-time tracking of economic and financial developments as they unfold.

There are a number of studies that exploit news coverage for predicting financial and economic data. Ammann et al. (2013) summarise the content of newspaper articles by means of word-count indices and use them to predict *DAX*[1] stock returns in and out of sample. Engelberg and Parsons (2011) show that press coverage is significantly correlated with the daily trading volume of retail investors. Tetlock (2007) and Doms and Morin (2004) find that news coverage is strongly associated with response by the financial markets. Uhl (2011b) finds out that sentiment indices constructed from Reuters news feeds are useful for predicting stock returns of the Dow Jones Industrials Index. Similarly, Uhl (2011a) reports that TV sentiment index, constructed using broadcast news on American TV, has predictive power for US private consumption.

As mentioned above, the World Wide Web represents another important source of information, which likewise could be successfully exploited for forecasting purposes. Preis et al. (2013) employs the dataset released by *AOL* on all search queries from its users and establish a correlation between attention allocation of US investors and their holding of foreign assets. Kholodilin and Siliverstovs (2012) report a strong correlation between official regional inequality measures in Germany with web-based regional indices constructed using personal car selling advertisements. There is also a growing literature reporting the usefulness of *Google Trends*, a data base which summarises web search queries in the *Google* search engine, for forecasting various economic phenomena like stock market returns (Preis et al., 2013), unemployment (Askitas and Zimmermann, 2009; D'Amuri and Marcucci, 2009), consumption (Kholodilin et al., 2010), number of travellers (Choi and Varian, 2009), *inter alia*.[2]

This paper belongs rather to the former stream of the literature exploiting the content of newspapers for tracking current economic conditions. To this end, we use constructed indices counting occurrence of the word "Recession" in the articles published in *Handelsblatt* (*HB*) in Germany and the *Neue Zürcher Zeitung* (*NZZ*) in Switzerland. The respective so-called Recession-word indices (in short, R-word index or RWI) are used for out-of-sample forecasting of GDP growth

[1] DAX stands for *Deutscher Aktien Index*, the German stock index.

[2] A comprehensive list of related references to the articles using *Google Trends* is provided in Choi and Varian (2012).

in the corresponding countries. As a robustness check, relating our work to the latter stream of the literature, we compare forecasting performance of the newspaper-based R-word indices with those extracted from the web or, more precisely, from *Google Trends*. The web-based R-word indices are derived from the volume of queries submitted from Germany and Switzerland through the Google search engine that contain the word "Recession".

The R-word index was launched in the early 1990s by the weekly Magazine *The Economist*. The idea behind the indicator is astonishingly simple: *The Economist* counts how many stories in the two newspapers *Financial Times* and the *Wall Street Journal* had used the word "Recession" in each quarter. The higher the number, the worse the economy is doing. For the US, the R-word index boasts its ability to pinpoint the start of recessions in 1981, 1990, and 2001, including the Great Recession. In fact, in the second half of 2007 the R-word index started to rise after a prolonged period of tranquillity. This rise prompted Economist (2008) to issue a warning that "*Recession talk is rising*" in January of 2008. This was one of the early signals about the impending Great Recession, which came out to be true. This concern was rightfully justified by the observation that a similar pattern in dynamics of the R-word index was spotted at the beginning of the previous three recessions.

In April 2001, the German *Hypo Vereinsbank* started a German version of the R-word index counting the occurrence of the word "Recession" in the German business newspaper *Handelsblatt*. The release of the German version of the RWI inevitably attracted attention of researchers (Breitung and Jagodzinski, 2001; Bandholz and Funke, 2003; Kholodilin and Siliverstovs, 2006), who tested practical importance of these indices for economic forecasting. These studies uniformly conclude that the R-word index cannot be regarded as a trustworthy indicator of the business cycle dynamics in Germany. On a more positive note, Grossarth-Maticek and Mayr (2008) argument that the poor forecasting performance of the R-word index recorded in previous studies can be explained by the fact that the R-word index by construction is better suited towards timely detecting recessions rather than general tracking of the business cycle. Grossarth-Maticek and Mayr (2008) record a favourable performance of a model augmented with the R-word index in timely detection of the business cycle turning points as compared to the benchmark model based on the interest-rate spread. The publication of the R-word index for Germany was abandoned in 2007, ironically, shortly before the Great Recession started to unfold.

The mixed results on the usefulness of the RWI for economic forecasting in Germany are in sharp contrast with the results reported in Iselin and Siliverstovs (2013). Iselin and Siliverstovs (2013) construct the R-word index for Switzerland and show that an autoregressive model, enhanced with the RWI, produces more accurate forecasts than the benchmark autoregressive model. As expected, the largest gains in forecast accuracy are observed in the period around the financial crisis after the collapse of *Lehman Brothers*.

In this paper we capitalise on these encouraging results and expand our initial analysis along the following dimensions. First, we conduct a comparative analysis on the usefulness of the RWI for forecasting GDP growth in Germany and Switzerland. The fact that the publication of the German RWI had been stopped shortly before the outburst of the global financial crisis prevents one to evaluate how the RWI would had fared in predicting economic conditions during the Great Recession in Germany. In order to do so, we obtained the German version of the RWI that now extends until the second quarter of 2012. Secondly, we compare the forecasting performance of the RWI-augmented models with models that include well-established economic indicators in the both countries. Third, we broaden the set of alternative forecasts by including expert forecasts as published by *Consensus Economics Inc.* Contrary to the model forecasts based on a single economic indicator, experts have a holistic view of economic conditions fed by information from different sources and more sophisticated models. As mentioned above, our fourth contribution to the literature is that we compare predictive content not only of the newspaper-based but also web-based R-word indices extracted from *Google Trends*.

Our paper is structured as follows. In Section 2 we present our data set. In Section 3 we discuss our econometric model, followed by the discussion of our results in Section 4. In Section 5 we verify the robustness of our results by allowing for more observations for parameter estimation and by applying expanding rather than rolling estimation window. In this section we also report the influence of the Box-Cox transformation of the newspaper based indices on their forecasting ability and we analyse if data taken from *Google* search queries do deliver similar results as our indices. The final section concludes.

2 Data

The data used in our paper come from different sources. First of all, we use real-time historical vintages of real GDP. These vintages come from the online real-time database of the *Bundesbank* for Germany and from the internal database of the *KOF Swiss Economic Institute at ETH Zurich*. The earliest vintage of GDP data for Germany stored in the database was released on September 6, 1995, and contains the data ending in 1995Q2. In sequel, we will refer to a particular vintage by the date of the last observation available. For Germany quarterly year-on-year GDP growth rates are computed from vintages of real GDP in constant prices available for the period from 1995Q2 until 2004Q4[3] and from vintages of real GDP in chain-linked prices released since 2005Q1[4]. For Switzerland the available real-time GDP vintages are from 1997Q4 until 2003Q4 reported in constant prices and since 2004Q1 these are reported in chain-linked prices. For both countries the latest GDP vintage used contains observations through 2012Q2. Since we work with growth rates rather than levels changes in definitions and base years, the typically introduced shifts in levels of the original time series are treated as any other revision (e.g., see Koenig et al., 2003). The typical publication lag of GDP data in Germany was about eight weeks until 2003. It was reduced to six weeks for vintages starting from 2003Q1. In Switzerland the publication lag is about eight weeks.

[3] Q.DE.Y.A.AG1.CA010.C.A Pfad: Volkswirtschaftliche Gesamtrechnungen [A] / Bruttoinlandsprodukt [AG1] / Gesamtwirtschaft (Inlandskonzept) [CA010] / in konstanten Preisen [C] / absolute Angaben [A] / vierteljhrlich [Q] / kalender- und saisonbereinigt [Y]

[4] Q.DE.Y.A.AG1.CA010.A.I Pfad: Volkswirtschaftliche Gesamtrechnungen [A] / Bruttoinlandsprodukt [AG1] / Gesamtwirtschaft (Inlandskonzept) [CA010] / in verketteten Vorjahrespreisen [A] / Index [I] / vierteljhrlich [Q] / kalender- und saisonbereinigt [Y]

Two versions of the R-word index for Germany were provided by the *Handelsblatt* newspaper. The first version is based on the number of articles mentioning the word "Recession" (the German word "Rezession" is used) and the second version counts the number of articles that mention both words "Recession" and "Germany" (the German word "Deutschland" is used) in the *Handelsblatt* newspaper. We refer to these versions of the R-word index as *HB-RW* and *HB-RW-GE*, respectively. The R-word indices cover the period from 1986Q1 until 2012Q2. We construct the Swiss counterparts of the R-word index using the keyword search in the *Neue Zürcher Zeitung* (*NZZ*), stored electronically in the Swiss Media Database *Schweizerischer Mediendienst (SMD)*. We denote by *NZZ-RW* and *NZZ-RW-CH* the corresponding R-word indices based on counts of articles containing the word "Recession" and joint occurrence of words "Recession" and "Switzerland". The R-word indices for Switzerland cover the period 1998Q1—2012Q2. In contrast to GDP data the nature of newspaper publishing process ensures that the time series of R-word indices are not subject to any revisions.

The simple idea of mechanical counting of newspaper articles containing a particular keyword without further analysis of the relevant context allows for a fast and cheap way to summarise the prevailing sentiment about the performance of the economy. At the same time, such a mechanical counting can be perceived as a shortcoming. For example, the keyword "Recession" can also appear in an article featuring not only a gloomy recessionary but also a more upbeat context. This would be the case when, for example, the German Chancellor Angela Merkel would point out something like "The recession is over" receiving many-fold citations in German newspapers and therefore contributing positively to keyword count, even though the original meaning of the outspoken message intends to dampen recession fears. Another example would be a popular phrase "Recession fears are overblown" with which politicians attempt to soothe the general public typically shortly before a recession takes its toll. The validity of such phrases can only be assessed *ex post*, but at the time they are made public they reflect inherent uncertainty shared by politicians regarding the course of the economy. Even though we cannot rule out such cases when the word recession appears out of general context, the very fact that it has been mentioned is likely to reflect the genuine concern about the state of the economy which our index intends to pick up.

We compare the predictive content of our newspaper-based indicator with that of the most closely watched economic indicators in the two countries. These include the *Ifo Business Climate Index*[5] as well as the *ZEW Indicator of Economic Sentiment*[6] for Germany and the *KOF Economic Barometer*[7] and the *Purchasing Managers Index (PMI)* in manufacturing[8] for Switzerland. The first three indicators are published by the *Ifo*, the *ZEW*, and the *KOF* economic institutes,

[5]The *Ifo Business Climate Index* is based on around 7000 monthly survey responses from firms in manufacturing, construction, wholesaling and retailing. The firms are asked to give their assessments of the current business situation and their expectations for the next six months.

[6]The *ZEW Indicator of Economic Sentiment* is constructed by asking up to 350 financial experts on their expectations for the economic development in Germany in six months.

[7]The *KOF Economic Barometer* is comprised of 25 individual indicators sub-divided into core-GDP, banking, and construction modules. The core-GDP module is further split into three sub-modules capturing dynamics of consumption, manufacturing, and exports.

[8]The *PMI* is constructed by asking 200 purchasing managers in industrial companies in Switzerland.

respectively. The last indicator is published by the *Credit Suisse* bank. These indicators are available at monthly frequency since 1991M1 (the *Ifo Business Climate Index* and the K*OF Economic Barometer*), 1991M12 (the *ZEW Indicator of Economic Sentiment*), and 1995M1 (the *Purchasing Managers Index in manufacturing*). All indicators were downloaded from *Datastream*. In contrast to the R-word index, these indicators undergo some revisions caused by updates of seasonal factors due to changes in estimation sample and updates of underlying firms' surveys due to the incorporation of late responses, for example. These revisions, however, are of relatively minor magnitude compared to those of GDP. Since we do not possess all historical vintages of these indicators, we are forced to approximate the availability of these indicators in the past by appropriately truncating them.[9] We refer to these indicators by the names of the corresponding economic institutes that release them, i.e., *IFO*, *ZEW*, and *KOF*, with exception of the *PMI Purchasing Manager Index* time series to which we refer as *PMI*. Since the economic indicators are available at a monthly frequency but our target variable is quarterly, we convert them to a quarterly frequency by taking average values of all months in a particular quarter.

We also utilise alternative information that are provided by *Consensus Economic Inc.* which publishes consensus forecasts reflecting sentiment and opinion of a number of professional forecasters on the current and future prospects of the economies in question. The consensus forecasts of our interest are typically published in the beginning of the third month of the reference quarter, shortly after official GDP data for the previous quarter are released. Naturally, once these forecasts are released these are not subject to any revision. The consensus forecasts are available for the period from 1999Q1 until 2012Q2.

Last but not least, we augment the set of alternative forecast sources by including the web-based R-word indices for Germany and Switzerland, based on the number of search queries that contain the word "Recession" that originated from these two countries. This information is provided by *Google Trends* that summarises all search queries submitted through the Google search engine since January 2004. Table 1 presents the summary of data definitions and sources.

[9] In doing so, we treat this indicators in a pseudo real-time framework similar to (Drechsel and Scheufele, 2012) and (Marcellino and Schumacher, 2010), for example, who is even more restrictive then us by using final data vintages not only for economic indicators but also for time series they forecast like industrial production and GDP, respectively. We, however, chose to retain the real-time aspect of forecasting as much as data availability allows us.

Table 1: Data definitions and sources

Indicators/Forecasts	Country	Abbreviation	Source	Frequency	Release Timing
R-word Index (Recession)	Germany	HB-RW	Handelsblatt	Quarterly	Next quarter (1st business day)
R-word Index (Recession and Germany)	Germany	HB-RW-GE	Handelsblatt	Quarterly	Next quarter (1st business day)
R-word Index (Recession)	Switzerland	NZZ-RW	Swiss Media Database	Quarterly	Next quarter (1st business day)
R-word Index (Recession and Switzerland)	Switzerland	NZZ-RW-CH	Swiss Media Database	Quarterly	Next quarter (1st business day)
Ifo Business Climate Index	Germany	IFO	Datastream (BDCNFBUSQ)	Monthly	Current month (end)
ZEW Indicator of Economic Sentiment	Germany	ZEW	Datastream (BDZEWECSR)	Monthly	Current month (end)
KOF Economic Barometer	Switzerland	KOF	Datastream (SWECOBARR)	Monthly	Current month (end)
Purchasing Managers Index in manufacturing[a]	Switzerland	PMI	Datastream (SWPURCHSQ)	Monthly	Next month (1st business day)
Consensus forecasts	Germany	CE-GE	Consensus Economics Inc.	Quarterly	Current quarter (3rd month)
Consensus forecasts	Switzerland	CE-CH	Consensus Economics Inc.	Quarterly	Current quarter (3rd month)
R-word Index (Recession), web-based	Germany	GOOGLE-RW-GE	Google Trends	Weekly	Next quarter (1st business day)
R-word Index (Recession), web-based	Switzerland	GOOGLE-RW-CH	Google Trends	Monthly	Next quarter (1st business day)
Quarterly year-on year real GDP growth	Germany	GDP-GE	Bundesbank database	Quarterly	Next quarter (2nd month)
Quarterly year-on year real GDP growth	Switzerland	GDP-CH	KOF database	Quarterly	Next quarter (2nd month)

[a] The PMI is published by Credit Suisse in Switzerland.

3 Econometric methodology

3.1 Forecasting models

We employ autoregressive distributed lag (ARDL) models of the form:

$$y_t^{\upsilon} = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i}^{\upsilon} + \sum_{j=0}^{q} \beta_j X_t^i + \epsilon_t, \tag{1}$$

where in the most general case we allow for up to p lags of the dependent variable y_t^{υ} and up to q lags of the explanatory variable X_t^i. As described in Section 2 the dependent variable is quarterly year-on-year real GDP growth for either Germany or Switzerland. The superscript $^{\upsilon}$ indicates that at each point of time we use the respective historical GDP vintage, that was actually available to a forecaster at the respective forecast origin, for model estimation and forecast generation. The explanatory variable X_t^i represents one variable out of the available indicators for Germany, $HB-RW, HB-RW-GE, IFO, ZEW$, and for Switzerland, $NZZ-RW, NZZ-RW-CH, KOF, PMI$, that are appropriately truncated in order to reflect their availability in the past. We allow for a contemporaneous value of the explanatory variable X_t^i, i.e., $j = 0$, in order to reflect the fact that economic indicators as well as R-word indices are available one quarter earlier than official GDP releases.

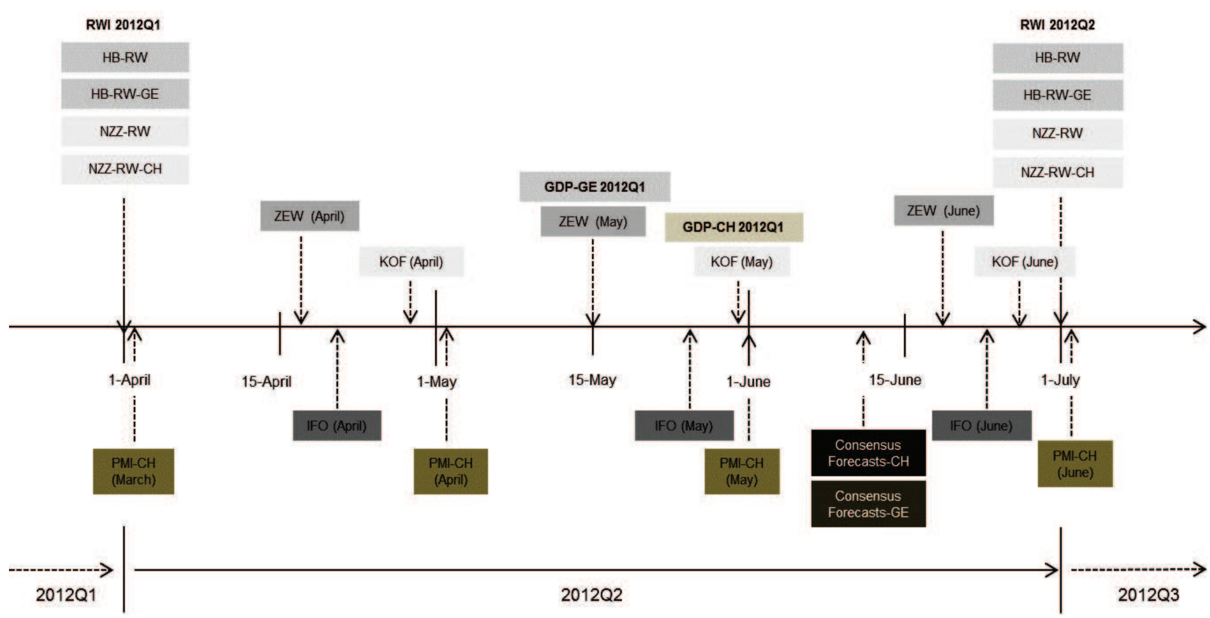

Figure 1: Timeline of data releases.

The release timing of the alternative indicators is shown in Figure 1, where we illustrate the availability of GDP data as well as indicators for production of forecasts for 2012Q2. The GDP vintages that contain data up to the previous quarter, 2012Q1, were released on 15th of May and 31st of May 2012 for Germany and Switzerland, respectively. The *IFO*, *ZEW*, *KOF*, and

PMI indicators that contain data up to June 2012 were released on 26th, 15th, 30th, and 2nd May, respectively. In order to produce a value of the respective indicators for 2012Q2 we take the average of their values observed in April, May, and June. The R-word recession indices were computed on the first business day of July by appropriate keyword searches in the *Handelsblatt* and *Neue Zürcher Zeitung* newspapers. Consensus forecasts for 2012Q2 were released on 11th of June for both countries. This release timing implies that as early as on the first business day of July 2012 we possess all the information necessary to compute out-of-sample forecasts for 2012Q2. We produce forecasts maintaining this data release timing for every quarter in our forecast sample (2005Q1—2012Q2). As noted above, in order to reflect the real-time information flow we use historical GDP vintages that were available in the past in order to estimate parameters of Equation (1) and generate forecasts. For example, in order to produce forecasts for 2005Q1 we make use of the GDP vintage released in this quarter (on 15th of February 2005 for Germany and 3rd of March 2005 for Switzerland) that contains observations up to 2004Q4. The time series of alternative indicators were truncated at 2005Q1 in order to reflect their availability on the first business day of April 2005.

We assess the forecast accuracy of competing models with respect to first-released as well as last-available estimates of GDP growth. First-releases provide timely estimates of the current state of the economy and therefore are crucial for decisions made by the government and the private business sector in real time. As it is well-known, initial GDP data releases undergo several revisions. Under the assumption that revised data do more accurately reflect the true but unobserved state of the economy it is of further interest to compare the forecasting performance of competing approaches also with respect to revised data. Thus, for example, forecasts made for 2005Q1 are compared with actual values for this quarter that were released in the next quarter 2005Q2 on 12th of May 2005 for Germany and 2nd of June 2005 for Switzerland. These forecasts are also compared with actual values of GDP growth taken from the last GDP vintage in our sample released on 14th of August and 4th of September 2012, for Germany and Switzerland, respectively.

We compare forecasts obtained with help of additional information with those generated by a simple univariate autoregressive model:

$$y_t^v = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i}^v + \epsilon_t. \tag{2}$$

Also in this case, models parameters were estimated and forecasts generated using historical vintages of GDP data available at the forecast origin. We restrict ourselves to comparison of one-step ahead forecasts, i.e., for the latest quarter for which values of the R-word index are available.

3.2 Tests of equal forecast accuracy

The formal assessment of statistical differences in the accuracy of forecasts is carried out within the testing framework suggested in Diebold and Mariano (1995), when forecasts were generated from non-nested models, and in Clark and West (2007) for forecasts from nested models.

Denote a pair of h-steps ahead competing forecasts as $(f_{1,t}, f_{2,t})$, such that the corresponding forecast errors are $(y_t - f_{1,t}, y_t - f_{2,t}) = (e_{1,t}, e_{2,t})$. The test for equal predictive ability is based on

$$E[(e_{1,t})^2 - (e_{2,t})^2] = 0,$$

assuming the quadratic loss function, the same is used for in-sample parameter estimation of forecasting models. Defining the loss differential, $d_t = (e_{1,t})^2 - (e_{2,t})^2$, the expectation operator in the above equation is replaced by its empirical counterpart:

$$\overline{d} = \frac{1}{P}\sum_{t=1}^{P}[(e_{1,t})^2 - (e_{2,t})^2] = \frac{1}{P}\sum_{t=1}^{P} d_t, \tag{3}$$

where P specifies the number of forecasts available for a given horizon h. Diebold and Mariano (1995) prove that under appropriate regularity conditions the test for equal predictive ability is based on the following test statistic:

$$\frac{\overline{d}}{\widehat{\sigma}_P/\sqrt{P}} \xrightarrow{d} N(0,1), \tag{4}$$

where $\widehat{\sigma}_P^2$ is a heteroskedasticity- and autocorrelation-consistent (HAC) estimator of the asymptotic variance, $\sigma_P^2 = var(\sqrt{P}\ \overline{d})$. Since our forecast evaluation sample is of a moderate size, we apply the modified Diebold-Mariano test suggested in Harvey et al. (1997).

As pointed out in McCracken (2007), when comparing the forecasting ability of two nested linear regression models by means of the Diebold-Mariano test statistic, reported in Equation (4), its asymptotic distribution under the null hypothesis of equal predictive accuracy is nonstandard. Under the null hypothesis that data are generated by a parsimonious nested model, the estimation of a larger nesting model introduces noise into its forecasts as additional model parameters are estimated whose population values are zero. Hence, under the null hypothesis it is expected that the smaller model will produce more accurate forecasts than the larger model, i.e., a smaller Mean Squared Forecast Error (MSFE). In order to circumvent this problem of inferring critical values of a nonstandard limiting distribution we use the approach suggested in Clark and West (2007). Their test statistic has an approximately normal distribution under the null hypothesis.

Clark and West (2007) suggest to adjust the loss differential in Equation (3) as follows:

$$\overline{d}^{adj.} = \frac{1}{P}\sum_{t=1}^{P}[(e_{1,t})^2 - [(e_{2,t})^2 - (f_{1,t} - f_{2,t})^2] = \frac{1}{P}\sum_{t=1}^{P} d_t^{adj.}, \tag{5}$$

where in this case $f_{1,t}$ and $f_{2,t}$ denote forecasts from nested and nesting models, respectively. The test statistic has a familiar form

$$\frac{\overline{d}^{adj.}}{\widehat{\dot{\sigma}}_P/\sqrt{P}} \xrightarrow{d} N(0,1), \tag{6}$$

where $\widehat{\dot{\sigma}}_P^2$ is a heteroskedasticity- and autocorrelation-consistent (HAC) estimator of the asymptotic variance, $\dot{\sigma}_P^2 = var(\sqrt{P}\,\overline{d}^{adj.})$. Since the null hypothesis implies that $\text{MSFE}_1 < \text{MSFE}_2$, Clark and West (2007) suggest to use one-tailed tests.

4 Results

This section contains description of the empirical results according to the following plan. First, we compare first-released with last-available estimates of GDP growth in both countries during the forecast period. This forecast period can be subdivided into three phases: the pre-crisis period from 2007, the period of the Great Recession, and the post-crisis period extending to the middle of 2012. Second, we compare indicators of interest for each country. Finally, we describe the results of the out-of-sample forecasting of GDP growth using alternative information sources.

4.1 Descriptive analysis

The first- and last-available estimates of GDP growth are presented in the left and the right panels of Figure 2 for Switzerland and Germany, respectively. There is a systematic upward revision of GDP growth in the pre-crisis period observed for both countries. During the Great Recession the German economy contracted at a faster rate than did the Swiss economy. According to the last-available estimate in 2009Q2 the Swiss economy contracted at a rate of 3.7%, while for Germany the maximum contraction rate of 6.8% was observed in 2009Q1. It is also remarkable that while for Germany there is a very minor revision of the initially estimated depth of the trough (from -6.9% to -6.8% in 2009Q1), for Switzerland there is a substantial downwards revision of the first estimate (-2.4% in 2009Q1 to -3.7% in 2009Q2).

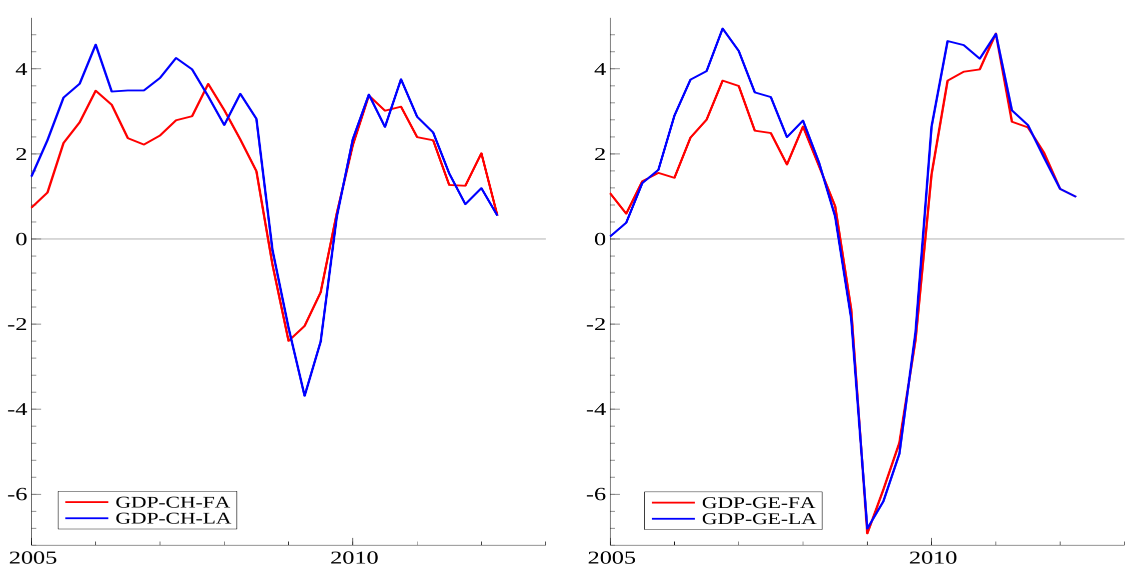

Figure 2: Quarterly year-on-year real GDP growth: first- (FA) vs last-available (LA) estimates; Switzerland - left panel, Germany - right panel.

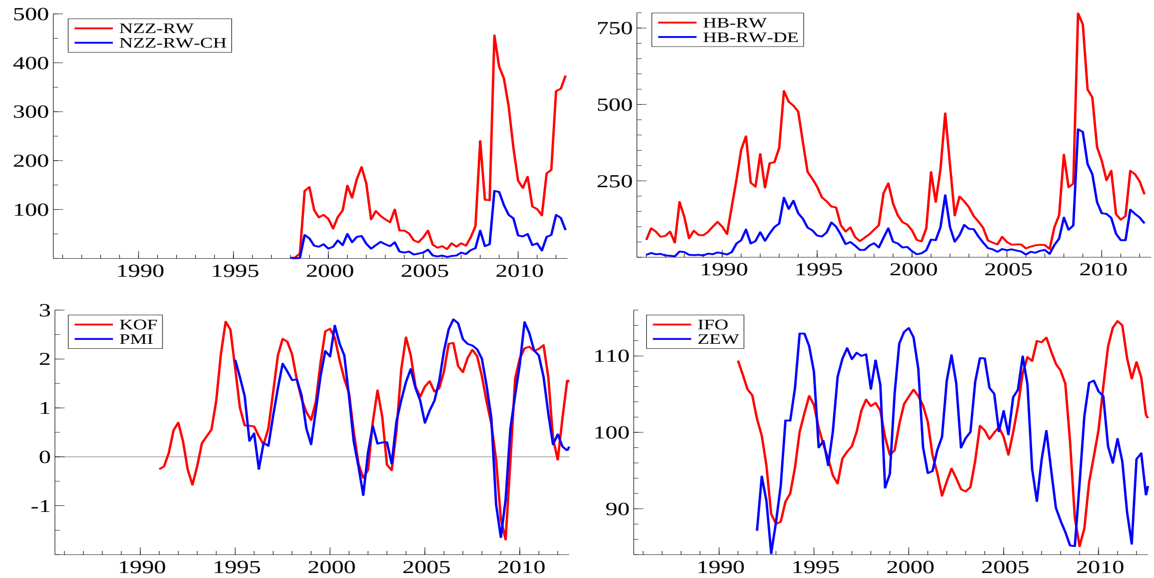

Figure 3: Indicators; Switzerland - upper- and lower-left panels, Germany - upper- lower right panels. Indicators on the lower panels (KOF and PMI) and (IFO and ZEW) were adjusted to have the same mean and range for the sake of better comparison.

The alternative indicators are presented in Figure 3. The upper panels contain different versions of the R-word index for Switzerland on the left and Germany on the right. In the lower panels economic indicators for Switzerland (*KOF* and *PMI*, on the left) and Germany (*IFO* and *ZEW*, on the right) are displayed. All these indicators were converted from the original monthly to quarterly frequency by averaging their monthly values in each quarter. In order to facilitate a comparison, each pair of indicators were adjusted to have the same mean and range.

Given the close geographical and cultural proximity as well as very closely interconnected economies of both countries, it is not surprising that both *NZZ*- and *HB*-based R-word indicators display very close co-movement. Both versions of the index surge in 1999 and 2000, possibly related to the burst of the Dotcom bubble. As a first warning signal of the emanating global crisis, both indices surge in 2008Q1 indicating presumably increasing economic worries caused by the bankruptcy of *Bear Sterns* and its eventual sell to *JP Morgan Chase*. Both indices reach all-time heights in 2008Q4 and 2009Q1, reflecting the meltdown of the global financial system, triggered by the collapse of *Lehman Brothers*. Due to the longer time period for which the *HB*-based indicators are available, we are able to track its response to economic news in the early 1990s. This surely includes the economic turmoil during and after German reunification. Furthermore, the US invasion in Iraq could have contributed to the spike in 1991Q2 (and the one in 1993Q3) due to fears related to the oil price development. These observations conform the basic idea behind the simple construction of the recession-word index, namely, the more acute events are the broader the coverage they receive in press. The intensity of negative events, in the end, is summarising the prevailing (negative) sentiments about the state of the economy. It also turns out that, apart from short-run fluctuations, an R-word index based on a single key word or its more restrictive version based on the combination of two keywords move very synchronously. Thus, the question on which version can better track the business cycle remains to be resolved empirically.

The *KOF* and *PMI* economic indicator for Switzerland also display very high degree of co-movement. The largest divergence occurs in the end of our sample in the year 2012, when the *KOF Economic Barometer* exhibits a high upward growth, while the *PMI* goes slightly down. There is though much less coherence between the *Ifo* and *ZEW* economic indicators. By examining the turning points of both the indicators one may conclude that the *ZEW* indicator tend to lead the *Ifo Business Climate Index*. For example, the *ZEW* indicator started its sharp descent already in 2006 when the *Ifo* indicator was at its historical height. Because of the fact that most politicians and economists were taken by a great surprise by the Great Recession and its impact on the worldwide economy, it is still open to speculation whether this descent can be interpreted as a rightful anticipation of the upcoming Great Recession long before the crisis started unfolding itself or whether it was just a spurious result.

4.2 Empirical analysis: Rolling estimation window

Based on data availability and in order to ensure comparable samples for both countries we have chosen to use the following data sample starting from 1998Q1 and ending in 2012Q2.[10] The starting date coincides with the earliest observation of the R-word index for Switzerland. The end date coincides with the last date for which the R-word index for Germany is available. We split this sample into two parts of approximately equal size. The first part from 1998Q1 until 2004Q4, comprising 28 observations, is used as an initial estimation sample, striking a balance between a reasonably long sample for precise model parameter estimation and allowing enough data points for out-of-sample evaluation of forecast accuracy. The resulting forecast evaluation sample is 2005Q1—2012Q2, consisting of 30 observations.

We use estimated model parameters in order to generate forecasts for 2005Q1. Then we shift both the starting and ending points of the estimation sample by one quarter resulting in the new estimation sample from 1998Q2 until 2005Q1 and use the estimated model parameters in order to generate forecasts for 2005Q2. We repeat this procedure based on a rolling estimation window until a forecast for 2012Q2 is generated using estimation sample 2005Q2—2012Q1.

The results of the out-of-sample forecast evaluation of the alternative indicators are summarised in Table 2. The benchmark model is specified in Equation (2), where we fixed the lag length $p = 1$ and $p = 2$. As the AR(2) produced lower RMSFE in all cases than the AR(1) model, we maintain $p = 2$ for the benchmark forecasts obtained without any extraneous information. All forecasts from the indicator-augmented models were generated by an ARDL(p,q) model specified in Equation

[10] Since one may argue that by restricting the starting date for the model estimation in 1998Q1 we are likely to artificially impair the forecasting performance of the models augmented with economic indicators for which more data points are available, we report the results of the robustness check in Section 5, where we allow for maximum possible estimation samples.

Table 2: Out-of-sample forecast accuracy: Rolling estimation window, 2005Q1–2012Q2

Switzerland				Germany			
	First-available GDP RMSFE	Ratio	p-value[a]		First-available GDP RMSFE	Ratio	p-value
AR(1)	0.944	1.25		AR(1)	1.652	1.01	
AR(2)	0.756	1.00		AR(2)	1.631	1.00	
NZZ-RW	0.650	0.86	0.023	HB-RW	1.411	0.87	0.058
NZZ-RW-CH	0.626	0.83	0.035	HB-RW-GE	1.386	0.85	0.050
KOF	0.768	1.02	0.040	IFO	1.343	0.82	0.075
PMI	0.644	0.85	0.005	ZEW	1.657	1.02	0.300
CE-CH	0.776	1.03	0.569	CE-GE	0.799	0.49	0.038
	Last-available GDP RMSFE	Ratio	p-value		Last-available GDP RMSFE	Ratio	p-value
AR(1)	1.255	1.08		AR(1)	1.924	1.06	
AR(2)	1.166	1.00		AR(2)	1.822	1.00	
NZZ-RW	0.969	0.83	0.000	HB-RW	1.673	0.92	0.086
NZZ-RW-CH	0.990	0.85	0.002	HB-RW-GE	1.661	0.91	0.078
KOF	0.973	0.83	0.000	IFO	1.460	0.80	0.042
PMI	0.969	0.83	0.001	ZEW	1.868	1.02	0.510
CE-CH	1.084	0.93	0.147	CE-GE	1.147	0.63	0.035
FA-GDP-CH	0.862	0.74		FA-GDP-GE	0.670	0.37	

[a] The column entries are the p-values of the tests of equal predictive ability between indicator-augmented models and the benchmark AR(2) model, described in Section 3.2. The test of Diebold and Mariano (1995) is used to compare forecasting accuracy of the consensus forecasts with that of the benchmark model. The test of Clark and West (2007) is used to test the null hypothesis that each of the remaining indicator-based models has equal predictive ability as the AR(2) model.

(1), where we fixed the lag length $p = 2$ and $q = 0$.[11] Such a choice of fixed specification has an advantage that it exactly fits into the testing framework of Clark and West (2007), specifically developed for nested models. The corresponding forecasts from Consensus Economics Inc. were taken from their bulletins. In order to compare the forecast accuracy between the consensus and benchmark model forecasts we use the Diebold-Mariano test, as mentioned above in Section 3.

The first observation to be made is that the comparison of forecast accuracy reveals that both AR and ARDL models are much better in forecasting first releases of GDP growth rather than revised figures. The reason for this is that, as shown in Figure 2, in the pre-crises period estimates of

[11] We experimented with alternative criterion for model specification choice, like minimising the Schwarz Information Criterion, and longer lags of the explanatory variables, but it did not result in any systematic improvement in forecast accuracy.

GDP growth in both countries were systematically revised upwards, indicating that first estimates were too pessimistic. On the contrary, in the period of the Great Recession the first estimates of Swiss GDP growth were mostly revised downwards, indicating too optimistic initial estimates. Altogether, if we use the first GDP estimate as a predictor of the revised GDP data then the resulting RMSFEs are 0.86 and 0.67, as reported in Table 2 by entries FA-GDP-CH and FA-GDP-GE for Switzerland and Germany, respectively. This indicates a rather large discrepancy between these two estimates of GDP growth, which apparently was rather difficult to foresee given the information available to forecasters in the past.

The left panel of Table 2 summarises the outcome for Switzerland. Compared with the benchmark model, the models augmented with *NZZ-RW*, *NZZ-RW-CH*, and *PMI* indicators deliver a similar forecast accuracy with the reduction in RMSFE of about 15%. For these indicators we can reject the null hypothesis of equal predictive accuracy of univariate AR(2) and indicator-augmented ARDL(2,0) models in favour of the latter. Observe that even though the ARDL model augmented with the *KOF* indicator delivers a very similar RMSFE with that of the AR(2) model, the test of Clark and West (2007) rejects the null hypothesis of equal forecast accuracy also in this case. This happens because the adjustment term $(f_{1,t} - f_{2,t})^2$ in the expression of the Clark-West loss function in Equation (5) is sufficiently large than it is implied by the null hypothesis. An alternative explanation of this, at first glance surprising test outcome, is that as noted in Clark and West (2007), their test statistic is identical to that used in the forecast encompassing test of Harvey et al. (1998). Hence, this outcome has the following interpretation that the null hypothesis that forecasts of the AR(2) model encompass those of the ARDL model in question can be rejected at appropriately chosen significance level. Finally, when comparing accuracy of the consensus forecasts (*CE-CH*) with that of the benchmark model, we cannot reject the null hypothesis that these are equal.

Similar conclusions can be reached when comparing the forecast accuracy of the last-available GDP vintage. In this case, all indicator augmented models produce reduction in the RMSFE of a similar magnitude (about 17%). The corresponding null hypothesis can be rejected at the 1% significance level. As before, we cannot reject the null hypothesis that the reduction in the RMSFE brought by the consensus forecasts (*CE-CH*) of 7% is statistically significant.

The results of the forecast comparison for Germany are shown in the right panel of Table 2. In this case the consensus forecasts display by far superior performance to the rest of the models. When compared to the predictive performance of the univariate model, the reduction in the RMSFE is about 51% and 37% for first-released and last-available GDP data. These reductions in RMSFE are statistically significant at the 5% level. The next most informative indicator is the *Ifo* indicator. Using this indicator results in reduction in RMSFE of about 20%, which is significant at 10% and 5% for first- and last-available GDP estimates, respectively. Forecasts generated with the help of the *Handelsblatt* indicators result in reduction of the RMSFE of about 13-15% and 8-9% for first- and last-available GDP estimates, which are found to be significant at the 10% level. It is interesting to observe that we cannot reject the null hypothesis when evaluating forecasts from the model augmented with the *ZEW* indicator.

The forecasting performance of the alternative models is illustrated in Figures 4—13. The upper

Table 3: Out-of-sample forecast accuracy during the Great Recession: Rolling estimation window

	First-available GDP		Last-available GDP	
	RMSFE	Ratio	RMSFE	Ratio
	Switzerland[a]			
AR(2)	1.484		1.249	
NZZ-RW	0.941	0.63	0.241	0.19
NZZ-RW-CH	0.761	0.51	0.468	0.37
	Germany[b]			
AR(2)	2.651		2.663	
HB-RW	1.930	0.73	1.942	0.73
HB-RW-DE	1.765	0.67	1.770	0.66

[a] The period of the Great Recession was determined using the business cycle chronology in Siliverstovs (2013, Table 3) for Switzerland (2008Q4—2009Q2).

[b] For Germany we used the business chronology of the Economic Cycle Research Institute (ECRI) `http://www.businesscycle.com/` (2008Q2—2009Q1).

panels display official GDP growth estimates (first releases on the left and last-available estimates on the right) together with the benchmark model forecasts and forecasts produced with the help of extraneous information. The lower panels report the absolute forecast error for the corresponding models.

A referee suggested to investigate the hypothesis whether during the period of the Great Recession relative improvement in forecasting accuracy of the models augmented with the RWI is especially pronounced with respect to the benchmark autoregressive model. The results of such a comparison are presented in Table 3. The period for the Great Recession was taken from the business cycle chronology of Siliverstovs (2013, Table 3) for Switzerland (2008Q4—2009Q2) and from the business cycle chronology the *Economic Cycle Research Institute* (*ECRI*). Since the latter chronology is provided at the monthly frequency (peak in 2008M04 and trough in 2009M01) we converted it to the quarterly frequency (2008Q2—2009Q1). Even though the absolute values of the reported RMSFEs are higher than those estimated for the whole forecast period, see Table 2, the RMSFE ratio are lower. For example, it is a remarkable finding that the ARDL models with *NZZ-RW* and *NZZ-RW-CH* variables yield substantial reductions in RMSFE. The corresponding RMSFE ratios for the last-available GDP vintage are 0.19 and 0.37, respectively. This supports the hypothesis that gains in forecast accuracy due to inclusion of the information provided by the Recession-word indices are more pronounced during periods of economic hardship.

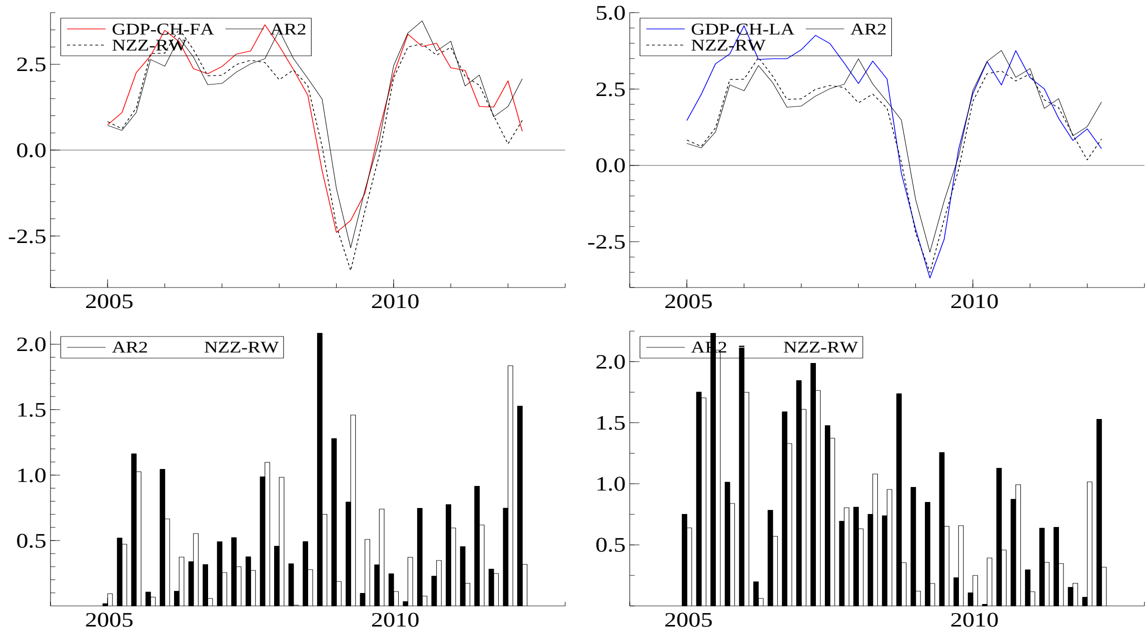

Figure 4: Quarterly year-on-year real GDP growth: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-NZZ-RW models

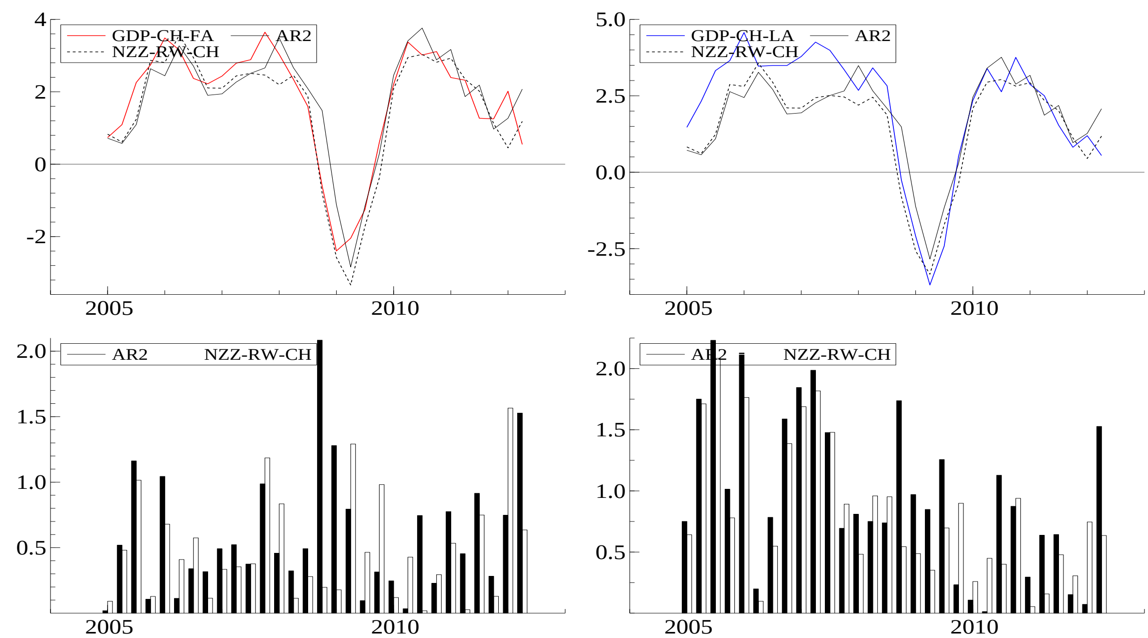

Figure 5: Quarterly year-on-year real GDP growth: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-NZZ-RW-CH models

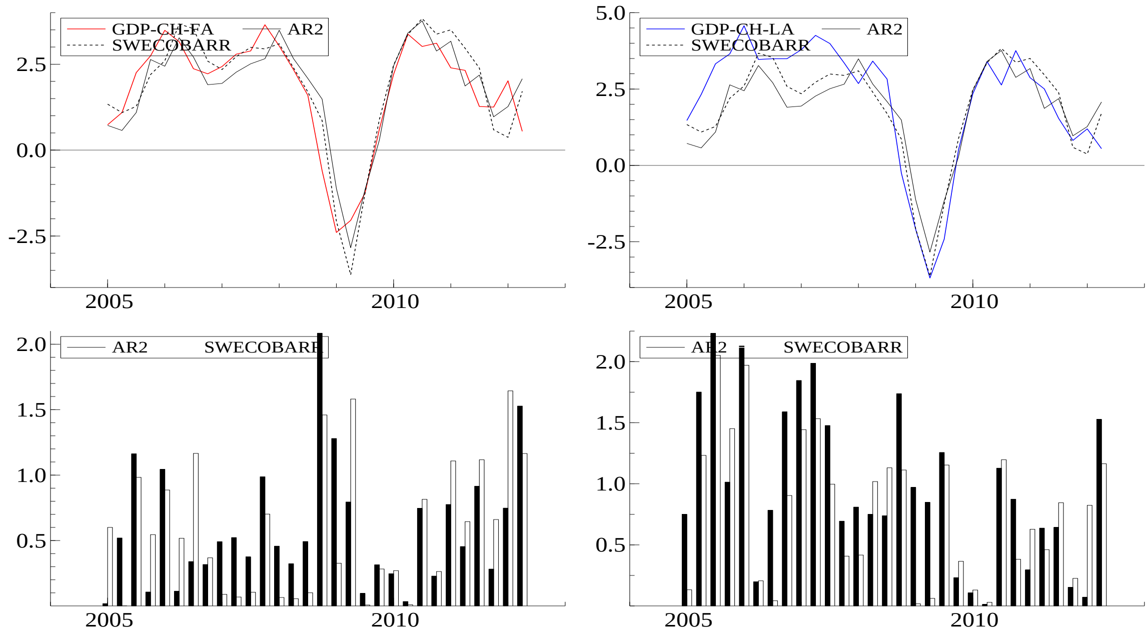

Figure 6: Quarterly year-on-year real GDP growth: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-KOF models

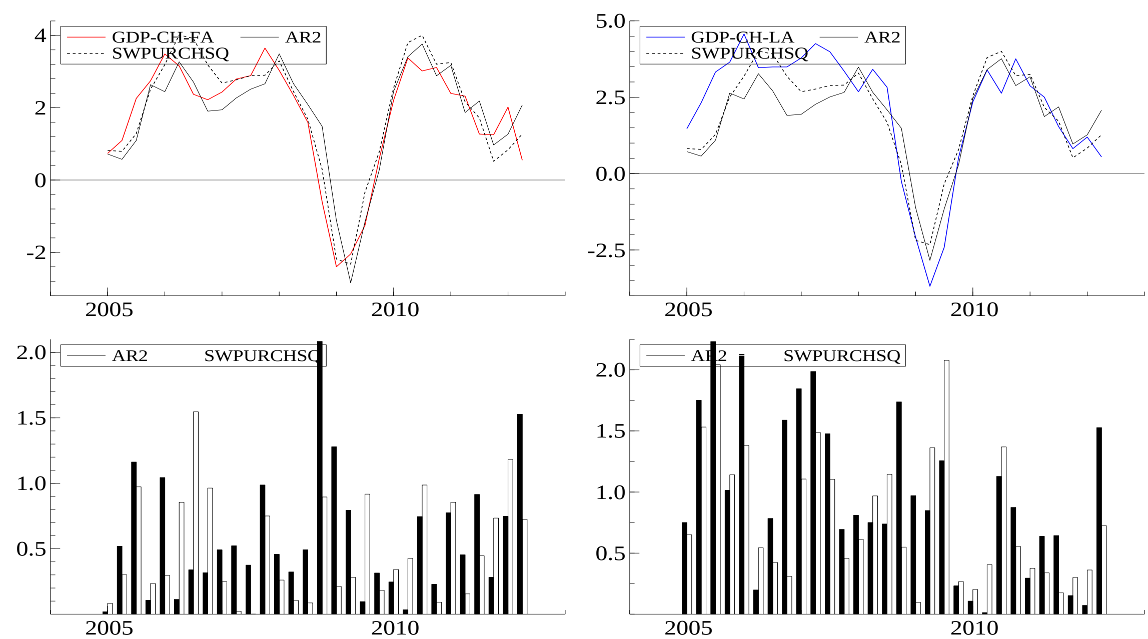

Figure 7: Quarterly year-on-year real GDP growth: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-PMI models

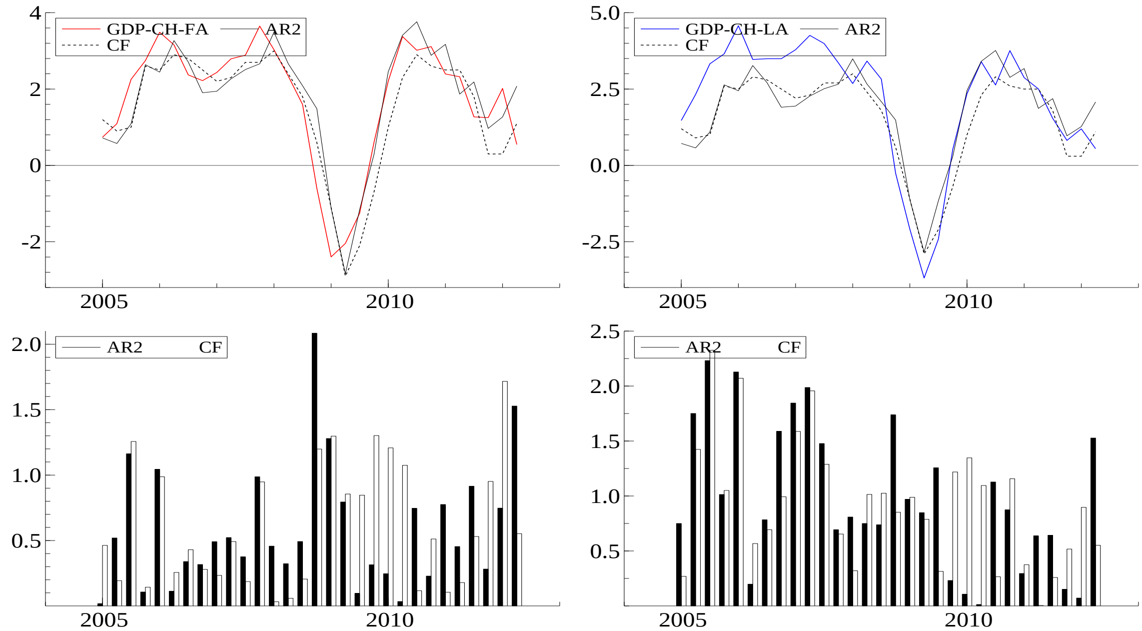

Figure 8: Quarterly year-on-year real GDP growth: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-CE-CH models

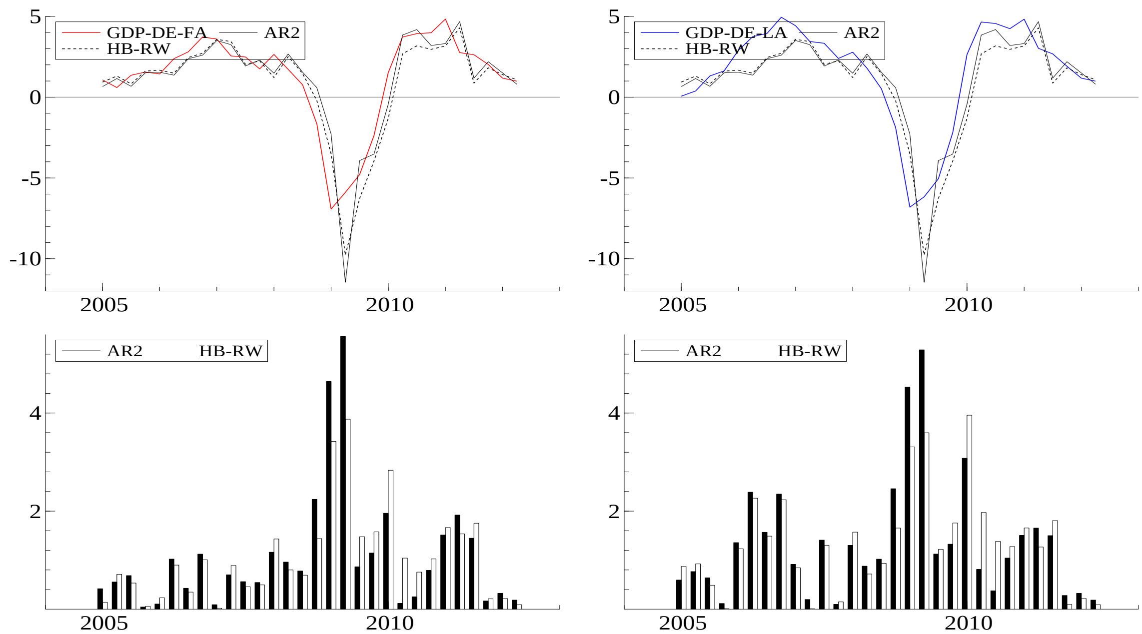

Figure 9: Quarterly year-on-year real GDP growth: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-HB-RW models

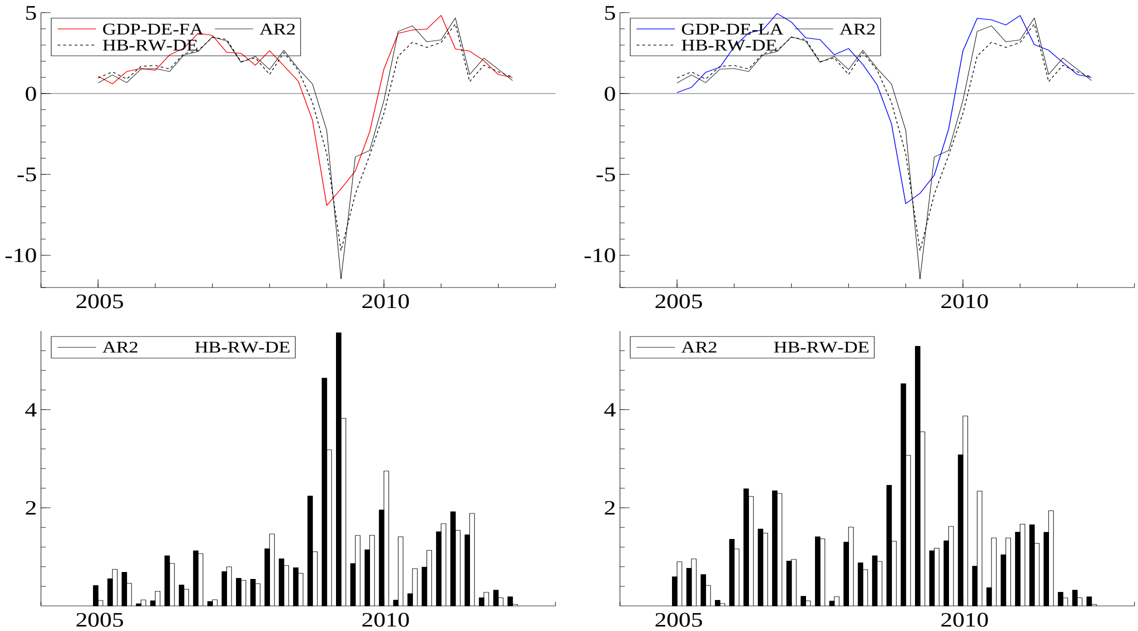

Figure 10: Quarterly year-on-year real GDP growth: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-HB-RW-GE models

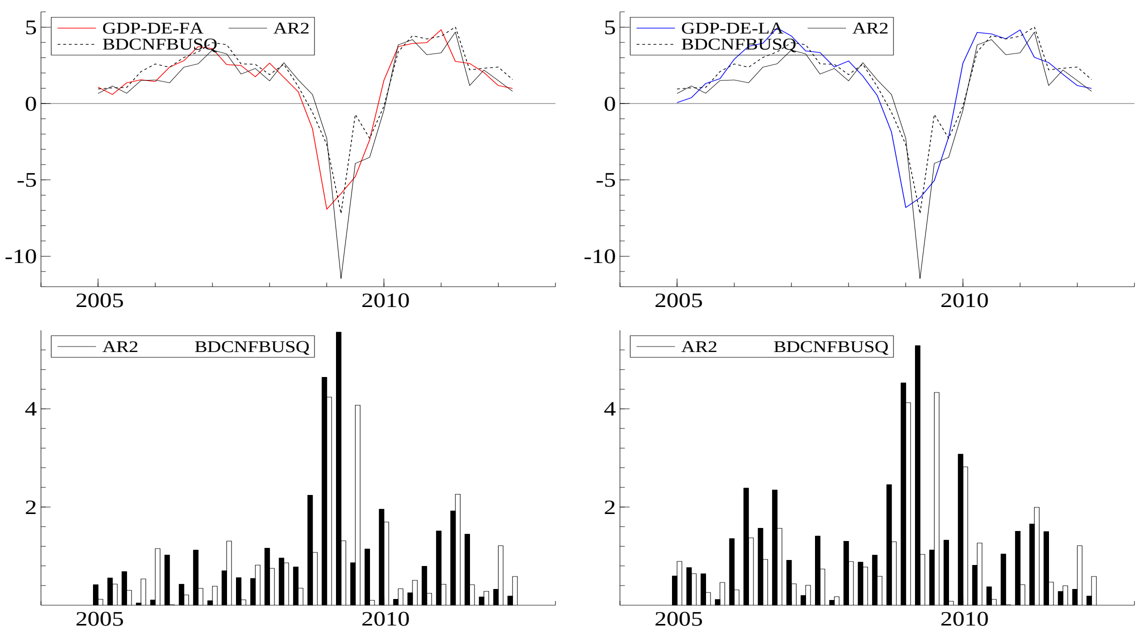

Figure 11: Quarterly year-on-year real GDP growth: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-IFO models

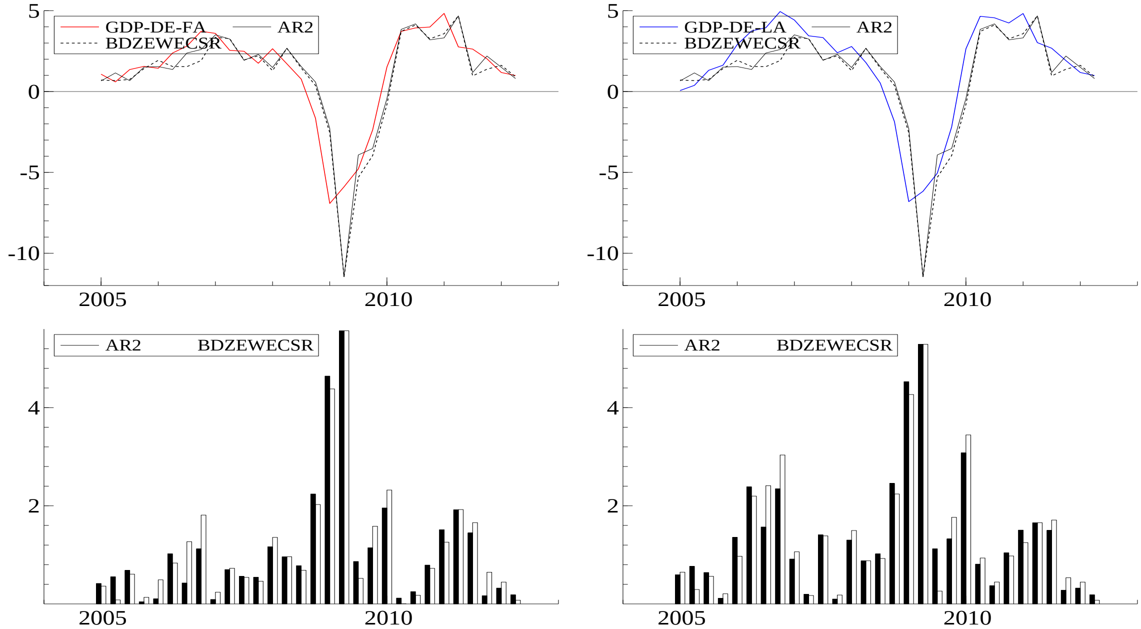

Figure 12: Quarterly year-on-year real GDP growth: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-ZEW models

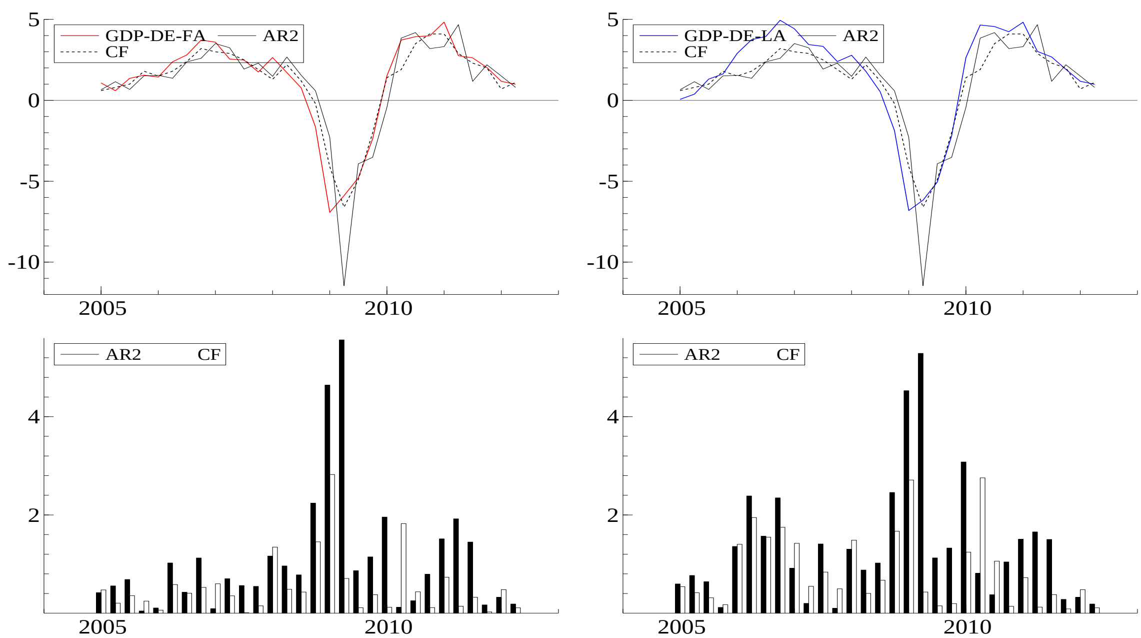

Figure 13: Quarterly year-on-year real GDP growth: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-CE-GE models

5 Robustness testing

In this section we verify the robustness of the results reported in Section 4. In Subsection 5.1 we address the potential criticism that by restricting initial estimation sample to begin in 1998Q1 we artificially may impair forecasting ability of models that use variables for which observations are available earlier than this date. We also employ expanding rather than rolling estimation window. In Subsection 5.2 we check whether a Box-Cox transformation of the R-word index may improve its forecasting performance, since the simple count that we use in Section 4 may not be optimal. In Subsection 5.3 we compare the predictive performance of our newspaper-based R-word indices with that of the web-based R-word indices extracted from *Google Trends*.

5.1 Expanding estimation window

In this subsection we report the results of the forecasting exercise in a modified setting. First, we allow for more data points for parameter estimation. In particular, we set the earliest period for estimation sample to start in 1992Q1. This date is chosen because the vintages of German GDP downloaded from the Bundesbank online database are available since 1991Q1. As four observations are lost after computation of the quarterly year-on-year growth rate, we start estimation sample in 1992Q1. For Germany this initial estimation date is used for all indicators (*HB-RW, HB-RW-GE, IFO,* and *ZEW*), because all of them are available since this date or some of them even earlier. For Switzerland, we use 1992Q1 as a starting estimation period when generating forecasts from the model augmented with the *KOF* indicator. Due to the fact that the PMI indicator is available since 1995Q1, we use this period as a staring date for estimating parameters of the respective ARDL model. For the *NZZ*-based R-word indices the starting period of the estimation sample remains the same as in Section 4, i.e., 1998Q1.

As an additional modification to the forecasting setup, we use recursively expanding rather than rolling window for parameter estimation. At the same time we retain the same forecast evaluation sample, 2005Q1—2012Q2, as in Section 4. Hence, whenever possible, model parameters are estimated using the initial sample from 1992Q1 until 2004Q4.[12] The estimated model parameters are used in order to generate a forecast for 2005Q1. Then we retain the initial estimation period but shift the ending points of the estimation sample by one quarter resulting in the new estimation sample from 1992Q1 until 2005Q1 and use the estimated model parameters in order to generate forecast for 2005Q2. We repeat this procedure based on expanding estimation window until forecast for 2012Q2 is generated using the estimation sample 1992Q1—2012Q1.

The results of this forecasting exercise are reported in Table 4. The use of the expanding rather than recursive estimation window generally results in increased forecast accuracy not only for the benchmark model but also for indicator augmented models. Nevertheless, the general conclusions from results reported in Section 4 hold through. For Switzerland, the three indicators (*NZZ-RW, NZZ-RW-CH*, and *PMI*) have similar forecast accuracy characterised by decrease in RMSFE of

[12] As noted above, for the *PMI* and *NZZ*-based R-word indices the starting point is 1995Q1 and 1998Q1, respectively.

Table 4: Out-of-sample forecast accuracy: Expanding estimation window, 2005Q1–2012Q2

Switzerland				Germany			
	First-available GDP				First-available GDP		
	RMSFE	Ratio	p-value[a]		RMSFE	Ratio	p-value
AR(1)	0.913	1.22		AR(1)	1.581	1.11	
AR(2)	0.748	1.00		AR(2)	1.424	1.00	
NZZ-RW	0.608	0.81	0.023	HB-RW	1.324	0.93	0.085
NZZ-RW-CH	0.589	0.79	0.031	HB-RW-GE	1.264	0.89	0.103
KOF	0.665	0.89	0.017	IFO	1.332	0.94	0.002
PMI	0.581	0.78	0.002	ZEW	1.444	1.01	0.278
CE-CH	0.776	1.04	0.613	CE-GE	0.799	0.51	0.038
	Last-available GDP				Last-available GDP		
	RMSFE	Ratio	p-value		RMSFE	Ratio	p-value
AR(1)	1.274	1.05		AR(1)	1.886	1.11	
AR(2)	1.212	1.00		AR(2)	1.701	1.00	
NZZ-RW	0.971	0.80	0.000	HB-RW	1.633	0.96	0.124
NZZ-RW-CH	0.981	0.81	0.000	HB-RW-GE	1.595	0.94	0.139
KOF	1.022	0.84	0.000	IFO	1.426	0.84	0.000
PMI	0.952	0.79	0.000	ZEW	1.770	1.04	0.636
CE-CH	1.084	0.89	0.029	CE-GE	1.147	0.67	0.029

[a] See notes for Table 2.

about 20% compared to that of the AR(2) model. We also notice a slight increase in forecasting performance of the ARDL model augmented with the *KOF* indicator.

For Germany the same conclusion of superior forecasting performance of the consensus forecasts is confirmed. Similarly as before, the *Ifo* indicator delivers the second-best result. For the *HB*-based R-word indices we find that using recursive estimation scheme resulted in a decrease in reported RMSFE compared with results reported in Table 2. However, due to the fact that also for the benchmark model reported RMSFE decreased, the relative forecasting performance seems to be less impressive than before. We record at most a 10% decrease in RMSFE when compared to that of the AR(2) model. But still the R-word indices fare better than the model augmented with the *ZEW* indicator, for which the null hypothesis cannot be rejected at the usual significance level.[13]

The forecasting performance of the alternative models is illustrated in Figures 14—21. The

[13] In order to see whether this results is due to restrictive lag structure that we allow in the ARDL(2,0) model, we repeated the forecasting exercise by allowing more rich dynamic structure by allowing up to four lags of the *ZEW* indicator. For each estimation window the preferred model was selected by minimising the Schwarz Information Criterion. The resulting RMSFE from this ARDL(2,4) model are 1.449 and 1.799 for first- and last-available data, respectively.

upper panels display official GDP growth estimates (first releases on the left and last-available estimates on the right) together with the benchmark model forecasts and forecasts produced with the help of extraneous information. The lower panels report the absolute forecast error for the corresponding models.

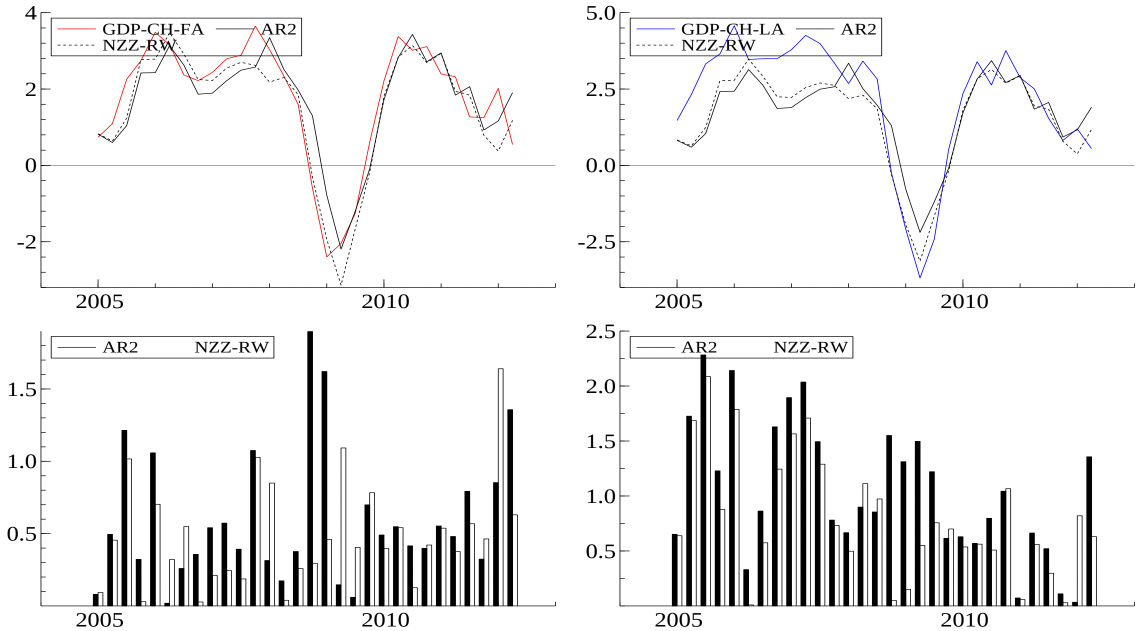

Figure 14: Quarterly year-on-year real GDP growth; Expanding estimation window: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-NZZ-RW models

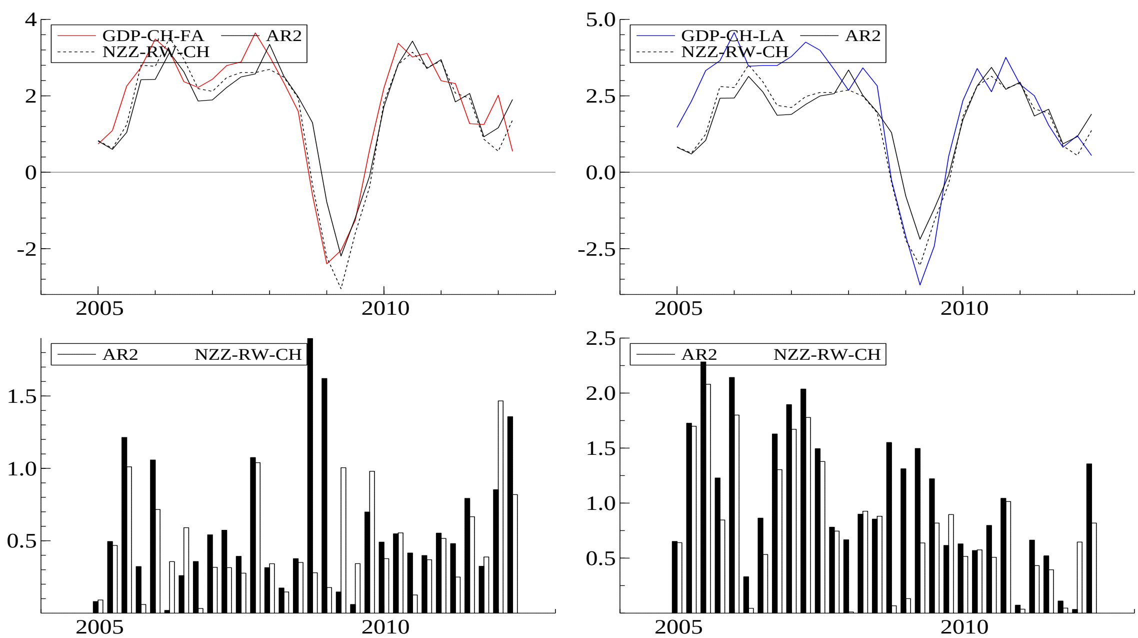

Figure 15: Quarterly year-on-year real GDP growth; Expanding estimation window: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-NZZ-RW-CH models

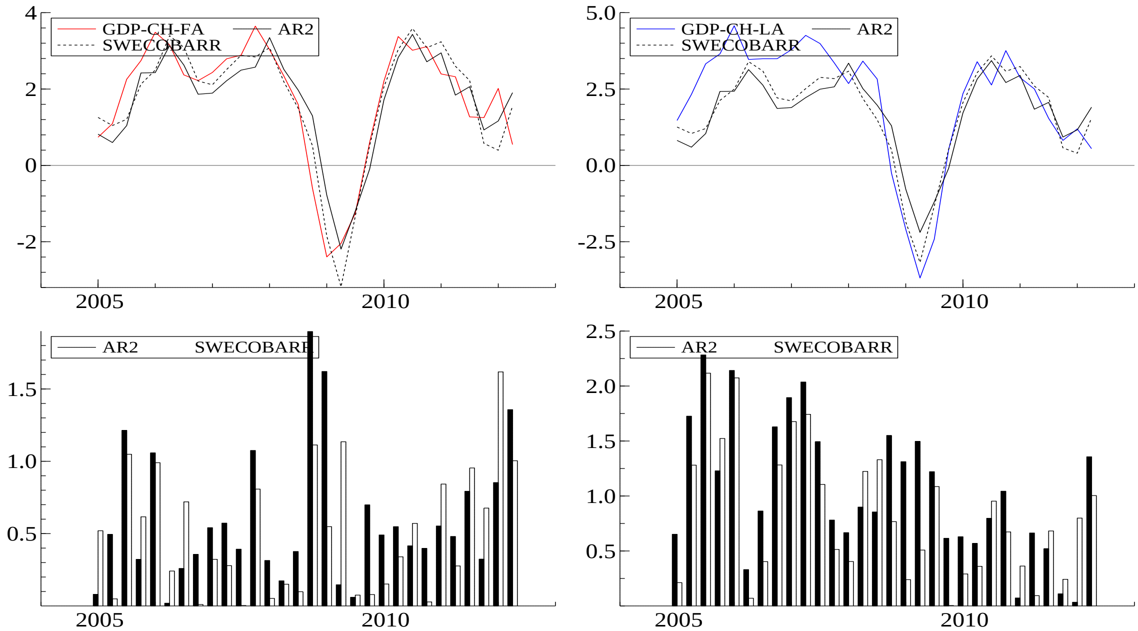

Figure 16: Quarterly year-on-year real GDP growth; Expanding estimation window: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-KOF models

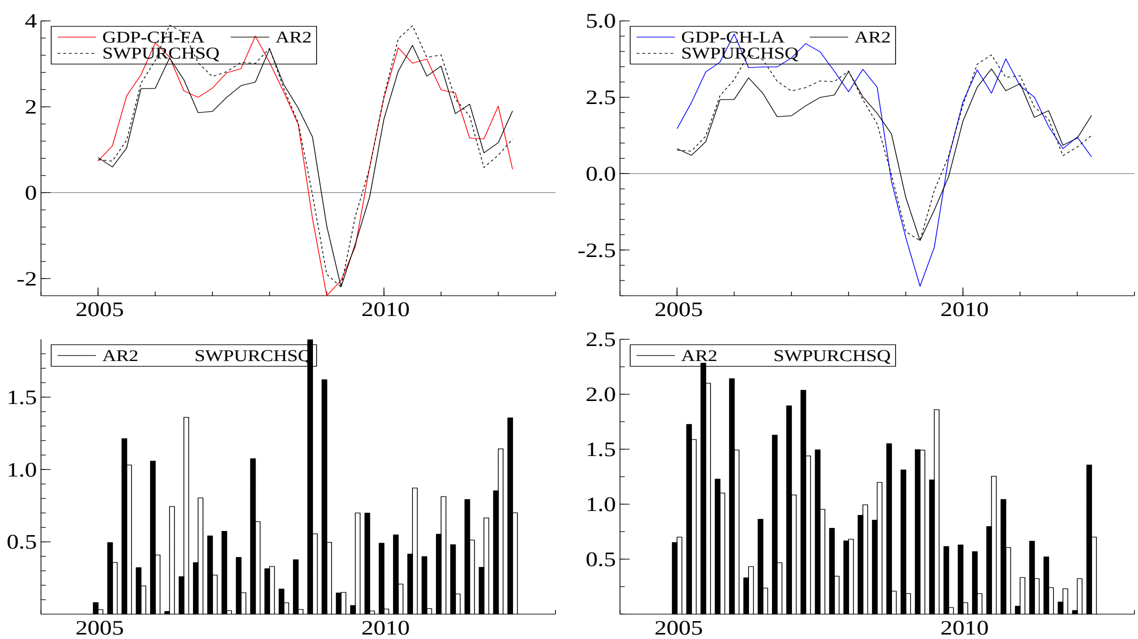

Figure 17: Quarterly year-on-year real GDP growth; Expanding estimation window: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-PMI models

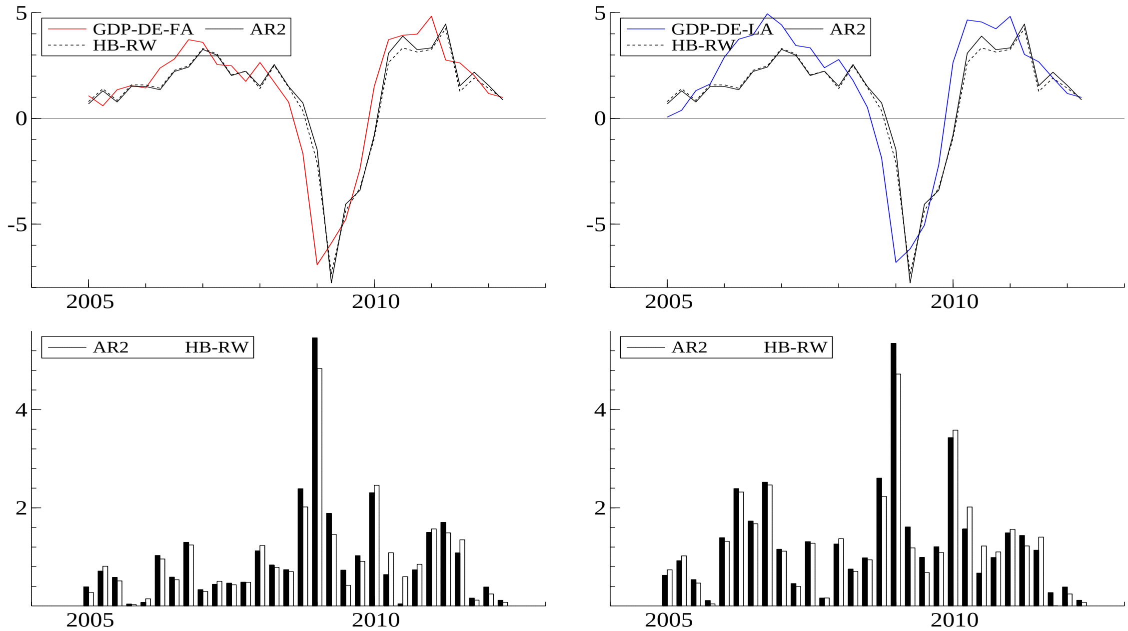

Figure 18: Quarterly year-on-year real GDP growth; Expanding estimation window: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-HB-RW models

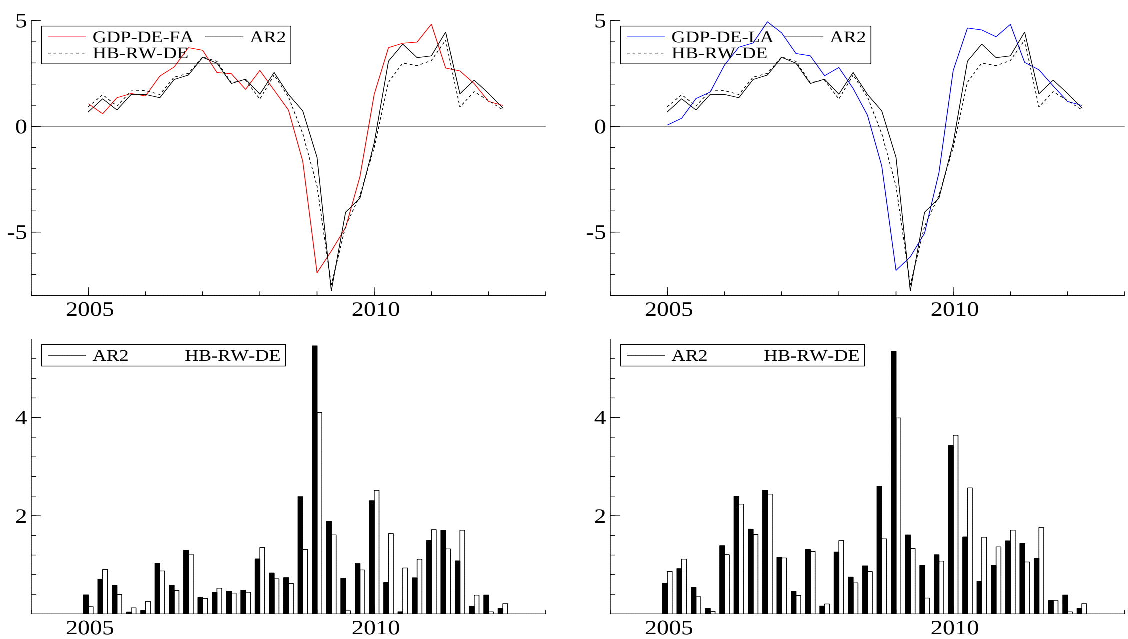

Figure 19: Quarterly year-on-year real GDP growth; Expanding estimation window: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-HB-RW-GE models

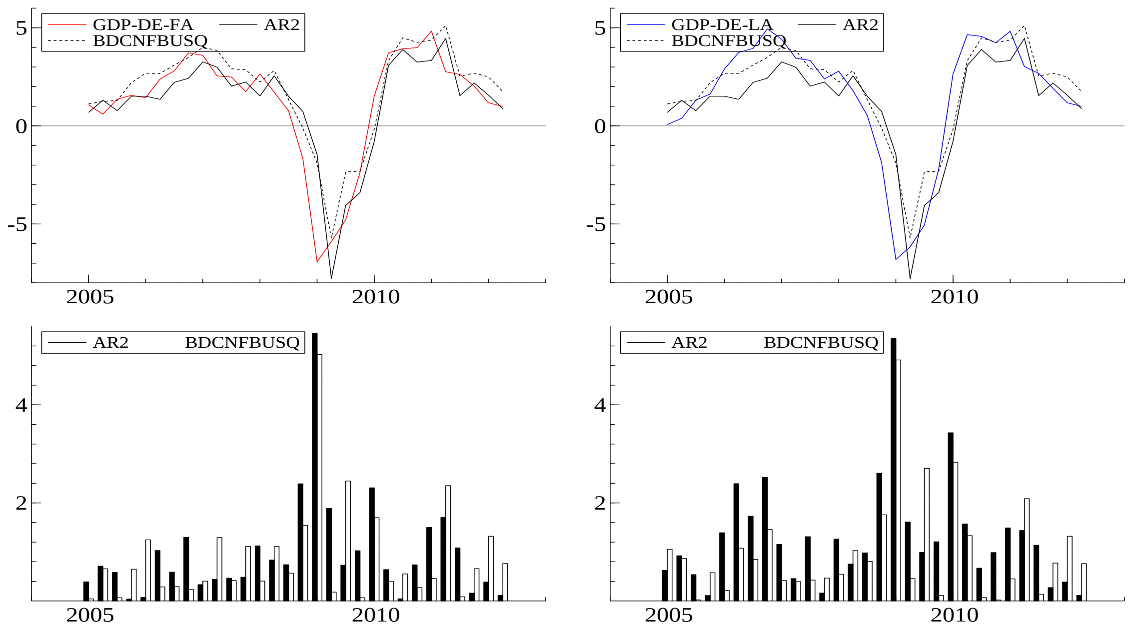

Figure 20: Quarterly year-on-year real GDP growth; Expanding estimation window: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-IFO models

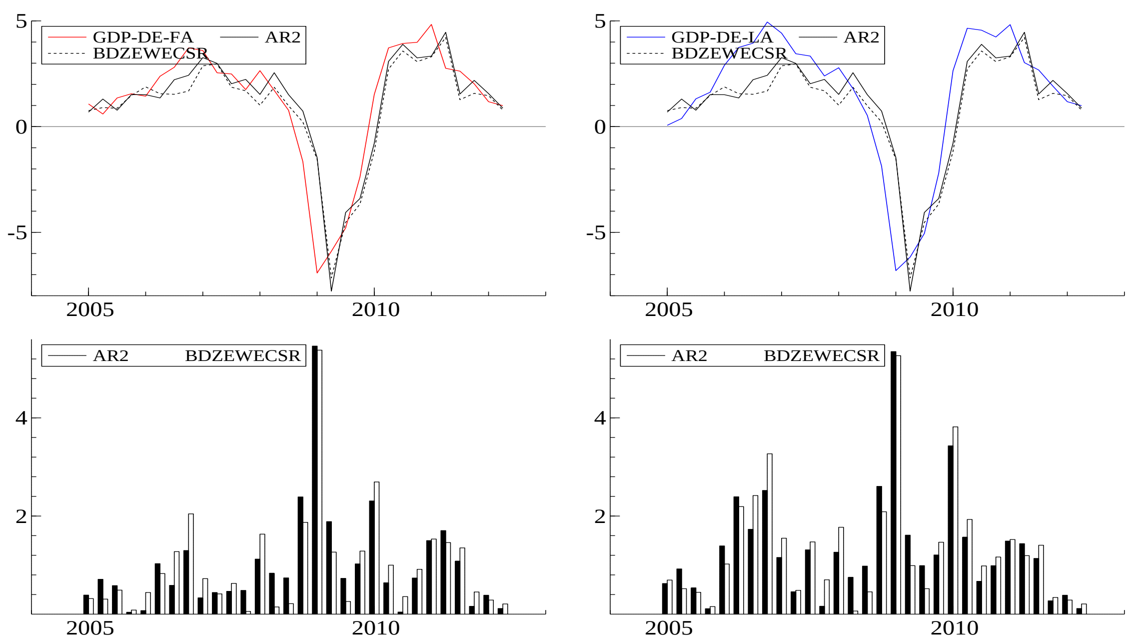

Figure 21: Quarterly year-on-year real GDP growth; Expanding estimation window: Forecasts and absolute forecast errors from AR(2) and ARDL(2,0)-ZEW models

5.2 Box-Cox transformation

In this subsection we explore the possibility of an alternative representation of the R-word index based on the Box-Cox transformation:[14]

$$z_t^{(\lambda)} = \begin{cases} \dfrac{z_t^{\lambda} - 1}{\lambda} \text{ if } \lambda \neq 0 \\ \log z_t \text{ if } \lambda = 0. \end{cases} \quad (7)$$

This Box-Cox transformation, depending on values of λ, shrinks (for $\lambda < 1$) or stretches ($\lambda > 1$)the original time series. For $\lambda = 1$ the original time series is preserved.

We replicated the forecasting exercise in Section 4 for each of our R-word indices *(NZZ-RW, NZZ-RW-GE, HB-RW, HB-RW-GE*) by varying values of λ in the range $[0, 3]$ for the Swiss and $[0, 4]$ for the German time series. The resulting RMSFEs, scaled by the value of RMSFE reported in Table 2, are shown in Figures 22 and 23 for Switzerland and Germany, respectively. For Switzerland we observe very limited improvement over the results reported above in Table 2. The largest effect we observe for *NZZ-RW* when comparing forecast accuracy using first-available (FA) data. However, even in this case the improvement in RMSFE is very small — 0.645 observed for $\lambda = 1.7$ against 0.650 reported in Table 2. This allows us to conclude that the Box-Cox transformation is not necessary when predicting GDP growth in Switzerland with the R-word index.

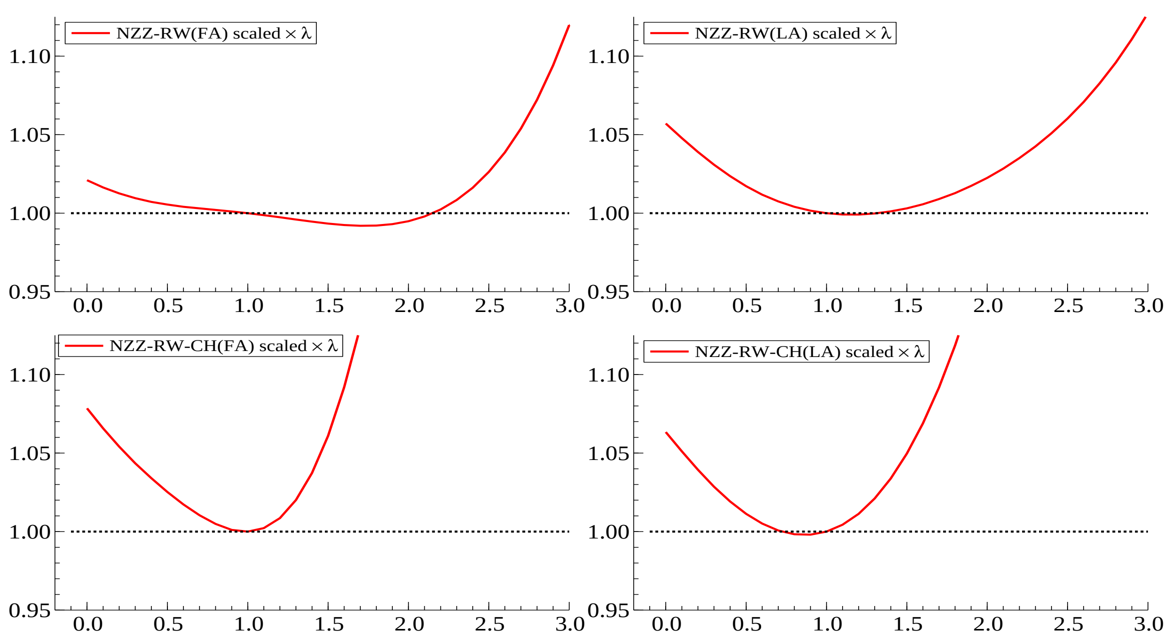

Figure 22: Box-Cox transformation for NZZ-based R-word indices: RMSFE for $\lambda \in [0, 4]$, scaled by the respective value of RMSFE from Table 2, corresponding to $\lambda = 1$.

[14]Since this transformation requires positive keywords counts, we substituted zero counts with one.

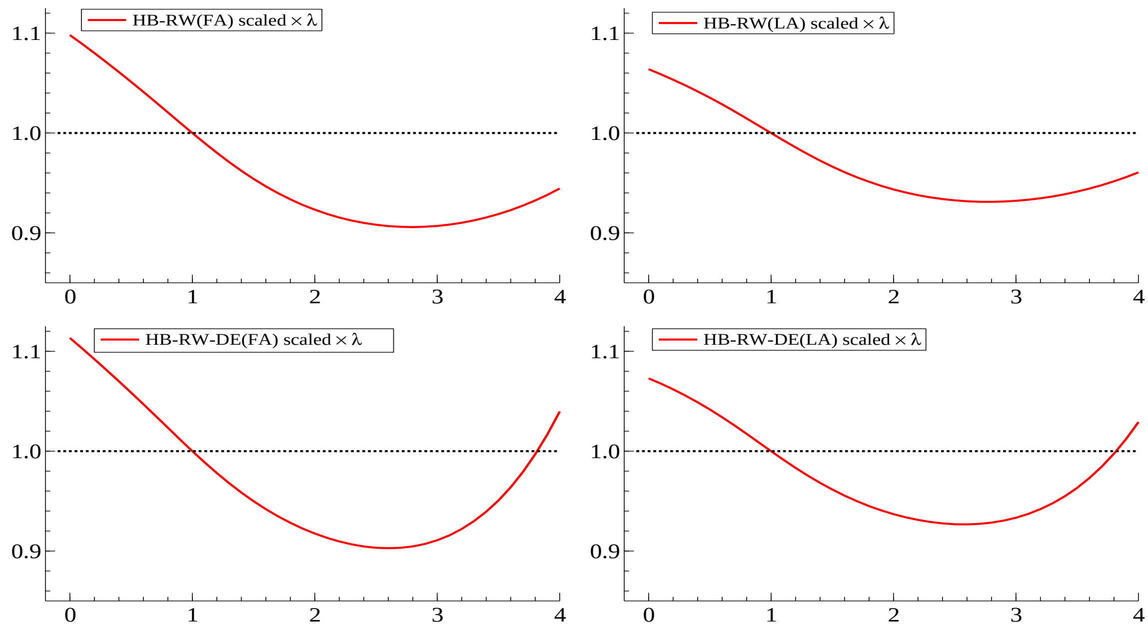

Figure 23: Box-Cox transformation for HB-based R-word indices: RMSFE for $\lambda \in [0, 4]$, scaled by the respective value of RMSFE from Table 2, corresponding to $\lambda = 1$.

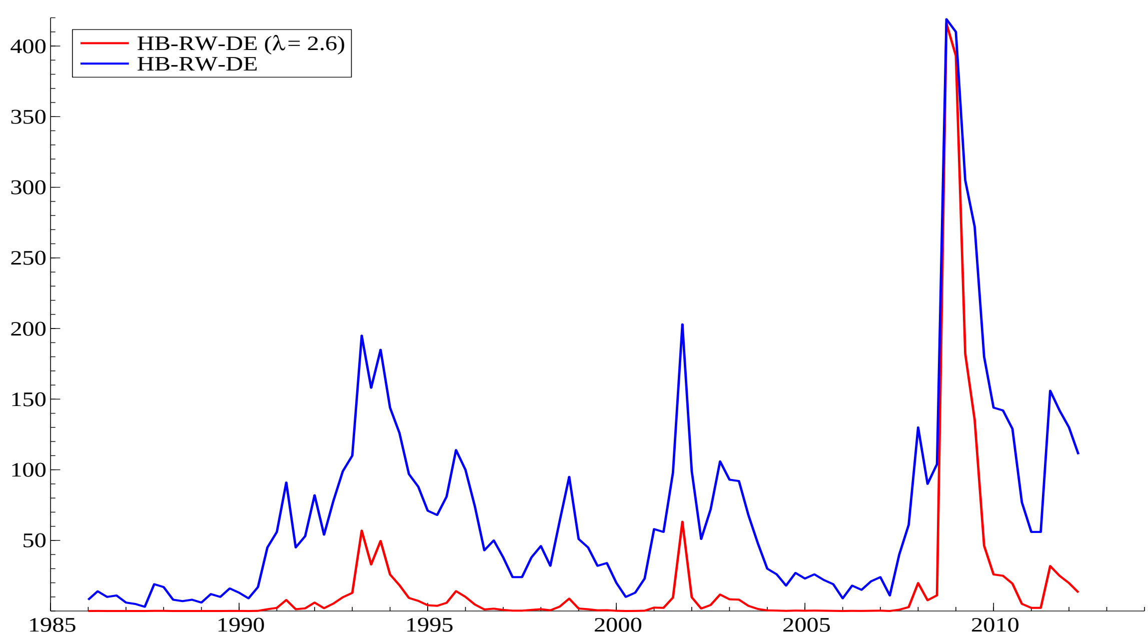

Figure 24: Comparison of original time series HB-RW-GE and after application of the Box-Cox transformation with $\lambda = 2.6$. Both time series adjusted to have the same range.

The outcome is, however, completely different for Germany. For values of λ between two and three we observe further reduction in RMSFE up to 10%. For example, for HB-RW-GE for $\lambda = 2.6$

the achieved RMSFE is 1.252 in comparison to 1.386 reported in Table 2, when comparing forecast accuracy using first-available (FA) GDP estimates. Both time series (untransformed and Box-Cox transformed with $\lambda = 2.6$) are shown in Figure 24, where we, for the sake of comparison, adjusted the time series to have the same range. In comparison to the untransformed time series the Box-Cox transformation in question emphasises observations with relatively large values at the expense of observations with relatively small values. The effect of this transformation can be interpreted within a model with two regimes. In the first regime, when the economy is in the expansion phase, the influence of variation in the R-word indicator is minimised. But in the second (recessionary) regime, the variation of the indicator contributes significantly to regression explanatory power and, as a result, to accuracy of out-of-sample forecasts. Based on this finding, it would be of a further interest to fit non-linear models like Threshold Autoregressive models (see, e.g., Hansen, 2011, for a recent review) in order to formally investigate our conjecture. But we leave it for our further research.

5.3 Google Trends

The rise of the Internet spurred a growing body of literature which uses this rich source of information for forecasting economic variables (see, e.g., Choi and Varian, 2012, for a recent review). Since January 2004 aggregated results of queries using the *Google* search engine are made publicly available at the following homepage `http://www.google.com/trends/`. These results are typically published at weekly frequency after the number of searches for a particular keyword exceeds some unspecified threshold. In order to control for increasing search volumes, the number of searches including a particular keyword is divided by the total number of queries that were submitted within past week. Then, the data are scaled by the maximum over the selected period. The query data can be differentiated according to geographical areas from which the internet searches originated.

In this subsection we investigate whether the queries for the keyword "Recession" (in German, "Rezession") submitted from Germany and Switzerland are useful for tracking the business cycle dynamics like the newspaper-based indices. The respective web-based indices are reported at a weekly frequency for Germany and a monthly frequency for Switzerland due to the comparatively small search volume. Both time series were aggregated to the quarterly frequency to match that of the GDP data. These indices are displayed at Figure 25. Apart from minor short-run deviations, both time series move very synchronously featuring two peaks in 2008Q1 and 2008Q4. There is also a surge in the second half of 2011 reflecting growing concerns on the state of the economy and the escalating sovereign debt crisis.

For the sake of comparison we placed both newspaper- and web-based R-word indices in Figure 26. For Switzerland the recorded peaks of the web-based index are well above those for the *NZZ*-based index, indicating more intense reaction to negative news. But once the web-based index reaches its height in 2008Q4 it also falls fast to the levels before peak whereas the *NZZ*-based index succeeds much slower. For Germany the relative sizes of the peaks are very similar, but also in this case we observe much faster return of the web-based index to its pre-crisis level. One could

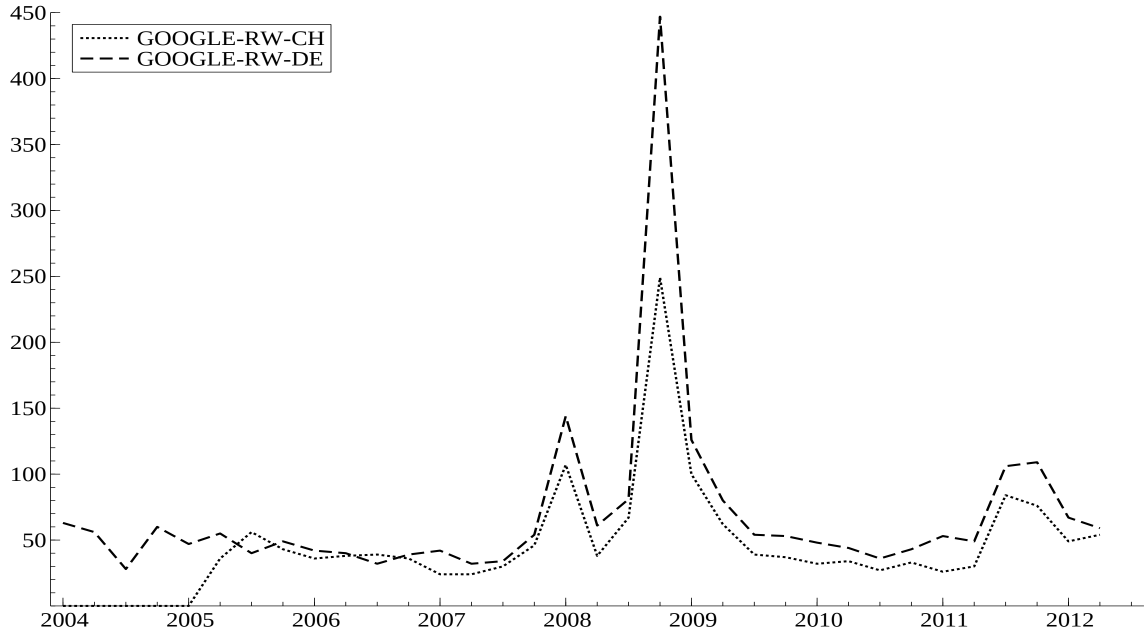

Figure 25: Google Trends indices based on the web searches of "Rezession" for Germany and Switzerland, aggregated to the quarterly frequency.

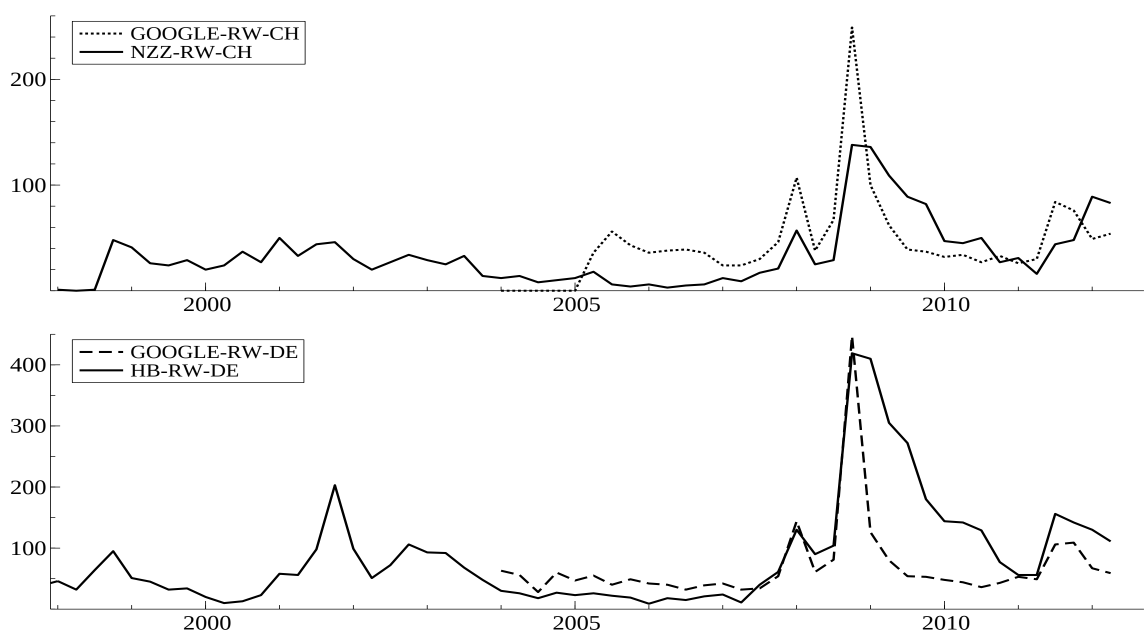

Figure 26: Newspaper-based and Google Trends R-word indices for Switzerland (upper panel) and Germany (lower panel).

speculate that a such of pattern reflects the fact the quality newspapers provide more thorough and comprehensive coverage of ongoing troubles in the economy, whereas a web search inquiry might only be triggered by the reported first headlines mentioning a recession.

Since the Google Trends indices are available from 2004Q1, in order to allow for a sufficient number of observations for parameter estimation, we shortened forecast evaluation sample to 2007Q1—2012Q2. The first estimation window for ARDL model augmented with the web-based R-word indices is 2004Q1—2006Q4. Because of the relatively short initial estimation window, we apply the recursively expanding parameter estimation scheme, described in Subsection 5.1.

The results are reported in Table 5. Both for Switzerland and Germany, the RMSFEs of the Google Trends indices are larger than the corresponding values of the RMSFE of the benchmark AR(2) model. This holds for the actual values of these indices as well as for the Box-Cox transformation applied to them, which brings about a slight increase in forecast accuracy. The optimal values of the Box-Cox component λ^* are close to zero, when it corresponds to the logarithmic transformation, see Equation (7). At the same time, the RMSFEs associated with the models augmented with the newspaper-based R-word indices are lower than those observed for the benchmark model, conforming with the earlier reported results in Section 4 and Subsection 5.1. Thus, our earlier results are also robust to changes in the forecast evaluation sample.

Table 5: Out-of-sample forecast accuracy: Expanding estimation window, 2007Q1–2012Q2

Switzerland			Germany		
	First-available GDP RMSFE	Ratio		First-available GDP RMSFE	Ratio
AR(2)	0.788	1.00	AR(2)	1.606	1.00
GOOGLE-RW-CH	1.077	1.37	GOOGLE-RW-GE	2.233	1.39
GOOGLE-RW-CH ($\lambda^* = -0.2$)	1.033	1.31	GOOGLE-RW-GE ($\lambda^* = -0.2$)	2.049	1.28
NZZ-RW-CH	0.609	0.77	HB-RW-DE	1.421	0.88
	Last-available GDP RMSFE	Ratio		Last-available GDP RMSFE	Ratio
AR(2)	1.083	1.00	AR(2)	1.762	1.00
GOOGLE-RW-CH	1.240	1.14	GOOGLE-RW-GE	2.296	1.30
GOOGLE-RW-CH ($\lambda^* = 0.1$)	1.145	1.06	GOOGLE-RW-GE ($\lambda^* = -0.3$)	2.109	1.20
NZZ-RW-CH	0.833	0.77	HB-RW-DE	1.640	0.93

6 Conclusion

In this paper we investigate the predictive power of R-word indices for forecasting real GDP growth both in Germany and Switzerland. These R-word indices are based on a keyword search in the newspapers *Handelsblatt* and *Neue Zürcher Zeitung*, respectively. We compare their predictive ability with that of a univariate autoregressive benchmark model as well as forecasts obtained with help of the most popular economic indicators in both countries: the *Ifo Business Climate Index* and the *ZEW Indicator of Economic Sentiment* for Germany, and the *KOF Economic Barometer* and the *Purchasing Manager Index in manufacturing* for Switzerland. In addition, forecasts generated with models augmented with R-word indices are compared with consensus forecasts published by *Consensus Economics Inc.* The forecasting performance of alternative indicators is under investigation during the sample period from 2005Q1 until 2012Q2, that includes the period of the financial crisis of 2008/2009. We compare the forecasting ability of different indicators using both rolling and expanding estimation windows. In addition, we verify whether the application of the Box-Cox transformation of keyword counts brings about any boost in forecast accuracy of the newspaper-based indicators. And we also compare our models with similar keyword searches in *Google Trends*.

For Switzerland, we find that the R-word indices significantly improve upon forecasting accuracy of the univariate model, confirming the results of Iselin and Siliverstovs (2013). More importantly, their forecast accuracy is comparable to that provided by the traditional economic indicators. In addition, the newspaper-based forecasts are more accurate than the consensus forecasts for Switzerland and they also fare much better than the internet-based recession-word index obtained from *Google Trends*.

For Germany, we find weaker support for the forecasting ability of the R-word indices over that of the univariate model. Using a rolling estimation window set-up the corresponding null hypothesis could be rejected at the 10% level. For the expanding estimation window setup we find that the null hypothesis can be barely rejected at the 10% significance level when comparing forecast accuracy of first-released estimates of GDP growth. When we attempt to forecast last-available estimates of GDP growth the corresponding null hypothesis could not be rejected. An interesting result is the fact that for Germany the consensus forecasts offer a by far superior forecasting ability than any other model considered in this paper. The second best forecast accuracy is provided by information contained in the *Ifo Business Climate Index*. To our surprise, the *ZEW Indicator of Economic Sentiment* fares relatively poor in our forecasting contest. One possible explanation is that the *ZEW* indicator targets rather turning points of stock market indices rather than business cycles. Similarly, the index based on *Google Trends* data fares worse than newspaper-based RWI and the traditional economic indicators.

In sum, we demonstrate that the newspaper-based recession-word indices provide an alternative and useful source of information to that already available from traditional economic indicators. After all, professional forecasters shape their vision of the current and future economic developments, among other things, on the basis of stories appearing in the press. Thus, the R-word indices enrich the palette of available indicators to forecasters and help shape a holistic view of ongoing economic processes. As Baumohl (2013, p. 7) puts it: "[In the US] at least four key economic indicators are released on a weekly basis, 43 every month, and nine each quarter. Do we really need so many measures? Absolutely... No single indicator can provide a complete picture of what the economy is up to. [...] At best, each indicator can give you a snapshot of what conditions are like within a specific sector of the economy at a particular point in time. Ideally, when you piece all these snapshots together, they should provide a clearer picture of how the economy is faring and offer clues on where it is heading."

References

Ammann, M., R. Frey, and M. Verhofen (2013). Do newspaper articles predict aggregate stock returns? *Journal of Behavioral Finance*, forthcoming.

Askitas, N. and K. F. Zimmermann (2009). Google econometrics and unemployment forecasting. *Applied Economics Quarterly (formerly: Konjunkturpolitik) 55*(2), 107–120.

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

Baumohl, B. (2013). *The Secrets of Economic Indicators* (3rd edition ed.). Wharton School Publishing.

Breitung, J. and D. Jagodzinski (2001). Prognozeneigenschaften alternativer Indikatoren für die Konjunkturentwicklung in Deutschland. *Konjunkturpolitik 47*(4), 292–313.

Choi, H. and H. Varian (2012). Predicting the present with Google Trends. *Economic Record 88*(s1), 29.

Choi, H. and H. R. Varian (2009). Predicting the present with Google Trends. *Google Research Blog*.

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.

D'Amuri, F. and J. Marcucci (2009). 'Google it!' Forecasting the US unemployment rate with a Google job search index. ISER Working Paper Series 2009-32, Institute for Social and Economic Research.

Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics 13*(3), 253–63.

Doms, M. and N. J. Morin (2004). Consumer sentiment and the media. *FRBSF Economic Letter* (Oct 22).

Doornik, J. A. (2007). *Object-Oriented Matrix Programming Using Ox*. London: Timberlake Consultants Press and Oxford: www.doornik.com.

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.

Economist (2008). Warning lights. The R-word index. Recession talk is rising. http://www.economist.com/node/10498907, Jan 10th 2008.

Engelberg, J. E. and C. A. Parsons (2011). The causal impact of media in financial markets. *Journal of Finance 66*(1), 67–97.

Grossarth-Maticek, J. and J. Mayr (2008). Medienberichte als Konjunkturindikator. *Ifo Schnelldienst 61*(7), 17 – 29.

Hansen, B. E. (2011). Threshold models in economics. *Statistics and Its Interface 4*(2), 123–127.

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

Harvey, D. I., S. J. Leybourne, and P. Newbold (1998). Tests for forecast encompassing. *Journal of Business & Economic Statistics 16*(2), 254–59.

Iselin, D. and B. Siliverstovs (2013). The R-word Index for Switzerland. *Applied Economics Letters 20*(11), 1032–1035.

Kholodilin, K. A., M. Podstawski, and B. Siliverstovs (2010). Do Google searches help in nowcasting private consumption? A real-time evidence for the US. KOF Working papers 10-256, KOF Swiss Economic Institute, ETH Zurich.

Kholodilin, K. A. and B. Siliverstovs (2006). On the forecasting properties of the alternative leading indicators for the German GDP: Recent evidence. *Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik) 226*(3), 234–259.

Kholodilin, K. A. and B. Siliverstovs (2012). Measuring regional inequality by internet car price advertisements: Evidence for Germany. *Economics Letters 116*(3), 414–417.

Koenig, E. F., S. Dolmas, and J. Piger (2003). The use and abuse of real-time data in economic forecasting. *The Review of Economics and Statistics 85*(3), 618–628.

Marcellino, M. and C. Schumacher (2010). Factor MIDAS for nowcasting and forecasting with ragged-edge data: A model comparison for German GDP. *Oxford Bulletin of Economics and Statistics 72*(4), 518–550.

McCracken, M. W. (2007). Asymptotics for out of sample tests of Granger causality. *Journal of Econometrics 140*(2), 719–752.

Preis, T., H. S. Moat, and H. E. Stanley (2013). Quantifying trading behavior in financial markets using Google Trends. *Scientific Reports 1684*(3), 1–6.

Siliverstovs, B. (2013). Dating business cycles in historical perspective: Evidence for Switzerland. *Jahrbücher für Nationalökonomie und Statistik (forthcoming).*

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. (2011a). Nowcasting private consumption with TV sentiment. KOF Working Papers 293, KOF Swiss Economic Institute.

Uhl, M. W. (2011b). Reuters sentiment and stock returns. KOF Working papers 288, KOF Swiss Economic Institute, ETH Zurich.