International Journal of Forecasting 28 (2012) 428–445

Contents lists available at SciVerse ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

The performance of short-term forecasts of the German economy before and during the 2008/2009 recession

Katja Drechsel [a], Rolf Scheufele [b,*]

[a] Halle Institute for Economic Research, Kleine Maerkerstr. 8, 06108 Halle, Germany

[b] Swiss National Bank (SNB), Boersenstrasse 15, P.O. Box, CH-8022 Zurich, Switzerland

ARTICLE INFO

Keywords:
Macroeconomic forecasting
Evaluating forecasts
Combining forecasts
Structural change

ABSTRACT

The paper analyzes the forecasting performance of leading indicators for industrial production in Germany. We focus on single and pooled leading indicator models both before and during the financial crisis. Pairwise and joint significant tests are used to evaluate single indicator models, as well as forecast combination methods. In addition, we investigate the stability of forecasting models during the most recent financial crisis. We find that only a small number of single indicator models were performing well before the crisis. Pooling can substantially increase the reliability of leading indicator forecasts. During the crisis, the relative performances of many leading indicator models (e.g. using surveys, term and risk spreads) improved.

1. Motivation

The most recent financial and economic recession differed from other economic downturns in many ways. Germany experienced its strongest cut in production by far since the Second World War. The GDP in 2009 (q1) was 7% lower than in the first quarter of 2008, and industrial production shrunk even more, by 20%. Despite the exceptional scale of the recession, many professional forecasters failed to predict it.

This paper analyzes the out-of-sample forecasting performances of leading indicator models before and during the financial crisis of 2008–2009. Most of the literature on leading indicator performances in forecasting industrial production (and GDP) in Germany originated after 2000 (see, among others, Breitung & Jagodzinski, 2001, and Fritsche & Stephan, 2002, for single equation leading indicator models; as well as Kholodilin & Siliverstovs, 2006, Kuzin, Marcellino, & Schumacher, 2009, and Schumacher & Breitung, 2008, using dynamic factor models). However, while they made extensive use of leading indicators for extracting information relating to future economic development, none of the authors pointed specifically to the forecasting properties of leading indicators during a pronounced recession. As far as the US is concerned, there was some work on leading indicator properties during the recession 2001/2002 (see Clements & Galvão, 2006, chap. 2, 2009; Stock & Watson, 2003b). We therefore decided to take the recession of 2008/2009 and the years before that as the starting point of our investigation.

In forecasting output growth, we concentrate on industrial production (IP), which is the main monthly coincident indicator available for the German economy.[1] The average share of total industry in total gross value added amounts to about 25% over the last twenty years. The industrial sector exhibits the most volatility in terms

* Corresponding author. Tel.: +41 44 631 37 44.
E-mail addresses: katja.drechsel@iwh-halle.de (K. Drechsel), rolf.scheufele@snb.ch (R. Scheufele).

[1] We also analyzed the forecast ability of leading indicators for quarterly GDP. For the sake of convenience, we solve the mixed frequency problem in the case of GDP by taking three month averages of the available monthly indicator (see Clements & Galvão, 2009, for more sophisticated techniques for dealing with this kind of problem). In general we find that the relative forecasting performances of leading indicators for the two reference variables IP or GDP are very similar.

0169-2070/$ – see front matter
doi:10.1016/j.ijforecast.2011.04.003

of output during a business cycle peak to trough, and has a sizeable impact on the rest of the German economy. In this study, we investigate a very large set of 120 leading indicators (1–12 months ahead) in the light of the latest recession. While our data set comprises survey-based measures, financial market indicators, real activity variables and composite leading indicators, we focus in particular on financial indicators as predictors for real activity, since the recession is often seen as originating in the financial sector (see Stock & Watson, 2003a, for a literature review).

Another strand of the literature (for details, see Timmermann, 2006, chap. 4) shows that forecast combination leads to significant improvements relative to forecasts based on individual indicators. Hence, the second contribution of our analysis implements forecast combination schemes.

We apply several weighting schemes for combining leading indicator forecasts for IP: simple averaging schemes (mean and median forecast), trimmed mean (owing to past out-of-sample performances), forecasts based on in-sample criteria (AIC, R^2), weights computed by relative mean square forecast errors, OLS weights, and shrinkage techniques (motivated by Bayesian averaging) (see, among others, Drechsel & Maurin, 2011).

To assess the forecasting performance in detail, we compute root mean squared forecasting errors relative to a benchmark autoregressive forecast in a pseudo out-of-sample experiment after 2000. In addition, we use a Giacomini and White (2006) pairwise test of equal forecast ability to decide which of the models perform significantly better than the benchmark model. We also conduct a joint significance test, as suggested by White (2000), to test the adequacy of leading indicator forecasts in general.

In order to ensure that the results will be robust, we further divide our forecasting sample into a pre-crisis period and a crisis period, to analyze how the forecasting performance changed during the recession. We use an end-of-sample instability test, as proposed by Andrews (2003), to investigate whether the financial crisis led to a break in the relative forecasting performances of leading indicator forecasts. This approach is unique in the forecasting setting, and makes adequate testing for the stability of the forecast quality at the end of the sample possible.

During the pre-crisis period, until 2007, only certain single indicator models show satisfying forecasting properties. These are mainly survey-based measures (ifo surveys and economic sentiment indicators provided by the EU Commission). Joint significance tests even suggest that no single indicator models beat a univariate benchmark model within this period. However, employing forecast combination schemes based on the individual models results in better and more reliable forecasting performances. In particular, weights based on the forecasting performance in the recent past (discounted MSFE) show sizeable, significant improvements in forecasting accuracy. In addition, other pooling strategies, such as AIC weights, the trimmed mean and Bayesian Model Averaging schemes, provide significant improvements over the benchmark model.

We generally find that average forecasting errors increased dramatically during the recession. While most of the indicators indicated a slowdown, none of them predicted the sharpness of the downturn accurately. Interestingly, while the overall forecasting performance worsens during the crisis, the relative performances of certain individual indicator forecasts improves substantially.

Further, most of these indicators show relatively good forecasting properties during the recession period. During the crisis, the number of leading indicator forecasts which perform better than the univariate AR model increased significantly. Stability tests indicate that many indicator forecasts perform significantly better than before the outbreak of the crisis. The relative forecast accuracy of indicator models, consisting of term spreads, risk spreads and composite indicators, improves substantially during the crisis period. Surveys and model averaging schemes display a relatively stable behavior during the two subperiods.

The paper is structured as follows. The next section provides an overview of the leading indicators which we use for our forecast analysis, and gives the selection criteria for the individual forecast equations. In addition, the forecast pooling methods we apply to aggregate the individual forecasts are described. Section 3 explains the assessment of the relative predictive power of the forecasts. Section 4 presents the results of indicator forecasts (single and pooled) during the pre-crisis and crisis periods. Section 5 summarizes and concludes.

2. Forecast models

In this section, we present our data set, discuss selected leading indicators, and explain the applied methodology and the various weighting schemes used for pooling the forecasts.

2.1. Leading indicators

The paper analyzes a large set of 120 leading indicators which are commonly used in the literature. Because we are interested solely in the leading properties of these indicators, we have left out coincident indicators such as retail sales, which might be useful for nowcasting exercises but are published with delay. All of the indicators are available at a monthly frequency, so we can use them for monthly IP forecasts. Broadly speaking, our analyzed indicators can be grouped as follows: (i) financial indicators, (ii) surveys, (iii) real economy, and (iv) composite leading indicators.

As the source of the latest recession is linked to the financial sector, we consider several financial market indicators as predictors for real activity. In their seminal paper, Stock and Watson (2003a) provide a review of the forecasting performances of financial market indicators. Similarly, we use six interest rates with different maturities (see Table 3). Further, term spreads defined as the difference between interest rates on long and short maturity debts are used. Numerous studies have shown that these indicators may provide useful information for future economic activity (see, for example, Estrella & Hardouvelis, 1991; Estrella, Rodrigues,