Kredit und Kapital, 45. Jahrgang, Heft 1, Seiten 1–26
Abhandlungen

The Financial Crisis from a Forecaster's Perspective

By Katja Drechsel, Halle, and Rolf Scheufele, Zürich

I. Motivation

The financial crisis of 2008/2009 and the followed recession in Germany are very distinct from past recessions. It turned out having been the most severe recession since World War II. Production declined by about 7% within one year measured by GDP. Industrial production was hit even stronger and shrunk by 20% during the same period. The origins of this slump can be found in the US financial and banking sector in 2007. The following credit crunch basically had implications on all industrialized countries. Germany, a country that can be characterized as an export oriented industry, had been heavily affected by the shrinking demand and thus saw one of the most pronounced drop in production among all developed countries.

Despite the exceptional magnitude of the recession, many professional forecasters did not foresee the current recession. Thus professionals have been highly criticized for not anticipating the huge downturn neither in time nor in extent for a long time (see e.g. *Koll* et al. (2009) for a discussion). Because many professionals use leading indicators to assess the current and future situation of the economy, we ask how leading indicator forecasts perform during this exceptionally heavy recession. Therefore, we analyze how econometric models that use leading indicator information have performed during the crisis.

The literature on the performance of leading indicators for Germany is large (see *Kholodilin/Siliverstovs* (2006) and the references therein). However, none of the authors draw special attention on the forecasting properties of leading indicators during a pronounced recession. In contrast, there is also some literature on forecasting recessions with non-linear models such as probit models (see *Fritsche/Kuzin* (2005)) that concentrates on the probability of turning into a recession. However, this approach does not provide a quantitative forecast of output growth which is more informative.

The first contribution of this paper is to document how professional forecasters performed during the financial crisis. We document that no one anticipated the recession early and furthermore, all professional forecasters underestimated the impact on production. Motivated by the work of *Stock/Watson* (2003a) who analyzed the performance of leading indicators during the 2001 recession in the US, we ask whether leading indicators provide useful information before and during the crisis and hence, can be conducive to an adequate policy making.

We investigate a set of prominent leading indicators for Germany in the emergence of the recession, consisting of survey based measures, financial market indicators, real activity variables and composite leading indicators. We analyze the performance of each indicator in forecasting both (i) GDP and (ii) industrial production (IP) from 1 to 4 quarters ahead. Since the origin of the recession sprouted out in the financial sector, we particularly analyze financial indicators as predictors for real activity (for a literature review see *Stock/Watson* (2003b)). One central contribution we make is that we consider not only linear models for output growth, but also non-linear models that take into account a threshold effect (threshold leading indicator models). Furthermore, we augment our analysis to forecast combinations. Since in practice individual indicators are not used in isolation, forecast combination schemes provide an efficient way to summarize the results given by many different models. Finally, we compare leading indicator forecasts (single and pooled) with forecasts from professional forecasters. To evaluate the resulting forecasts we apply a non-parametric test based on signed-ranks (with a modification also suited for autocorrelated errors) that can deal with the small out-of-sample forecast period in our case.

The paper is structured as follows: The next section briefly describes the 2008/2009 recession in Germany and investigates the professional forecasts during the crisis episode. Section III. provides an overview on the leading indicators we use for our forecast analysis and the model set up for the forecast experiment. Results based on linear and nonlinear models are discussed as well. Section IV. presents the performance of the pooled forecasts. Section V. compares leading indicator forecasts with those of professional forecasters. Section VI. summarizes and concludes.

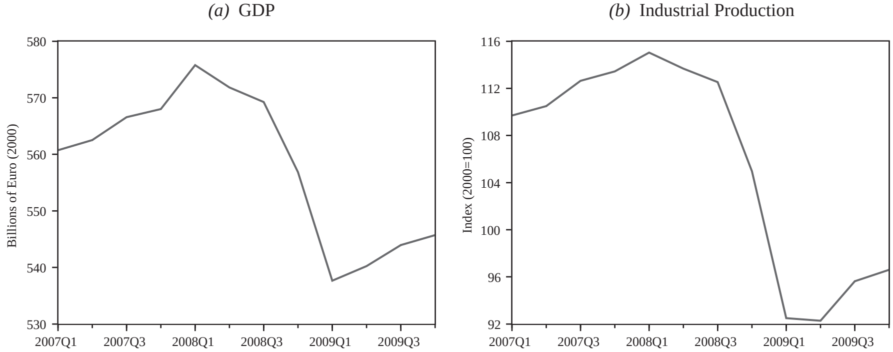
II. The 2008/2009 Recession and Evidence from the Consensus Economics Forecasters

Figure 1 shows GDP and industrial production for the German economy during the crisis period. Both series peaked in the first quarter of 2008, then output declined over four consecutive quarters.

With the most sizable downturn in output since decades, GDP and IP have seen the biggest slump during the two winter quarters. In the second quarter of 2009, GDP shows some recovery and again a positive quarterly growth rate. At the same time IP dropped slightly further, but also has shown signs of a recovery since May 2009. Despite some positive signs after the first quarter in 2009 the average growth rate of GDP is strongly negative and in the range of –5 %. Since the manufacturing sector is much more affected by this slowdown than any other sector, IP was expected to fall more severe – on average, forecasts were around –17 % for 2009.

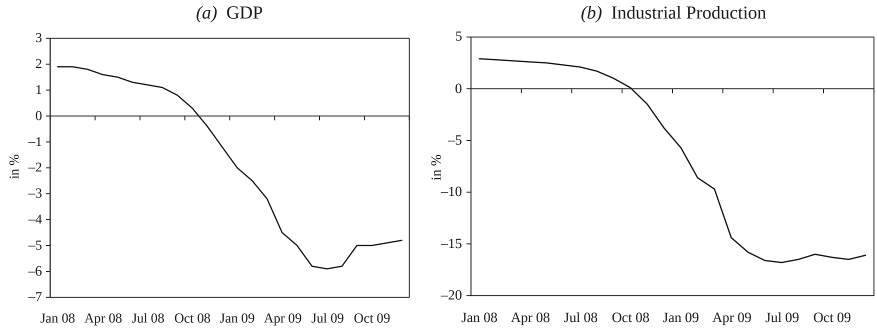
During the year 2010 the German economy continued to recover. Although there is no official business cycle committee in Germany, one can simply define the recession from peak to trough of production. Hence, the recession would be judged from 2008q1 through 2009q1. However, in this paper we take a broader view and consider some additional quarters before and after this narrow definition as our period of interest; namely we analyse the period between 2007q1 and 2009q4.

Each month, Consensus Economics surveys a large panel of financial and economic experts about their estimates on important macroeconomic



Source: Fachserie 18, Reihe 1.3, release November 2010, Federal Statistical Office Germany.

Figure 1: Key Indicators



Note: Annual GDP and industrial production Consensus Forecasts (average) for 2009 for Germany are shown.
Source: Consensus Economics (2009).

Figure 2: Consensus Forecasts for 2009

variables such as growth, inflation and interest rates. This survey is known as the *Consensus Forecast*. For Germany, about 30 institutions participate in this poll - mainly banks and economic research institutes. The monthly poll asks for the forecast of various macroeconomic variables for the current and following year.

Figure 2 shows the mean point forecast for the growth rates of GDP and IP for 2009. In January 2008, the mean GDP forecast for 2009 was slightly below 2 %. This indicates that professional forecasts did not take into account first hints of the upcoming financial crisis for their yearly growth projections. Until summer 2008, the mean GDP forecast was only revised down slightly to 1 %. The conventional view was that the world economy is experiencing a small temporary weakness which has also effects on the German economy. Things changed dramatically when Lehman went bankrupt at the end of September 2008. In November 2008, the mean GDP forecast turned negative and was further revised to –6 % in summer 2009. A similar pattern is also found for IP, where in November 2008 the mean forecast was below zero and was then gradually revised down to about –17 %.

This picture is supported by looking at year-on-year forecasts for each quarter. Table 1 shows for each survey date (in rows) all quarterly forecasts made up to the end of 2009. While in 2008q1 all forecasts were relatively homogeneous between 1.3 % and 1.9 %, in the second quarter a weakness was expected for the first two quarters 2009. In 2008q3, a few weeks before the Lehman breakdown, panelists reported a negative year-

Table 1

Quarterly GDP Forecasts

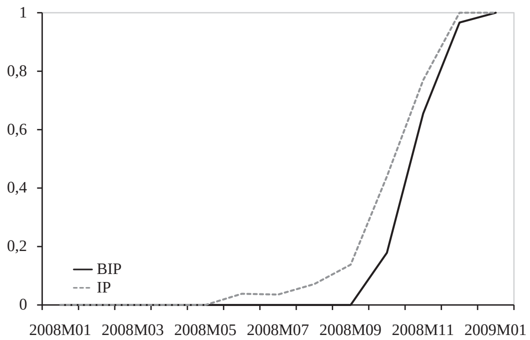
			Forecast horizon											
			2007				2008				2009			
			Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Time period	2007	Q1	3.0	2.2	1.9	1.5	1.8	2.0	2.0	2.0				
		Q2	*3.6*	2.9	2.6	2.3	2.2	2.2	2.2	2.1				
		Q3		*2.5*	2.5	2.1	2.0	2.4	2.2	2.1	2.3	2.1		
		Q4			*2.5*	1.9	1.6	1.8	1.6	1.6	1.8	1.8		
	2008	Q1				*1.8*	1.3	1.6	1.3	1.4	1.6	1.7	1.9	1.9
		Q2					*2.6*	2.2	1.7	1.7	0.6	1.2	1.6	1.8
		Q3					*2.6*	*1.7*	1.2	0.9	–0.1	0.7	1.0	1.3
		Q4					*2.7*	*1.9*	*0.8*	–0.2	–1.9	–1.8	–1.2	–0.3
	2009	Q1							*0.8*	*–1.7*	–4.1	–3.6	–3.0	–1.0
		Q2							*0.8*	*–1.8*	*–6.9*	–6.6	–6.0	–3.7
		Q3									*–6.7*	*–5.9*	–4.9	–2.1
		Q4									*–6.7*	*–5.8*	*–4.8*	–2.0

Note: Quarterly expected and realized year-on-year percentage growth rates of real GDP are shown. The official release is given in italics. Figures are working-day adjusted.

Source: Consensus Economics (2009), Federal Statistical Office Germany (2007–2009).

on-year growth rate for 2009q1, but afterwards a relatively fast recovery. In the next subsequent quarters, the economic outlook dramatically worsened and a negative growth rate was reported for all upcoming quarters. However, for 2008q4 and the first half of 2009, the first numbers released clearly exceed the so far predicted figures. For instance, in 2008q4 the consensus forecasters expected a GDP growth of –1.9 % for the first quarter of 2009, which turned out to be –6.7 % based on the final release by the German Statistical Office. Analyzing the recession probability by the fraction of panelists who report a negative growth rate of GDP or IP for the year 2009, we find that while none of the participating institutions had expected a negative growth rate for 2009 until September 2008, this fraction increased rapidly until December 2008, where all participating institutions expected a recession in 2009. Looking at industrial production, some participants anticipated the recession earlier.[1]

[1] Interestingly, it is Lehman Brothers that already forecast negative growth for industrial production for 2009 in June 2008.



Note: Fraction of forecasters that predict a negative rate for annual GDP and industrial production growth for 2009 for Germany. The calculation takes into account the different number of the institutions participating at the Consensus Forecast.

Source: Consensus Economics (2009), own calculation.

Figure 3: Fraction of Panelists Expecting a Negative Growth Rate for 2009

Taken together, the professional forecasts indicate several facts: First, before the Lehman breakdown nobody expected a sharp slowdown. If anything, then a temporary weakness for the second half of 2008 or in the beginning of 2009 was anticipated. Secondly, after Lehman's bankruptcy, forecasters revised down their forecasts quickly, but still underestimated the severity of the recession. More recently we have seen some tendency that forecasters have started to revise up their growth figures. However, the aim of our study is not the analysis of the performance of Consensus Forecasts per se (see *Ager/Kappler/Osterloh* (2009)), but to show how they perform compared to selected leading indicators during this recession.

III. Forecasts Based on Individual Leading Indicators

It is well known that many institutions commonly use leading indicators in judging' the current and future situation of the economy. Thus, we also employ these indicators to produce forecasts for real economic activity. This procedure quasi mimics the process of forecasting by the professional forecasters. In what follows, we investigate a huge set of indicators and analyze which indicator has signaled the slowdown in production and which has not. Therefore, we use specifications within the class of linear as well as non-linear models.

1. Linear Models of Output Growth

For constructing leading indicator forecasts we follow standard practice (see e.g. *Stock/Watson* (2003b)) and estimate dynamic models where each model includes one single indicator (with potential lagged values). More specifically, we regress one to four quarters of seasonally adjusted output growth on its past growth rates and on lags of a candidate indicator (e.g. interest rates) over the period 1992q1-2006q4-h+1. Let $Y_t = \Delta \ln Q_t$ where Q_t is the level of output (either the level of real GDP or the index of IP) and let X_t be a candidate predictor.[2] As indicated by standard ADF unit root tests, the indicator variables can be all characterised by stationary behavior (see Table 6 in the Appendix). Y^h_{t+h} is the output growth over the next h periods (quarters) in terms of an annualized rate.[3]

Forecasts are based on a h-step ahead regression model:

$$Y^h_{t+h} = \alpha + \sum_{i=1}^{p} \beta_i Y_{t-i} + \sum_{j=k}^{q} \gamma_j X_{t-j} + \varepsilon^h_{t+h}, \tag{1}$$

where ε^h_{t+h} is an error term and α, β and γ are the regression coefficients to be estimated. Different from other studies we take into account the timely availability of the indicators (reflected in k). Depending on the publication lag of the candidate predictor, k varies from 0 to 1 for quarterly data.[4] The optimal number of lags in the quarterly analysis is restricted to $1 \leq p \leq 4$ and $0 \leq q \leq 4$ and are selected by the Schwarz criterion (SIC).

For the quasi real-time out-of-sample forecasting experiment we estimate eq.(1) only using data prior to the forecasting date by applying a recursive scheme.[5] The recursive estimation scheme implies that for each

[2] We take the data set as it was available in January 2011. All subsequent analyses are based on this publication date including the forecast evaluation step. We construct a quarterly IP series by taking monthly averages.

[3] $Y^h_t = (400/h)\ln(Q_t/Q_{t-h})$ for real GDP and industrial production, respectively.

[4] In order to guarantee comparability to the consensus forecast we consider all information for the ongoing quarter until the beginning of the respective third month.

[5] However, the simulated real-time forecasting scheme does not consider revisions of the data. This problem is of minor importance for the indicator variable, since financial market indicators or survey measures are hardly ever revised. For the dependent variables GDP and IP this can be an issue. In particular IP revisions can be substantial and therefore the performance can appear better than it

forecasting round we include one additional observation. One to four steps ahead forecasts are made for the period 2007q1 to 2009q4.

2. Non-linear Models of Output Growth

We also augment our analysis by including non-linear models which is novel in the context of leading indicator models of output growth for Germany. International evidence suggests that for some indicators it is more realistic to assume a non-linear relationship (see e.g. *Galbraith/Tkacz* (2000)). This seems to be evident particularly for interest rate spreads. Therefore, we follow *Clements/Galvao* (2006) and consider threshold models as originally proposed by *Tong* (1983). The resulting threshold leading indicator regressions can be formulated as

$$
\begin{aligned}
Y^h_{t+h} = {} & \left[\alpha_1 + \sum_{i=1}^{p} \beta_{1i} Y_{t-i} + \sum_{j=k}^{q} \gamma_{1j} X_{t-j}\right] I(z_{i,\,t-d} \le r) \\
& + \left[\alpha_2 + \sum_{i=1}^{p} \beta_{2i} Y_{t-i} + \sum_{j=k}^{q} \gamma_{2j} X_{t-j}\right] \left[1 - I(z_{i,\,t-d} \le r)\right] + \varepsilon^h_{t+h},
\end{aligned}
\tag{2}
$$

where $I(.)$ is an indicator function equal to 1 when $z_{i,\,t-d} \le r$, and equal to zero otherwise. d is the time delay and r the threshold value. Estimates for d, r, α_1, $\beta_{11},\ldots,\gamma_{11},\ldots,\alpha_2$, $\beta_{21},\ldots,\gamma_{21},\ldots,\gamma_{2q}$ are obtained by conditional least squares. This implies that conditional on the estimates of r and d, the remaining parameters are estimated by least squares. The parameters of r and d are defined as the values that minimize the sum of squared residuals over a grid of possible values.[6] For the sake of simplicity we take the same number of lags for the leading indicator and output growth which are chosen by SIC of the linear model.

3. Data Set

In this paper we consider several leading indicators that have been suggested in the literature.[7] The most prominent indicators used are sur-

might be in real time. For Germany, *Benner/Meier* (2005) as well as *Schumacher/Breitung* (2008) compare the performance of leading indicators with both real-time data and final revised data in a similar setting and conclude that the relative performance of indicators remains stable (also the absolute precision is somewhat lower with real-time data).

[6] The limits of the grid for the delay d are 1 (lower) and 2 (upper). The limits for the threshold r are such that each regime has at least 30% of the observations.

vey based measures such as the ifo business cycle climate index or the ZEW sentiment indicator. Another important group of leading indicators considered in this paper consists of financial market indicators. Since the origins of the analyzed recession can be found in the financial sector, we might expect some early warning signals particularly from these indicators.[8] The advantage of both financial market indicators and survey measures is their early availability and their mostly forward-looking character. In addition, these indicators are not revised.

Our dataset comprises 42 leading indicators from different categories: surveys, financial variables and real activity measures (new orders, labor market indicators and prices).[9] Seasonally adjusted series are used whenever available. All variables are made stationary if necessary.[10] Additionally, we apply stability tests for every linear indicator model. Since *Kholodilin/Silverstovs* (2006) document some instabilities in the forecasting performance of leading indicators and identify a break in 2001, we therefore calculated the F test for stability of the parameters against the alternative of a single break at unknown date. The supremum test (or Quandt-Andrews test) is used for this purpose (*Andrews* (1993)). The test employed for the first in-sample period (1992q1–2006q4) indicates that only for a small fraction of leading indicator models, i.e. less than 10% at the 5% level of significance, the stability tests reject the null which implies that instabilities are of minor importance for the sample under consideration (see Table 7 in the Appendix).

4. *Forecast Evaluation*

To assess the forecasting performance in detail, we investigate the forecast errors of the different models. More precisely, the relative root mean squared forecast error (RMSFE) of a candidate forecast i is compared with the univariate benchmark model. Let $\hat{Y}^h_{i,t+h|t}$ be the forecast of the

[7] There is a large literature on leading indicators for Germany, both for GDP and IP (see among others *Döpke/Krämer/Langfeldt* (1995); *Breitung/Jagodzinski* (2001); *Fritsche/Stephan* (2002); *Kholodilin/Silverstovs* (2006) or *Drechsel/Scheufele* (2012)). A more detailed description of the leading indicators can be also found in these references.

[8] Financial indicators as leading indicators for Germany have been discussed and analyzed by *Ragnitz* (1994), *Kirchgaessner/Savioz* (2001), *Sauer/Scheide* (1995), *Fritsche/Kuzin* (2005) and *Burgstaller* (2009).

[9] See Appendix Table 5 for an overview.

[10] Table 6 in the Appendix includes the results of the standard ADF unit root test.

realization Y^h_{t+h}, computed using data up to time t, based on the i^{th} indicator. $\hat{Y}^h_{0,t+h|t}$ is the corresponding benchmark autoregressive forecast. The relative RMSFE can then be expressed as

$$\text{relative RMSFE} = \frac{\sqrt{\sum_{t=T_1+h}^{T_2} \left(Y_t^h - \widehat{Y}^h_{i,t|-h}\right)^2}}{\sqrt{\sum_{t=T_1+h}^{T_2} \left(Y_t^h - \widehat{Y}^h_{0,t|t-h}\right)^2}}, \tag{3}$$

where $T_1 + h$ and T_2 are respectively the first and the last date for the forecasting exercise. Over the period 2006q4+h to 2009q4 the forecast models are evaluated. A value of the relative RMSFE less than one indicates that the candidate model has a smaller root mean square forecast error than the benchmark model. However, a value smaller than one could simply occur due to sampling variability. Furthermore, the RMSFE does not indicate whether this result is statistically significant. For this purpose, we apply the test for equal predictability (against the alternative that the candidate model has smaller forecast errors). Under squared loss we can define the loss differential as $d_{i0} = (e_i)^2 - (e_0)^2$ where e_i are the forecast errors of indicator model i and the benchmark model 0, respectively. Generally, when models are nested, standard tests are inappropriate since they do not take the estimation uncertainty of the parameters into account (see *West* (1996)). In our setting, the proportion of the sample for the out-of-sample experiment relative to the estimation sample is very small, thus we can ignore the effect of parameter estimation uncertainty (see *West* (2006)).

In order to handle the extremely small sample with only $12 - h$ observations, we make use of a non-parametric rank test — the Wilcoxon signed-rank test. This test is an exact test even in finite samples and does not require the normality condition. *Diebold/Mariano* (1995) document the favorable properties of this approach for testing the null of equal accuracy of two competing forecasts. However, the original test is only valid under the restrictive iid assumption. Since we also analyze multi-step ahead forecasts (when $h > 1$), where the forecast errors follow an MA($h - 1$) process per construction, we take the resulting autocorrelation pattern into account. *Diebold/Mariano* (1995) suggest to split the sample into h parts in order to have h subsamples where the individual observations are independent of each other. Under the assumption that the loss differential is ($h - 1$)-dependent, each of the following h sets of loss differentials will be free of serial correlation: $\{d_{i0,1}, d_{i0,1+h}, d_{i0,1+2h}, ...\}$, $\{d_{i0,2}, d_{i0,2+h}, d_{i0,2+2h}, ...\}, \ldots, \{d_{i0,h}, d_{i0,2h}, ...\}$. A test with size bounded by α can be obtained by performing h tests, each of size α/h on each of the h loss differential sequences and rejecting the null hypothesis if the null is rejected for any of the h samples.[11]

5. *Results*

Tables 2 and 3 reveal the evaluation of the individual leading indicator forecasts both for GDP as well as for industrial production one to four quarters ahead. Obviously, the average forecast errors are extremely large in absolute size. For GDP (and IP) the RMSFEs of the benchmark models range between 6.08 (19.14) and 4.56 (13.15) depending on the forecasting horizon. This is a result of the exceptional recession in 2008/2009 and the fact that forecast errors are largest at turning points (see e.g. *Zarnowitz* (1992), Section 13).

Using leading indicator models may result in a considerable gain in average forecasting performance as one might have expected (see Table 2). For the best linear models the RMSFE for both GDP and IP is about 35-40% lower as compared to the benchmark and in some cases the forecast errors are significantly smaller compared to the univariate model. This difference is huge since after the year 2000 it has been previously found that the forecasting performance of leading indicators for Germany has deteriorated remarkably and that they do not offer much gain against a univariate benchmark model (see e.g. *Kholodilin/Silverstovs* (2006); *Kuzin/Marcellino/Schumacher* (2009) and *Drechsel/Scheufele* (2012)).

Generally, we find that survey based forecasts dominate in forecast accuracy. For GDP, Purchasing Managers' Index for manufacturing, the confidence indicators provided by the European Commission and the ifo indicators provide the smallest forecasting errors (although only the ifo expectations in the manufacturing sector offer significant improvements). Also financial indicators, in particular risk spreads and the DAX provide relatively good forecasting performance. For industrial production at the

[11] Due to the small number of observations we can perform the rank test only for $h = 1$ and $h = 2$. We apply the one-sided test in order to investigate whether the forecast errors from leading indicator model i are smaller than the ones from the univariate benchmark model. The critical values for the Wilcoxon test in small samples are tabulated (see e.g. *Büning/Trenkler* (1994), Table H).

Table 2

Forecast Results for GDP and IP during the Crisis – Linear Models

	GDP				IP			
	h = 1	h = 2	h = 3	h = 4	h = 1	h = 2	h = 3	h = 4
	RMSFE				*RMSFE*			
AR	6.08	5.16	4.82	4.56	19.14	16.61	14.07	13.15
	RMSFE relative to AR Model				*RMSFE relative to AR Model*			
Interest Rates								
IS-3M	0.97	1.00	1.00	0.99	0.93	0.98	1.01	0.99
DIL-10	0.90	0.95	0.97	0.98	0.87**	0.95	1.00	0.99
Interest Rates Spreads								
SPR-10Y-3M	1.00	1.00	1.00	1.00	0.96	0.92	0.95	0.91
SPR-C-G	0.80	0.87	0.84	0.77	0.98	0.91	1.02	0.97
SPR-B-G	0.88	0.96	0.92	0.75	0.89**	0.99*	0.93	0.83
SPR-BF-G	1.35	2.01	2.65	1.21	1.51	2.03	2.34	0.94
Monetary Aggregates								
DLNM1	1.00	0.92	0.92	1.01	1.00	0.90	0.94	0.89
DLNM1R	1.00	0.92	0.91	0.89	1.01	0.90	0.95	0.91
DLNM2	1.00	1.01	1.00	1.00	1.00	1.00	1.00	1.00
DLNM2R	1.00	1.05	1.05	1.00	1.00	1.00	1.00	1.00
DLNM3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DLNM3R	1.00	1.07	1.09	1.07	1.00	1.01	1.00	1.00
Other Financial Indicators								
DLNDAX	0.80	0.82	0.85	0.89	0.89*	0.85*	0.90	0.89
VOLA1	0.98	0.95	1.00	1.00	0.90	0.95*	1.00	1.01
DLNEX	1.04	1.04	1.01	1.01	1.00	1.05	1.07	1.02
DLNEXR	1.00	1.01	1.01	1.01	1.00	1.08	1.05	1.01
DLNHWWI	0.99	1.04	1.03	1.06	0.77	0.87	0.99	0.98
DLNHWWIEX	0.87	0.93	0.96	0.96	0.82	0.90	0.97	0.96
DLNOIL	1.05	1.02	1.02	1.00	0.89	0.94	1.00	1.00
Survey Indicators								
IFO-C	0.73	0.77	0.80	0.85	0.75**	0.70	0.77	0.81
IFO-EXP	0.71	0.72	0.75	0.83	0.67***	0.70*	0.75	0.81
IFOM-C	0.72	0.76	0.81	0.86	0.66***	0.70*	0.78	0.82
IFOM-EXP	0.73*	0.75*	0.80	1.00	0.71**	0.70*	0.77	0.82
IFO-WC	0.74	0.75	0.79	0.89	0.82*	0.81*	0.79	0.83
IFO-WEXP	0.91	0.96	1.00	1.00	0.86	0.89	0.94	0.99
ZEW-EXP	0.82	0.86	0.85	0.99	0.99	0.86	0.98	0.96
ESI	0.69	0.79	0.83	0.88	0.74*	0.75*	0.85	0.86
ESI-INDU	0.68	0.80	0.85	0.90	0.62**	0.73*	0.84	0.86
ECCS99	0.78	0.88	0.95	0.97	0.87	0.81*	0.91	0.94
PMI	0.66	0.87	0.95	1.00	0.66**	0.75	0.89	0.93

	GDP				IP			
	h = 1	h = 2	h = 3	h = 4	h = 1	h = 2	h = 3	h = 4
	RMSFE				*RMSFE*			
AR	6.08	5.16	4.82	4.56	19.14	16.61	14.07	13.15
	RMSFE relative to AR Model				*RMSFE relative to AR Model*			
Real Economic Indicators								
DLNIP-VORL	0.99	1.05	1.00	1.00	0.93	0.98	1.03	0.99
DLNORD	0.75	1.03	0.98	0.88	0.71*	0.87	0.89	0.84
DLNORD-C	1.00	1.06	1.07	1.00	1.00	0.99	1.05	1.00
DLNORD-I	0.74	1.00	0.94	0.92	0.81	1.00	0.93	0.88
CAPA	0.77	0.94	1.00	1.00	0.73	0.89	0.93	0.91
DLNEW	1.01	1.02	1.00	1.03	1.00	1.00	1.05	1.05
DALQ	1.00	1.02	1.00	1.00	1.00	1.04	1.05	1.05
DLNVAC	0.91	0.98	0.97	0.95	1.02	1.06	1.00	0.97
DLNWHOUR	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DLNCPI	1.00	0.94	1.00	1.00	1.00	1.00	1.01	0.98
DLNCPI-EX	1.00	1.00	1.00	1.00	1.00	1.01	1.04	1.02
Composite Leading Indicators								
COM	0.97	0.93	0.93	0.94	1.04	0.82	0.97	0.94

Note: The entry in the first line is the RMSFE for the AR model forecast, in percentage growth rates at an annual rate. The remaining entries are the relative RMSFE of the forecast based on the individual indicator, relative to the RMSFE of the benchmark AR forecast. The forecast period is 2007q1 to 2009q4. The abbreviation of leading indicators are outlined in Table 5. ***: 1%, **: 5% and *: 10% indicating the significance level of the modified Wilcoxon signed-rank test for h = 1 and h = 2 as proposed by *Diebold/Mariano* (1995).

short horizon the general performance of leading indicator models is even slightly better and some more forecasts turn out to be significantly better than the benchmark. Monetary aggregates do not turn out to be helpful in this recession. Only narrow money (nominal and real M1) reports forecast errors slightly smaller than the benchmark; however they are not significant.

When we turn to non-linear models (see Table 3), we find that some of the indicators further improved in terms of forecast accuracy. In particular for financial variables a threshold effect seems to be evident (which is in line with the literature, see e.g. *Clements/Galvao* (2006)). We find improvements for the term spread, stock prices and stock price volatilities by considering non-linearities. For survey indicators the gains from using non-linear models are less evident; only for expectation measures some improvements can be observed. For other indicators (e.g. prices of commodities and goods) the effect of employing non-linearities is ambiguous.

Table 3

Forecast Results for GDP and IP during the Crisis – Non-linear Models

	GDP				IP			
	h = 1	h = 2	h = 3	h = 4	h = 1	h = 2	h = 3	h = 4
	RMSFE				*RMSFE*			
AR	6.08	5.16	4.82	4.56	19.14	16.61	14.07	13.15
	RMSFE relative to AR Model				*RMSFE relative to AR Model*			
Interest Rates								
IS-3M	1.01	0.99	0.98	0.97	1.07	0.99	1.00	0.98
DIL-10	0.81**	0.92	0.91	0.93	0.60**	1.03	0.97	0.96
Interest Rates Spreads								
SPR-10Y-3M	0.91*	0.99	1.02	1.03	0.71**	0.90	0.93	0.88
SPR-C-G	0.86	0.99	0.86	0.75	1.12	0.92*	1.00	0.91
SPR-B-G	0.76	1.01	0.92	0.74	0.94	1.51	1.03	0.79
SPR-BF-G	1.39	2.14	2.95	1.02	2.39	2.33	2.93	1.20
Monetary Aggregates								
DLNM1	1.17	1.00	0.92	0.98	1.39	0.93	0.95	0.85
DLNM1R	1.16	0.93	0.90	0.89	1.45	0.88	0.91	0.85
DLNM2	1.06	1.01	1.03	1.04	1.05	0.91	0.99	0.94
DLNM2R	1.05	1.03	1.07	1.05	0.69*	1.15	1.00	1.03
DLNM3	0.99	1.01	1.02	1.04	0.72*	0.90	0.99	0.94
DLNM3R	1.08	1.01	1.16	1.16	0.75	1.13	1.02	1.09
Other Financial Indicators								
DLNDAX	0.78*	0.83	0.83	0.86	0.69***	0.92*	0.85	0.87
VOLA1	0.96	0.90	0.95	0.93	1.26	0.81	0.96	0.95
DLNEX	1.03	1.04	1.04	1.01	0.90	1.11	1.05	1.03
DLNEXR	1.07	1.03	1.01	1.01	1.30	1.12	1.04	1.03
DLNHWWI	1.02	1.14	1.09	1.10	0.58	0.95	1.00	1.02
DLNHWWIEX	1.02	0.94	0.99	0.99	0.51	0.96	0.97	1.01
DLNOIL	1.22	1.04	1.10	1.04	0.98	1.06	1.05	1.05
Survey Indicators								
IFO-C	0.86	0.80	0.80	0.84	0.90	0.68	0.69	0.76
IFO-EXP	0.62	0.75	0.73	0.79	0.65*	0.70*	0.72	0.78
IFOM-C	0.74	0.84	0.84	0.92	0.62*	0.69	0.79	0.84
IFOM-EXP	0.64**	0.72	0.78	0.98	0.61**	0.63*	0.77	0.79
IFO-WC	0.70	0.78	0.86	0.84	0.82*	0.83*	0.84	0.86
IFO-WEXP	0.96	0.91	0.94	0.96	0.92	0.97	0.94	0.93
ZEW-EXP	0.75	0.81	0.75	0.90	1.09	0.80	0.94	0.87
ESI	0.74	0.90	0.84	0.96	1.08	0.78*	0.79	0.83
ESI-INDU	0.81	0.81	0.86	0.90	0.68*	0.75	0.85	0.95
ECCS99	0.76	0.88	0.98	1.03	1.04	0.78*	0.89	0.92
PMI	0.82	0.85	0.86	1.02	0.47***	0.90	0.91	0.99

	GDP				IP			
	h = 1	h = 2	h = 3	h = 4	h = 1	h = 2	h = 3	h = 4
	RMSFE				*RMSFE*			
AR	6.08	5.16	4.82	4.56	19.14	16.61	14.07	13.15
	RMSFE relative to AR Model				*RMSFE relative to AR Model*			
Real Economic Indicators								
DLNIP-VORL	1.22	1.30	0.95	0.97	1.09	0.89	1.00	1.01
DLNORD	0.80	1.31	1.09	0.94	0.72*	1.39	1.05	0.84
DLNORD-C	1.16	0.99	0.98	0.96	1.15	0.91	1.11	0.96
DLNORD-I	0.74*	1.01	0.99	0.88	0.82	1.42	1.25	0.80
CAPA	0.79	1.06	0.95	0.94	0.74	0.81	0.90	0.86
DLNEW	0.95	0.98	1.07	1.05	0.89	0.96	1.05	1.05
DALQ	0.99	1.03	1.08	1.09	1.22	0.98	1.07	1.03
DLNVAC	0.88	0.94	0.98	0.89	1.11	1.13	1.00	0.97
DLNWHOUR	0.95	0.97	1.00	1.00	0.86	1.00	0.99	1.00
DLNCPI	1.26	1.02	0.97	0.98	1.49	0.98	0.97	0.96
DLNCPI-EX	1.17	0.96	0.95	0.95	0.75	0.93	1.07	1.02
Composite Leading Indicators								
COM	0.92	0.92	0.91	0.93	1.20	0.77	0.98	0.93

Note: The entry in the first line is the RMSFE for the AR model forecast, in percentage growth rates at an annual rate. The remaining entries are the relative RMSFE of the forecast based on the individual indicator, relative to the RMSFE of the benchmark AR forecast. The forecast period is 2007q1 to 2009q4. The abbreviation of leading indicators are outlined in Table 5. ***: 1%, **: 5% and *: 10% indicating the significance level of the modified Wilcoxon signed-rank test for h = 1 and h = 2 as proposed by *Diebold/Mariano* (1995).

IV. Forecast Combination

Since the seminal work by *Bates/Granger* (1969), the literature on forecast pooling has conclusively shown that the forecasting performance of forecast combination is much more stable than that of single indicator models.[12] In general, it has been shown that even very simple combination schemes do well in terms of forecasting.

The pooling of individual indicators via combination schemes offers the possibility to take various sources of information into account. Due to estimation uncertainty, the aggregation of information in one model is practically challenging. To circumvent this problem the literature has

[12] See, *Timmermann* (2006), for literature overview; for the US (*Stock/Watson* (2004)), the euro area (*Drechsel/Maurin* (2011)) and also for Germany before the outbreak of the crisis (*Kuzin/Marcellino/Schumacher* (2009); *Drechsel/Scheufele* (2012)).

proposed techniques such as dynamic factor models and shrinkage methods. The attractive feature of forecast combination methods is their simplicity and the fact that their performance can still be attributed to their constitute models (which is helpful in the interpretation of the results). In this paper we consider three simple forecast combination schemes to analyze their performance for GDP as well as IP during the economic crisis 2007–2009.[13] We therefore differentiate two strategies. First of all, we only use linear models as it is done in most of the literature. Secondly, we augment the pooling approach to include also the non-linear models. In general, the weight $\omega_{i,t}^{h}$ that is assigned to each indicator forecast is based on the i^{th} individual equation described by eq.(1). Accordingly, the total forecast of output growth is

$$\tilde{Y}_{t,t+h}^{h} = \sum_{i=1}^{n} \omega_{i,t}^{h} \hat{Y}_{i,t+h}^{h} \quad \text{with} \sum_{i=1}^{n} \omega_{i,t}^{h} = 1. \tag{4}$$

The first pooling method that is quite standard and often used as a benchmark is the equal weighting scheme. Simply to calculate, it is hard to beat by more complicated methods. Furthermore, this is the weighting scheme that is used to produce the consensus forecast. Secondly, besides mean forecasts, where the weights are the same for each period, we use the median forecast to take the effect of outliers into account. We also use the in-sample fit to calculate individual weights. In the literature, Bayesian Model Averaging (BMA) has received much attention because it can be an attractive way of dealing with model uncertainty. As shown by Hansen (2008), BMA (under the assumption of diffuse priors) can be easily approximated by calculating weights along the Schwarz criteria (SIC) which is also known as Bayesian Information Criteria (BIC).[14] Finally, we consider also the use of R^2 as an alternative to the SIC which also takes into account the error variance of each indicator model (see *Drechsel/Maurin* (2011)).

The results based on forecast combination indicate that model averaging schemes improve the forecast accuracy compared to the benchmark (see Table 4). The findings for the weighting schemes presented are very similar, however, for many of them the differences compared to the benchmark are even statistically significant. Some individual leading indicator forecasts provide more accurate results than the combination of

[13] For an overview of several pooling methods, see *Drechsel/Scheufele* (2012).

[14] These weights are calculated as $\omega_{t,i}^{SIC} = \exp\left(-0.5 \cdot \Delta_{t,i}^{SIC}\right) / \sum_{i=1}^{n} \exp\left(-0.5 \cdot \Delta_{t,i}^{SIC}\right)$, with $\Delta_{t,i}^{SIC} = SIC_{t,i} - SIC_{t,\min}$.

Table 4

Relative RMSFEs of Combination Forecasts

	GDP				IP			
	h = 1	h = 2	h = 3	h = 4	h = 1	h = 2	h = 3	h = 4
Linear Models								
Equal weights	0.77**	0.87*	0.88	0.87	0.79**	0.87	0.91	0.89
Median weights	0.73**	0.86*	0.87	0.88	0.82**	0.88*	0.88	0.89
SIC weights	0.78**	0.88*	0.89	0.87	0.79**	0.88	0.92	0.89
R^2 weights	0.78**	0.87*	0.87	0.87	0.79**	0.87	0.91	0.89
Linear & Non-linear Models								
Equal weights	0.84	0.85	0.85	0.86	0.81**	0.85	0.85	0.81
Median weights	0.84**	0.85	0.86	0.88	0.84***	0.84	0.84	0.83
SIC weights	0.83	0.86	0.86	0.85	0.81**	0.87	0.86	0.81
R^2 weights	0.84	0.85	0.85	0.86	0.82**	0.85	0.85	0.89

Note: Relative RMSFE of the forecast based on pooling of individual indicators is shown. ***: 1% , *: 5% and *: 10% significance level of the modified Wilcoxon signed-rank test for h = 1 and h = 2 as proposed by *Diebold/Mariano* (1995).

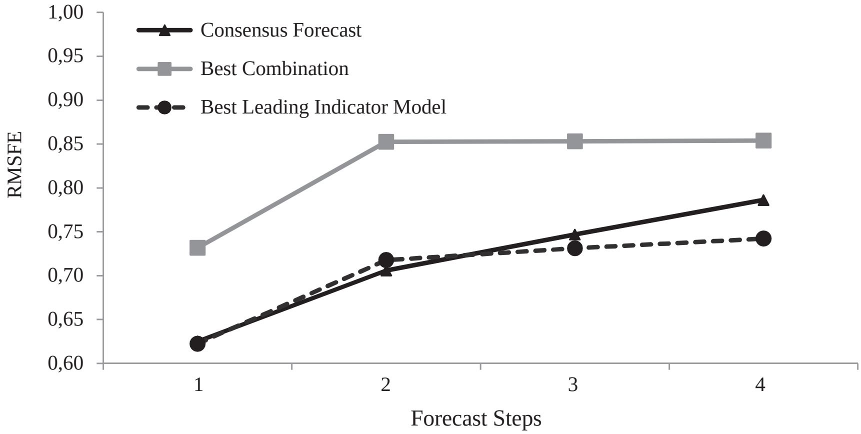
the bundle of forecasts. It is also interesting that the inclusion of non-linear models into the pooling does not always lead to an improvement in forecasting accuracy. Only for a longer forecast horizon, the inclusion of non-linear models leads to lower forecast errors of the combination schemes (although the differences remain small).

V. Comparison between Leading Indicator Forecasts and Professional Forecasters

Using the quarterly forecasts by the professional forecasters, we create a forecast dataset that is comparable with the forecasts of the annualized growth rate given by the individual leading indicators and the forecast combination. Therefore, we have to transform year-on-year to quarterly annualized GDP growth rates.[15]

We find that the forecasts by the professionals display good forecasting properties and at each horizon beat the univariate benchmark (see Figure 4). Professionals do also well compared to leading indicator mod-

[15] Which is done by using past real-time GDP series. Unfortunately, Consensus Economics does not provide quarterly growth rates for IP with fixed forecasting horizon. Thus we have to solely rely on GDP forecasts.



Source: Consensus Economics (2009) and own calculations.

Figure 4: Performance of the Professional Forecasters

els and tend to perform better than the forecast combination schemes. The forecast errors are extremely close to those of the best leading indicator model. This may imply that during the recession professional forecasters processed information very fast and thus might have also used qualitative information not explicitly considered in econometric models. It has to be kept in mind that most forecasters of the consensus economics work for banks and other financial companies which might be aware of the crisis earlier compared to other people in the economy. Overall, the mean forecast from Consensus Economics did relatively well during the recession and kept up with the best econometric models.

VI. Discussion and Conclusion

In this paper we analyzed the regression in 2008/2009 from a forecaster's perspective. In a first attempt we analyzed the forecasts from Consensus Economics before and during the recession. For Germany, we find that before the crash of Lehman the crisis was not predicted by the professionals. After the bankruptcy, forecasters heavily revised their forecasts for the upcoming year and even tended to exceed the actual value.

From the investigation of leading indicators we can learn several things. Generally, we can confirm that forecasts based on leading indicators provide some warning signals before the outbreak of the recession.

In particular, survey indicators (sentiment indicators, ifo expectations, PMI) and financial indicators (risk spreads, stock prices) give early warnings. In contrast to other studies, we also take into account non-linear leading indicator models. We find that non-linearities are only helpful for some indicators (including financial variables, survey expectations and for some price variables). The partial success of financial variables can be attributed to the origins of the recession in the financial sector. In particular, risk spreads (i.e. the spread between corporate and government bond yields) which did not signal subsequent recessions (*Fritsche/Kuzin* (2005)) reflect some of the causes of this recession.

When we compare leading indicator forecasts with those of the professionals, we find that the professionals did relatively well. This implies that this recession was not foreseeable with a comprehensive forecast knowledge based on experiences during prior recessions, in particular in its exceptional magnitude.

References

Ager, P./*Kappler,* M./*Osterloh,* S. (2009): The accuracy and efficiency of the Consensus Forecasts: A further application and extension of the pooled approach. International Journal of Forecasting, 25(1): 167–181. – *Andrews,* D. W. K. (1993): Tests for parameter instability and structural change with unknown change point. Econometrica, 61(4): 821–856. – *Bates,* J. M./*Granger,* C. W. J. (1969): The Combination of Forecasts. Operations Research Quarterly. Operational Research Society, 20(4): 451–468. – *Benner,* J./*Meier,* C.-P. (2005): Was leisten Stimmungsindikatoren für die Prognose des realen Bruttoinlandsprodukts in Deutschland? Eine Echtzeit-Analyse. Die Weltwirtschaft, 2005/3, 341–355. – *Breitung,* J./*Jagodzinski,* D. (2001): Prognoseeigenschaften alternativer Indikatoren für die Konjunkturentwicklung in Deutschland. Konjunkturpolitik, 47: 292–314. – *Büning,* H./*Trenkler,* G. (1994): Nichtparametrische statistische Methoden. Berlin: Walter de Gruyter. – *Burgstaller,* J. (2009): Financial Predictors of Real Activity and the Propagation of Aggregate Shocks. Kredit und Kapital, 42(1), 1–23. – *Clements,* M. P./*Galvão,* A. B. (2006): Combining predictors and combining information in modeling: Forecasting US recession probabilities and output growth. In: Milas, C.; and Rothman, P.; and van Dijk, D. Nonlinear Time Series Analysis of Business Cycles. Contributions to Economic Analysis Series, Elsevier, 55–73. – *Consensus Economics* (2009): Consensus forecasts. Technical report, 2009. – *Diebold,* F. X./*Mariano,* R. S. (1995): Comparing predictive accuracy. Journal of Business and Economic Statistics, 13(3): 253–263. – *Döpke,* J./*Krämer,* J. W./*Langfeldt,* E. (1994): Konjunkturelle Frühindikatoren in Deutschland. Applied Economics Quarterly (Konjunkturpolitik), 40(2): 135–153. – *Drechsel,* K./*Maurin,* L. (2011): Flow of conjunctural information and forecast of euro area economic activity. Journal of Forecasting, 30(3): 336–354. – *Drechsel,* K./*Scheufele,* R. (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. – *Galbraith,* J. W./*Tkacz,* G. (2000): Testing for asymmetry in the link between the yield spread and output in the G-7 countries. Journal of International Money and Finance, 19(5): 657–672. – *Fritsche,* U./*Kuzin,* V. (2005): Prediction of Business Cycle Turning Points in Germany Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik), 225(1): 22–43. – *Fritsche,* U./*Stephan,* S. (2002): Leading indicators of German business cycles: An assessment of properties. Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik), 222(3): 289–315. – *Hansen,* B. E. (2008): Least-squares forecast averaging. Journal of Econometrics, 146(2): 342–350. – *Kapetanios,* G./*Labhard,* V./*Price,* S. (2008): Forecasting Using Bayesian and Information-Theoretic Model Averaging: An Application to U.K. Inflation. Journal of Business & Economic Statistics, 26(1): 33–41. – *Kholodilin,* K./*Siliverstovs,* B. (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. – *Kirchgässner,* G./*Savioz,* M. (2001): Monetary Policy and Forecasts for Real GDP Growth: An Empirical Investigation for the Federal Republic of Germany. German Economic Review, 2(4): 339–365. – *Koll,* W./*Klüh,* U./*Schwonke,* C./*Zimmermann,* U./*Heilemann,* U./*Kirchgässner,* G./*Stahlecker,* P. (2009): Welche Rolle spielen Prognosen? Wirtschaftsdienst, 89(2): 79–100. – *Kuzin,* V./*Marcellino,* M./*Schumacher,* C. (2009): Pooling versus model selection for nowcasting with many predictors: An application to German GDP. C.E.P.R. Discussion paper No. 7197. – *Ragnitz,* J. (1994): Zinsstruktur und Wirtschaftswachstum. Kredit und Kapital, 27(1): 11–29. – *Sauer,* C./*Scheide,* J. (1995): Money, interest rate spreads, and economic activity. Review of World Economics, 131(4): 708–722.– *Schumacher,* C./*Breitung,* J. (2008): Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data. International Journal of Forecasting, 24(3): 386–398. – *Stock,* J. H./*Watson,* M. W. (2003 a). How did leading indicator forecasts perform during the 2001 recession? Federal Reserve Bank of Richmond, Economic Quarterly, 89(3): 71–90. – *Stock,* J. H./*Watson,* M. W. (2003 b). Forecasting output and inflation: The role of asset prices. Journal of Economic Literature, 47(3): 788–829. – *Stock,* J. H./*Watson,* M. W. (2004): Combination forecasts of output growth in a seven-country data set. Journal of Forecasting, 23(6): 405–430. – *Timmermann,* A. (2006): Forecast Combinations, Volume 1 of Handbook of Economic Forecasting, chapter 4, 135–196. Elsevier. – *Tong,* H. (1983): Threshold Models in Non-linear Time Series Analysis. Heidelberg: Springer. – *West,* K. D. (1996): Inference about Predictive Ability. Econometrica, 64(5): 1067–1084. – *West,* K. D. (2006): Forecast Evaluation, Volume 1 of Handbook of Economic Forecasting, chapter 3, 99–134. Elsevier. – *Zarnowitz,* V. (1992): Business Cycles: Theory, History, Indicators, and Forecasting, Studies in Business Cycles, Volume 27, The University of Chicago Press.

Appendix

Table 5

Definition of Indicators

Label	Name	Source
Dependent variable		
	GDP, real	Destatis
	Industrial production	Buba
Interest Rates		
IS-3M	3-month-money market rate	Buba
DIL-10	Long term government bond yield – 9–10 years	Buba
Interest Rates Spreads		
SPR10Y-3M	Term spread (10y – 3-month-money market rate)	Buba
SPR-C-G	Corporate bond-government bonds	Buba
SPR-B-G	Spread corporate BBB- government bonds	Buba/ML
SPR-BF-G	Spread corporate financial BBB-government bonds	Buba/ML
Monetary Aggregates		
DLNM1	M1	Buba
DLNM1R	M1, real	Buba
DLNM2	M2	Buba
DLNM2R	M2, real	Buba
DLNM3	M3	Buba
DLNM3R	M3, real	Buba
Other Financial Indicators		
DLNDAX	DAX share price index	Boerse
VOLA	DAX volatility	Boerse
DLNEX	Nominal effective exchange rate	Buba
DLNEXR	Real effective exchange rate	Buba
DLNHWWI	HWWI index of world market prices of raw mats.	HWWI
DLNHWWI-EX	HWWI index, excl. Energy łn	HWWI
DLNOIL	Oil prices (euros per barrel)	ECB
Survey Indicators		
IFO-C	Ifo climate index	ifo
IFO-EXP	Ifo expectations index	ifo
IFOM-C	Ifo climate index, manufacturing	ifo
IFOM-EXP	Ifo expectations index, manufacturing	ifo
IFO-WC	World economic climate index	ifo

(Continue next page)

(Table 5: Continued)

Label	Name	Source
Survey Indicators		
IFO-WEXP	World economic expectations index	ifo
ZEW-EXP	ZEW economic expectations	ZEW
ESI	Economic sentiment indicator (average)	EC
ESI-INDU	Industrial confidence indicator	EC
ECCS99	Economic confidence indicator (average)	EC
PMI	Markit survey, PMI: manufacturing	Markit
Real Economic Indicators		
DLNIP-VORL	Intermediate goods production	Buba
DLNORD	Manufacturing orders	Buba
DLNORD-C	Manufacturing orders – consumer goods	Buba
DLNORD-I	Manufacturing orders – capital goods	Buba
DCAPA	Capacity utilization	ifo
DLNEW	Employed persons (work-place concept)	BfA
DALQ	unemployment rate	BfA
DLNVAC	Vacancies	Buba
DLNWHOUR	Hours worked	Destatis
DLNCPI	Consumer price index	Buba
DLNCPI-EX	Core CPI	Buba
Composite Leading Indicators		
COM	Early Bird indicator	Commerzbank

Note: The data is used in levels unless the label starts with D, indicating the use of first differences or DLN for logged differences. The data is published with a lag of 0 or 1 quarters. The sources are labeled as follows: Buba – Deutsche Bundesbank, ML – Merrill Lynch, EC – European Commission, Destatis – Federal Statistical Office Germany, BfA – Bundesagentur für Arbeit.

Table 6

Unit Root Test Results

Name	t-stat	lag	Name	t-stat	lag
Key Variables			**Survey Indicators**		
DLNGDP	–6.56***	[0]	IFO-C	–4.75***	[1]
DLNIP	–8.41***	[0]	IFO-EXP	–5.37***	[1]
Interest rates			IFOM-C	–4.93***	[1]
IS-3M	–2.79*	[1]	IFOM-EXP	–5.33***	[1]
DIL-10	–5.99***	[0]	IFO-WC	–4.01***	[1]
Interest rates Spreads			IFO-WEXP	–4.17***	[1]
SPR-10Y-3M	–3.20**	[1]	ZEW-EXP	–4.36***	[1]
SPR-C-G	–2.67*	[1]	ESI	–4.55***	[1]
SPR-B-G	–2.93**	[1]	ESI-INDU	–5.18***	[1]
SPR-BF-G	–3.01**	[1]	ECCS99	–3.82***	[1]
Monetary Aggregates			PMI	–4.43***	[1]
DLNM1	–5.64***	[0]	**Real Economic Indicators**		
DLNM1R	–5.73***	[0]	DLNIP-VORL	–5.63***	[1]
DLNM2	–5.25***	[0]	DLNORD	–5.20***	[1]
DLNM2R	–5.61***	[0]	DLNORD-C	–7.61***	[0]
DLNM3	–3.79***	[0]	DLNORD-I	–5.03***	[0]
DLNM3R	–4.43***	[0]	DCAPA	–4.91***	[0]
Other financial indicators			DLNEW	–3.53***	[0]
DLNDAX	–5.35***	[0]	DALQ	–4.65***	[0]
VOLA1	–3.08**	[0]	DLNVAC	–3.34**	[0]
DLNEX	–7.86***	[0]	DLNWHOUR	–6.10***	[3]
DLNEXR	–6.74***	[0]	DLNCPI	–4.96***	[0]
DLNHWWI	–6.67***	[1]	DLNCPI-EX	–4.91***	[0]
DLNHWWAEX	–6.31***	[0]	**Composite Leading Indicators**		
DLNOIL	–7.02***	[0]	COM	–4.03***	[1]

Note: ADF-test results are shown. Significance levels are defined by ***: 1% , **: 5% and *: 10% . Lags are chosen according to SIC.

Table 7

Break Test Results for GDP Models

Name	h = 1		h = 2		h = 3		h = 4	
Interest Rates								
IS-3M	–	–	–	–	–	–	–	–
DIL-10	–	–	–	–	–	–	–	–
Interest Rates Spreads								
SPR-10Y-3M	*	2004Q2	***	2004Q1	–	–	–	–
SPR-C-G	–	–	–	–	*	1994Q2	–	–
SPR-B-G	*	2004q1	***	2004q2	***	2004q3	***	2004q4
SPR-BF-G	–	–	**	2005q2	**	2005q2	**	2005q3
Monetary Aggregates								
DLNM1	–	–	–	–	–	–	–	–
DLNM1R	–	–	–	–	–	–	–	–
DLNM2	–	–	–	–	–	–	–	–
DLNM2R	–	–	–	–	–	–	–	–
DLNM3	–	–	–	–	–	–	–	–
DLNM3R	–	–	–	–	–	–	–	–
Other Financial Indicators								
DLNDAX	–	–	–	–	–	–	–	–
VOLA1	–	–	–	–	–	–	–	–
DLNEX	–	–	–	–	–	–	–	–
DLNEXR	–	–	–	–	–	–	–	–
DLNHWWI	–	–	–	–	–	–	–	–
DLNHWWIEX	–	–	–	–	–	–	–	–
DLNOIL	–	–	–	–	–	–	–	–
Survey Indicators								
IFO-C	–	–	–	–	–	–	**	2002q4
IFO-EXP	–	–	–	–	–	–	–	–
IFOM-C	–	–	–	–	–	–	*	2002q4
IFOM-EXP	–	–	–	–	–	–	–	–
IFO-WC	–	–	–	–	–	–	–	–
IFO-WEXP	–	–	–	–	–	–	–	–
ZEW-EXP	–	–	–	–	–	–	–	–
ESI	–	–	–	–	–	–	–	–
ESI-INDU	–	–	–	–	*	2002q3	–	–
ECCS99	–	–	–	–	–	–	–	–
PMI	–	–	–	–	*	2005q3	–	–

Name	h = 1		h = 2		h = 3		h = 4	
Real Economic Indicators								
DLNIP-VORL	–	–	–	–	–	–	–	–
DLNORD	–	–	–	–	–	–	–	–
DLNORD-C	–	–	–	–	–	–	–	–
DLNORD-I	–	–	–	–	–	–	–	–
CAPA	–	–	–	–	–	–	–	–
DLNEW	–	–	–	–	–	–	–	–
DALQ	–	–	–	–	–	–	–	–
DLNVAC	–	–	–	–	–	–	–	–
DLNWHOUR	–	–	–	–	–	–	–	–
DLNCPI	–	–	–	–	–	–	–	–
DLNCPI-EX	–	–	–	–	–	–	–	–
Composite Leading Indicators								
COM	–	–	***	2004q1	***	2004q3	***	2001q1
Percentage of models significant at								
10 % level	4.8 %		9.5 %		14.3 %		11.9 %	
5 % level	0.0 %		9.5 %		7.1 %		9.5 %	

Note: Results of the Quandt-Andrews breakpoint test are shown along the most likely break point. Significance levels are defined by ***: 1 %, **: 5 % and *: 10 %. All results are based on the maximum *F*-Test. The trimming level is 15 %.

Summary

The Financial Crisis from a Forecaster's Perspective

This paper analyses the recession in 2008/2009 in Germany. This recession is very different from previous recessions in particular regarding their causes and magnitude. We show to what extent forecasters and forecasts based on leading indicators fail to detect the timing and the magnitude of the recession. This study shows that large forecast errors for both expert forecasts and forecasts based on leading indicators resulted during this recession which implies that the recession was very difficult to forecast. However, some leading indicators (survey data, risk spreads, stock prices) have indicated an economic downturn and hence, beat univariate time series models. Although the combination of individual forecasts provides an improvement compared to the benchmark model, the combined forecasts are worse than several individual models. A comparison of expert forecasts with the best forecasts based on leading indicators shows only minor deviations. Overall, the range for an improvement of expert forecasts in the crisis compared to indicator forecasts is small. (JEL E37, C53)

Zusammenfassung

Die Finanzkrise aus Sicht eines Prognostikers

Dieser Beitrag untersucht die Rezession der Jahre 2008/2009 in Deutschland. Diese Rezession hebt sich in ihrer Ursache und Schwere deutlich von früheren Rezessionen ab. Es wird gezeigt, inwieweit Prognostiker und Prognosen basierend auf Frühindikatoren das Timing und die Stärke dieser Rezession verfehlt haben. Diese Studie deutet darauf hin, dass aufgrund der großen Prognosefehler bei Expertenprognosen und bei Prognosen basierend auf Frühindikatoren die Rezession sehr schwer zu prognostizieren war. Allerdings gibt es einige Frühindikatoren (Umfragedaten, Risikoaufschläge, Aktienpreise), die eine Wachstumsabschwächung prognostiziert haben und damit deutlich besser abschneiden als univariate Zeitreihenmodelle. Jedoch wurde insbesondere die Stärke nicht richtig eingeschätzt. Die Kombination einzelner Prognosemodelle bietet zwar eine Verbesserung zur Benchmarkprognose, schneidet aber schlechter ab als einige Einzelindikatormodelle. Vergleicht man die Expertenprognosen mit den besten Prognosen auf Basis von Frühindikatoren, so ist der Abstand relativ klein. Insgesamt ist der Spielraum einer Verbesserung der Expertenprognose in der Krise im Vergleich zu Indikatormodellen relativ gering.