Institut für Wirtschaftsforschung Halle
Halle Institute for Economic Research

Should We Trust in Leading Indicators? Evidence from the Recent Recession

Katja Drechsel
Rolf Scheufele

April 2010 No. 10

IWH-Diskussionspapiere
IWH-Discussion Papers

Should We Trust in Leading Indicators? Evidence from the Recent Recession

Katja Drechsel
Rolf Scheufele

April 2010 No. 10

Authors: *Katja Drechsel*
Department of Macroeconomics

Phone: +49/(0)345 77 53 836
Fax: +49/(0)345 77 53 799
Email: Katja.Drechsel@iwh-halle.de

Rolf Scheufele
Department of Macroeconomics

Phone: +49/(0)345 77 53 728
Fax: +49/(0)345 77 53 799
Email: Rolf.Scheufele@iwh-halle.de

Herausgeber:

INSTITUT FÜR WIRTSCHAFTSFORSCHUNG HALLE (IWH)
Prof. Dr. Dr. h. c. Ulrich Blum (Präsident), Dr. Hubert Gabrisch (Forschungsdirektor)
Das IWH ist Mitglied der Leibniz-Gemeinschaft

Hausanschrift: Postfach 11 03 61, 06017 Halle (Saale)
Postanschrift: Kleine Märkerstraße 8, 06108 Halle (Saale)

Telefon: +49 345 7753 60
Telefax: +49 345 7753 820

Internetadresse: http://www.iwh-halle.de

Should We Trust in Leading Indicators?

Evidence from the Recent Recession

Zusammenfassung

Dieses Papier untersucht die Prognosegüte konjunktureller Frühindikatoren für das Bruttoinlandsprodukt sowie die Industrieproduktion in Deutschland vor und während der Krise. Die Prognosegüte einzelner und durch verschiedene Gewichtungsschemata kombinierter Prognosen basierend auf Frühindikatoren wird durch gemeinsame Signifikanztests bewertet. Des Weiteren geben End-of-sample Instabilitätstests Auskunft über die Stabilität der Prognosemodelle während der aktuellen Finanzkrise. Es wird gezeigt, dass nur wenige Einzelindikatoren vor der Krise genauere Prognosen liefern als das AR-Modell. Durch Kombination kann die Prognosegüte von Frühindikatoren erheblich verbessert werden. Während Umfragedaten für die Kurzfristprognose die Prognosegüte erheblich verbessern, liefern Finanzmarktdaten, wie bspw. Zinsspreads und Risikoaufschläge, bessere Prognosen als die Benchmark für längerfristige Prognosehorizonte.

Schlagwörter: Frühindikatoren, Prognosegüte, Prognosekombination, Strukturbrüche
JEL-Klassifikation: E37, C22, C53

Should We Trust in Leading Indicators? Evidence from the Recent Recession*

Abstract

The paper analyzes leading indicators for GDP and industrial production in Germany. We focus on the performance of single and pooled leading indicators during the pre-crisis and crisis period using various weighting schemes. Pairwise and joint significant tests are used to evaluate single indicator as well as forecast combination methods. In addition, we use an end-of-sample instability test to investigate the stability of forecasting models during the recent financial crisis. We find in general 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 performance of many leading indicator models increased. At short horizons, survey indicators perform best, while at longer horizons financial indicators, such as term spreads and risk spreads, improve relative to the benchmark.

Keywords: Leading Indicators, Forecast Evaluation, Forecast Pooling, Structural Breaks
JEL classification: E37, C22, C53

* We would like to thank Elena Andreou, Raffaella Giacomini and Klaus Wohlrabe as well as the participants of the 10th IWH-CIREQ Workshop on "Recent Advances in Macroeconomic Forecasting", the DIW macroeconometric workshop 2009, the DAGStat 2010 and the Scottish Economic Society 2010 annual conference for comments and suggestions. We also thank Ralph Solveen and IfW for providing some of the data sources.

Should We Trust in Leading Indicators? Evidence from the Recent Recession

1 Motivation

The recent financial and economic recession differed in many ways from other economic downturns. Germany, experienced by far the strongest cut in production since the Second World War. In comparison with the first quarter of 2008, GDP in 2009 (Q1) was 7% lower. During the same period, industrial production shrunk even more, by 20%. Despite the exceptional scale of the recession, many professional forecasters failed to foresee the current recession.

This paper analyzes the out-of-sample forecasting performance of leading indicator models before and during the financial crisis of 2008-2009. Most of the literature on leading indicator performance in forecasting GDP and industrial production in Germany originated after 2000 (see, among others, Breitung and Jagodzinski (2001) and Fritsche and Stephan (2002) for single equation leading indicator models as well as Kholodilin and Siliverstovs (2006), Schumacher and Breitung (2008) and Kuzin et al. (2009), using dynamic factor models). However, while they made excessively use of leading indicators to extract information for future economic development, none of the authors pointed specifically to the forecasting properties of leading indicators during a pronounced recession.

We investigate a very large set of leading indicators for both German GDP (1 to 4 quarters ahead) and industrial production (1 to 12 months ahead) in the light of the recent 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 origin of the recession is often viewed in the financial sector (see Stock and Watson, 2003a, for a literature review).

Another strand of literature (for details see Timmermann, 2006) shows that forecast combination leads to significant improvements in comparison to forecasts based on individual indicators. Hence, the second contribution of our analysis is to make extensively use of forecast combination schemes. We apply several weighting schemes to combine leading indicator forecasts for GDP and IP: simple averaging schemes (mean and median forecast), the trimmed mean (owing to past out-of-sample performance), forecast based on in-sample criteria (AIC, R^2), weights computed by relative mean square forecast errors, OLS weights as well as shrinkage techniques (motivated by Bayesian averaging) (see, among others, Drechsel and Maurin, 2010).

To assess the forecasting performance in detail, we compute relative root mean squared forecasting errors and relative mean absolute forecasting errors relative to a benchmark autoregressive forecast in a pseudo out-of-sample experiment from 2000-2009. In addition, we use Giacomini and White's (2006) pairwise test of equal forecast ability to decide which of the models do 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.

To yield robust results, 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 performance of leading indicator forecasts. This approach is unique in the forecasting setting and makes it possible to test adequately for the stability of forecasting quality at the end of the sample.

In the pre-crisis period, 2001-2007, only certain single indicator models show favorable forecasting properties. These are: survey based measures (ifo business climate and expectations and the economic sentiment indicator provided by the EU Commission) and stock market returns. Many forecast combination schemes (such as AIC weight, the median, discounted MSFE weights) often outperformed the benchmark model significantly. Joint tests indicate that there is basically no single indicator model that significantly outperforms the benchmark AR model. However, considering forecast combination schemes yields significant improvements.

We generally find that average forecasting errors increased dramatically during the recession. While most of the indicators indicated a slowdown, none has adequately recognized the sharpness of the downturn. Interestingly, while the total forecasting performance worsens during the crisis, the relative performance of individual indicator forecasts increases substantially. Further, most of these indicators show relatively good forecasting properties during the recession period. During the crisis, the number of leading indicator forecasts that perform better than the univariate AR model has increased notably. The relative forecast accuracy of indicator models consisting of term spreads, risk spreads and survey indicators improve substantially during the crisis period. Break tests indicate that many indicator forecasts do significantly better compared with a simple benchmark model (particularly when mean squared error loss is assumed).

The paper is structured as follows: The next section provides an overview of the leading indicators we use for our forecast analysis and presents the selection criteria for the individual forecast equations. In addition, the forecast pooling methods we applied to aggregate the individual forecasts are described. Section 3 presents the results of indicator forecasts (single and pooled) during the pre-crisis and crisis period. Finally, section 4 summarizes and concludes.

2 Forecasts based on Leading Indicators

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. Finally, we explain the assessment of the relative predictive power of the forecasts.

2.1 Leading Indicators

A large set of leading indicators that are commonly used in the literature are analyzed in this paper. Because we are interested solely in the leading properties of these indicators, we have left out coincidence indicators, such as retail sales, which might be useful for nowcasting exercises but are published with delay. Most of the indicators are available at monthly frequency so we can use them for both quarterly GDP forecasts and monthly IP forecasts. Broadly speaking, our analyzed indicators can be grouped as follows: (i) Financial indicators, (ii) Surveys, (iii) Real economy, (iv) Prices and wages and (v) Composite leading indicators.

As the source of the current 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 performance of financial market indicators. Similarly we use six interest rate measures: the monetary policy instrument, the overnight rate, the three-month money market rate and government bond yields (with maturities of 3-5, 5-8 and 9-10 years, respectively).[1] Further, term spreads are defined as the difference between interest rates on long and short maturity debt are used. It has been shown in numerous studies that these indicators may provide useful information for future economic activity (see, for example, Estrella and Hardouvelis, 1991; Estrella et al., 2003; Wheelock and Wohar, 2009). Our spread measures consist of five term spreads including government bond yields (9-10 years) minus policy instrument, government bond yields (9-10 years) minus overnight rate, government bond yields (9-10 years) minus three-month money market rate, three-month money market rate minus overnight rate and overnight rate minus monetary policy rate. In addition, we consider default spreads as predictors of real growth (inspired by Gertler and Lown, 1999). The spreads between corporate and government bond yields, between AA and BBB rated corporate bonds (financial and nonfinancial cooperations), between BBB corporate bonds and gov-

[1] Kirchgässner and Savioz (2001) show that short-run interest rates provide a very good out-of-sample forecast performance for real GDP growth in Germany.

ernment bonds as well as a high yield ("junk bond") spread are therefore analyzed (see Table 5 for the exact definition).[2]

Besides interest rates, we also employ monetary aggregates, in both nominal and real (deflated by the CPI excluding energy) terms. Sims (1972) provides evidence of a causal relationship (in a Granger sense) between money and income, which runs from money to income but not vice versa. This implies that money provides useful information for future output. Although there is some recent evidence for the predictive content of money for growth (see Swanson, 1998; Brand et al., 2003), this relationship is mostly found to be unreliable in out-of-sample forecasting setups (see, for example, Stock and Watson, 2003a).[3] Moreover, the use of German monetary aggregates as leading indicators is complicated by the fact that, owing to the transition into the EMU, a continuous definition does not exist within our sample period.

Since stock prices reflect the expected discount value of future earnings, stock returns should provide useful information for predicting earnings and therefore future output growth. While this theoretical relationship is well established, the empirical evidence for stock prices as a reliable leading indicator for future output growth is ambiguous. Besides stock returns, volatility of stock returns is also considered (see Campbell et al., 2001). Moreover, commodity prices are used as additional indicators. We use real oil prices and aggregate indexes of commodity prices (including and excluding energy). This is motivated by the fact that some recessions, namely those in the 1970s and early 1980s, were associated with a dramatic increase in oil prices, which is regarded as the origin of these recessions. We also saw a large increase in commodity prices in 2008, so it is natural to include these variables as potential leading indicators. Further, we investigate both the effective nominal exchange rate (defined as the exchange rate against a trade-weighted basket of countries) and the real effective exchange rate, which can be interpreted as a measure of domestic competitiveness. In comparison with other studies on German leading indicators, we provide the most complete set of financial variables as leading indicators (at least as far as we are aware).

The second group of indicators consists of survey-based measures. One common feature of both financial market indicators and survey-based indicators is their early availability in time. While most financial variables are immediately available, survey indicators are usually available before the end of a particular month.[4] Survey-based

[2] Some default spreads are not available for the whole sample. However, we use them when they are available (which includes the entire out-of-sample period).

[3] For Germany, Fritsche and Stephan (2002) conclude that the out-of-sample predictive content of monetary aggregates is very pure.

[4] See appendix for the timely availability of leading indicators.

measures are extremely popular coincident and leading indicators in Germany. This study also considers a variety of survey measures: ifo Business Climate and Business Expectations for the headline series as well as for some subcomponents[5], ifo World Climate and World Business Expectations, ZEW Economic Sentiment Indicator, PMI for manufacturing, GFK income expectations and business cycle expectations, as well as business and consumer sentiment indicators collected by the European Commission.[6] While the relative performance of various survey indicators is documented in many studies, no consensus has emerged concerning their relative forecast performance.[7] In a recent study for IP by Robinzonov and Wohlrabe (2009), this is attributed to the different settings for each study, which complicates comparisons. The results depend on the sample periods and datasets as well as on whether further restrictions on the parameters are employed or whether equations are updated at each point in time.

The next variable set consists of real economy indicators such as labor market variables, prices and new orders. Typically new orders indicate the strength of foreign and domestic demand. New orders today will result in higher production in the future and will thus provide useful information for output growth. We further differentiate between new orders for consumer and investment goods. In addition, labor market indicators may also be useful. Owing to labor turnover costs, dismissals are costly and labor demand decisions should be forward-looking as well. In our paper we use different labor indicators in our paper: the unemployment rate, the number of employed persons as well as the number of vacancies.

We also look at inflation rates, since, according to the New Keynesian Phillips curve (NKPC), inflation is forward-looking and is determined by future marginal costs. Consequently, higher marginal costs are associated with excessive demand (as motivated by Galí and Gertler, 1999); inflation may thus contain information on output dynamics (see Scheufele, 2010, for the empirical relevance of the NKPC in Germany). The inflation rates considered here are: CPI, core CPI (excluding energy) and wage inflation (measured as negotiated wage).

Finally, we consider composite leading indicators such as the Early Bird (Commerzbank), FAZ (Frankfurter Allgemeine Zeitung) indicator and the leading indicators published by the OECD. Those measures are already a combination of the

5 The headline series is defined as climate and expectations in industry and trade, which includes manufacturing, construction, wholesaling and retailing.

6 See appendix for the exact indicator definition used in this analysis. Since the specific characteristics of these indicators have been discussed elsewhere, we skip the characterization of each indicator here (see Breitung and Jagodzinski, 2001; Hüfner and Schröder, 2002).

7 See, among others, Breitung and Jagodzinski (2001), Hüfner and Schröder (2002) and Benner and Meier (2004).

indicator measures presented above. Typical choices are: ifo climate index, the stock market index DAX, interest rates and spreads, exchange rates and/or orders inflow.[8]

2.2 Individual models

We conduct leading indicator forecasts for both quarterly GDP and monthly IP data. Using IP as an additional output indicator has particular advantages. IP is available at monthly frequency so no aggregation to quarterly data is needed for the indicators, implying a loss of information. Furthermore the number of observations (and hence the degrees of freedom) increases considerably if monthly information is used. Additionally, industrial production is available earlier. Although IP measures only a small fraction of total GDP, it is a good proxy for GDP in Germany.[9]

The indicator forecasts are computed in a simulated out-of-sample forecasting environment for the period 1991Q1-2009Q2.[10] The first half of this sample (37 quarterly and 111 monthly observations) is used to construct the initial estimation period, and the remaining sample is used for collecting forecasts. 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. Y_{t+h}^h is the output growth over the next h periods (months or quarters) in terms of an annualized rate.[11] Forecasts are based on an h-step ahead regression model:

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

where ε_{t+h}^h is an error term and α, β and γ are regression coefficients to be estimated. Unlike other studies, we take into account the timely availability of the indicators by the indices l and k which are, in the case of quarterly data, $l = 2$ and for monthly data $l = 3$. Depending on the publication lag of the candidate predictor, k varies from 0 to 1 for quarterly data and from 0 to 2 for monthly data. The optimal number of lags in the quarterly analysis is restricted to $1 \leq p \leq 4$ and $0 \leq q \leq 4$

8 See Breitung and Jagodzinski (2001) and Hüfner and Schröder (2002) for assessments of the Early Bird and the FAZ indicator for Germany. A recent comparison of their composition is given by Robinzonov and Wohlrabe (2009).

9 The average share of total industry in total gross value added is 25.4% in the period 1991m1-2009m6.

10 While most of the data is available prior to 1991, the literature generally includes only the data for the post-unification period.

11 $Y_t^h = (400/h)\ln(Q_t/Q_{t-h})$ for real GDP and $Y_t^h = (1200/h)\ln(Q_t/Q_{t-h})$ for industrial production, respectively.

($1 \leq p \leq 12$ and $0 \leq q \leq 12$ in the monthly exercise) and is selected by the Akaike criterion.[12]

Figure 1: Forecast Design

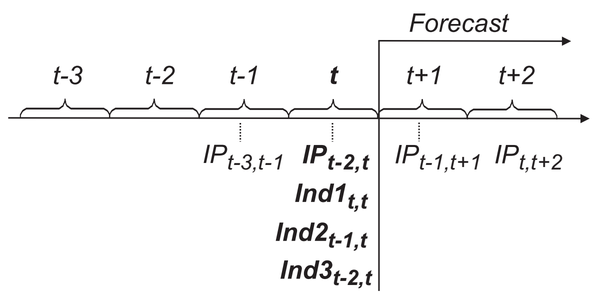


Figure 1 illustrates the interaction of various indicators and monthly IP data, which is released with a delay of approximately 45 days after the reference month. For the forecast we use all the information available by the end of period t, so we take into account the indicators that are released by the end of the reference month (Ind1), a second subset of indicators (Ind2) published with a delay of 15-30 days after the reference month and a third subset (Ind3) available at the earliest 40 days after the reference month but before IP data is published.

The simulated real-time forecast scheme depends on the estimation step. Here equation (1) is estimated using only data from prior to the forecast date. This means, for example, that when a forecast for the fourth quarter has to be made at the beginning of October, the equation is estimated until the second quarter (note that GDP is not instantly available for the third quarter) with indicators for the second quarter and before.[13] The forecast is then made using these estimated coefficients and knowing the indicator for the third quarter (when available, otherwise only the indicator from the second quarter can be used) as well as additional lags of the endogenous variable available in the second quarter and before. The advantage of this procedure is that no future information enters into the forecasting step in order to keep the setting as close as possible to real forecasting situations.[14]

---

[12] For the sake of completeness the Schwarz information criteria (SIC) is also considered. The results are similar to those given by a lag structured based on AIC.

[13] The GDP flash estimate is released approximately 45 days after the reference quarter and the IP flash 45 days after the reference month, respectively.

[14] However, the simulated real-time forecast scheme does not consider revisions of the data. This problem is of minor importance for the indicator variable, since financial markets indicator

2.3 Pooling of leading indicators

Recently there have been various attempts to enrich simple and parsimonious time series models (such as ARIMA or leading indicator models) with more information. This makes the forecasting process more realistic in relation to the real world application, where hundreds (or even thousands) of data series are available and have been investigated. One strand of literature models such variables using a dynamic factor structure (see, for example, Stock and Watson, 2002; Forni et al., 2003; Schumacher and Breitung, 2008). Yet another way of incorporating a great deal of information is to pool single indicator models (as discussed by Timmermann, 2006; Drechsel and Maurin, 2010). The forecasting combination approach has been successfully applied following the seminal work of Bates and Granger (1969), and results mostly in a more favorable and stable forecasting performance than that of single indicator models. An advantage of forecast combination as opposed to factor models is that their performance can still be attributed to their constitute models (which is often helpful in interpreting the results). Despite its growing success, the literature on forecast combination for leading indicators in Germany is extremely scarce. Most results for Germany on forecast combination are available from Stock and Watson (2003b, 2004) in a multi-country comparison.[15] We therefore intend to complete this gap by providing evidence on forecast combination after 2000 and during the economic crisis 2008-2009 in particular.

The total forecast of output growth $\tilde{Y}^h_{t,t+h}$ is based on the pooling of the individual indicator forecasts $\widehat{Y}^h_{i,t+h}$:

$$\tilde{Y}^h_{t,t+h} = \sum_{i=1}^{n} \omega^h_{i,t} \widehat{Y}^h_{i,t+h} \quad \text{with} \quad \sum_{i=1}^{n} \omega^h_{i,t} = 1 \tag{2}$$

Where $\omega^h_{i,t}$ is the weight assigned to each indicator forecast, that is based on the i^{th} individual equation described by eq.(1). The re-estimation of this equation after

or survey measures are hardly 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 might be in real time. For Germany, both Benner and Meier (2005) and Schumacher and Breitung (2008) compare the performance of leading indicators with both real time data and final revised data in a setting similar to ours. Both studies conclude that the relative performance of indicators remains approximately the same (also the absolute precision is somewhat lower with real time data).

15 Dreger and Schumacher (2005) and Robinzonov and Wohlrabe (2009) provide the only country-specific literature on forecasting combination in Germany (at least to the best of our knowledge). However, they consider only a limited number of leading indicator models ($n < 10$) and they provide results only for IP.

each period implies that the weight associated with this indicator also differs from time to time. So we generally allow the weights to be time varying (although the degree of time variation depends heavily on the specific averaging scheme).

Several pooling methods are applied to optimize the performance of the forecast. In the forecast pooling literature, it is common to use (i) equal weights as a benchmark. This is mainly because these simple weights can easily be calculated and the contribution of each indicator to the pooled forecast is straightforward.[16] Because the mean-combination forecast depends only on the number of variables used, the weights will be the same in each forecasting period. A similar weighting scheme is (ii) median-combination forecast. In comparison with the equal weights, these account for forecast outliers and they differ for each period.

Taking into account the error variance of each leading indicator model, we can use information criteria (iii) for constructing weights. The AIC criterion is therefore used (see Atkinson, 1980; Kapetanios et al., 2008). The highest weights are assigned to models with the lowest AIC value. Finally, for robustness, the R-squared approach is used.[17]

With reference to the forecast errors, we also analyze weighting schemes incorporating information given by the variance covariance matrix of the in-sample forecast errors. A natural choice is to construct weights by minimizing the sum of squared residuals from all the candidate leading indicator models. Per construction, this automatically leads to the smallest mean squared error (at least in-sample). From a theoretical point of view, this should lead to the optimal combination weights (as discussed and applied by Granger and Ramanathan, 1984). However, in practice, this approach often suffers from overparameterization when the number of predictors is high in relation to the sample size and it tends to be very sensitive to breaks in the relative model performance. Nevertheless, like Granger and Ramanathan (1984), we apply (iv) a restricted OLS estimator. We therefore use the optimal weight vector, which is the linear projection of Y (the realization) onto the vector of individual forecasts subject to two constraints: the weights sum to one and an intercept is not included.

As already stated, when the number of candidate models is relatively large in comparison with the sample size, covariance structures are difficult to estimate due to collinearity. One way of dealing with this problem is to rely solely on the variances,

16 However, the condition that the indicator forecasts have the same variance and similar correlations is often neglected.

17 At each forecasting step the weights are calculated as $\omega_{t,i}^{AIC} = \exp\left(-0.5 \cdot \Delta_{t,i}^{AIC}\right) / \sum_{i=1}^{n} \exp\left(-0.5 \cdot \Delta_{t,i}^{AIC}\right)$ with $\Delta_{t,i}^{AIC} = AIC_{t,i} - AIC_{t,\min}$ and $\omega_{t,i}^{R^2} = \exp\left(-0.5 \cdot \Delta_{t,i}^{R^2}\right) / \sum_{i=1}^{n} \exp\left(-0.5 \cdot \Delta_{t,i}^{R^2}\right)$ with $\Delta_{t,i}^{R^2} = R_{t,\max}^2 - R_{t,i}^2$.

which can be done by using information criteria (see above). Another attempt is rationalized by using a Bayesian framework. Diebold and Pauly (1990) suggest shrinking towards equal weights (v) and obtaining a simple expression which equals

$$\overline{\omega}_t = \omega_0 + \frac{\hat{\omega}_t - \omega_0}{1 + g_t}, \tag{3}$$

where $\overline{\omega}_t$ is the OLS estimate shrunk towards the uniform prior (which corresponds to the equal weighting scheme), $\hat{\omega}_t$ is the vector of OLS weights and ω_0 is the equal weights vector. The value of g determines the degree of shrinkage and the larger this value the more shrinkage is attained towards the mean. Diebold and Pauly (1990) employ empirical Bayes methods to estimate g_t depending on two parameters σ_t^2 and τ_t^2 which can be estimated with

$$\hat{\sigma}_t^2 = \frac{(Y_t - \hat{\omega}_t \hat{Y}_t)'(Y_t - \hat{\omega}_t \hat{Y}_t)}{T} \quad \text{and} \quad \hat{\tau}_t^2 = \frac{(\hat{\omega}_t - \omega_0)'(\hat{\omega}_t - \omega_0)}{\operatorname{tr}(\hat{Y}_t' \hat{Y}_t)^{-1}} - \hat{\sigma}_t^2,$$

where Y is the vector of realizations, $\hat{\omega}_t \hat{Y}_t$ are the OLS weighted individual forecasts and T is the number of observations. Note that if $\hat{\sigma}_t^2 / \hat{\tau}_t^2 \rightarrow 0$, $\overline{\omega}_t$ equals the OLS estimator, while if $\hat{\sigma}_t^2 / \hat{\tau}_t^2 \rightarrow \infty$, the arithmetical average is obtained.

Another way of combining models in a Bayesian framework for forecasting purpose is proposed by Wright (2008, 2009) (vi). Weights are constructed in proportion to the posterior probability of each model, which can be calculated as

$$\omega_{i,t} \propto (1 + \phi)^{-p_{i,t}/2} S_{i,t}^{-T}, \tag{4}$$

where $S_{i,t}^2 = Y_t' Y_t - Y_t' \hat{Y}_{i,t} \frac{\phi}{1+\phi}$. $\hat{Y}_{i,t}$ is the vector of model i's in sample predictions, $p_{i,t}$ denotes the number of parameters in model i and T is the number of in-sample observations. Parameter ϕ controls the degree of shrinkage. The smaller ϕ is, the stronger the degree of shrinkage (which makes the prior more informative). If ϕ is large, one moves away from the model prior in response to what the data say.[18] As noted by Wright (2008) it is not clear what the optimal degree of shrinkage is for the purpose of obtaining good forecasts. Like Kapetanios et al. (2008), we also consider three variants in the degree of shrinkage: $\phi = 0.5$ (high shrinkage), $\phi = 2$ (medium shrinkage) and $\phi = 20$ (low shrinkage).

[18] Note that the Wright (2008) weighting scheme (assuming low shrinkage) is related to information theoretic weighting schemes. Both take into account the in-sample model fit and penalize the model complexity (i.e. the number of estimated parameters).

So far, the model combination schemes have been constructed using in-sample information. This is appropriate as long as the estimated relationships are not too affected by structural instabilities. However, there is evidence that structural breaks can distort the relationship between in-sample and out-of-sample forecasting performance.[19] In this case, it might be better to use out-of-sample information for constructing combination weights. We purposely construct the out-of-sample weights in the same quasi-real-time setting in which we construct our forecasts. This implies that we can use the information in past forecast errors only when they can be observed (so we consider a relevant information lag). For instance, we cannot observe GDP at t when the forecast $t+h$ is made because GDP is unknown. We can therefore only include forecast errors until $t-1$ (a similar argument holds for IP). This aspect has often been overlooked or not stated explicitly in the available literature. It also implies that, for the first few runs, when there is no out-of-sample information available, we use the equal weighting scheme until the first past forecasts can be compared with their corresponding realization.

A simple and often very effective combination scheme is the trimming approach (vii), which discards a subset of indicators (see, for example, Timmermann, 2006). In general, these outliers are the indicators with the worst performance. The performance measure is given by the recursively computed mean squared forecast error, which is calculated up to that point in time when the latest forecast error can be observed. According to the literature, we scrap an indicator if the individual indicator forecast belongs to the 25%, 50% or 75% of the worst performers.[20] The remaining indicators are pooled by equal weights.

Following Stock and Watson (2003b, 2004) and Costantini and Pappalardo (2009), we incorporate weighting based on discounted MSFEs (viii). This means that current weights are inversely proportional to the forecast errors of the recent past. This obviously implies that the most recent best indicators obtain a relatively high weight. This approach follows that of Bates and Granger (1969), who successfully applied similar techniques. Discount mean square forecast error weights are based on

$$w_{i,t} = \frac{\lambda_{it}^{-1}}{\sum_{j=1}^{n} \lambda_{jt}^{-1}} \tag{5}$$

[19] Stock and Watson (2003a), among others, show that in-sample predictability evaluated by Granger causality provides a poor guide for a model's out-of-sample performance.

[20] There is no consensus in the literature as to which share should be discarded. Armstrong and Collopy (1992) even suggest discarding both the high and low errors, which they refer to as "winsorizing".

where $\lambda_{it} = \sum_{s=T_0}^{t-h} \delta^{t-h-s} \left(\hat{e}_{i,s}^h\right)^2$ with δ being the discount factor and $\hat{e}_{i,s}^h$ the forecast error of model i. Note that imposing $\delta = 1$ (no discounting) implies long memory, meaning that all estimation errors in the sample are equally important. The other extreme is $\delta = 0$, where only the most recent best performance is considered. The literature tends to set δ relatively high between 0.9 and 1 (see Stock and Watson, 2004; Costantini and Pappalardo, 2009). However, there is also evidence that high discounting (lower δ's) produces more accurate forecasts (see Timmermann, 2006, section 7.5). We also experiment with different values of δ and find that a low value ($\delta = 0.3$) performs best for quarterly and monthly time series.

2.4 Forecast Evaluation

To analyze the forecast performance of our indicator models, we examine the forecast errors for the specified out-of-sample period. We concentrate on the mean squared forecast error (MSFE) as a benchmark loss function. More precisely, we compute root mean squared forecast errors (RMSFE) of a candidate forecast relative to a benchmark model. The latter is a forecast from a univariate autoregression model which corresponds to forecasts from eq(1), where no further indicator X is specified. We denote $\widehat{Y}_{i,t+h|t}^h$ as the forecast with indicator i and $\widehat{Y}_{0,t+h|t}^h$ as the benchmark forecast. Comparing the realization Y_{t+h}^h with the forecast results in the corresponding forecast errors $\widehat{e}_{i,t+h}^h = Y_{t+h}^h - \widehat{Y}_{i,t+h|t}^h$ and $\widehat{e}_{0,t+h}^h = Y_{t+h}^h - \widehat{Y}_{0,t+h|t}^h$. The h-step ahead relative RMSFE of model i relative to the benchmark is then equal to

$$relative\ RMSFE = \frac{\sqrt{\sum_{t=T_1}^{T_2-h} \left(Y_{t+h}^h - \widehat{Y}_{i,t+h|t}^h\right)^2}}{\sqrt{\sum_{t=T_1}^{T_2-h} \left(Y_{t+h}^h - \widehat{Y}_{0,t+h|t}^h\right)^2}} = \frac{\sqrt{\sum_{t=T_1}^{T_2-h} \left(\widehat{e}_{i,t+h}^h\right)^2}}{\sqrt{\sum_{t=T_1}^{T_2-h} \left(\widehat{e}_{0,t+h}^h\right)^2}}, \qquad (6)$$

where T_1 indicates the first date of the pseudo out-of-sample forecast and T_2 is the last date, where the last forecast is observed. Whenever the average performance of the indicator forecast is better than the AR forecast, the relative RMSFE is smaller than one. Further, we also employ the mean absolute forecast error (MAFE) as an alternative.

2.4.1 Pairwise Comparisons

However, the RMSFE and MAFE measures provide no evidence whether the difference is statistically significant. A more formal test procedure to decide which models are preferable relative to a simple AR model is necessary. Although some studies on forecasting performance of indicator variables for Germany explicitly test for equal forecasting performance (see Benner and Meier, 2004; Dreger and Schumacher, 2004), these tests are all based on the Diebold and Mariano (1995) test of equal predictive ability. However, this procedure ignores the consequences of parameter uncertainty when forecasts are made by regression models (see West, 1996). The so-called *asymptotic irrelevance* applies only with subject to certain assumptions, then inferences can be based on the normality assumption (for a general overview see also West, 2006).

In our setting, forecast evaluation is complicated by the fact that the benchmark model may be nested in the indicator model. Since we have chosen a rolling window and may select different models from time to time, there is the likelihood that we will have to evaluate forecasts that are mixtures from nested and nonnested models. Although methods of comparing nested models exist (see e.g. Clark and McCracken, 2001), these do not apply to different forecasting models in time. Because of these complications we choose the Giacomini and White (2006) test of conditional predictive ability. Taking a perspective different from those analyzed by West (1996), the proposed test has a number of advantages. First, it is possible to compare both nested and nonnested models, which allows the comparison of models that change from time to time. Second, we may also evaluate forecast combination schemes.

More formally, we define $\Delta L^i_{m,t+h}$ as the loss difference of the indicator model i and the benchmark model (the AR model), which is equal to

$$\Delta L^i_{m,t+h} = \left(\widehat{e}^h_{i,t+h}\right)^2 - \left(\widehat{e}^h_{0,t+h}\right)^2$$

for mean squared loss.[21] To test the null of equal conditional predictive ability,[22] the Giacomini and White (2006) test statistic is a Wald-type and can be formulated as

$$GW^{(i,0)}_h = m\left(\frac{1}{m}\sum_{t=T_1}^{T_2-h} g_t \Delta L^i_{m,t+h}\right)' \hat{\Omega}^{-1} \left(\frac{1}{m}\sum_{t=T_1}^{T_2-h} g_t \Delta L^i_{m,t+h}\right), \tag{7}$$

where $m = T_2 - T_1 - h + 1$ is the sample size and g_t is a $q \times 1$ measurable test function, which we set to $g_t = [1\ \Delta L_t]$, as suggested by Giacomini and White (2006). The covariance matrix $\hat{\Omega}$ is an HAC-type matrix like that proposed by Newey and West (1987). Under some standard regularity conditions, $GW^{(i,0)}_h \overset{\alpha}{\sim} \chi^2_q$.

21 Similarly, the loss difference in absolute can be defined as $\Delta L^i_{m,t+h} = \left|\widehat{e}^h_{i,t+h}\right| - \left|\widehat{e}^h_{0,t+h}\right|$.

22 With a loss difference $\Delta L_{m,t+h}$, test function h_t and a information set G_t the null is (H_0 : $E[g_t \Delta L_{m,t+h}\,|G_t] = 0$).

2.4.2 Joint Tests

While pairwise comparisons are helpful in deciding which indicator models are useful in forecasting GDP and IP, it is not completely certain whether the indicator models taken together provide any information relative to the benchmark model. To be precise, in pairwise comparisons we do not take problems of multiple testing into account. However, to answer the question of whether any indicator forecast is better than a simple AR model, we rely on the White (2000) reality check for data snooping. Basically, we can state the null hypothesis for this problem as

$$H_0: \; E(\Delta L^1_{m,t+h}) = E(\Delta L^2_{m,t+h}) = \ldots = E(\Delta L^n_{m,t+h}) \leq 0, \tag{8}$$

where $E(\Delta L^i_{m,t+h})$ is the expected loss difference of indicator model i. The null hypothesis is that no indicator model outperforms the benchmark. The test statistic is then equal to

$$T_m = \max\left(m^{1/2}\Delta\bar{L}^1, \ldots, m^{1/2}\Delta\bar{L}^n\right), \tag{9}$$

where m is the sample size of the out-of-sample forecast period, n the number of models and $\Delta\bar{L}^i = \frac{1}{m}\sum_{t=T_1}^{T_2-h}\Delta L^i_{t+h}$. Owing to the complexity of this inference and problems stemming from the need to control for the full set of alternatives, bootstrap techniques are employed to calculate corresponding p-values. Although this test was originally proposed in the framework where asymptotic irrelevance occurs, namely when models are nonnested, it can be related to the Giacomini and White (2006) framework, which is a multivariate extension that can be used under a rolling estimation window. In addition, we use the modification of this statistic proposed by Hansen (2005), which is more powerful and less sensitive to the inclusion of poor and irrelevant alternatives.

2.4.3 Stability

Besides testing for average predictive ability for the whole out-of-sample interval, we are interested in the relative performance of indicator forecasts during the crisis. Since it is well documented that indicator models may be unstable over time, we evaluate the relative forecasting properties of indicators before and during the crisis. We therefore split the sample into a pre-crisis and a crisis period, in which the latter comprise all the forecasts that have been made for the period 2008m1-2009m6. All forecasts before that are consequently pre-crisis forecasts. Note that this definition involves an exogenous determination of the break date. We choose this date according to the official recession announcement by CEPR (2009).[23] Obviously our results will depend on the pre-specified break data. Although there are methods available for dealing with endogenous breaks at unknown times (see Giacomini and Rossi, 2009), these procedures are inapplicable at the end of the out-of-sample period.

Instead we propose a generalization of the well-known Chow test which was put forward by Andrews (2003) to test for instabilities during the recent economic crisis. While this methodology has been successfully applied for testing the stability of coefficients in a standard regression framework, we are, at least to the best of our knowledge, the first who use this methodology in testing for end-of-instability in forecasting performance. The Andrews (2003) methodology is designed specifically for instabilities at the end of a sample. It is robust to autocorrelation and heteroskedasticity, and is easy to compute. Critical values and p-values can be obtained by using a subsample technique.

For the implementation of this method we use the regression version of the Diebold-Mariano test as discussed, for example, by West (2006). Here, the loss difference (indicator model minus benchmark) ΔL^i_{t+h} is regressed on a constant and inference is conducted by using a t-test (with HAC adjustment).[24] For the end-of-sample stability test applied to the relative forecasting performance, we split the sample of forecasts $t = T_1+1, \ldots, T_2$ into the first T' and the last $p = T_2 - T' + 1$ observations. The starting point is the regression model with the loss difference as dependent variable and a constant as the only regressor.

$$\Delta L_t = \begin{cases} \beta_0 + u_t, & t = T_1+1, T_1+2, \ldots, T' \\ \beta_{1t} + u_t, & t = T'+1, \ldots, T_2, \end{cases} \tag{10}$$

The null hypothesis of interest is then stability of the model, i.e. $\beta_0 = \beta_{1t}$ for all $t \in \{T'+1, \ldots, T_2\}$ (as well as stationarity of u_t for $t = T_1, \ldots, T_2$). The alternative hypothesis is $\beta_0 \neq \beta_{1t}$ for some $t \in \{T'+1, \ldots, T_2\}$ and / or the distribution of $\{u_{T'+1}, \ldots, u_{T_2}\}$ differs from that of $\{u_t, \ldots, u_{t+p-1}\}$ for $t = T_1, \ldots, T' - p + 1$.

To set up the test statistic (called S statistic), the following steps are necessary. First, estimate the equation to be tested over the whole forecast period ($t = T_1 + 1, \ldots, T_2$) and let $\widehat{\beta}_{T_1+1--T_2}$ be the LS estimate of this parameter. In our context, this

[23] The CEPR Euro Area Business Cycle Dating Committee announced the beginning of the recession in January 2008 where they found the peak in economic activity. Accordingly the period 2008Q1 and 2008m1 marks the beginning of the crisis period.

[24] This test is equivalent to the unconditional test for equal predictive ability as suggested by Giacomini and White (2006)

corresponds to the mean error loss over the whole forecasting period $\Delta\bar{L}$. Second, the error covariance matrix is estimated as

$$\hat{\Sigma} = (T' - T_1 + 1)^{-1} \sum_{j=T_1}^{T'+1} \hat{U}_{j,j+p-1} \hat{U}'_{j,j+p-1}, \tag{11}$$

where $\hat{U}_{j,j+p-1} = (\hat{u}_j, \ldots, \hat{u}_{j+p-1})$ is a vector of residuals computed as $\hat{u}_j = \Delta L_j - \widehat{\beta}_{T_1+1--T_2}$. Finally the statistic S is defined as

$$S = \hat{U}'_{T'+1,T_2} \hat{\Sigma}^{-1} \hat{U}_{T'+1,T_2}. \tag{12}$$

This expression can be interpreted as the sum of squared transformed post-change residuals (where the transformed residuals correspond to $\hat{\Sigma}^{-1/2} \hat{U}_{T'+1,T_2}$).

For calculating appropriate p-values of the test, Andrews (2003) propose a parametric subsampling technique instead of large-sample asymptotics. This procedure works as follows: For the first subset, estimate the equation using observations $T_1 + [p/2]$–T' and then compute the sum of squared transformed residuals for period $T_1 + 1$–$T_1 + p + 1$ denoted by d_1. For the next subset, estimate the equation using observation $T_1 + 1$ and $T_1 + [p/2] + 1$–T' and calculate again the sum of squared transformed residuals for period $T_1 + 2$–$T_1 + p + 2$ saved as d_2. Taken together, $T' - T_1 - p + 1$ subsets can be computed like this and all corresponding d_1 to $d_{(T'-T_1-p+1)}$ sum of squared transformed residuals are saved. Andrews calls this technique "leave-[p/2]-out" estimator. We set p equal to five. Next we sort all d_i's by size and then observe where S falls within the distribution of d_i. The p-value is then given simply by the percentage of the d_i values that lie above S.

3 Estimation results

This section summarizes the results for forecasts of growth in GDP and industrial production. Forecasts for GDP growth are made for one to four quarters ahead, and IP forecasts for one-, four-, eight-, and twelve-months ahead.[25] For both indicators, we distinguish between a pre-crisis period (until the end of 2007) and a crisis period (ranging from 2008q1 to 2009q2). Two standard loss functions are used: quadratic error loss and absolute error loss. Accordingly, our out-of-sample performance measures are root mean squared forecast errors (RMSFE) and mean absolute forecast errors (MAFE).

[25] The results for the remaining months are available upon request.

Table 1: Ranking of Indicators for GDP before the crisis: models with greatest forecast accuracy

	I. RMSFE			
	h=1	h=2	h=3	h=4
1	msfe***	msfe**	msfe***	msfe**
2	ECCS99*	ECCS6*	ECCS6*	ECCS6*
3	DIFOWH-C*	DIFOWH-EXP	DLNDAX	trim75
4	IFO-EXP	DLNDAX	DECCS1***	DECCS10
5	DIS-3M	DECCS4	DECCS99**	IFOWH-EXP**
6	DECCS4	DECCS1	DESI-TRADE***	trim50**
7	DIFO-UNCER	ECCS10	DLNHWWA-EX	DLNDAX**
8	DESI-TRADE	DESI-TRADE	trim50	DLNEX
9	DIL-3	DECCS5	trim75	DLNVAC
10	DIFO-C	DECCS8	ECCS10**	trim25*
11	ECCS10	ECCS99	DECCS8	DLNM2R
12	DECCS1	DECCS99	Wright2	DESI-TRADE
13	IFOWH-EXP	DECCS3	DLNM2	DECCS1
14	IFOMV-EXP	DLNM2R*	DLNVAC	DLNHWWA-EX**
15	r^2	ECCS1	DIFOWH-C	Wright20
	II. MAFE			
	h=1	h=2	h=3	h=4
1	msfe***	msfe***	msfe***	ECCS6**
2	ECCS99	ECCS6	ECCS6**	msfe***
3	IFOWH-EXP	DIS-D	DLNDAX	IFOWH-EXP***
4	DIFOWH-C**	ECCS99	DLNHWWA-EX	IFO-C
5	ECCS10	DIFOWH-EXP	DDOILR	trim75
6	DIFOWH-EXP	DLNDAX	DESI-TRADE***	IFOMV-C**
7	DIFO-C*	DECCS99	DECCS99	DLNDAX*
8	DIS-3M	DIFOMI-EXP	DECCS1***	DLNHWWA**
9	DIFOMI-C	med	DIS-3M	DIFO-C
10	IFOMI-C	DLNM2R	IFO-C**	IFO-EXP***
11	DECCS7	DECCS1	ECCS2	ECBS2
12	DECCS4	DIFO-C	trim75*	DLNEX
13	DIFO-UNCER	DECCS4	DIFOWH-EXP***	trim50
14	ECCS4	DESI-TRADE	DIFO-C*	DECCS10
15	DESI-TRADE	DDPBIP	DLNVAC	DLNVAC

Note: The fifteen best leading indicators for real GDP before the crisis are shown (measured with relative Root Mean Square Forecast Errors and relative Mean Absolute Forecast Errors, respectively). A more detailed table can be found in the appendix (see Table 6). ***,**,* indicates whether the forecast ability is significant at the 1%, 5% and 10% level, respectively. The Giacomini-White test for conditional predictive ability is used for that purpose (benchmark model is the AR model).

3.1 Forecasts in the pre-crisis period

Table 1 gives a ranking of the best indicator models during the period 2000q2-2007q4 (a more detailed summary can be found in the appendix, Table 6). For the short horizon (one and two steps ahead), survey measures clearly dominate in terms of forecast accuracy. The Economic Confidence Indicator provided by the European Commission, the ifo business climate and business expectations indexes, as well as some consumer confidence measures perform better than the univariate AR model. ifo wholesale indexes and price expectations of consumers provide particularly good results. However, this difference is statistically significant for only a small proportion. At a longer forecasting horizon some survey-based indicators still do well, but stock prices and commodities also provide useful information for economic growth. The forecasting performance of other financial indicators is limited. Other prominent leading indicators like term spread measures did far worse during that period compared with the benchmark. Moreover, composite indicators do not offer much improvements. Only the OECD leading indicators do slightly better than the benchmark model one quarter ahead. Model averaging schemes improve forecast accuracy. These differences are often statistically significant. In particular, the weights obtained according to past MSFEs (msfe) show large and significant improvements. For three and four quarters ahead, weights based on trimmed forecasts and Bayesian model averaging also perform well in the out-of-sample experiment. For Bayesian weights, high and medium shrinkage does provide slightly better results than a low degree of shrinkage. R^2 and AIC weights have recently been performing well, but the difference relative to the equal weighting scheme is small. Weighting schemes that incorporate the complete covariance, such as the restricted OLS estimator or the Diebold-Pauly method, perform less well in our out-of-sample experiment. This can be attributed to the high number of predictors relative to the sample size.

So far we have concentrated only on pairwise predictive ability. However, this approach does not control for multiple test problems and disregards the correlation between the different models. Employing a joint test is thus generally more reliable. Table 2 presents the results based on the Hansen (2005) methodology. It can be seen that with single indicator models there is no evidence of superior predictive ability. This implies that single models do not significantly outperform the benchmark AR model. When we include the model averaging schemes as well, the results change completely and the test is significant for almost all horizons (excluding the four quarter ahead forecast) and loss functions. This implies that although the single indicators are basically not better than the benchmark, pooling of models results in superior predictive ability compared to the benchmark. This finding is compatible with the general view of D'Agostino et al. (2007) and Campbell (2007) that macroeconomic forecastability has noticeably declined since the 1980s (this is one byproduct of the great moderation). Results from Kholodilin and Siliverstovs

(2006) and Kuzin et al. (2009) suggest similar developments in Germany before the outbreak of the financial crisis. The advantage of forecast combination is that the weight of each indicator can be backtracked and that for each point in time the relative importance of each single indicator model can be assessed. Our large set of indicator forecasts allows us to merge them according to the pre-specified groups presented above (survey indicators, financial variables, ...) for each of the pooling methods. Figures 4 - 7 thus present the time-varying distributions of each indicator group. Naturally the equal weighting scheme serves as the relevant benchmark. Wright weights with low shrinkage ($\phi = 20$), weights based on mean squared forecast errors and including only the 25% and 50% best forecasts (trim75, trim50) yield the most volatile distribution of the blocks over the sample. For some periods the forecast is even based on only two or three blocks. AIC weights only show small time variation.

The results for industrial production forecasts are similar to those of GDP. For a selection of the forecast horizons, the results in Table 3 show the best 15 indicator forecasts based on relative root mean squared forecast errors and relative mean absolute errors during the pre-crisis. European Commission Surveys, the employment rate and ifo expectations, especially the sub-index expectations in manufacturing investment, yield good results. The ifo wholesale indices, both climate and expectations, are once again among the top performers for all forecast horizons. Further, both stock prices and commodities (HWWA indices, oil price) perform across horizons and error measures. For longer horizons, short-term interest rates are useful predictors. In addition, pooled forecasts display a good forecasting performance at all horizons (which are often statistically significant). Again MSFE weights dominate all other weighting schemes. Compared with GDP, even more pooled forecasts are under the best performers in the pre-crisis period. Table 7 in the Appendix provides the results for all indicators and pooled forecasts.

Table 2 also presents the Hansen (2005) SPA test results for IP. Once more, only by including model averaging schemes can the benchmark model be significantly outperformed (at least for steps 4, 8 and 12). Single indicator models are statistically different from simple univariate models only for a one-year horizon. In general we find that forecast combination improves forecasting accuracy during the period 2000 to 2007 and that no single indicator model gives reliable results.

Figures 8-9 show the volatility of the weights associated with each block over time. Due to the high frequency in comparison with to GDP, the weights are even more volatile. Interestingly, the volatility of weights based on mean squared forecast errors, and using only the best 75% of forecast, is more similar to equal, Akaike and R^2-weights for the short horizons.

Because a weighting scheme using high discount is the best forecasting model in our setting, it is obvious that model instabilities are an important issue in macroeco-

Table 2: Test for Superior Predictive Ability (SPA)

I. GDP		h=1			h=2			h=3			h=4		
		SPA_l	SPA_c	SPA_h	SPA_l	SPA_c	SPA_h	SPA_l	SPA_c	SPA_h	SPA_l	SPA_c	SPA_h
RMSFE													
SI	q=0.5	0.222	0.255	0.266	0.213	0.301	0.328	0.375	0.558	0.643	0.784	0.971	0.988
	q=0.25	0.236	0.263	0.280	0.243	0.347	0.379	0.256	0.353	0.408	0.718	0.946	0.968
ALL	q=0.5	**0.013**	**0.014**	**0.014**	**0.070**	**0.088**	**0.089**	**0.010**	**0.011**	**0.011**	0.274	0.384	0.479
	q=0.25	**0.009**	**0.009**	**0.009**	**0.035**	**0.047**	**0.047**	**0.004**	**0.005**	**0.005**	0.205	0.281	0.343
MAFE													
SI	q=0.5	0.377	0.440	0.455	0.375	0.552	0.589	0.754	0.920	0.948	0.752	0.896	0.937
	q=0.25	0.306	0.358	0.372	0.379	0.555	0.602	0.678	0.863	0.886	0.690	0.861	0.893
ALL	q=0.5	**0.006**	**0.006**	**0.006**	**0.039**	**0.059**	**0.059**	**0.009**	**0.012**	**0.012**	0.133	0.180	0.199
	q=0.25	**0.005**	**0.006**	**0.006**	**0.020**	**0.033**	**0.034**	**0.009**	**0.013**	**0.013**	**0.056**	**0.074**	**0.079**
II. IP		h=1			h=4			h=8			h=12		
		SPA_l	SPA_c	SPA_h	SPA_l	SPA_c	SPA_h	SPA_l	SPA_c	SPA_h	SPA_l	SPA_c	SPA_h
RMSFE													
SI	q=0.5	0.748	0.950	0.968	0.716	0.954	0.978	0.498	0.706	0.795	**0.015**	**0.020**	**0.020**
	q=0.25	0.558	0.794	0.839	0.692	0.961	0.971	0.456	0.673	0.722	**0.078**	0.107	0.110
ALL	q=0.5	0.760	0.952	0.971	**0.000**	**0.000**	**0.000**	**0.001**	**0.001**	**0.001**	**0.019**	**0.023**	**0.024**
	q=0.25	0.574	0.801	0.847	**0.000**	**0.000**	**0.000**	**0.001**	**0.001**	**0.001**	**0.073**	**0.095**	**0.099**
MAFE													
SI	q=0.5	0.865	0.979	0.987	0.784	0.956	0.976	0.298	0.481	0.538	**0.068**	0.104	0.121
	q=0.25	0.773	0.942	0.957	0.821	0.976	0.984	0.349	0.539	0.577	0.109	0.154	0.170
ALL	q=0.5	0.128	0.226	0.246	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.006**	**0.007**	**0.008**
	q=0.25	0.181	0.292	0.319	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.029**	**0.038**	**0.040**

Notes: The reported p-values correspond to the multiple test of equal predictive ability. q is a parameter that accounts for the dependencies in the relative loss series ($q = 0.5$ and $q = 0.25$ are reported). 1000 bootstrap replications are used. SPA_h is the upper bound of the test, which corresponds to reality check proposed by White (2000) . The lower bound SPA_l represents the probability value as proposed by Hansen (2005). The null hypothesis is that no model outperforms the benchmark (AR model). SI comprises all single indicator models. ALL additionally includes the model averaging schemes.

Table 3: Ranking of Indicators for IP before the crisis: models with greatest forecast accuracy

I. RMSFE

	h=1	h=4	h=8	h=12
1	msfe***	msfe***	msfe***	msfe***
2	trim75	IFOWH-EXP	IFOWH-EXP	IFOWH-C
3	trim50	DIFOWH-EXP	IFOWH-C	ECCS12**
4	trim25	DLNHWWA-EX	DIFOWH-C	DIFOWH-C
5	ECCS5**	DIFOWH-C	DIFOWH-EXP	IS-M
6	ESI	aic*	DIFOM-C*	IS-D
7	DLNEW*	eq*	DIFO-C	IFOWH-EXP
8	VOLA1	IFO-UNCER	IFO-C	ECCS4
9	DLNDAX	r^2	DIL-3	trim25
10	VOLA2	med	r^2 **	Wright0.5**
11	ECCS9	IFOMI-C	eq**	IFO-C*
12	IFOMI-EXP	Wright0.5	aic**	trim50
13	DDCPI	DIFO-C*	ESI-TRADE	Wright2***
14	DCOM	IFOWH-C	ECCS4*	IS-3M
15	DIFOWH-EXP	Wright2	DIL-5	DECCS10

II. MAFE

	h=1	h=4	h=8	h=12
1	msfe***	msfe***	msfe***	msfe***
2	trim75	DIFOWH-EXP	IFOWH-EXP	IS-M
3	trim50	IFOWH-EXP	DIFOWH-C	trim25
4	DLNDAX	DIFOWH-C	IFOWH-C	IS-D
5	VOLA1	DIFO-C	DIFOWH-EXP	ECCS12**
6	DECCS10	IFO-UNCER	GFK-EXP*	trim50
7	trim25	ESI	DIFO-C	DIFOWH-C
8	DECCS11	DLNHWWA-EX	DIL-3	IFOWH-C
9	ECCS5	IFOMI-C	DIFOM-C*	IS-3M
10	DIFOWH-EXP	aic**	DIL-5	IFOWH-EXP
11	ECCS3	DIFO-UNCER**	DLNM3R	trim75
12	DLNEW	eq**	ESI-TRADE	IFOMV-C
13	VOLA2	r^2 *	r^2 ***	ECCS4
14	DECCS9	ESI-SERV	IFOMV-C*	DLNM2*
15	DECCS5	IFO-EXP	eq***	Wright0.5

Note: The fifteen best leading indicators for real IP before the crisis are shown(measured with relative Root Mean Square Forecast Errors and relative Mean Absolute Forecast Errors, respectively). A more detailed table can be found in the appendix (see Table 6). ***,** ,* indicates whether the forecast ability is significant at the 1%, 5% and 10% level, respectively. The Giacomini-White test for conditional predictive ability is used for that purpose (benchmark model is the AR model).

Table 4: AR Forecast Errors

GDP	RMSFE				MAFE			
	1	2	3	4	1	2	3	4
precrisis	1.97	1.65	1.52	1.48	1.64	1.32	1.24	1.28
crisis	8.84	6.99	6.24	5.10	3.99	3.22	2.93	2.53
total	3.99	3.22	2.93	2.53	2.61	2.02	1.83	1.73
IP	RMSFE				MAFE			
	1	4	8	12	1	4	8	12
precrisis	18.71	5.28	3.87	3.22	15.27	4.39	3.13	2.63
crisis	47.63	25.95	20.01	14.77	34.82	18.62	14.97	10.79
total	25.71	11.64	9.03	6.91	18.44	6.77	5.18	4.10

Note: The Root Mean Squared Forecast Errors and Mean Absolute Forecast Errors for the AR benchmark forecasts are shown for the periods investigated.

nomic forecasting after 2000 and that the relative importance changes very rapidly. This might be one reason why model averaging yields better results than single indicator models.

3.2 Stability during the financial crisis

If we include the most recent period of the financial crisis in our analysis, the picture changes considerably. First, we find that average forecasting errors increased dramatically during the recession. Average forecast errors are about four times greater in the crisis in relation to those in the pre-crisis period (irrespective of whether the RMSFE or MAFE is compared). Table 4 gives an indication of this enormous increase. By considering the last six quarterly forecast errors (or 18 monthly forecast errors) the average forecasting performance of simple econometric models decreased considerably. Generally, while most of the indicator variables point to a slowdown, none of them has adequately recognized the sharpness of the downturn.

Interestingly, while the total forecasting performance worsened during the crisis, the relative performance of indicator forecasts substantially increased. This implies that most models using indicator information perform much better than the univariate AR model. This can be illustrated by Figures 12 and 13, which show the share of predictors displaying a better forecasting performance than that of the benchmark. While prior to the crisis only a small fraction showed lower relative error measures when compared with the AR model, even falling below the 10% level (particularly for longer horizons), over 60% did better in the crisis period for GDP, and even up to even 80% for IP. Over the forecast horizons the power of the indicator forecast decreases, but it still performs better than the benchmark. This indicates that leading indicator forecasts are much more helpful during unusual times (recessions, phases with high volatility) than during low volatility regimes. This emphasizes previous findings that univariate time series models have problems before and after turning points.

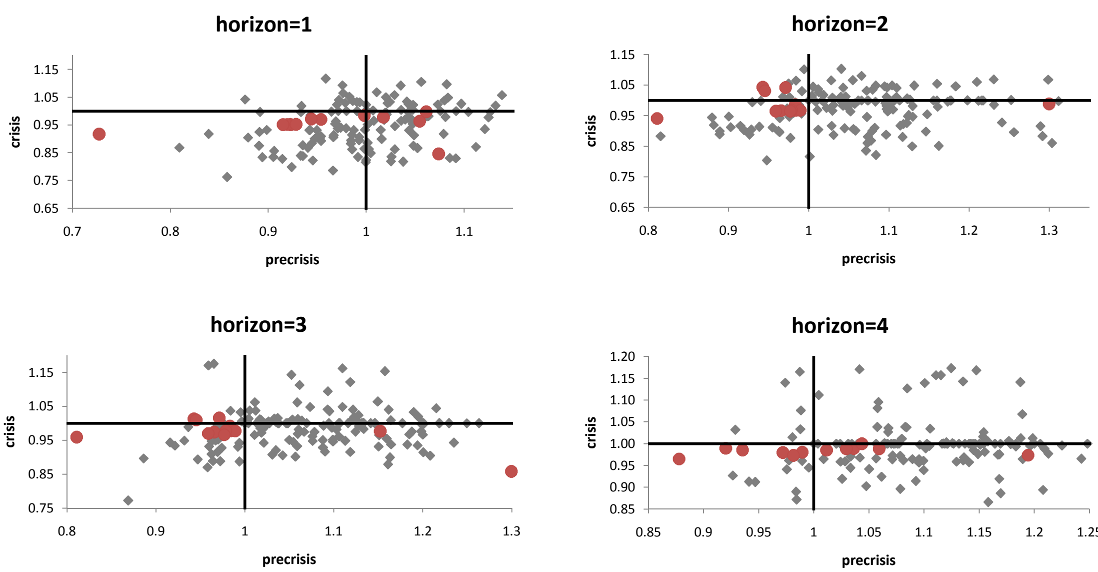
Moreover, Tables 8 and 9 show the leading indicators with the greatest improvements in forecast accuracy. A negative sign indicates that the average forecast error is smaller than the AR benchmark model. In addition, we provide Andrews-type break tests that indicate whether the relative forecasting performance has significantly changed during the crisis period.

The results for GDP indicate that survey indicators offer the greatest improvements for the short horizon. The ifo business expectation index and the sub-index for manufacturing show the smallest forecast errors at the one step ahead horizon. Also the harmonized indicators from the EU commission do well. At larger horizons ($h > 1$) the spread between corporate and government bond yields offers good results during the crisis period. This is in contrast with the pre-crisis results, where this indicator proved not to be particularly useful. Furthermore, the spread between BB-ranked financial cooperations and government bonds offer great improvements compared with the benchmark, which is not surprising since the origins of the recessions are assumed to be in the banking and financial sector. Over and above this, spreads between non-financial cooperations and government bonds perform better in the crisis period. Instead, monetary aggregates do not substantially improve over the univariate benchmark during the financial crisis.

For industrial production, OECD leading indicators do extremely well for shorter horizons during the crisis. With increasing forecast horizons, financial indicators are becoming increasingly important. In particular, spreads (between corporate (financial) and government bond yields as well as BBB-AA spreads) display high relative forecasting accuracy during the recession. The term spread, which has been widely accepted as a standard regression indicator, shows a good forecasting performance for longer horizons, too. The ifo surveys once more provide robust results. Manufacturing and wholesale climate and expectations are among the best indicators during the crisis. Test results indicate that for mean squared loss a significant relative forecasting gain can be obtained compared with the benchmark. Furthermore, with increasing horizon, the number of indicator models and pooled models with significant gain decreases for both mean squared loss and mean absolute loss (see Table 9). In deciding whether there is strong evidence of a break in the forecasting performance with the beginning of the recession depends to an extent on the individual loss function. For mean squared loss there is a stronger hint towards a break when compared with mean absolute loss.

In general, forecast combination schemes do only slightly better than the benchmark model. However, it is important to say that there are nevertheless single models that perform better than combinations (ex-post). Before the outbreak of the crisis there was no hint (e.g. via in-sample information) that one particular model should be used in the immediate future rather than combinations that had performed well previously. This relates to the conclusion by Hibon and Evgeniou (2005) that the advantage of combining forecasts is not that the best possible combinations are necessarily better than the best possible forecasts, but that it is less risky in practice to combine forecasts than to select an individual model (or method).

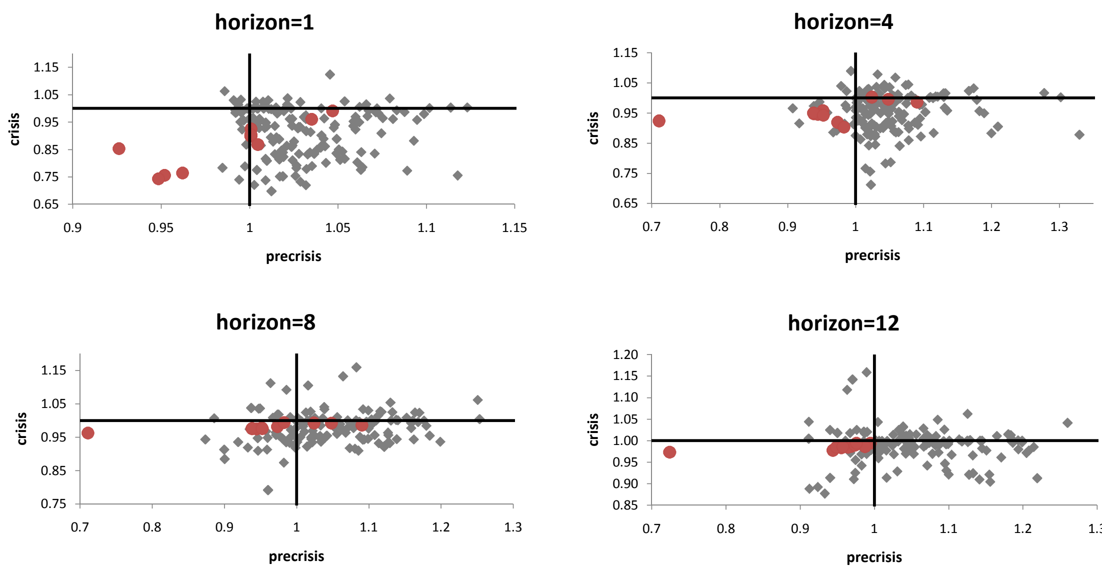
Figure 2: Out-of-sample stability for GDP



Note: Measured by the relative RMSFE, the performance of individual indicator forecasts (grey) compared to the pooled forecasts (red) during the pre-crisis and crisis period are shown for GDP. Some extreme outliers are discarded.

Finally Figures 2 and 3 illustrate the performance of the indicator forecasts and the pooled forecast before and during the crisis. It is obvious that some indicators that performed well before the crisis continued to provide useful information during the turbulent period. A large number of leading indicator models that performed less well in relation to the benchmark before the crisis did well during the crisis (this is in line with Figures 12 and 13). Generally, the performance of the pooled forecasts

Figure 3: Out-of-sample stability for IP



Note: Measured by the relative RMSFE, the performance of individual indicator forecasts (grey) compared to the pooled forecasts (red) during the pre-crisis and crisis period are shown for industrial production. Some extreme outliers are discarded.

in the crisis remained relatively stable. For IP in particular, most of the combined forecasts can be found in the lower left area of the figures, which indicates stability. For some averaging schemes we can even see relative improvements between the two sub-periods, for instance, for Wright weights and Trimmed means. MSFE-weights still do better than the benchmark, but lose to some extent their very dominant position in forecast accuracy during the crisis period. It is also interesting that there is no clear evidence for the increased dispersion of model forecasts during the crisis period despite the fact that visual inspection could lead one to conclude the opposite.

4 Conclusion

This paper analyzes the performance of leading indicator forecasts in the light of the recent recession. In a quasi real time out-of-sample environment, the forecast accuracy of various leading indicators (with special emphasis on financial indicators) is evaluated before and during the crisis. We find evidence that during the period 2000-2007 no single indicator model significantly outperformed the benchmark. However, pooling leading indicators shows promising results and yields significant improvements.

During the financial crisis 2008-2009, a large increase in average forecast errors can be observed. However, at the same time, leading indicators do much better than benchmark univariate time series models. For many indicator models there is evidence of a structural break during that period (in comparison with the pre-crisis period). For both GDP and industrial production, we find that survey indicators did well during the crisis, while at longer horizons financial variables such as term and risk spreads showed remarkable improvement. Model averaging schemes display a relatively stable performance, in comparison with the pre-crisis period.

Our results show that some indicators are useful at extreme turning points (the financial crisis of 2008-2009) which are not helpful in forecasting in tranquil periods, such as term or risk spreads. On the other hand, there are some indicators (mainly from qualitative surveys) that can be characterized by a relative stable performance in the two sub-periods. To some extent this can be attributed to the problems of the AR benchmark models, especially at turning points. Since our tests of stability mainly indicate that the relatively performance changed during the crisis period, it would be interesting to see whether this could be attributed to, for example, non-linearities that might be more important in extreme situations. Furthermore, whether the favorable performance of leading indicator forecasts during the recession implies a *return of leading indicator models* in the future is the subject of further work.

References

Andrews, D. W. (2003), "End-of-Sample Instability Tests". *Econometrica*, **71**(6), 1661–1694.

Armstrong, J. S. and Collopy, F. (1992), "Error measures for generalizing about forecasting methods: Empirical comparisons". *International Journal of Forecasting*, **8**(1), 69–80.

Atkinson, A. C. (1980), "A Note on the Generalized Information Criterion for Choice of a Model". *Biometrika*, **67**(2), 413–418.

Bates, J. M. and Granger, C. W. J. (1969), "The Combination of Forecasts". *Operations Research Quarterly*, **20**(4), 451–468.

Benner, J. and Meier, C.-P. (2004), "Prognosegüte alternativer Frühindikatoren für die Konjunktur in Deutschland". *Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik)*, **224**, 627–652.

Benner, J. and Meier, C.-P. (2005), "Was leisten Stimmungsindikatoren für die Prognose des realen Bruttoinlandsprodukts in Deutschland? Eine Echtzeit-Analyse". *Die Weltwirtschaft*, **2005/3**, 341–355.

Brand, C., Reimers, H.-E. and Seitz, F. (2003), "Forecasting Real GDP: What Role for Narrow Money?" Working Paper 254, ECB.

Breitung, J. and Jagodzinski, D. (2001), "Prognoseeigenschaften alternativer Indikatoren für die Konjunkturentwicklung in Deutschland". *Konjunkturpolitik*, **47**, 292–314.

Campbell, J. Y., Lettau, M., Malkiel, B. G. et al. (2001), "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk". *The Journal of Finance*, **56**(1), 1–43.

Campbell, S. D. (2007), "Macroeconomic Volatility, Predictability, and Uncertainty in the Great Moderation: Evidence From the Survey of Professional Forecasters". *Journal of Business & Economic Statistics*, **25**, 191–200.

CEPR (2009), "Euro Area Business Cycle Dating Committee: Determination of the 2008 Q1 Peak in Economic Activity". Tech. rep., C.E.P.R.

Clark, T. E. and McCracken, M. W. (2001), "Tests of equal forecast accuracy and encompassing for nested models". *Journal of Econometrics*, **105**(1), 85–110.

Costantini, M. and Pappalardo, C. (2009), "A hierarchical procedure for the combination of forecasts". *International Journal of Forecasting*, **In Press, Corrected Proof**, –.

D'Agostino, A., Giannone, D. and Surico, P. (2007), "(Un)Predictability and Macroeconomic Stability". Discussion Papers 6594, C.E.P.R.

Diebold, F. X. and Mariano, R. S. (1995), "Comparing Predictive Accuracy". *Journal of Business & Economic Statistic*, **13**, 253–263.

Diebold, F. X. and Pauly, P. (1990), "The use of prior information in forecast combination". *International Journal of Forecasting*, **6**(4), 503–508.

Drechsel, K. and Maurin, L. (2010), "Flow on conjunctural information and forecast of euro area economic activity". *Journal of Forecasting*, forthcoming.

Dreger, C. and Schumacher, C. (2004), "Estimating Large-Scale Factor Models for Economic Activity in Germany: Do They Outperform Simpler Models?" *Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik)*, **224**, 732–750.

Dreger, C. and Schumacher, C. (2005), "Out-of-sample Perfomance of Leading Indicators for the German Business Cycle: Single vs Combined Forecasts". *Journal of Business Cycle Measurement and Analysis*, **2**(1), 71–88.

Estrella, A. and Hardouvelis, G. (1991), "The Term Structure as a Predictor of Real Economic Activity". *Journal of Finance*, **46**(2), 555–576.

Estrella, A., Rodrigues, A. R. and Schich, S. (2003), "How Stable Is the Predictive PÜower of the Yield Curve? Evidence from Germany and the United States". *The Review of Economics and Statistics*, **85**(3), 629–644.

Forni, M., Hallin, M., Lippi, M. et al. (2003), "Do financial variables help forecasting inflation and real activity in the euro area?" *Journal of Monetary Economics*, **50**(6), 1243–1255.

Fritsche, U. and Stephan, S. (2002), "Leading Indicators of German Business Cycles: An Assessment of Properties". *Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik)*, **223**, 289–315.

Galí, J. and Gertler, M. (1999), "Inflation dynamics: A structural econometric analysis". *Journal of Monetary Economics*, **44**(2), 195–222.

Gertler, M. and Lown, C. S. (1999), "The Information in the High-Yield Bond Spread for the Business Cycle: Evidence and Some Implications". *Oxford Review of Economic Policy*, **3**(3), 132–150.

Giacomini, R. and Rossi, B. (2009),"Forecast Comparisons in Unstable Environments". *Journal of Applied Econometrics*, **forthcoming**.

Giacomini, R. and White, H. (2006),"Tests of Conditional Predictive Ability". *Econometrica*, **74**(6), 1545–1578.

Granger, C. W. and Ramanathan, R. (1984),"Improved Methods of Combining Forecasting". *Journal of Forecasting*, **3**, 197–204.

Hansen, P. R. (2005),"A Test for Superior Predictive Ability". *Journal of Business and Economic Statistics*, **23**, 365–380.

Hibon, M. and Evgeniou, T. (2005),"To combine or not to combine: selecting among forecasts and their combinations". *International Journal of Forecasting*, **21**, 15–24.

Hüfner, F. and Schröder, M. (2002),"Prognosegehalt von ifo-Geschäftserwartungen und ZEW-Konjunkturerwartungen: Ein ökonometrischer Vergleich". *Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik)*, **222**(3), 316–336.

Kapetanios, G., Labhard, V. and 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. A. and 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. and 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.

Kuzin, V., Marcellino, M. and Schumacher, C. (2009),"Pooling versus model selection for nowcasting with many predictors: An application to German GDP". Discussion Papers 7197, C.E.P.R.

Newey, W. K. and West, K. D. (1987),"A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix". *Econometrica*, **55**(4), 703–708.

Robinzonov, N. and Wohlrabe, K. (2009),"Freedom of Choice in Macroeconomic Forecasting". *CESifo Economic Studies*, **in press**.

Scheufele, R. (2010),"Evaluating the German (New Keynesian) Phillips curve". *The North American Journal of Economics and Finance*, **In Press**, doi:10.1016/j.najef.2009.06.001.

Schumacher, C. and 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.

Sims, C. A. (1972),"Money, Income, and Causality". *American Economic Review*, **62**, 540–552.

Stock, J. H. and Watson, M. W. (2002),"Macreconomic forecasting using diffusion indexes". *Journal of Business and Economic Statistics*, **20**, 147–162.

Stock, J. H. and Watson, M. W. (2003a),"Forecasting Output and Inflation: The Role of Asset Prices". *Journal of Economic Literature*, **47**(3), 788–829.

Stock, J. H. and Watson, M. W. (2003b),"How did leading indicator forecasts perform during the 2001 recession?" *Federal Reserve Bank of Richmond, Economic Quarterly*, **89**(3), 71–90.

Stock, J. H. and Watson, M. W. (2004),"Combination forecasts of output growth in a seven-country data set". *Journal of Forecasting*, **23**(6), 405–430.

Swanson, N. R. (1998),"Money and output viewed through a rolling window". *Journal of Monetary Economics*, **41**(3), 455–473.

Timmermann, A. (2006),*Forecast Combinations*, *Handbook of Economic Forecasting*, vol. 1, chap. 4, 135–196. Elsevier.

West, K. D. (1996),"Inference about Predictive Ability". *Econometrica*, **64**(5), 1067–1084.

West, K. D. (2006),"Forecast Evaluation". In G. Elliott, C.W.J. Granger and A. Timmermann (editors), "Handbook of Forecasting", vol. 1, chap. 3, 99–134., Elsevier, Amsterdam.

Wheelock, C., David and Wohar, M. E. (2009),"Can the Term Spread Predict Output Growth and Recessions? A Survey of the Literature". *Federal Reserve Bank of St. Louis Review*, **91**(5), 419–440.

White, H. (2000),"A Reality Check for Data Snooping". *Econometrica*, **68**(5), 1097–1126.

Wright, J. H. (2008),"Bayesian Model Averaging and exchange rate forecasts". *Journal of Econometrics*, **146**(2), 329–341.

Wright, J. H. (2009),“Forecasting US inflation by Bayesian model averaging”. *Journal of Forecasting*, **28**(2), 131–144.

Appendix

Table 5: Indicators and Labels

Block	Name	Label	Data transformation	Publication lag monthly	Publication lag quarterly	Source
Dependent variable						
	GDP, real	BIPR				Destatis
	Industrial production	IP				Buba
Financial						
	Money market rate (mth.avg.)	IS-M	L,D	0	0	Buba
	Discount rate / short term repo rate (mth.avg.)	IS-D	L,D	0	0	Buba
	3m-money market rate (mth.avg.)	IS-3M	L,D	0	0	Buba
	Yields on debt securities outstanding (mat.3-5 years)	IL-3	L,D	0	0	Buba
	Yields on debt securities outstanding (mat.5-8 years)	IL-5	L,D	0	0	Buba
	Long term government bond yield - 9-10 years	IL-10	L,D	0	0	Buba
	Term spread (10y - money market rate)	SPR-10Y-M	L	0	0	Buba
	Term spread (10y - discount rate)	SPR-10Y-D	L	0	0	Buba
	Term spread (10y - 3 month-money market rate)	SPR-10Y-3M	L	0	0	Buba
	Spread (discount rate -money market rate)	SPR-1D-M	L	0	0	Buba
	Corporate bond-government bonds	SPR-C-G	L	0	0	Buba
	Spread corp BBB-corp AA	SPR-B-A	L	0	0	ML
	Spread High Yield - corpAA	SPR-HY-A	L	0	0	ML
	Spread corp BBB- government bond	SPR-B-G	L	0	0	Buba / ML
	Spread corp financial BBB-government bond	SPR-BF-G	L	0	0	Buba / ML
	Spread High Yield - government bond	SPR-HY-G	L	0	0	Buba / ML
	Spread corpAA - government bond	SPR-A-G	L	0	0	Buba / ML
	Nominal effective exchange rate	EX	D ln	1	1	Buba
	Real effective exchange rate	EXR	D ln	1	1	Buba
	DAX share price index	DAX	D ln	0	0	Boerse
	DAX vola new	VOLA1	L,D	0	0	Boerse
	DAX vola old	VOLA2	L,D	0	0	Boerse
	M1	M1	D ln	1	1	Buba
	M1, real	M1R	D ln	1	1	Buba
	M2	M2	D ln	1	1	Buba
	M2, real	M2R	D ln	1	1	Buba
	M3	M3	D ln	1	1	Buba
	M3, real	M3R	D ln	1	1	Buba
	Hwwa index of world market prices of raw mats.,	HWWA	D ln, D^2 ln	1	1	HWWI
	Hwwa index $\sim$, real	HWWAR	D ln, D^2 ln	1	1	HWWI
	Hwwa index $\sim$, energy	HWWA-E	D ln, DD ln	1	1	HWWI
	Hwwa index $\sim$, energy, real	HWWA-ER	D ln, DD ln	1	1	HWWI
	Hwwa index $\sim$, excl. Energy	HWWA-EX	D ln, DD ln	0	0	HWWI
	Hwwa index $\sim$, excl. Energy, real	HWWA-EXR	D ln, DD ln	1	1	HWWI
	Oil prices (euros per barrel)	OIL	D ln, DD ln	0	0	ECB
	Oil prices (euros per barrel), real	OILR	D ln, DD ln	1	1	ECB
Surveys						
	Economic climate	IFO-WC	L,D	0	0	ifo
	Economic expectations	IFO-WEXP	L,D	0	0	ifo
	Ifo index climate	IFO-C	L,D	0	0	ifo
	Ifo expectations climate	IFO-EXP	L,D	0	0	ifo
	Ifo index manufacturing	IFOM-C	L,D	0	0	ifo
	Ifo expectationsmanufacturing	IFOM-EXP	L,D	0	0	ifo
	Ifo index capital goods	IFOMI-C	L,D	0	0	ifo
	Ifo expectationscapital goods	IFOMI-EXP	L,D	0	0	ifo
	Ifo index intermediate goods	IFOMV-C	L,D	0	0	ifo
	Ifo expectationsintermediate goods	IFOMV-EXP	L,D	0	0	ifo
	Ifo index wholesale	IFOWH-C	L,D	0	0	ifo
	Ifo expectations wholesale	IFOWH-EXP	L,D	0	0	ifo
	Ifo: sum of worse and same in expectations and assessment	IFO-UNCER	L,D	0	0	ifo
	GFK consumer climate survey- business cycle expectations	GFK-EXP	L,D	0	0	GfK
	ZEW economic sentiment	ZEW	L,D	0	0	ZEW
	Markit survey, PMI: manufacturing	PMI	L,D	1	1	Markit
	Assessment of order-book levels	ECBS2	L,D	0	0	EC
	Assessment of export order-book levels	ECBS3	L,D	0	0	EC
	Assessment of stocks of finished products	ECBS4	L,D	0	0	EC
	Production expectations for the months ahead	ECBS5	L,D	0	0	EC
	Selling price expectations for the months ahead	ECBS6	L,D	0	0	EC
	Employment expectations for the months ahead	ECBS7	L,D	0	0	EC
	Industrial confidence indicator (40%)	ESI-INDU	L,D	0	0	EC
	Services confidence indicator (30 %)	ESI-SERV	L,D	0	0	EC
	Consumer confidence indicator (20%)	ESI-C	L,D	0	0	EC
	Retail trade confidence indicator (5%)	ESI-TRADE	L,D	0	0	EC
	Construction confidence indicator (5%)	ESI-CTR	L,D	0	0	EC
	Economic sentiment indicator (average)	ESI	L,D	0	0	EC
	Economic Confidence Indicator (average)	ECCS99	L,D	0	0	EC

To be continued...

Block	Name	Label	Data transformation	Publication lag monthly	Publication lag quarterly	Source
	Financial situation over last 12 months	ECCS1	L,D	0	0	EC
	Financial situation over next 12 months	ECCS2	L,D	0	0	EC
	General economic situation over last 12 months	ECCS3	L,D	0	0	EC
	General economic situation over next 12 months	ECCS4	L,D	0	0	EC
	Price trends over last 12 months	ECCS5	L,D	0	0	EC
	Price trends over next 12 months	ECCS6	L,D	0	0	EC
	Unemployment expectations over next 12 months	ECCS7	L,D	0	0	EC
	Major purchases at present	ECCS8	L,D	0	0	EC
	Major purchases over next 12 months	ECCS9	L,D	0	0	EC
	Savings at present	ECCS10	L,D	0	0	EC
	Savings over next 12 months	ECCS11	L,D	0	0	EC
	Statement on financial situation of household	ECCS12	L,D	0	0	EC
Prices and wages						
	CPI	CPI	D ln, DD ln	0	0	Buba
	Core CPI	CPI-EX	D ln, DD ln	1	1	Buba
	Negotiated wage and salary level	TARIF	D ln, DD ln	1	1	Buba
	GDP deflator	PBIP	D ln, DD ln	1	1	Buba
Real Economy						
	Intermediate goods production	IP-VORL	D ln	1	1	Buba
	Manufacturing orders	ORD	D ln	1	1	Buba
	Manufacturing orders- consumer goods	ORD-C	D ln	1	1	Buba
	Manufacturing orders- capital goods	ORD-I	D ln	1	1	Buba
	Employed persons (work-place concept)	EW	D ln	1	1	BfA
	1+unemployment(% civilian labour)	ALQ	D	1	1	BfA
	Vacancies	VAC	D ln	1	1	Buba
	Capacity utilisation	CAPA	L,D	0	0	ifo
	Hours worked	WHOUR	L,D	1	1	Destatis
Composite Indicators						
	FAZ indicator	FAZ	D ln	1	1	IfW
	Early Bird indicator, Commerzbank	COM	D	1	1	Com
	Composite leading indicator (amplitude adjusted)	OECDL1	L,D	2	1	OECD
	Composite leading indicator (trend restored)	OECDL2	D	2	1	OECD
	Composite leading indicator (normalised)	OECDL3	L,D	2	1	OECD
Weights						
	Akaike	aic				
	R^2	r2				
	Trimming the 25% worst	trim25				
	Trimming the 50% worst	trim50				
	Trimming the 75% worst	trim75				
	Mean squared forecast error	msfe				
	OLS	rols				
	Diebold Pauly	dp				
	Wright with $\phi = 0.5$	Wright0.5				
	Wright with $\phi = 2$	Wright2				
	Wright with $\phi = 20$	Wright20				

Note: If the data is not used in levels (L), they are transformed in first differences (D), logged differences (DD ln) and/or second difference (DD ln). The data is published with a lag by 0, 1 or 2 months, and 0 or 1 quarters, respectively. The sources are labeled as follows: Buba - Deutsche Bundesbank, ML - Merrill Lynch, EC - European Commission, BfA - Bundesagentur für Arbeit.

Table 6: Forecast results for GDP based on pre-crisis subsample

	RMSFE h=1	RMSFE h=2	RMSFE h=3	RMSFE h=4	MAFE h=1	MAFE h=2	MAFE h=3	MAFE h=4
	Root Mean Squared Forecast Error				*Mean Absolute Forecast Error*			
AR	1.97	1.65	1.52	1.48	1.64	1.32	1.24	1.28
	RMSFE Rel. to AR Model				*MAFE Rel. to AR Model*			
Interest Rates								
IS-M	0.979	1.097	1.194	1.208	1.021	1.095	1.168	1.192
DIS-M	1.036	1.130	1.110	1.057	1.001	1.078	1.013	1.006
IS-D	0.985	0.975	0.994	1.081	0.966	0.982	1.007	1.101
DIS-D	0.987	0.940	0.969	1.033	0.997	0.908	0.992 *	0.998
IS-3M	0.972	1.018	1.106	1.169	0.980	1.041	1.089	1.160
DIS-3M	0.876	1.098	1.037	1.071	0.899	1.049	0.933	0.986
IL-3	0.970	1.110	1.250	1.254	1.005	1.180	1.231	1.157
DIL-3	0.892	1.028	1.054	1.129	0.925	1.085	1.025	1.032
IL-5	1.045	1.132	1.332	1.382	1.080	1.238	1.358	1.256
DIL-5	0.934	0.975	0.996	1.141	0.961	1.053	1.061	1.065
IL-10	1.097	1.210	1.377	1.392	1.136	1.316	1.403	1.286
DIL-10	0.965	0.984	1.073	1.100	0.992	1.066	1.137	1.054
Interest rates Spreads								
SP10Y-M	1.035	1.161	1.229	1.180	1.000	1.171	1.335	1.207
SP10Y-D	1.080	1.155	1.199	1.137	1.019	1.170	1.260	1.166
SP10Y-3M	1.007	1.066	1.201	1.142	0.936	1.063	1.246	1.174
SP1D-M	1.146	1.342	1.381	1.559	1.106	1.365	1.424	1.533
SPC-G	1.043	1.162	1.161	1.104	1.001	1.132	1.126	1.090
SPB-A	1.717	2.068	2.282	1.704	1.485	1.952	2.133	1.550
SPHY-A	1.218	1.040	1.188	1.802	1.273	1.062	1.217	1.530
SPB-G	1.398	1.572	1.734	2.466	1.309	1.641	1.737	2.139
SPBF-G	1.407	1.649	1.341	1.922	1.277	1.471	1.384	1.531
SPHY-G	1.191	1.046	1.185	2.384	1.238	1.065	1.195	1.839
SPA-G	1.547	1.361	1.215	1.526	1.329	1.298	1.277	1.343
Monetary Aggregates								
DLNM1	0.991	1.080	1.200	1.145	1.042	1.169	1.243	1.136
DLNM1R	1.019	1.105	1.198	1.135	1.052	1.220	1.271	1.168 *
DLNM2	1.020	1.014	0.959	0.987	1.030	0.991	0.967	0.982
DLNM2R	1.011	0.936 *	0.965	0.974	1.016	0.941	0.994	0.970
DLNM3	1.022	1.041	1.110	1.135	1.027	1.056	1.145	1.138
DLNM3R	0.976	0.994	1.052	1.041	0.929	1.039	1.101	1.029
Other financial indicators								
DLNDAX	0.952	0.882	0.886	0.941 **	0.957	0.927	0.895	0.904 *
VOLA1	0.983	1.018	1.017	1.168	0.993	0.989	1.044	1.139
DVOLA1	1.082	1.057	1.046	1.064	1.092	1.033	1.065	1.043
VOLA2	0.979	1.048	1.018	1.149	0.982	1.021	1.050	1.112
DVOLA2	1.056	1.057	1.038	1.076	1.077	1.035	1.053	1.054
DLNEX	1.062	0.976	0.967	0.947	1.083	1.007	1.041	0.930
DLNEXR	1.064	1.013	1.016	1.048	1.051	1.026	1.088	1.025
DLNOILR	1.165	1.043	0.995	1.067	1.161	1.056	0.994	1.011
DDOILR	1.333	1.177	0.976	1.004	1.290	1.130	0.918	0.975
DLNHWWA	1.257	1.231	1.093	1.068	1.254	1.168	1.042	0.910 **
DDHWWA	1.433	1.311	1.146	1.132	1.395	1.276	1.032	0.979
DLNHWWAR	1.218	1.298	1.122	1.106	1.233	1.247	1.073	0.969 **
DDHWWAR	1.466	1.390	1.132	1.200	1.408	1.372	0.974	1.069
DLNHWWA-E	1.127	1.104	1.005	1.083	1.125	1.121	1.008	1.011
DDHWWA-E	1.301	1.217	1.024	1.040	1.252	1.172	0.999	0.991
DLNHWWA-ER	1.156	1.084	1.012	1.084	1.175	1.093	1.020	1.011
DDHWWA-ER	1.303	1.198	1.027	1.031	1.237	1.154	0.988	0.977
DLNHWWA-EXR	1.020	1.095	1.088	1.105	1.036	1.055	1.032	1.105
DDHWWA-EXR	1.104	1.176	1.129	1.065	1.089	1.138	1.088	1.031
Survey Indicators								
IFO-WC	1.091	1.291	1.163	1.178	1.042	1.164	1.049	0.961
DIFO-WC	1.180	1.289	1.150	1.136	1.142	1.240	1.064	0.965 *
IFO-WEXP	0.993	1.073	1.164	1.141	1.007	1.043	1.127	1.083
DIFO-WEXP	0.971	1.231	1.134	1.072	0.948	1.195	1.079	0.994
IFO-C	0.945	0.977	0.970 **	0.988	0.992	1.008	0.935 **	0.883
DIFO-C	0.894	0.948	0.972 **	0.996	0.892 *	0.947	0.945 *	0.915
IFO-EXP	0.858	1.084	1.057	1.071	0.925	1.042	0.986 **	0.923 ***
DIFO-EXP	1.016	1.114	1.090	1.089	0.973	1.093	1.048	0.956 **
IFOM-C	0.924	1.073	1.064	1.101	0.971	1.102	1.059	1.009
DIFOM-C	0.986 *	1.071	1.031	1.095	0.969	1.053	0.986 *	0.996
IFOM-EXP	0.966	1.303	1.201	1.205	1.001	1.214	1.093	1.044
DIFOM-EXP	1.121	1.173	1.164	1.167	1.015	1.102	1.087	1.022
IFOMI-C	0.933	1.129	1.162	1.187	0.907	1.146	1.152	1.075
DIFOMI-C	0.941	1.001	1.118	1.143	0.902	1.008	1.036	1.003
IFOMI-EXP	1.048	1.130	1.128	1.119	1.005	0.996	1.075	1.028
DIFOMI-EXP	0.971	1.052	1.027	1.116	0.967	0.938	0.955 *	1.004
IFOMV-C	0.954	1.079	1.033	1.004	0.997	1.148	0.982	0.898 **
DIFOMV-C	1.049	1.117	1.059	1.081	1.047	1.139	1.039	1.001
IFOMV-EXP	0.914	1.058	1.076	1.101	0.952	1.044	1.042	0.974 *
DIFOMV-EXP	0.968	1.146	1.088	1.121	0.974	1.161	1.025	0.988

To be continued. . .

	RMSFE h=1	RMSFE h=2	RMSFE h=3	RMSFE h=4	MAFE h=1	MAFE h=2	MAFE h=3	MAFE h=4
	Root Mean Squared Forecast Error				*Mean Absolute Forecast Error*			
AR	1.97	1.65	1.52	1.48	1.64	1.32	1.24	1.28
	RMSFE Rel. to AR Model				*MAFE Rel. to AR Model*			
IFOWH-C	0.954	1.123	1.120	1.085	0.941	1.100	1.135	0.994 *
DIFOWH-C	0.839 *	0.962	0.963	1.050	0.868 **	1.021	1.007	1.045
IFOWH-EXP	0.906	0.981	0.963 **	0.929 **	0.859	0.995	0.986 ***	0.846 ***
DIFOWH-EXP	0.953	0.879	0.969 ***	1.028	0.891	0.921	0.943 ***	0.951 ***
ZEW	1.018	1.099	1.177	1.177	1.045	1.166	1.168	1.121
DZEW	0.999	1.096	1.107	1.103	0.982	1.113	1.076	1.056
PMI	1.008	1.499	1.445	1.531	1.013	1.347	1.334	1.366
DPMI	0.943	1.252	1.359	1.248	0.915	1.107	1.301	1.192
GFK-EXP	0.966	1.022	1.078	1.120	0.994	1.067	1.105	1.065
DGFK-EXP	0.951	0.991	1.050	1.078	0.993	1.054	1.080	1.051
IFO-UNCER	0.992	1.096	1.124	1.125	1.030	1.128	1.139	1.031
DIFO-UNCER	0.891	1.078	1.035	1.099	0.912	1.048	1.013	0.986
ECBS2	1.052	1.073	1.190	1.080	1.025	1.053	1.141	0.927
DECBS2	0.947	1.056	1.200	1.117	0.927	1.014	1.117	0.987
ECBS3	0.985	1.109	1.157	1.109	0.955	1.161	1.158	1.025
DECBS3	1.003	0.955	1.054	1.093	0.992	0.968	0.982	1.027
ECBS4	1.029	1.125	1.328	1.541	1.054	1.092	1.225	1.337
DECBS4	1.026	1.148	1.141	1.150	1.037	1.075	1.031	1.029
ECBS5	1.085	0.992	1.062	1.152	1.104	1.001	1.033	1.052
DECBS5	1.038	1.043	1.035	1.064	1.035	1.072	1.018	1.008
ECBS6	0.945	1.122	1.133	1.154	0.917	1.056	1.057	1.057
DECBS6	1.079	1.074	1.099	1.154	1.012	1.029	1.010	1.023
ECBS7	0.990	1.055	1.123	1.146	0.997	1.114	1.080	1.100
DECBS7	1.048	1.039	1.027	1.126	1.045	1.092	1.000	1.059
ESI-INDU	1.005	1.030	1.095	1.059	0.986	1.026	1.129	1.042
DESI-INDU	1.025	1.045	1.067	1.056	0.988	1.012	1.043	1.035
ESI-SERV	0.935	1.082	1.093	1.169	0.949	1.089	1.046	1.027
DESI-SERV	0.999	1.031	1.021	1.061	1.019	1.037	0.962	0.969
ESI-C	1.423	1.555	1.721	1.589	1.231	1.416	1.634	1.491
DESI-C	1.747	1.756	1.971	1.750	1.447	1.517	1.852	1.522
ESI-TRADE	0.969	0.976	1.098	1.183	0.989	1.033	1.089	1.133
DESI-TRADE	0.891	0.903	0.932 ***	0.976	0.914	0.950	0.929 ***	0.949
ESI-CTR	0.967	0.982	1.019	1.096	0.948	1.010	1.000	1.057
DESI-CTR	0.993	0.964	1.076	1.155	0.972	0.973	1.109	1.184
ESI	0.959	1.022	1.118	1.147	0.953	1.042	1.102	1.096
DESI	1.043	1.126	1.060	1.168	1.013	1.158	1.096	1.198
ECCS99	0.809 *	0.926	0.992 ***	1.042	0.815	0.919	1.002	0.972
DECCS99	0.922	0.929	0.921 **	1.009	0.949	0.933	0.930	0.947
ECCS1	0.999	0.938	1.099	1.209	0.989	0.988	1.115	1.157
DECCS1	0.905	0.889	0.916 ***	0.978	0.961	0.943	0.932 ***	0.970
ECCS2	1.040	1.005	0.973	1.036	1.049	1.023	0.938	1.016
DECCS2	0.991	1.123	1.124	1.095	1.021	1.129	1.134	1.159
ECCS3	1.015	1.048	1.113	1.243	0.994	1.043	1.127	1.213
DECCS3	1.036	0.930	1.011	1.024	1.054	0.981	1.031	1.044
ECCS4	0.935	0.964	1.017	1.058	0.913	1.026	1.040	1.051
DECCS4	0.879	0.889	0.964	0.984	0.908	0.949	0.989	1.001
ECCS5	1.001	1.029	1.059	1.188	1.001	1.093	1.114	1.141
DECCS5	1.000	0.915	0.974	0.989	0.995	0.962	0.971 *	0.981
ECCS6	0.993	0.815 *	0.869 *	0.903 *	0.960	0.823	0.814 **	0.810 **
DECCS6	1.132	1.217	1.263	1.351	1.126	1.170	1.273	1.339
ECCS7	1.047	1.177	1.443	1.571	1.037	1.193	1.339	1.352
DECCS7	0.933	1.012	1.011	0.988	0.908	1.033	1.021	0.955
ECCS8	1.029	1.011	1.195	1.189	1.041	1.039	1.160	1.153
DECCS8	0.998	0.920	0.958	1.026	1.011	0.963	0.951	1.028
ECCS9	0.943	1.009	1.042	1.078	0.935	1.017	1.049	1.050
DECCS9	1.041	1.160	1.060	1.158	1.026	1.167	1.095	1.203
ECCS10	0.899	0.900	0.947 **	0.984	0.876	0.953	0.962	0.998 *
DECCS10	0.942	0.967	1.088	0.927	0.955	1.045	1.126	0.935
ECCS11-	1.243	1.097	1.198	1.371	1.167	1.074	1.205	1.266
DECCS11-	1.178	1.031	1.013	1.070	1.167	0.986	1.044	1.053
ECCS12-	1.125	1.106	1.103	1.213	1.080	1.143	1.145	1.216
DECCS12-	1.048	0.968	1.021	1.033	1.084	1.030	1.085	1.060
Real Economic Indicators								
CAPA	0.976	1.043	1.048	1.046	0.915	1.000	1.082	1.039
DCAPA	0.986	1.030	1.146	1.127	0.954	0.979	1.142	1.088
WHOUR	1.001	1.215	1.218	1.454	1.004	1.230	1.160	1.378
DWHOUR	1.092	1.098	1.106	1.145	1.076	1.048	1.095	1.104
DLNIP-VORL	1.083	1.198	1.200	1.164	1.043	1.136	1.194	1.094
DLNORD	1.138	1.241	1.208	1.189	1.068	1.121	1.092	1.115
DLNORD-C	1.000	1.012	1.028	1.024	1.000	1.009	1.026	1.034
DLNORD-I	1.178	1.104	1.183	1.158	1.125	1.044	1.100	1.080
DLNEW	1.081	1.066	1.059	1.198	1.102	1.060	1.031	1.188
DALQ	1.061	1.062	1.025	1.103	1.063	1.038	1.038	1.102
DLNVAC	1.045	1.004	0.960	0.953	1.012	1.025	0.947	0.938
Prices and Wages								
DLNCPI	1.029	1.083	1.020	1.196	0.973	1.050	1.041	1.166

To be continued...

	RMSFE								MAFE							
	h=1		h=2		h=3		h=4		h=1		h=2		h=3		h=4	
	Root Mean Squared Forecast Error								*Mean Absolute Forecast Error*							
AR	1.97		1.65		1.52		1.48		1.64		1.32		1.24		1.28	
	RMSFE Rel. to AR Model								*MAFE Rel. to AR Model*							
DDCPI	1.024		1.031		1.004		1.020		0.994		1.025		1.009		1.015	
DLNCPI-EX	0.950		0.973		0.994		1.098		0.969		0.960		1.017		1.038	
DDCPI-EX	1.044		0.983		1.004		1.034		1.054		0.973		1.017		1.016	
DLNPBIP	1.043		1.097		1.128		1.153		0.999		1.099		1.045		1.080	
DDPBIP	0.980		1.016		1.069		1.077		0.955		0.951		1.039		1.018	
DLNTARIF	1.073		1.074		1.116		1.183		1.060		1.067		1.146		1.147	
DDTARIF	1.000		1.000		1.000		1.000		1.000		1.000		1.000		1.000	
Composite Leading Indicators																
DLNFAZ	1.215		1.118		1.151		1.121		1.186		1.020		1.082		1.069	
DCOM	1.076		1.212		1.219		1.169		1.041		1.257		1.177		1.108	
OECDL1	0.974		1.183		1.150		1.115		0.977		1.144		1.030		0.949	**
DOECDL1	1.001		1.124		1.159		1.151		0.998	*	1.073		1.096		1.004	
DOECDL2	1.112		1.256		1.235		1.225		1.053		1.187		1.153		1.060	
OECDL3	0.974		1.183		1.149		1.111		0.977		1.145		1.031		0.942	**
DOECDL3	1.001		1.124		1.159		1.151		0.999	*	1.074		1.096		1.005	
Model Averaging																
eq	0.923	*	0.963		0.984	*	1.036		0.941	*	0.967		0.980	**	0.990	
med	0.954	*	0.948		0.983		1.044		0.945	*	0.939		0.968		1.004	
aic	0.920	*	0.966		0.989	*	1.059		0.938	*	0.969		0.982	**	1.019	
r^2	0.915		0.967		0.979	**	1.030		0.941	*	0.966		0.974	**	0.992	
trim25	0.998		0.999		0.971		0.972	*	0.975		0.979		1.013		0.980	
trim50	1.017		0.999		0.943		0.935	**	0.999		0.964		0.969	**	0.935	
trim75	1.055		1.005		0.945		0.920		1.036		0.992		0.943	*	0.896	
msfe	0.727	***	0.760	**	0.811	***	0.878	**	0.707	***	0.715	***	0.783	***	0.840	***
rols	1.074		1.034		1.152		1.194		1.078		1.029		1.142		1.143	
dp	1.084		1.275		1.300		1.327		1.058		1.330		1.401		1.261	
Wright0.5	0.928		0.952		0.966	*	1.012		0.946		0.954		0.956	*	0.968	
Wright2	0.944		0.958		0.959		0.990	**	0.963		0.959		0.957		0.973	*
Wright20	1.061		1.147		0.977		0.982		1.056		1.194		0.999		0.980	

Note: The entry in the first line is the RMSFE and the MAFE of 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 ends in 2007Q4. The abbreviation of leading indicators are outlined in Table 5. ***: 1%, **: 5% and *: 10% significance level of equal conditional predictability of Giacomini-White.

Table 7: Forecast results for IP based on pre-crisis subsample

	RMSFE h=1	RMSFE h=4	RMSFE h=8	RMSFE h=12	MAFE h=1	MAFE h=4	MAFE h=8	MAFE h=12
	Root Mean Squared Forecast Error				*Mean Absolute Forecast Error*			
AR	18.71	5.28	3.87	3.22	15.27	4.40	3.13	2.63
	RMSFE Rel. to AR Model				*MAFE Rel. to AR Model*			
Interest Rates								
IS-M	1.017	1.028	1.092	0.924	1.029	1.010	1.118	0.897
DIS-M	1.009	1.043	1.117	1.073	1.015	1.040	1.117	1.074
IS-D	1.019	0.985	1.022	0.933	1.029	0.996	1.027	0.908
DIS-D	1.038	1.019	1.013	1.017	1.054	1.008	0.998	1.026
IS-3M	1.010	1.021	1.034	0.960	1.023	0.998	1.085	0.936
DIS-3M	1.024	1.032	1.038	1.071	1.040	1.006	1.047	1.076
IL-3	1.067	0.987	1.036	1.071	1.077	0.962	1.012	1.059
DIL-3	1.081	1.005	0.937	0.967	1.072	1.005	0.925	0.990
IL-5	1.066	1.050	1.099	1.199	1.071	1.024	1.081	1.203
DIL-5	1.067	1.016	0.947	0.995	1.058	1.015	0.930	1.003
IL-10	1.079	1.074	1.130	1.260	1.079	1.050	1.108	1.289
DIL-10	1.071	1.039	0.974	1.002	1.061	1.038	0.969	1.014
Interest rates Spreads								
SPR-10Y-M	1.075	1.087	1.129	1.155	1.053	1.093	1.105	1.162
SPR-10Y-D	1.016	1.092	1.114	1.095	1.002	1.097	1.103	1.093
SPR-10Y-3M	1.024	1.063	1.121	1.141	0.993	1.062	1.101	1.149
SPR-1D-M	1.045	1.059	1.152	1.335	1.048	1.037	1.153	1.435
SPR-C-G	1.012	1.036	1.121	1.100	1.014	1.026	1.109	1.165
SPR-B-A	1.034	1.209	2.357	2.191	1.024	1.134	2.140	2.209
SPR-HY-A	1.050	1.098	1.097	1.034	1.061	1.085	1.084	1.025
SPR-B-G	1.049	1.200	1.651	2.244	1.029	1.148	1.452	1.722
SPR-BF-G	1.233	1.329	1.693	2.435	1.160	1.221	1.464	2.124
SPR-HY-G	1.051	1.051	1.104	1.032	1.056	1.039	1.095	1.027
SPR-A-G	1.062	1.131	1.159	1.412	1.058	1.131	1.175	1.365
Monetary Aggregates								
DLNM1	1.072	1.109	1.068	1.114	1.101	1.109	1.075	1.131
DLNM1R	1.062	1.077	1.021	1.078	1.082	1.082	1.012	1.114
DLNM2	1.123	1.088	1.064	0.970	1.101	1.085	1.061	0.955 *
DLNM2R	1.114	1.065	1.016	0.963	1.088	1.065	1.010	0.962
DLNM3	1.102	1.078	1.083	1.090	1.094	1.051	1.061	1.091
DLNM3R	1.041	1.017	0.964	0.989	1.041	1.004	0.932	0.980
Other financial indicators								
DLNDAX	0.992	1.004	1.127	1.171	0.971	0.979	1.128	1.228
VOLA1	0.992	1.015	1.164	1.183	0.981	1.007	1.158	1.201
DVOLA1	1.002	1.060	1.087	1.145	1.004	1.055	1.086	1.178
VOLA2	0.993	1.035	1.170	1.180	0.990	1.020	1.161	1.185
DVOLA2	1.001	1.064	1.086	1.142	1.000	1.059	1.076	1.173
DLNEX	1.099	1.174	1.161	1.186	1.071	1.172	1.134	1.199
DLNEXR	1.083	1.167	1.140	1.193	1.044	1.163	1.146	1.226
DLNOILR	1.014	1.121	1.004	0.986	1.015	1.115	1.016	1.013
DDOILR	1.023	1.000	1.001	1.001	1.024	1.006	1.002	1.003
DLNHWWA	1.027	1.132	1.071	1.038	1.029	1.125	1.097	1.079
DDHWWA	1.005	1.011	1.131	1.066	1.007	1.009	1.145	1.080
DLNHWWAR	1.048	1.129	1.104	1.039	1.037	1.115	1.125	1.086
DDHWWAR	1.009	1.007	1.110	1.066	1.009	1.002	1.140	1.082
DLNHWWA-E	1.029	1.118	1.017	0.985	1.027	1.122	1.030	1.012
DDHWWA-E	1.032	1.015	1.002	1.001	1.036	1.017	1.002	1.001
DLNHWWA-ER	1.028	1.130	1.005	0.990	1.026	1.136	1.020	1.021
DDHWWA-ER	1.024	1.015	1.001	1.001	1.028	1.016	1.002	1.002
DLNHWWA-EXR	1.032	1.006	1.038	1.025	1.012	1.027	1.070	1.063
DDHWWA-EXR	1.021	1.031	1.138	1.070	1.014	1.013	1.187	1.105
Survey Indicators								
IFO-C	1.093	1.040	0.935	0.950 *	1.103	0.978	0.943	0.960
DIFO-C	1.049	0.972 *	0.932	0.989	1.059	0.947	0.923	1.005
IFO-EXP	1.064	1.007	1.001	1.050	1.058	0.959	1.000	1.052
DIFO-EXP	1.042	1.029	0.998	1.118	1.021	1.004	1.023	1.138
IFOM-C	1.037	1.000	0.971	1.056	1.062	0.969	0.946	1.044
DIFOM-C	1.024	1.029	0.919 *	1.086	1.027	1.014	0.926 *	1.113
IFOM-EXP	1.026	1.065	1.081	1.194	1.038	1.018	1.071	1.212
DIFOM-EXP	1.035	1.096	1.088	1.129	1.046	1.063	1.113	1.137
IFOMI-C	1.009	0.967	1.108	1.165	1.016	0.950	1.071	1.206
DIFOMI-C	1.019	1.020	1.073	1.145	1.024	1.011	1.087	1.219
IFOMI-EXP	0.994	1.078	1.033	1.083	0.999	1.043	1.054	1.102
DIFOMI-EXP	1.024	1.078	1.024	1.093	1.022	1.050	1.039	1.121
IFOMV-C	1.043	1.014	0.975	0.967	1.053	0.972	0.937 *	0.950
DIFOMV-C	1.021	1.024	0.972 *	1.048	1.026	1.007	0.953 **	1.060
IFOMV-EXP	1.017	1.071	1.056	1.094	1.026	1.020	1.044	1.065
DIFOMV-EXP	1.040	1.028	1.020	1.069	1.050	0.993	1.025	1.048
IFOWH-C	1.021	0.979	0.886	0.911	1.040	0.966	0.900	0.934
DIFOWH-C	1.029	0.947	0.900	0.912	1.027	0.928	0.887	0.932
IFOWH-EXP	0.997	0.908	0.874	0.940	0.994	0.920	0.858	0.944
DIFOWH-EXP	0.995	0.917	0.900	0.972	0.986	0.917	0.906	0.972

To be continued...

	RMSFE				MAFE			
	h=1	h=4	h=8	h=12	h=1	h=4	h=8	h=12
	Root Mean Squared Forecast Error				*Mean Absolute Forecast Error*			
AR	18.84	5.33	3.88	3.23	15.38	4.43	3.14	2.6354
	RMSFE Rel. to AR Model				*MAFE Rel. to AR Model*			
ZEW	1.059	1.115	1.199	1.208	1.054	1.078	1.166	1.251
DZEW	0.999	1.045	1.077	1.100	1.000	1.004	1.095	1.158
PMI	1.054	1.178	1.253	1.353	1.054	1.150	1.254	1.437
DPMI	1.063	1.135	1.180	1.187	1.077	1.130	1.204	1.248
GFK-EXP	1.059	1.064	0.980	1.044	1.038	1.063	0.913 *	1.055
DGFK-EXP	1.073	1.103	1.039	1.128	1.052	1.112	1.006	1.182
IFO-UNCER	1.009	0.959	0.977	1.033	1.023	0.947	0.985	1.026
DIFO-UNCER	1.004	1.000	0.960	1.040	1.012	0.955 **	0.963	1.058
ECBS2	1.014	1.075	1.116	1.047	1.027	1.065	1.127	1.036
DECBS2	1.028	1.011	1.060	1.088	1.041	1.018	1.093	1.094
ECBS3	1.063	1.064	1.047	1.099	1.080	1.053	1.022	1.100
DECBS3	1.071	1.054	1.041	1.053	1.078	1.071	1.020	1.054
ECBS4	1.035	1.049	1.251	1.473	1.043	1.050	1.276	1.475
DECBS4	1.061	1.027	1.126	1.151	1.075	1.046	1.132	1.181
ECBS5	1.075	1.003	1.025	0.994 *	1.064	0.980	1.025	1.018
DECBS5	1.060	1.019	1.025	1.054	1.047	1.011	1.030	1.061
ECBS6	1.004	1.049	1.009	1.042	1.009	1.022	1.018	1.059
DECBS6	1.018	1.059	1.033	1.066	1.019	1.045	1.044	1.099
ECBS7	1.060	1.277	1.113	1.104	1.035	1.171	1.078	1.057
DECBS7	1.048	1.302	1.132	1.183	1.028	1.171	1.109	1.118
ESI-INDU	1.027	1.053	0.997	0.974	1.039	1.033	0.998	0.958
DESI-INDU	1.026	1.057	1.022	1.012	1.048	1.050	1.016	1.030
ESI-SERV	1.032	1.004	0.998	1.038	1.062	0.959	0.990	1.067
DESI-SERV	1.041	1.015	0.990	1.048	1.062	1.009	0.985	1.060
ESI-C	1.095	1.441	1.583	1.962	1.071	1.367	1.462	1.995
DESI-C	1.088	1.413	1.629	2.098	1.091	1.303	1.400	1.828
ESI-TRADE	1.004	1.000	0.946	1.005	1.009	1.003	0.935	1.005
DESI-TRADE	1.010	0.987	0.992	1.042	1.011	0.979	0.991	1.049
ESI-CTR	1.003	1.027	1.100	1.127	1.016	1.029	1.119	1.121
DESI-CTR	1.011	1.056	1.159	1.205	1.019	1.067	1.170	1.204
ESI	0.986	0.993	0.986	1.126	0.997	0.948	0.968	1.148
DESI	1.004	1.027	0.991 **	0.974	1.012	1.017	0.978 **	0.966
ECCS99	1.033	1.022	1.184	1.132	1.057	0.971	1.213	1.186
DECCS99	1.013	1.037	1.005	1.042	1.026	1.022	1.016	1.066
ECCS1	0.998	1.024	0.982	0.982	1.001	1.036	0.965	0.968
DECCS1	1.006	1.010	0.976	1.024	1.006	1.018	0.953	1.039
ECCS2	1.012	1.068	1.052	0.985	1.015	1.074	1.073	0.993
DECCS2	1.016	1.003	1.011	0.994	1.027	0.988	1.015	1.010
ECCS3	0.999	1.047	1.056	1.086	0.988	1.045	1.031	1.086
DECCS3	1.015	1.011	1.031	1.045	1.013	1.018	1.019	1.053
ECCS4	1.015	1.022	0.947 *	0.940	0.997	1.013	0.943	0.950
DECCS4	1.017	1.017	0.959	0.982 **	1.005	1.012	0.953	0.988 **
ECCS5	0.985 **	1.041	0.990	1.008	0.984	1.053	0.969	1.008
DECCS5	0.997	1.020	0.986 **	1.016	0.992	1.017	0.971	1.024
ECCS6	1.004	1.031	1.086	1.156	1.007	1.043	1.131	1.195
DECCS6	1.006	1.045	1.022	1.016	1.012	1.021	1.016	1.013
ECCS7	1.004	1.042	1.379	1.536	1.005	1.045	1.365	1.423
DECCS7	1.020	1.025	1.003	1.002	1.015	1.035	1.003	1.003
ECCS8	1.017	0.997	0.958	1.026	1.016	1.004	0.948	1.018
DECCS8	1.019	0.984	0.975	1.006	1.024	0.978	0.949	1.002
ECCS9	0.994	1.110	1.063	1.219	1.001	1.094	1.088	1.303
DECCS9	0.996	1.058	1.168	1.097	0.991	1.055	1.122	1.098
ECCS10	1.009	1.182	1.139	0.990	0.998	1.161	1.085	1.013
DECCS10	1.000	1.189	1.162	0.961	0.982	1.162	1.119	0.975
ECCS11-	1.042	1.054	1.104	1.214	1.028	1.029	1.106	1.247
DECCS11-	1.001	1.032	1.005	1.010	0.983	1.015	1.008	1.025
ECCS12-	1.008	1.062	1.019	0.911 **	1.007	1.063	1.013	0.911 **
DECCS12-	1.023	1.063	1.029	1.010	1.019	1.053	1.026	1.031
Real Economic Indicators								
DLNIP-VORL	1.052	1.033	1.082	1.045	1.040	1.012	1.094	1.052
DLNORD	1.089	1.095	1.126	1.096	1.097	1.049	1.151	1.173
DLNORD-C	1.019	1.001	1.004	1.001	1.009	1.000	1.004	1.003
DLNORD-I	1.043	1.053	1.094	1.101	1.048	1.041	1.106	1.145
DLNEW	0.991 *	1.028	1.002	1.038	0.989	1.018	1.011	1.077
DALQ	1.054	1.006	1.015	1.046	1.055	1.020	1.010	1.046
DLNVAC	1.008	1.042	0.967	1.006	1.005	1.022	0.987	1.018
Prices and Wages								
DLNCPI	0.996	1.014	1.020	1.125	1.002	1.014	1.025	1.141
DDCPI	0.995	1.003	1.005	1.002	0.995	1.003	1.004	1.003
DLNCPI-EX	1.011	1.058	1.079	1.133	1.014	1.068	1.092	1.165
DDCPI-EX	1.002	1.001	1.003	1.001	1.000	1.001	1.003	1.003
DLNTARIF	1.000	1.114	1.176	1.199	1.000	1.087	1.178	1.261
Composite Leading Indicators								
DLNFAZ	1.118	1.078	1.099	1.082	1.078	1.043	1.125	1.132
DCOM	0.995	1.082	1.092	1.125	0.994	1.063	1.130	1.167
OECDL1	1.003	1.015	1.014	1.066	0.997 *	0.961	0.993	1.082

To be continued...

	RMSFE								MAFE							
	h=1		h=4		h=8		h=12		h=1		h=4		h=8		h=12	
	Root Mean Squared Forecast Error								*Mean Absolute Forecast Error*							
AR	18.84		5.33		3.88		3.23		15.38		4.43		3.14		2.6354	
	RMSFE Rel. to AR Model								*MAFE Rel. to AR Model*							
DOECDL1	1.017		1.022		0.982	*	1.003		0.999		0.989		0.993	*	1.011	
DOECDL2	1.032		1.023		0.960		1.031		1.038		1.005		0.995		1.013	
OECDL3	1.012		1.052		1.042		1.085		1.006		0.990		1.027		1.098	
DOECDL3	1.029		1.044		0.998		1.046		1.012		1.016		1.013		1.061	
Model Averaging																
eq	1.001		0.958	*	0.938	**	0.969		1.000		0.956	**	0.937	***	0.999	**
med	1.001		0.963		0.952	**	0.994	***	0.999		0.970		0.945	***	1.017	
aic	1.001		0.957	*	0.944	**	0.987		1.000		0.954	**	0.943	**	1.025	
r^2	1.001		0.959		0.938	**	0.964	**	1.000		0.958	*	0.936	***	0.992	***
trim25	0.962		0.984		1.024		0.944		0.982		0.981		1.030		0.901	
trim50	0.952		0.991		1.048		0.955		0.971		0.990		1.059		0.930	
trim75	0.948		0.997		1.090		0.974		0.967		0.990		1.120		0.946	
msfe	0.926	***	0.797	***	0.711	***	0.724	***	0.904	***	0.763	***	0.631	***	0.638	***
rols	1.481		1.534		1.867		1.361		1.408		1.382		1.581		1.417	
dp	1.646		1.378		1.443		1.403		1.629		1.382		1.506		1.441	
Wright0.5	1.005		0.972		0.952	**	0.949	**	1.011		0.974		0.947	**	0.956	
Wright2	1.035		0.983		0.973		0.957	***	1.051		0.983		0.960		0.961	**
Wright20	1.047		1.026		0.983		0.976		1.055		1.037		0.973		0.978	

Note: The entry in the first line is the RMSFE and the MAFE of 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 ends in 2007Q4. The abbreviation of leading indicators are outlined in Table 5. ***: 1%, **: 5% and *: 10% significance level of equal conditional predictability of Giacomini-White.

Table 8: Performance and Stability of leading indicator forecasts for GDP during the crisis

	h=1			h=2			h=3			h=4		
		Av. gain	Stability p-Value		Av. gain	Stability p-Value		Av. gain	Stability p-Value		Av. gain	Stability p-Value
RMSFE												
1	IFO-EXP	-5.72	0.000	SPR-BF-G	-5.44	0.111	SPR-BF-G	-4.97	0.000	ECCS6	-3.57	0.000
2	IFOM-EXP	-5.47	0.000	DIFO-C	-4.15	0.000	ECCS6	-3.96	0.000	SPR-BF-G	-3.51	0.240
3	IFOM-C	-5.32	0.000	DIFOMI-C	-4.04	0.000	dp	-3.20	0.000	DLNCPI-EX	-3.48	0.000
4	DESI-SERV	-5.12	0.000	IFO-EXP	-3.98	0.000	DECCS8	-3.08	0.000	SPR-C-G	-2.83	0.000
5	DECBS5	-5.09	0.000	DIFOM-C	-3.83	0.000	SPR-C-G	-2.98	0.000	DLNVAC	-2.78	0.000
6	DESI-INDU	-5.08	0.000	IFOMI-C	-3.68	0.000	DLNCPI-EX	-2.88	0.000	IS-D	-2.76	0.000
7	IFO-C	-5.03	0.000	SPR-C-G	-3.67	0.185	DIFO-C	-2.87	0.000	dp	-2.55	0.000
8	DLNORD-C	-5.02	1.000	IFOM-C	-3.56	0.000	DECCS4	-2.86	0.000	DECCS9	-2.55	0.000
9	IFOMV-EXP	-4.96	0.000	IFOM-EXP	-3.56	0.259	ECCS10	-2.80	0.000	DECCS4	-2.49	0.000
10	IFO-WC	-4.94	0.000	DIFO-UNCER	-3.46	0.000	DLNDAX	-2.77	0.000	IS-3M	-2.36	0.200
11	ECBS5	-4.92	0.000	IFO-C	-3.35	0.000	IFO-WC	-2.73	0.192	ECCS10	-2.32	0.000
12	DIFO-C	-4.88	0.000	DIFOMV-C	-3.31	0.000	IFO-EXP	-2.70	0.000	IS-M	-2.28	0.120
13	DECCS1	-4.87	0.000	DECBS6	-3.30	0.000	SPR-B-G	-2.68	0.077	ECCS9	-2.26	0.000
14	IFO-UNCER	-4.87	0.000	ECCS6	-3.28	0.000	DLNORD	-2.66	0.000	DLNEXR	-2.19	0.000
15	DECBS4	-4.86	0.000	IFO-WC	-3.27	0.000	ECCS9	-2.65	0.000	DLNEX	-2.09	0.000
16	ESI-SERV	-4.82	0.000	DESI-SERV	-3.23	0.000	DIFOM-C	-2.62	0.000	DLNDAX	-2.08	0.000
17	IFOMI-C	-4.80	0.000	IFO-UNCER	-3.21	0.000	DIFOWH-C	-2.60	0.115	DECBS3	-2.06	0.200
18	rols	-4.72	0.000	ECCS99	-3.21	0.000	SPR-B-A	-2.59	0.269	SPR-B-A	-2.03	0.280
19	ESI-INDU	-4.68	0.000	DECCS4	-3.19	0.000	IFOM-EXP	-2.53	0.231	DESI	-2.01	0.120
20	IFOWH-EXP	-4.56	0.000	DOECDL2	-3.10	0.185	DIFOMI-C	-2.51	0.154	DECCS8	-2.00	0.000
21	DIFOM-C	-4.49	0.000	ECCS4	-3.10	0.000	DIFOWH-EXP	-2.51	0.000	DESI-SERV	-1.95	0.080
22	ECBS4	-4.48	0.000	DIFOMV-EXP	-3.07	0.000	DIFO-EXP	-2.50	0.000	DECCS10	-1.91	0.000
23	DECBS3	-4.47	0.000	DECCS1	-3.05	0.000	DIFO-UNCER	-2.44	0.000	ESI-C	-1.90	0.200
24	DECCS8	-4.45	0.000	IFOMV-C	-3.05	0.000	DECCS5	-2.34	0.000	DIL-10	-1.75	0.000
25	ECBS2	-4.41	0.000	DIFOWH-C	-3.04	0.000	DECCS10	-2.32	0.000	DESI-TRADE	-1.72	0.000
MAFE												
1	IFO-EXP	-1.43	0.000	SPR-BF-G	-2.51	0.111	SPR-BF-G	-1.80	0.000	ECCS6	-1.09	0.000
2	IFOM-EXP	-1.42	0.071	IFO-EXP	-1.22	0.000	ECCS6	-1.01	0.269	SPR-BF-G	-0.96	0.240
3	rols	-1.34	0.000	DIFOMI-C	-1.14	0.000	SPR-C-G	-0.68	0.231	DLNCPI-EX	-0.81	0.000
4	ECCS10	-1.31	0.000	IFOMI-C	-1.01	0.000	dp	-0.63	0.038	SPR-C-G	-0.77	0.080
5	ECCS9	-1.30	0.000	DIFO-C	-0.97	0.000	DECCS4	-0.61	0.000	IS-D	-0.73	0.200
6	DECBS5	-1.29	0.000	IFOM-C	-0.96	0.000	IFO-EXP	-0.55	0.808	dp	-0.69	0.040
7	IFO-WC	-1.23	0.286	SPR-C-G	-0.92	0.185	DLNDAX	-0.53	0.000	IS-M	-0.61	0.280
8	ECCS6	-1.19	0.000	IFOM-EXP	-0.88	0.259	IFO-WC	-0.50	0.423	DLNVAC	-0.61	0.240
9	DESI-INDU	-1.13	0.000	IFOMV-EXP	-0.87	0.000	DIFO-EXP	-0.49	0.077	DECCS9	-0.57	0.000
10	IFOM-C	-1.12	0.000	IFO-WC	-0.87	0.259	ECCS99	-0.49	0.346	ECCS9	-0.55	0.120
11	DLNORD-C	-1.12	1.000	DIFOM-C	-0.85	0.148	DIFO-C	-0.49	0.000	DECCS4	-0.48	0.000
12	DECCS1	-1.10	0.000	ECCS5	-0.85	0.037	DECCS8	-0.49	0.000	ECCS10	-0.45	0.000
13	IFO-UNCER	-1.08	0.179	ECCS99	-0.84	0.000	DLNORD	-0.48	0.000	DESI-TRADE	-0.44	0.000
14	IFOWH-EXP	-1.08	0.929	IFO-C	-0.80	0.222	DLNVAC	-0.45	0.154	IS-3M	-0.40	0.440
15	ESI-C	-1.06	0.107	ECBS6	-0.76	0.000	ECCS9	-0.43	0.000	DIL-10	-0.38	0.400
16	DESI-SERV	-1.06	0.071	IFOMV-C	-0.72	0.037	ECCS1	-0.43	0.038	DGFK-EXP	-0.38	0.320
17	DGFK-EXP	-1.01	0.000	ESI-SERV	-0.71	0.000	IFOM-EXP	-0.42	0.500	DECBS3	-0.37	0.440
18	DIFO-C	-0.97	0.000	DECBS6	-0.70	0.148	DESI-TRADE	-0.41	0.038	DLNDAX	-0.34	0.400
19	ESI-CTR	-0.96	0.000	DLNDAX	-0.67	0.000	DIFOM-C	-0.41	0.423	DECCS8	-0.34	0.000
20	IFO-C	-0.95	0.071	DECCS8	-0.65	0.037	DECCS5	-0.41	0.000	DESI-SERV	-0.33	0.440
21	IFOMV-EXP	-0.95	0.000	DIFOMV-EXP	-0.63	0.037	DECCS99	-0.41	0.231	msfe	-0.32	0.150
22	trim25	-0.91	0.000	ECCS6	-0.61	0.000	DIFO-WC	-0.40	0.423	DESI	-0.32	0.120
23	DECCS8	-0.89	0.000	DLNFAZ	-0.61	0.444	IS-M	-0.36	0.231	DIFO-C	-0.32	0.840
24	DECCS10	-0.88	0.000	DIFOMV-C	-0.60	0.259	ECCS10	-0.36	0.000	DIFO-UNCER	-0.31	0.480
25	trim75	-0.87	0.038	ECCS1	-0.59	0.000	DIFO-UNCER	-0.36	0.192	DIFO-EXP	-0.30	0.840

Note: The second and sixth column display the average difference between the indicator performance and the AR benchmark model (both measures RMSFE and MAFE are calculated). Columns three and seven show the p-value of the end of sample instability test. p-values are calculated by a parametric subsample technique. This test checks whether the forecast performance of the indicator forecast compared with the benchmark model stays constant during the crisis period. Only 25 indicator forecast per horizon are displayed.

Table 9: Performance and Stability of leading indicator forecasts for IP during the crisis

	h=1	Av. gain	Stability p-Value	h=4	Av. gain	Stability p-Value	h=8	Av. gain	Stability p-Value	h=12	Av. gain	Stability p-Value
RMSFE												
1	OECDL3	-34.10	0.000	DOECDL2	-18.23	0.000	SPR-BF-G	-12.60	0.078	IS-3M	-8.08	0.000
2	DOECDL2	-33.06	0.000	DOECDL1	-17.00	0.000	DOECDL2	-12.22	0.000	SPR-BF-G	-7.61	0.329
3	OECDL1	-33.00	0.000	OECDL1	-16.67	0.000	DOECDL1	-9.72	0.000	IS-D	-7.08	0.000
4	DOECDL3	-32.41	0.000	DOECDL3	-16.17	0.000	DIFOWH-EXP	-9.33	0.000	DIFOWH-C	-6.78	0.000
5	IFOMI-C	-32.14	0.000	OECDL3	-16.02	0.000	SPR-B-G	-8.78	0.039	IS-M	-6.66	0.000
6	IFOMI-EXP	-32.04	0.000	IFOM-EXP	-14.04	0.000	ECCS4	-8.31	0.000	SPR-B-G	-6.54	0.233
7	trim75	-31.87	0.000	DIFOMI-C	-14.00	0.000	ECCS6	-8.30	0.000	ECCS6	-6.29	0.274
8	DOECDL1	-31.42	0.000	DIFO-EXP	-13.88	0.000	DIFOWH-C	-8.18	0.000	SPR-10Y-3M	-6.12	0.000
9	DIFOMI-C	-31.38	0.000	DIFOMI-EXP	-13.25	0.000	DECCS4	-8.17	0.000	DIFOWH-EXP	-6.09	0.000
10	IFOM-EXP	-31.30	0.000	IFO-EXP	-13.14	0.000	IFOM-EXP	-8.04	0.000	ECCS9	-6.02	0.000
11	DLNFAZ	-31.21	0.000	DIFOM-C	-12.70	0.000	DIFOMI-C	-7.96	0.000	ECCS4	-6.00	0.000
12	trim50	-31.20	0.000	DIFOM-EXP	-12.69	0.000	DOECDL3	-7.93	0.000	DECCS6	-5.99	0.000
13	trim25	-30.73	0.000	ECCS4	-12.65	0.000	DIFO-C	-7.87	0.000	ESI-CTR	-5.75	0.000
14	DLNORD	-30.23	0.000	IFOM-C	-12.55	0.000	SPR-C-G	-7.79	0.000	SPR-10Y-M	-5.74	0.000
15	ECBS6	-30.07	0.000	SPR-BF-G	-12.42	0.173	OECDL1	-7.74	0.000	SPR-C-G	-5.74	0.000
16	ECBS3	-30.06	0.000	DECCS4	-12.39	0.000	SPR-10Y-3M	-7.71	0.000	DLNCPI-EX	-5.63	0.000
17	ESI-SERV	-29.87	0.000	DIFOMV-EXP	-12.33	0.000	SPR-B-A	-7.62	0.442	rols	-5.61	0.000
18	DESI-INDU	-29.69	0.000	SPR-B-G	-12.19	0.000	DIFO-EXP	-7.58	0.000	DESI	-5.60	0.000
19	ECCS5	-29.58	0.000	ZEW	-12.16	0.000	DLNORD	-7.28	0.000	DOECDL2	-5.47	0.000
20	IFO-EXP	-29.49	0.000	IFOMI-C	-12.03	0.000	SPR-10Y-D	-7.27	0.000	DECCS9	-5.40	0.000
21	DECCS5	-29.27	0.000	DECBS6	-12.03	0.000	SPR-10Y-M	-7.25	0.000	SPR-B-A	-4.97	0.589
22	ECBS4	-29.21	0.000	SPR-10Y-3M	-12.01	0.000	ECCS5	-7.19	0.000	DECCS4	-4.94	0.000
23	ESI-INDU	-29.09	0.000	DECCS5	-12.00	0.000	DECCS5	-7.12	0.000	IL-3	-4.93	0.000
24	DLNORD-I	-28.49	0.000	DESI-INDU	-11.83	0.000	IS-D	-7.12	0.000	SPR-10Y-D	-4.75	0.000
25	DECBS4	-28.36	0.000	DIFO-C	-11.76	0.000	IFO-EXP	-7.08	0.000	ESI-INDU	-4.37	0.000
MAFE												
1	DOECDL2	-9.20	0.000	DOECDL2	-4.35	0.000	SPR-BF-G	-3.55	0.364	DIFOWH-C	-1.80	0.630
2	DOECDL3	-8.90	0.000	DOECDL1	-3.74	0.000	DOECDL2	-3.30	0.091	DESI	-1.67	0.041
3	OECDL3	-8.46	0.000	DOECDL3	-3.35	0.000	DIFOWH-EXP	-2.57	0.312	IS-3M	-1.66	0.904
4	DOECDL1	-8.22	0.000	OECDL1	-3.17	0.000	DECCS4	-2.01	0.701	DIFOWH-EXP	-1.53	0.822
5	IFOM-EXP	-7.69	0.000	DECCS5	-3.10	0.000	ECCS4	-2.01	0.442	SPR-BF-G	-1.50	0.630
6	OECDL1	-7.60	0.000	DIFO-EXP	-2.83	0.160	DOECDL1	-1.90	0.286	IS-D	-1.23	0.808
7	ECCS5	-7.46	0.000	OECDL3	-2.79	0.000	DIFOWH-C	-1.88	0.623	IS-M	-1.14	0.904
8	trim75	-7.26	0.000	IFOM-EXP	-2.74	0.272	DIFO-C	-1.83	0.519	ECCS6	-1.14	0.685
9	DLNFAZ	-7.16	0.000	ECCS4	-2.74	0.000	SPR-C-G	-1.82	0.468	ESI-CTR	-1.03	0.397
10	DIFOMI-C	-7.16	0.000	DECCS4	-2.64	0.000	DIFO-EXP	-1.71	0.610	ECCS4	-1.02	0.479
11	DECCS5	-7.08	0.000	IFO-EXP	-2.62	0.247	ECCS6	-1.68	0.532	DOECDL2	-1.02	0.630
12	DESI-INDU	-6.97	0.000	SPR-BF-G	-2.55	0.160	SPR-B-G	-1.63	0.403	SPR-B-G	-0.96	0.740
13	PMI	-6.93	0.000	ECCS5	-2.52	0.000	IS-3M	-1.60	0.169	ESI	-0.87	0.055
14	IFOMI-EXP	-6.82	0.000	DECCS3	-2.41	0.000	IFOM-EXP	-1.58	0.506	SPR-C-G	-0.86	0.630
15	trim50	-6.81	0.000	DLNCPI	-2.39	0.000	IFOWH-EXP	-1.56	0.792	DECCS6	-0.84	0.000
16	trim25	-6.61	0.000	ECCS1	-2.33	0.160	DECCS99	-1.52	0.610	msfe	-0.78	0.864
17	IFOMI-C	-6.52	0.000	ECCS3	-2.33	0.000	DIFOM-C	-1.49	0.519	SPR-10Y-3M	-0.74	0.233
18	DECBS4	-6.39	0.000	ECCS6	-2.26	0.259	ESI-TRADE	-1.46	0.545	DLNCPI-EX	-0.70	0.452
19	IFO-EXP	-6.32	0.000	ZEW	-2.15	0.519	ECCS99	-1.45	0.844	DECCS4	-0.70	0.247
20	SPR-HY-G	-6.22	0.000	DIFOWH-EXP	-2.11	0.123	DGFK-EXP	-1.45	0.494	ECCS9	-0.69	0.027
21	DLNORD	-6.13	0.000	DGFK-EXP	-2.07	0.000	GFK-EXP	-1.39	0.468	DIFO-C	-0.69	0.877
22	ECBS4	-6.13	0.000	DECCS1	-2.06	0.062	IS-D	-1.38	0.675	IL-3	-0.67	0.753
23	ECBS3	-6.11	0.000	SPR-B-G	-2.01	0.000	DIFOMI-C	-1.35	0.519	SPR-10Y-M	-0.58	0.260
24	ESI-INDU	-5.96	0.000	DIFO-C	-1.99	0.667	msfe	-1.34	0.507	ECCS99	-0.55	0.849
25	ECBS6	-5.74	0.000	msfe	-1.90	0.373	DIFO-UNCER	-1.33	0.571	ESI-INDU	-0.54	0.055

Note: The second and sixth column display the average difference between the indicator performance and the AR benchmark model (both measures, RMSFE and MAFE are calculated). Columns three and seven show the p-value of the end of sample instability test. p-values are calculated by a parametric subsample technique. This test checks whether the forecast performance of the indicator forecast compared with the benchmark model stays constant during the crisis period. Only 25 indicator forecast per horizon are displayed.

Figure 4: Weights allocated to each block for GDP (h=1)

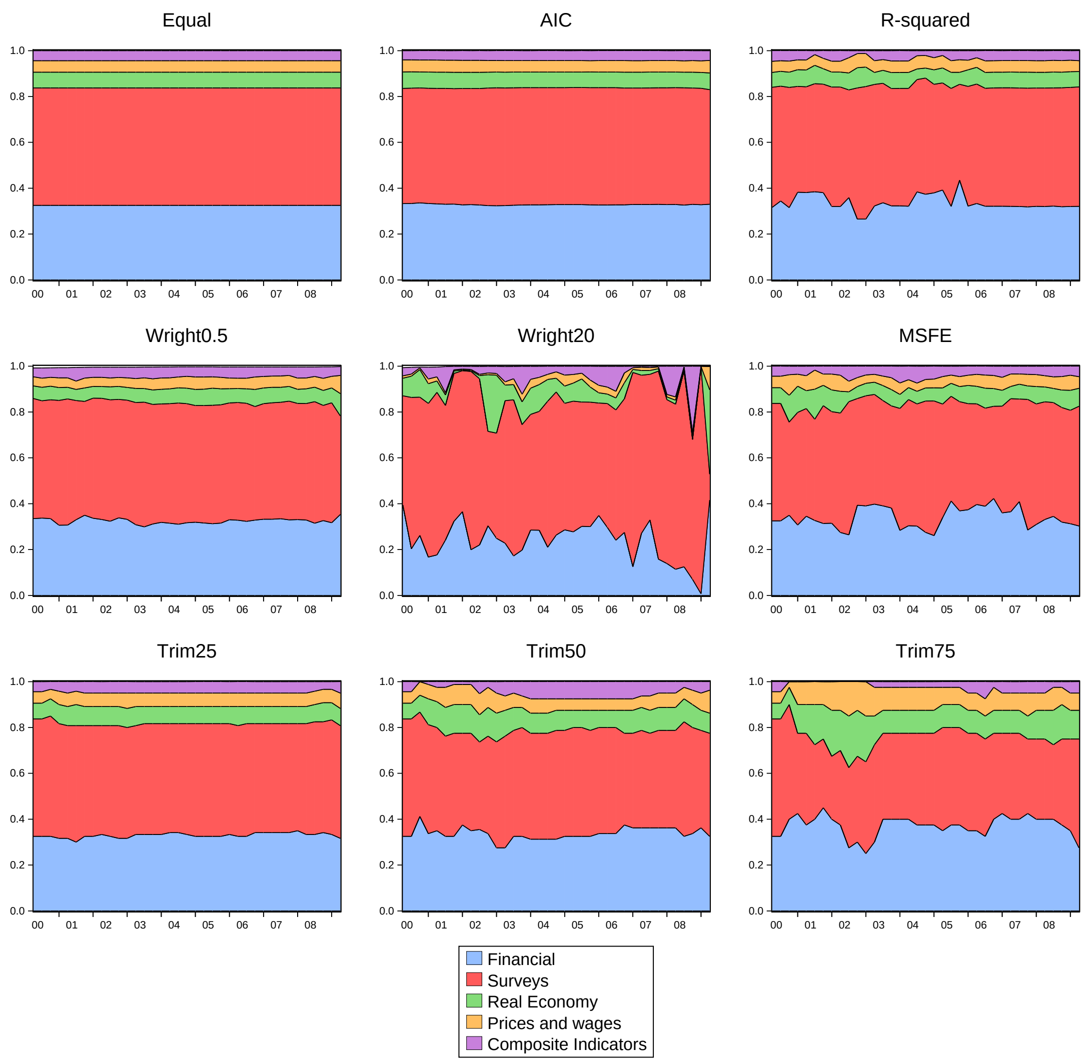


Figure 5: Weights allocated to each block for GDP (h=2)

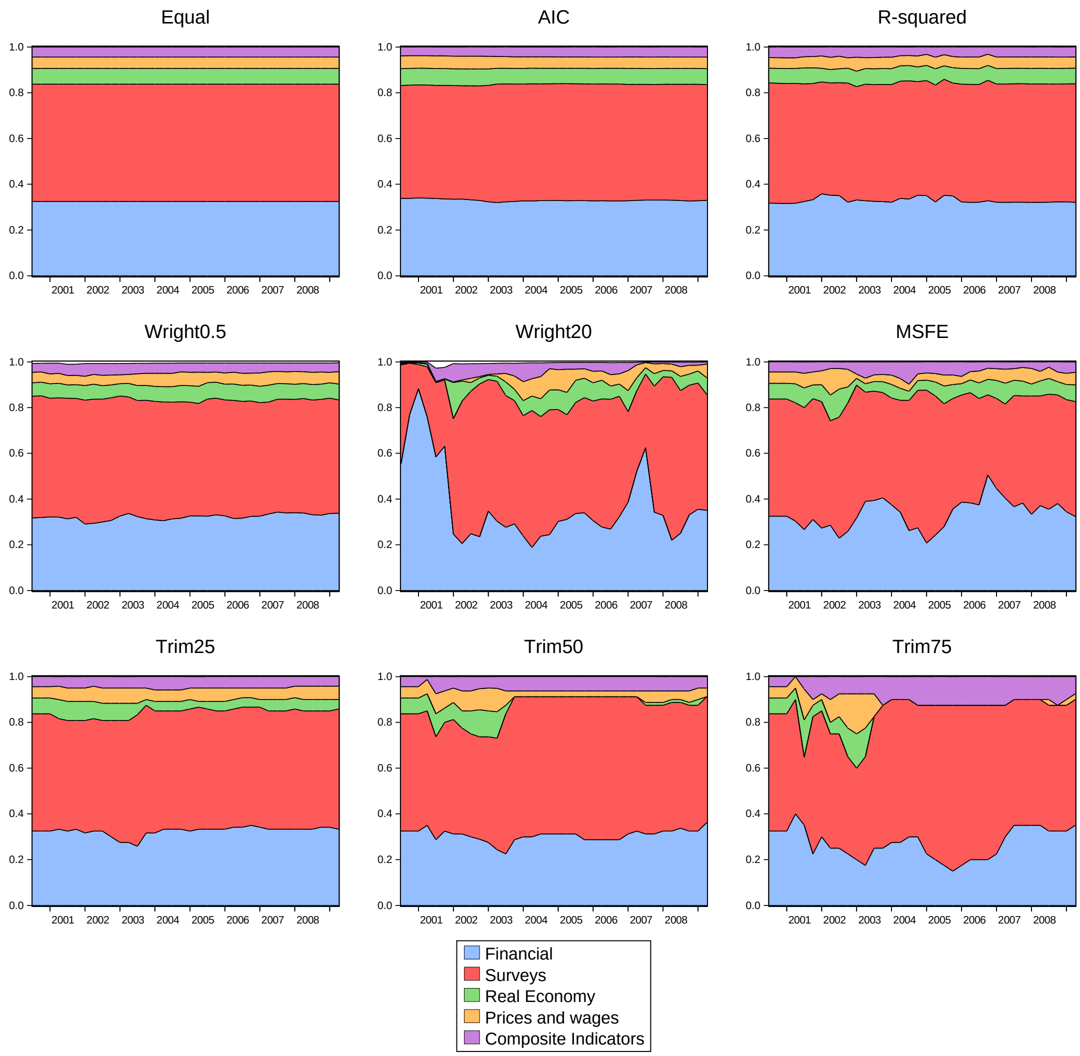


Figure 6: Weights allocated to each block for GDP (h=3)

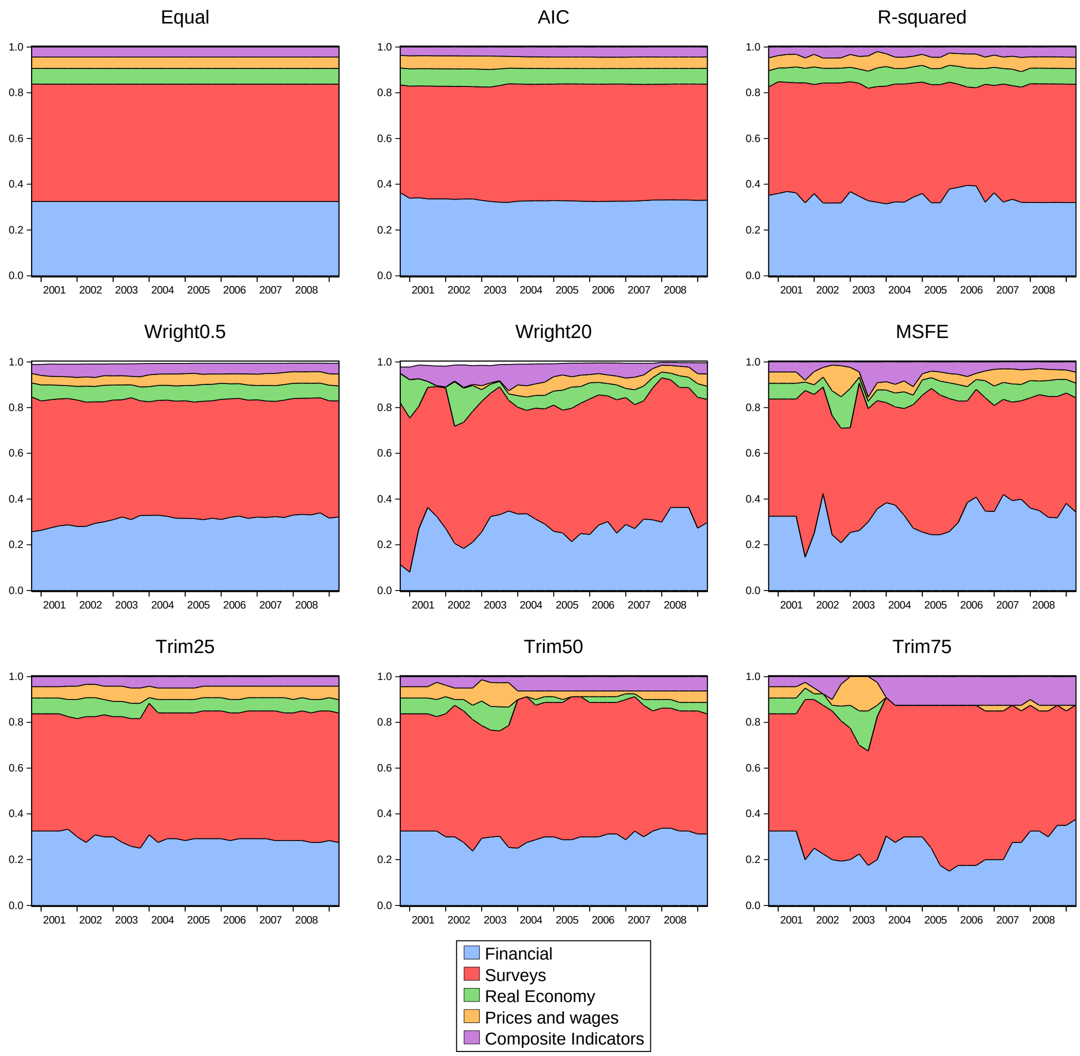


Figure 7: Weights allocated to each block for GDP (h=4)

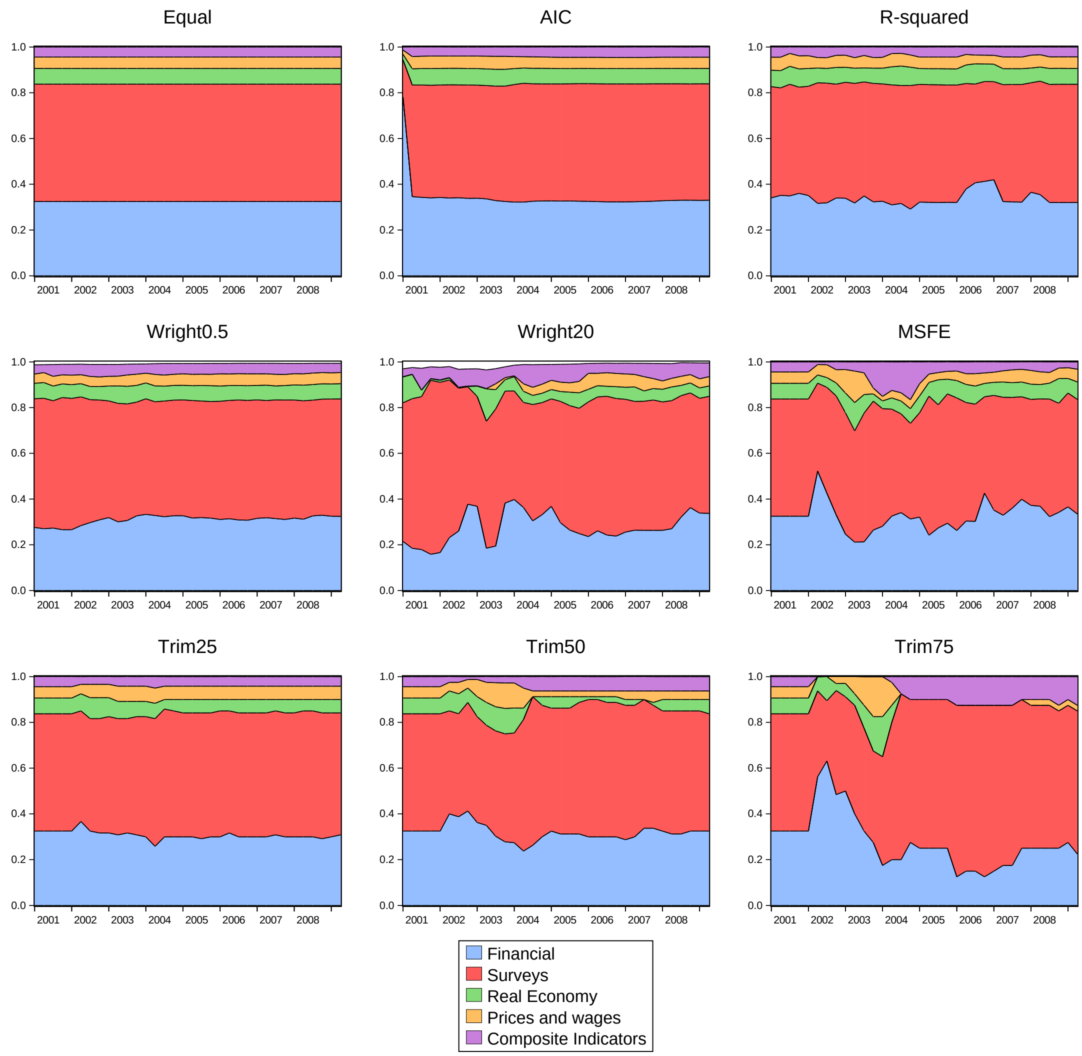


Figure 8: Weights allocated to each block for IP (h=1)

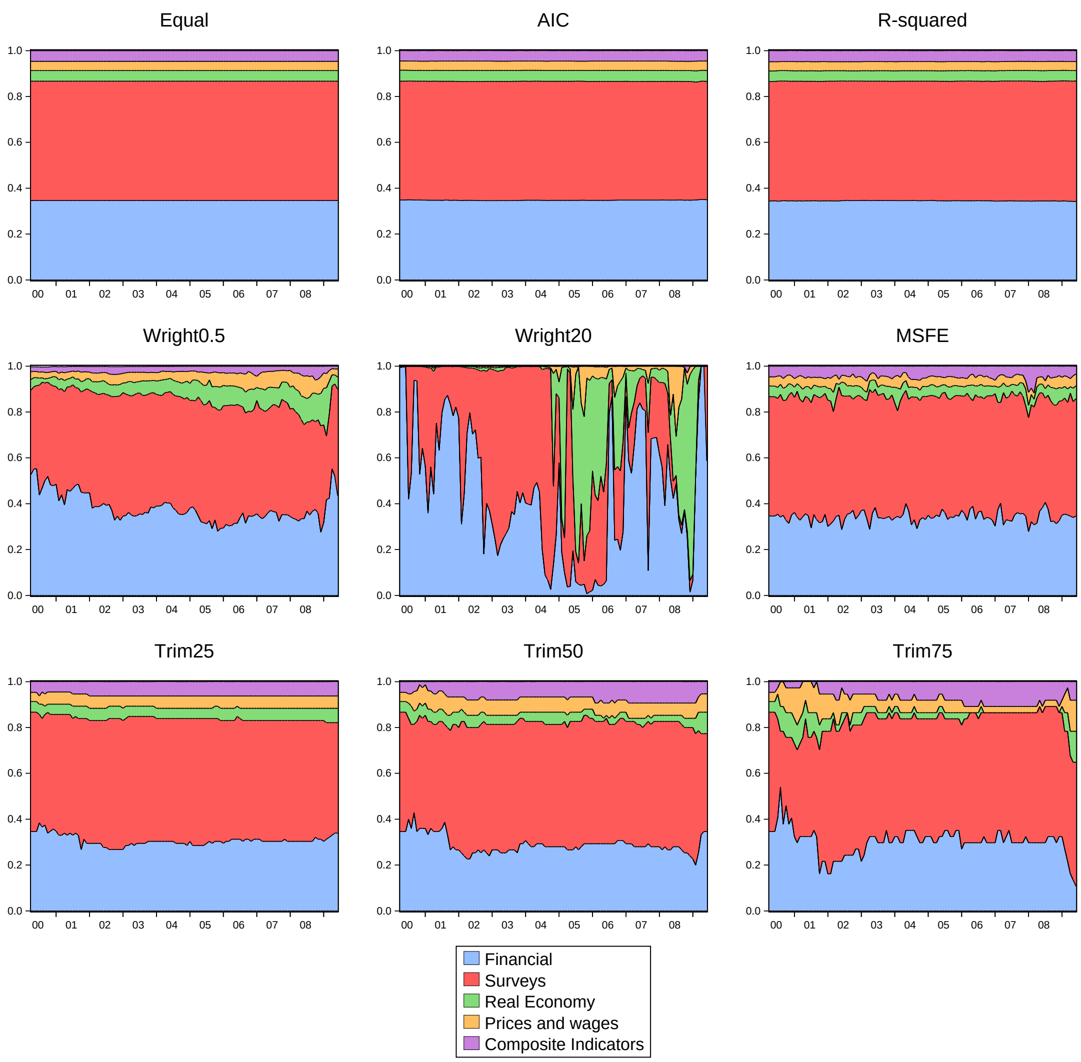


Figure 9: Weights allocated to each block for IP (h=4)

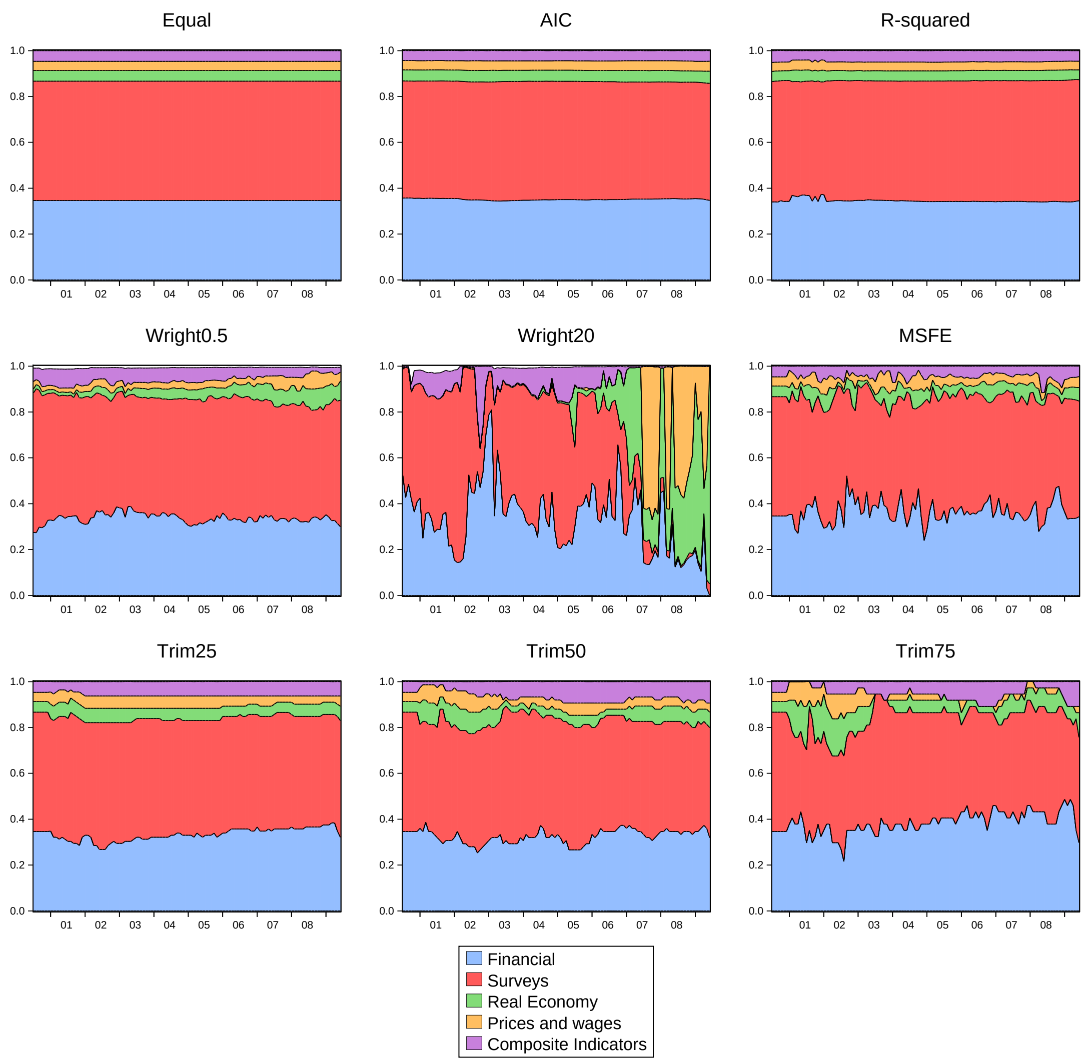


Figure 10: Weights allocated to each block for IP (h=8)

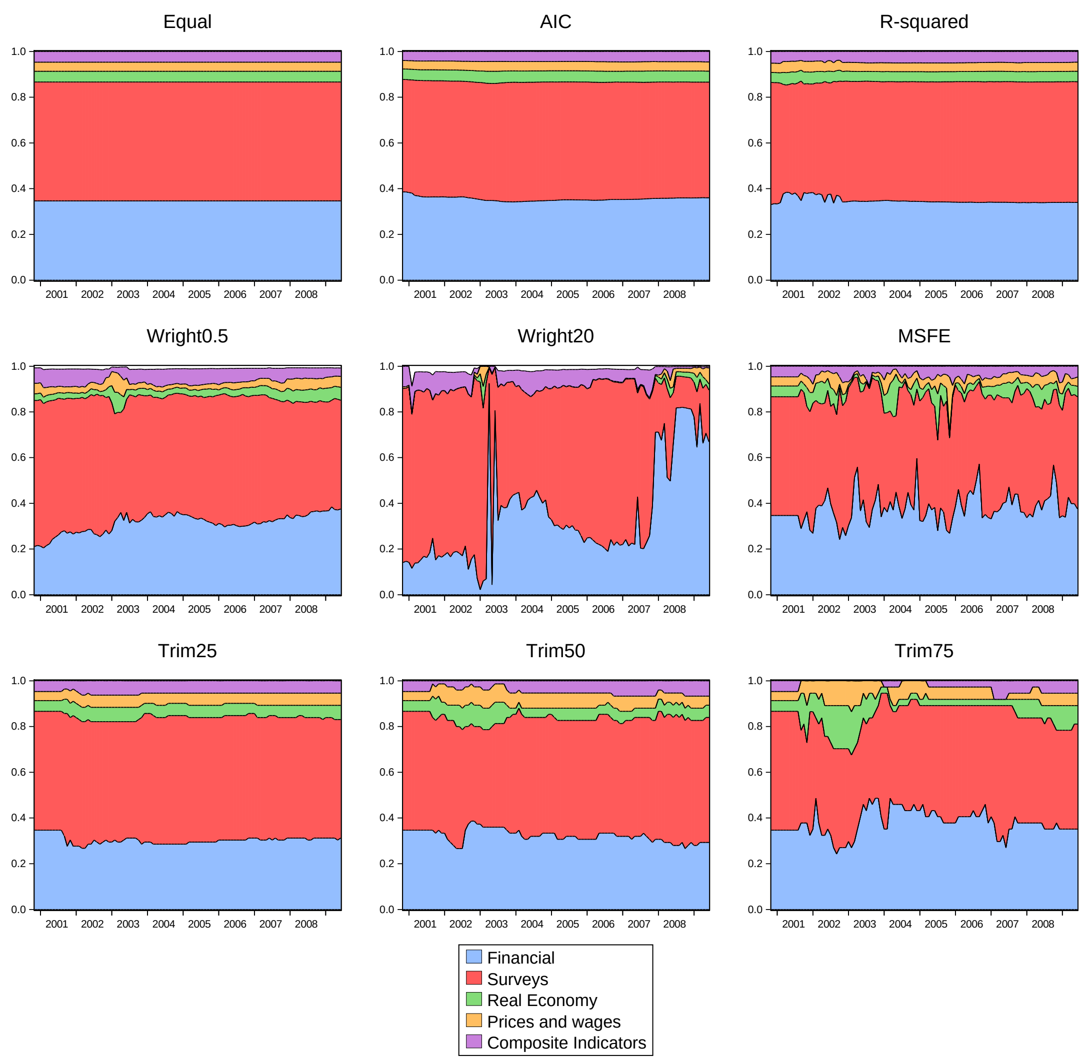


Figure 11: Weights allocated to each block for IP (h=12)

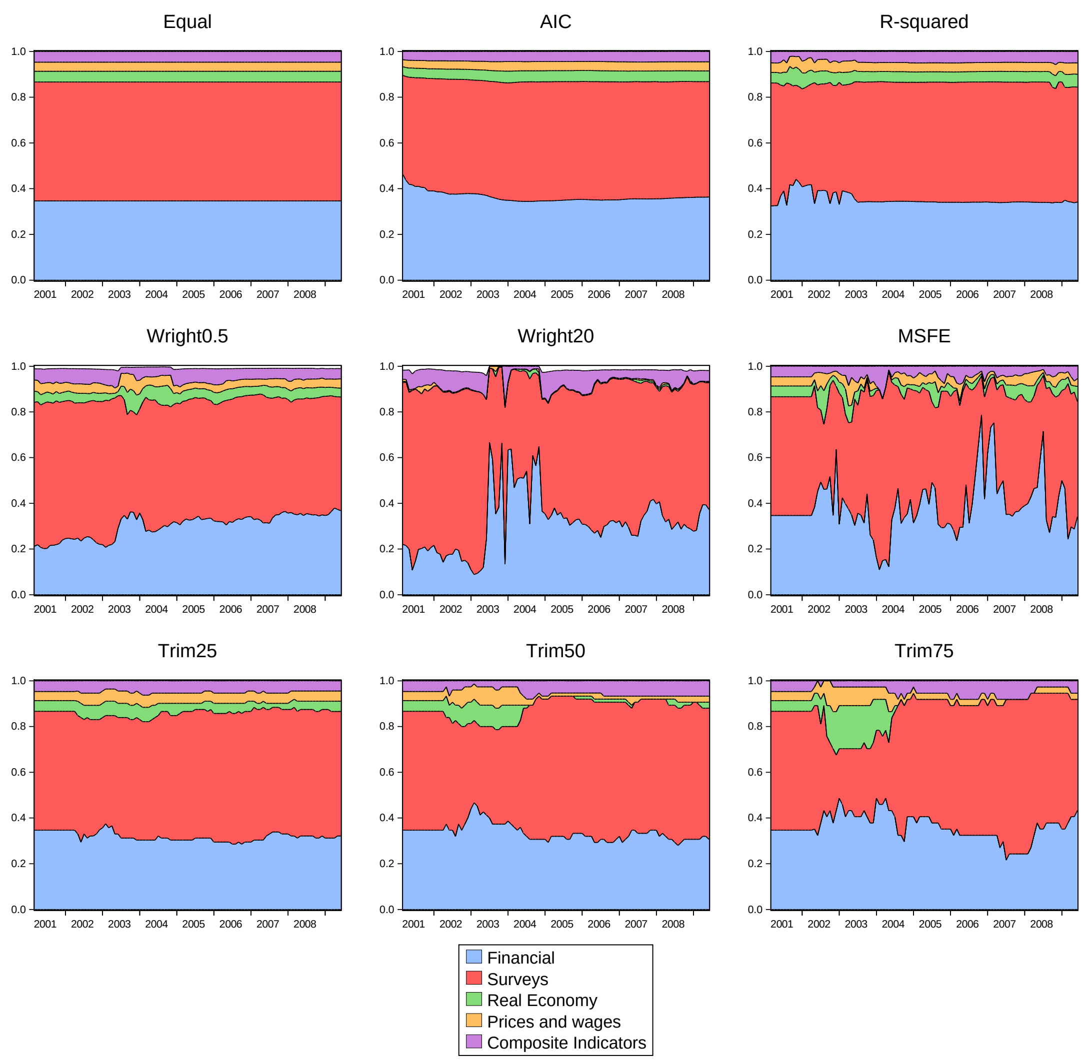
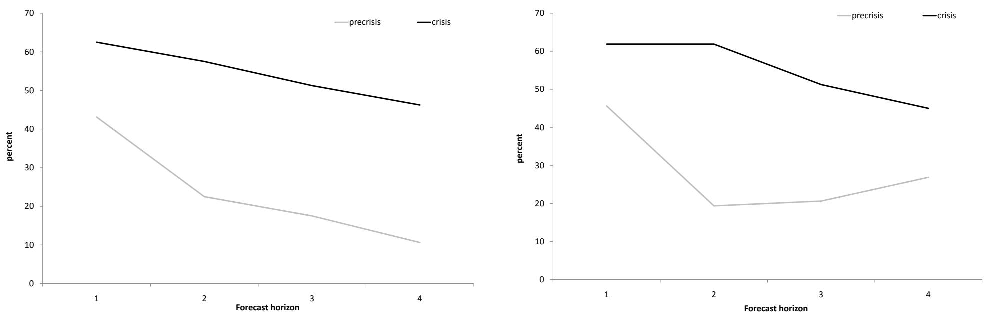


Figure 12: Performance of GDP Indicator Forecasts

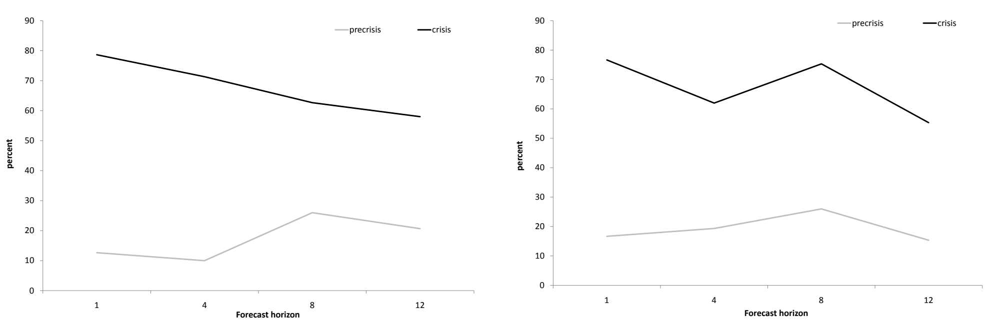
(a) Root Mean Squared Forecast Error (b) Mean Absolute Forecast Error



Note: Share of individual indicator forecasts better than the benchmark AR forecast is shown.

Figure 13: Performance of IP Indicator Forecasts

(a) Root Mean Squared Forecast Error (b) Mean Absolute Forecast Error



Note: Share of individual indicator forecasts better than the benchmark AR forecast is shown.