Empirical Economics
DOI 10.1007/s00181-005-0050-5

ORIGINAL PAPER

Jörg Döpke · Ulrich Fritsche

Growth and inflation forecasts for Germany

A panel-based assessment of accuracy and efficiency

Received: 7 April 2004 / Accepted: 14 September 2005
© Springer-Verlag 2005

Abstract We analyse forecasts of professional forecasters for Germany regarding the time span from 1970 to 2004. This novel panel data set renders it possible to assess the accuracy and efficiency of growth and inflation forecasts more efficiently than in previous studies. We argue that the forecasts are, on average, unbiased and weakly—but not strongly—efficient. Using model confidence sets suggested by Hansen et al. (2004), we find that, besides the effect of diverging forecasting dates, no other substantial differences in forecasting quality among forecasters exist. Nevertheless, on the basis of a direction-of-change analysis we argue that it is not always advisable to listen to the majority of forecasters.

Keywords Forecast error evaluation · Germany

JEL Classification E32 · E37

1 Introduction

A lot of studies have assessed the accuracy of professional business cycle forecasts in Germany. However, virtually all studies refer to individual forecasts or very

U. Fritsche (✉)
German Institute for Economic Research (DIW Berlin), Königin-Luise-Strasse 5,
14195 Berlin, Germany
E-mail: ufritsche@diw.de

U. Fritsche
Department Economics and Politics, University Hamburg, Von-Melle-Park 9,
20146 Hamburg, Germany

J. Döpke
Deutsche Bundesbank, Economics department, Wilhelm-Epstein-Strasse 14,
60047 Frankfurt, GermanyE-mail: joerg.doepke@bundesbank.de

small groups of institutions.[1] Until now, there has not been any study which takes into account a larger group of forecasting institutions for Germany. The present study tries to fill this gap by taking into account forecasts of (to our knowledge) all publicly available forecasts for Germany for a long time span, i.e. from (at least) the early 1970s onwards.

This novel data set renders it possible to analyse the accuracy of the predictions of German professional forecasters in a more appropriate way, since it is possible to exploit the cross-section information of the data as well. In particular, we use German annual data for growth and inflation forecasts from 1970 to 2004 stemming from up to 17 different forecasts, published at regular intervals by 14 different institutions. Regarding the forecasts' accuracy we find that the forecasts are usually unbiased, weakly efficient, and have significant information content for future turning points. However, a stronger version of the rational expectations hypothesis is usually rejected. Moreover, based on the "model confidence sets" technique suggested by Hansen et al. (2004) we find that, besides the effect of diverging forecasting dates, no other substantial differences in forecasting quality among forecasters exist. Nevertheless, on the basis of a direction-of-change analysis, we argue that it is not always advisable to listen to the majority of forecasters.

The paper is organised as follows: Section 2 describes the data set used and reports some standard descriptive statistics on the accuracy of German business cycle forecasts. Section 3 presents results for individual forecasts regarding measures of forecast accuracy and efficiency. Section 4 exploits the panel dimension of the data to gain more efficient tests for these hypotheses. In Section 5, we ask whether the forecasting quality differs among the forecasts under investigation. The following section makes use of a direction-of-change analysis to check whether the forecasters predict the underlying tendency of the business cycle correctly. Moreover, we analyse whether this tendency is better gauged by the majority or by the minority of the forecasters. The final section offers some conclusions.

2 Data

In the following section we evaluate forecasts of several institutions which deliver forecasts regarding the German economy. Details of the forecast under investigation can be found in the Appendix. Note, however, that the forecasters do not all publish at the same time. Below, we refer to forecasts in the winter semester from September to March/April at the latest in order to capture the typical forecasting situation of the majority of institutions.[2] Figure 1 shows the average timing of the forecasts during the year.

Of course, the varying forecasting dates have direct consequences for the accuracy of the forecasts since a later forecasting date allows the forecaster to take additional information into account.

---

[1] A by no means complete list may include the following examples: Büttner and Horn (1993), Döpke and Langfeldt (1995), Neumann and Buscher (1980), Heilemann (1998), Nierhaus (1998), Kirchgässner (1984), Hinze (1996), Heilemann and Stekler (2003).

[2] This is also the reason why some institutions appear twice in our data set, while others do not. In some other cases (e.g. the Kiel institute), the forecasting frequency has also changed over time from twice a year to four times a year.

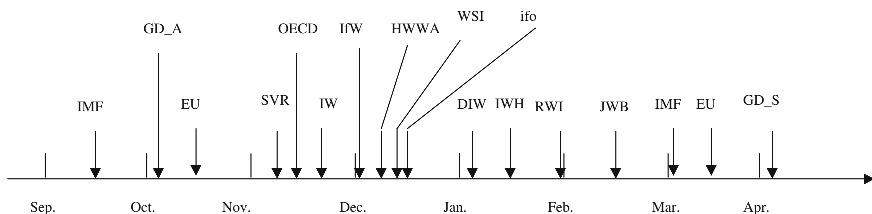


Fig. 1 Forecasting institutions and dates. IMF: International Monetary Fund, GD_A: joint forecast, autumn, EU: European Commission, SVR: Council of Economic Advisors, IW: Institute of the German Economy (IW), OECD: Organization for Economic Co-operation and Development, Paris, IfW: Kiel Institute for World Economics, HWWA: HWWA Institute of Economic Research, WSI: Trade Union Institute (WSI), ifo: Ifo Institute Munich, DIW: German Institute of Economic Research, Berlin, IWH: Institute for Economic Research, Halle, RWI: Rhine-Westphalia Institute for Economic Research, Essen, JWB: economic report of the German federal government, GD_S: joint forecast, spring

For all institutions, we have collected the growth and inflation forecasts. The growth forecast is the predicted growth rate of real GNP (for the time span 1983 to 1989) and of real GDP (for all other years).[3] In the case of interval forecasts the average is used. The numbers refer to western Germany up to 1992, and to the whole of Germany from 1993 to the present. As a measure of the inflation forecast, we use the predicted change of the deflator of private consumption in cases where this figure was available. In some instances, however, no explicit reference was given as to whether a mentioned inflation forecast referred to the consumption deflator or to the CPI. In such cases, we assume that no distinction between the figures was intended by the forecaster and used the available inflation forecast.[4] As regards the actual outcome, it is possible to refer to the last available revised data or to the first published ("real-time") data. As it is common in the analysis of business cycle forecasts, we make use of the latter type of numbers, i.e., we compare the forecasts made at the end of a given year *x* or at the beginning of the following year t+1 with the first published figure for the year t+1.

3 Forecast accuracy and efficiency: evidence for individual forecasts

To begin with, Table 1a and b presents a number of standard measures of forecast accuracy for each institution separately. As a natural "benchmark" we also considered the case of a "no change" forecast here. In particular, we calculate the following statistics:[5]

1. The mean error $\mathrm{ME} = \frac{1}{T}\sum_{t=1}^{T} e_{t+1,t}$, where $e_{t+1,t} = y_{t+1} - y_{t+1,t}$ is the forecast error in each period, defined as actual (in t+1) minus predicted (in t for period t+1)

[3] This choice is motivated by the "headline figure" of the statistical office for the respective year. Note, however, that the forecasts frequently refer to "growth" rather than explicitly to either GDP or GNP or to one figure only. In such cases, we have assumed that the forecasters made no distinction between the concepts and had the same forecast for both figures.
[4] The data used in this study may be downloaded from http://www.ulrich-fritsche.net.
[5] Unless otherwise stated, our notation follows the textbook of Diebold (1998).

value of the variable y. Thus, a positive (negative) value of the mean error corresponds to an under (over-) estimation of the growth rate. t is the time index.

2. The mean absolute error $\mathrm{MAE} = \frac{1}{T}\sum_{t=1}^{T} |e|_{t+1,t}$.
3. The root mean squared error $\mathrm{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T} e_{t+1,t}^2}$.
4. Furthermore, we present the results of statistical tests. First, we test for the unbiasedness of the forecasts. To this end it is possible to make use of the Mincer–Zarnowitz regression, i.e., running the following regression: $y_{t+1} = \beta_0 + \beta_1 y_{t+1,t} + u_{t+1}$ and testing the hypothesis $H_0 : \begin{cases} \beta_0 = 0 \\ \beta_1 = 1 \end{cases}$ with a F-test. The relevant column in the table reports the marginal significance level of this test. To take into consideration a possible moving average (MA) error structure (due to the fact that the forecast semester under consideration runs from October to March), we included an MA(1) term.
5. As a test for efficiency of the forecasts we refer to a test for autocorrelation of the forecast errors. For an optimal forecast one should be unable to find any variable which helps to forecast the errors. As a consequence, the lagged forecast errors should also be non-informative for the ex ante error. This is sometimes referred to as a test for weak efficiency (Kirchgässner 1984). Again, we report the marginal significance level of the test.
6. The former test is a test for weak rationality; i.e., the incorporated information set is limited to the lagged endogenous variable. A stronger test version of the hypothesis of rational expectation stipulates that the forecast errors are uncorrelated to *any* variables known to the forecasters at the time of the forecast. While weak rationality is usually confirmed for German growth forecasts, some authors challenge the hypothesis of stronger concepts of rationality (see Kirchgässner and Savioz 2001). However, as is emphasised by Tichy (1994), tests that rely on exogenous variables are problematic since it is unclear what the forecasters *exactly* know about the future stance of the cycle. For example, the GDP growth rate of the current year is usually unknown to the forecasters when they forecast the growth rate for the next year. GDP figures are —at best—available up to the third quarter of the current year, when it comes time to calculate the forecast. Thus, we refer to seasonally adjusted quarterly data which are not subject to major revisions and only use data up to the second quarter of the previous year in our tests of strong rationality. In particular, we use the following variables: i) Short-term interest rates as a proxy for the course of monetary policy. ii) The real external value of the domestic currency to capture possible exchange rate shocks. iii) The change of the oil price as a proxy for supply side shocks. iv) The change over previous year of industrial production in all OECD countries to take into account demand fluctuations outside Germany. With these variables at hand we estimate the equation $y_{t+1} = \beta_0 + \beta_1 y_{t+1,t} + \beta_2 X_{t-1} + u_{t+1}$ where X_{t-1} represents the relevant exogenous variable in the second quarter of the year in which the forecast is made. We again use a standard Wald-Test to test the null hypothesis $H_0 : \begin{cases} \beta_0 = 0 \\ \beta_1 = 1 \\ \beta_2 = 0 \end{cases}$ as

Table 1 Measures of forecast accuracy

Institution	Mean error	Mean absolute error	Root mean squared error	Test for unbiasedness (p value)	Test for MA(1) term (p value)	Test for weak rationality (p value)	Test for strong rationality based on short-term interest rates (coefficient/ p value)	Test for strong rationality based on the real external value of the currency (coefficient/p value)	Test for strong rationality based on the change of the oil price (coefficient/ p value)	Test for strong rationality based on OECD industrial production (coefficient/ p value)
a. Growth forecasts 1970 to 2004[a)]										
IMF, autumn	−0.71	1.39	1.83	**0.08**	0.96	0.96	**−0.29/0.01**	−3.90/0.14	**−0.01/0.09**	−0.08/0.13
Joint forecast, autumn	−0.39	1.21	1.65	0.33	0.86	0.86	**−0.23/0.06**	−4.84/0.29	−0.01/0.28	−0.08/0.37
European Commission, autumn	−0.46	1.22	1.67	0.26	0.81	0.88	**−0.26/0.02**	−3.29/0.33	−0.01/0.21	−0.11/0.24
Council of Economic Advisors	−0.36	1.02	1.45	0.35	0.96	0.95	**−0.19/0.09**	−4.77/0.31	−0.01/0.54	−0.09/0.34
OECD	−0.35	1.09	1.58	0.42	0.56	0.60	**−0.22/0.08**	−3.37/0.49	−0.01/0.30	−0.12/0.29
Employer's Institute	−0.34	1.17	1.71	0.30	0.17	0.14	**−0.24/0.09**	−1.27/0.48	−0.01/0.30	−0.05/0.42
Kiel Institute	−0.26	1.08	1.55	0.35	0.96	0.95	−0.18/0.13	−4.62/0.30	−0.01/0.32	−0.08/0.36
Hamburg Institute	−0.22	0.98	1.38	0.67	0.64	0.57	−0.16/0.29	−3.44/0.61	−0.01/0.56	−0.07/0.63
Trade Union's Institute	−0.14	1.15	1.58	0.51	0.77	0.71	−0.22 / 0.12	−5.24 / 0.44	−0.01 / 0.33	−0.11 / 0.38

Table 1 (continued)

Institution	Mean error	Mean absolute error	Root mean squared error	Test for unbiasedness (p value)	Test for MA(1) term (p value)	Test for weak rationality (p value)	Test for strong rationality based on short-term interest rates (coefficient/ p value)	Test for strong rationality based on the real external value of the currency (coefficient/p value)	Test for strong rationality based on the change of the oil price (coefficient/ p value)	Test for strong rationality based on OECD industrial production (coefficient/ p value)
Munich Institute	−0.13	0.96	1.33	0.86	0.54	0.44	−0.15/0.40	−2.31/0.86	−0.01/0.50	−0.08/0.64
Berlin Institute	0.02	1.03	1.51	0.50	0.70	0.62	**−0.21/0.09**	−4.02/0.46	−0.01/0.18	−0.07/0.50
Halle Institute	−0.40	0.75	1.00	0.42	0.31	0.17	−0.01/0.64	7.58/0.23	−0.00/0.61	−0.26/0.28
Essen Institute	−0.30	1.02	1.34	0.40	0.72	0.65	−0.17/0.15	−2.71/0.51	−0.01/0.55	−0.06/0.56
Government's economic report	−0.26	0.95	1.40	0.56	0.47	0.38	−0.15/0.29	−3.37/0.58	−0.01/0.35	−0.05/0.66
IMF, spring	−0.19	0.88	1.22	0.40	0.28	0.12	−0.13/0.20	−2.58/0.49	−0.01/0:21	−0.01/0.65
European Commission, spring	−0.23	0.83	1.11	0.42	0.58	0.44	−0.13/0.24	−1.77/0.67	−0.01/0.25	−0.02/0.76
Joint forecast, spring	−0.24	0.84	1.18	0.48	0.94	0.91	−0.14/0.20	−1.47/0.66	−0.01/0.24	−0.05/0.58
No change forecast	−0.09	1.80	2.49	**0.00**	**0.00**	**0.08**	**−0.43/0.00**	**4.04/0.00**	**−0.00/0.00**	**0.02/0.00**

Table 1 (continued)

Institution	Mean error	Mean absolute error	Root mean squared error	Test for unbiasedness (*p* value)	Test for MA(1) term (*p* value)	Test for weak rationality (*p* value)	Test for strong rationality based on short-term interest rates (coefficient/ *p* value)	Test for strong rationality based on the real external value of the currency (coefficient/*p* value)	Test for strong rationality based on the change of the oil price (coefficient/ *p* value)	Test for strong rationality based on OECD industrial production (coefficient/ *p* value)
b. Inflation forecasts 1970 to 2004[a)]										
IMF, autumn	−0.11	0.71	0.95	0.59	**0.01**	**0.09**	−0.04/0.73	2.96/0.58	0.00/0.63	0.00/0.79
Joint forecast, autumn	0.03	0.83	1.07	0.23	**0.01**	**0.00**	0.04/0.29	2.37/0.26	0.00/0.27	0.05/0.18
European Commission, autumn	0.07	0.73	0.97	0.13	**0.03**	**0.00**	0.03/0.22	1.57/0.22	0.00/0.19	0.05/0.12
Council of Economic Advisors	0.06	0.74	0.97	**0.09**	**0.07**	**0.01**	−0.01/0.20	**4.41/0.07**	0.00/0.12	0.02/0.18
OECD	−0.05	0.58	0.79	0.85	0.45	0.30	0.04/0.89	2.63/0.74	0.00/0.95	0.01/0.95
Employer’s Institute	−0.03	0.55	0.77	0.27	0.33	0.26	−0.08/0.27	2.57/0.29	−0.00/0.44	−0.01/0.44
Kiel Institute	0.19	0.81	1.20	0.22	**0.00**	**0.00**	0.07/0.16	1.40/0.21	0.00/0.20	0.02/0.18
Hamburg Institute	0.05	0.68	0.95	0.31	**0.04**	**0.01**	−0.05/0.46	4.11/0.23	0.00/0.39	0.01/0.49
Trade Union’s Institute	0.01	0.83	1.09	0.92	**0.00**	**0.00**	0.05/0.93	4.01/0.68	0.01/0.74	0.03/0.88

Table 1 (continued)

Institution	Mean error	Mean absolute error	Root mean squared error	Test for unbiasedness (p value)	Test for MA(1) term (p value)	Test for weak rationality (p value)	Test for strong rationality based on short-term interest rates (coefficient/ p value)	Test for strong rationality based on the real external value of the currency (coefficient/p value)	Test for strong rationality based on the change of the oil price (coefficient/ p value)	Test for strong rationality based on OECD industrial production (coefficient/ p value)
Munich Institute	−0.02	0.63	0.87	0.66	0.31	0.18	−0.03/0.77	3.20/0.50	0.00/0.73	−0.01/0.76
Berlin Institute	0.21	0.72	0.95	0.17	0.11	**0.02**	0.04/0.33	2.13/0.29	0.01/0.27	0.04/0.31
Halle Institute	0.04	0.53	0.74	0.22	**0.06**	0.42	**0.29/0.00**	4.13/0.22	−0.00/0.42	−0.07/0.30
Essen Institute	0.13	0.62	0.86	**0.03**	**0.01**	**0.00**	**−0.01/0.08**	**4.33/0.02**	**0.00/0.02**	**0.02/0.07**
Government's economic report	0.16	0.63	0.91	0.42	0.10	**0.03**	−0.02/0.70	2.84/0.50	0.00/0.75	−0.00/0.80
IMF, spring	−0.16	0.61	0.86	0.52	**0.00**	**0.01**	−0.09/0.46	5.37/0.12	0.01/0.49	−0.03/0.59

Table 1 (continued)

Institution	Mean error	Mean absolute error	Root mean squared error	Test for unbiasedness (*p* value)	Test for MA(1) term (*p* value)	Test for weak rationality (*p* value)	Test for strong rationality based on short-term interest rates (coefficient/*p* value)	Test for strong rationality based on the real external value of the currency (coefficient/*p* value)	Test for strong rationality based on the change of the oil price (coefficient/*p* value)	Test for strong rationality based on OECD industrial production (coefficient/*p* value)
European Commission, spring	−0.05	0.38	0.48	0.62	0.19	0.15	−0.01/0.78	0.29/0.80	0.00/0.66	−0.02/0.59
Joint forecast, spring	0.05	0.45	0.60	0.52	0.25	0.19	−0.04/0.61	2.72/0.33	−0.00/0.86	−0.03/0.55
No change forecast	−0.08	0.85	1.02	0.37	**0.07**	**0.06**	−0.02/0.57	4.60/0.24	0.01/0.26	0.07/0.19

See text for details. Bold numbers denote a violation of the hypothesis of rational expectations (at least at the 10% level), i.e. of no bias or of weak or of strong efficiency, respectively. The variables additionally used in the test for strong rationality have been standardized to avoid a biased estimation of the constant.

[a] Or longest time span available. See Appendix table for details.

suggested by Holden and Peel (1990). If this hypothesis cannot be rejected, the forecast has to be considered as rational in a stronger sense.[6]

Before turning to our empirical results we wish to emphasise that our analysis relies on a traditional understanding of forecast evaluation, i.e., on the implicit assumption of a symmetric loss function. While this assumption is line with what most forecasters would probably see as their main goal in forecasting, some recent papers have considered the possibility of forecasts that are biased but nevertheless "optimal". Christoffersen and Diebold (1996) for example analyse the optimal prediction under specific asymmetric loss functions. Moreover, Elliott et al. (2004) analyse the loss function under which a particular bias observed in the data may be considered to be rational. However, all statements by forecasters of we are aware point to an unbiased forecast as the objective of the forecast, which is consistent with a symmetric loss function. This, of course, might not hold for all kinds of forecasts. For example, Artis and Marcellino (2001) consider asymmetric loss functions in the case of fiscal (budget) forecasting, where it might well be case that forecasters wish to be "on the safe side" in order to guide policy.

Turning to our growth results, in a nutshell, the findings in Table 1a and b confirms the findings of a lot of previous studies. To begin with, the mean error of the growth forecasts is negative in all but one case, which suggests, on average, a slight tendency of the forecasters to be too optimistic. However, the bias is not significant. A notable exception is the DIW Berlin with a mean error that is even slightly positive. Both the mean absolute error and the root mean squared error are, with very few exceptions, much higher than 1% point. Though the magnitude of these numbers sounds very familiar to economists in this field, it is noteworthy that they still far exceed the expectations of the public regarding the accuracy of forecasts (see, for example, the statements surveyed in Döpke 2000). Moreover, the results come close to what may be considered as a limit of forecast accuracy (Heilemann 2004).

The test of (weak) efficiency is unable to reject the null of an efficient use of the available information. However, it should be borne in mind that the information set represented by the endogenous lagged variable is rather limited. Consequently, the picture presented by the tests on (strong) rationality based on lagged exogenous variables gives a much less favourable impression of forecast quality. In fact, the hypothesis of a rational forecast is rejected frequently.

A surprising result emerged: past short-term interest rates—which are definitely known to all forecasters—help to explain a substantial number of forecast errors and, thus, violate the rationality of the forecasts (see Kirchgässner and Savioz 2001). Even a brief examination of the forecasters' publications will make apparent that virtually all forecasters agree on the importance of monetary policy for the business cycle. As a consequence, the forecasters closely monitor short-term interest rates. If the coefficients are also taken into account, it becomes apparent that high interest rates or strong oil price increases tend to lead to an overestimation of growth. This points to the possibility that forecasters tend to underestimate the dampening effects of both events. The findings regarding real exchange rates and industrial production in OECD countries are, by contrast, less clear-cut. The often-

[6] The exogenous variables were standardised to avoid biases in the estimation of the constant.

stated hypothesis that the alleged poor quality of Germany's business cycle forecasts is due merely to external shocks is therefore not supported, at least not by this exercise.

Turning to the inflation forecasts, no tendency towards overestimation or underestimation was observed. The absolute errors and the RMSE's suggest that inflation forecasts are, on average, somewhat better than growth forecasts. However, this conclusion would be premature since the variance of the actual outcome is much lower in the case of the inflation rate than the relevant number for the growth forecasts. Moreover, although the tests suggest unbiased forecasts, the predictions turn out to be inefficient.[7] As regards the test for strong rationality, the null is rejected less frequently than in case of growth forecasts.

4 Forecast accuracy and efficiency—using the panel dimension

It has been argued that pooling the forecasts might allow more efficient testing of the hypothesis of rationality (Keane and Runkle 1990). Thus, we also estimate the panel $y_{t+1,i} = \beta_0 + \beta_1 y_{t+1,t,i} + u_{t+1,i}$ with i denoting the number of forecasters. To our knowledge, this is the first time that the panel dimension has been used for German data. Furthermore, since some of the empirical techniques used here require a balanced panel, we have left the Halle Institute out of the analysis and have restricted the sample to 1974 to 2004.

In a first step, we estimate the equation using OLS-without and with fixed effects. This might, however,be incorrect, as is argued by Keane and Runkle (1990). These authors argue that an estimation technique that allows for heteroscedasticity should be used in order to capture the influence of shocks common to all forecasters. Thus, we also present results based on GLS estimators taking into account heteroscedasticity in the spirit of Keane and Runkle (1990). However, we extend their approach and discuss results that take into account both heteroscedasticity and possible correlation between the individual forecasts. Cross-section SUR weighted least squares on this specification (sometimes referred to as the Parks estimator) is simply the feasible GLS estimator for systems where the residuals are both cross-sectionally heteroscedastic and contemporaneously correlated. The latter assumption seems reasonable since the forecast errors might be strongly correlated across forecasters. Last but not least, we also perform the Holden/Peel-type tests on forecast rationality using panel data.

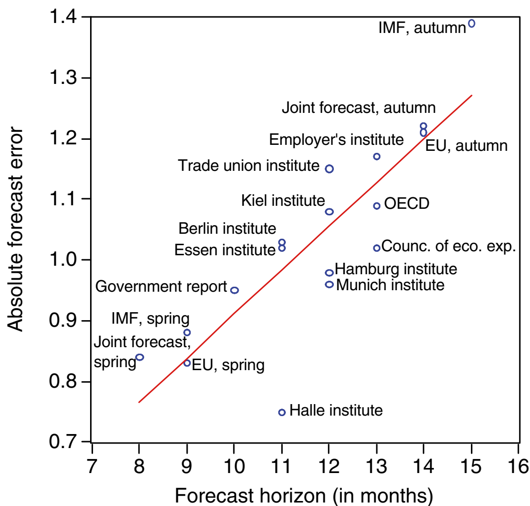
The results of these tasks are summarised in Table 2. The results broadly confirm the findings of the analysis undertaken for single forecasters. However, in contrast to the results regarding individual forecasters, the hypothesis of an unbiased forecast is rejected by means of the Mincer–Zarnowitz (unbiasedness) regression. If one allows, however, for some correlation across the forecasters (i.e., using the SUR estimator), this results no longer holds, since the standard errors become wider. Nevertheless, even using this estimator, the null of strict rationality is rejected in case of the short-term interest rates.

[7] Note that we treat the inflation rate as a stationary variable. We base this judgement on a standard test for unit roots, which are not presented in the paper, but available upon request from the authors.

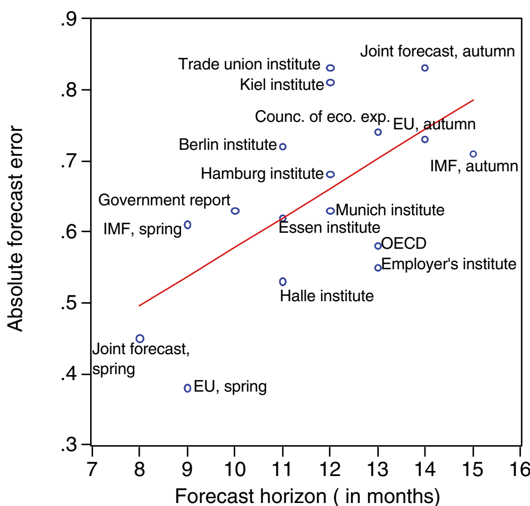
Table 2 Test for forecast rationality—panel evidence

Institution	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\beta}_2$	Test for rationality
	Growth forecasts			
OLS estimation, pooled data	−0.053	0.902	(–)	12.90***
	(−0.43)	(20.32)***		
OLS estimates, random effects	−0.053	0.902	(–)	12.67***
	(−0.43)	(20.14)***		
OLS with fixed effects	−0.076	0.912	(–)	12.20***
	(−0.62)	(20.26)***		
Estimation with cross section SUR	−0.076	0.912	(–)	0.90
and fixed effects	(−0.18)	(5.93)***		
Estimation with cross section SUR	0.289	0.745	−0.193	2.62**
and short-term interest rates	(0.68)	(4.71)***	(−5.89)***	
Estimation with cross section SUR	−0.095	0.912	−3.236	0.96
and real exchange rate	(−0.22)	(5.78)***	(−0.88)	
Estimation with cross section SUR	−0.131	0.933	−1.003	1.48
and oil price changes	(−0.31)	(5.94)***	(−1.60)	
Estimation with cross section SUR	−0.158	0.945	−6.625	0.94
and with industrial production abroad	(−0.72)	(5.76)***	(−1.03)	
Estimation with cross section SUR	0.047	0.898	(–)	0.61
and with control dummies for forecast date	(0.10)	(5.44)***		
Estimation with cross section SUR	0.417	0.719	−0.192	2.30**
and with control dummies for forecast date	(0.93)	(4.19)***	(−2.27)**	
	Inflation forecasts			
OLS estimation, pooled data	−0.315	1.112	(–)	11.88***
	(−3.82)***	(47.36)***		
OLS estimates, random effects	−0.315	1.112	(–)	11.73***
	(−3.79)***	(47.05)***		
OLS with fixed effects	−0.323	1.115	(–)	
	(−3.89)***	(47.08)***		12.25***
Estimation with cross section SUR	−0.323	0.537	(–)	
and fixed effects	(−1.22)	(14.98)***		1.24
Estimation with cross section SUR	−0.353	1.117	0.001	0.86
and short-term interest rates	(−1.09)	(33.47)***	(0.001)	
Estimation with cross section SUR	−0.162	1.055	3.116	1.45
and real exchange rate	(−0.56)	(12.71)***	(1.40)	
Estimation with cross section SUR	−0.310	1.090	0.300	1.08
and oil price changes	(−1.16)***	(14.93)***	(9.83)	
Estimation with cross section SUR	−0.381	1.113	0.792	0.45
and with industrial production abroad	(−1.50)	(15.55)***	(0.24)	
Estimation with cross section SUR	−0.298	1.113	(–)	0.49
and with control dummies for forecast date	(−0.85)	(11.17)***		

t-values in brackets
***(**,*) denotes rejection of the null hypothesis at the 1 (5, 10)% level. The variables additionally used in the test for strong rationality have been standardised to avoid a biased estimation of the constant



Panel a : Growth forecasts



Panel b : Inflation forecasts

Fig. 2 Date of forecast and forecast error 1970 to 2004. **a** Growth forecasts, **b** Inflation forecasts

One might argue against our panel approach that not all forecasts have the same forecast horizon. We are aware of this problem. To obtain an informal impression of the magnitude of this effect, we plotted the absolute forecast error together with the forecast horizon in a scatter plot (Fig. 2). For the growth forecasts, it is quite obvious that there is a strong relationship between both variables; for the inflation forecasts, the case is less clear-cut. However, we kept the data set, since we expect all forecasts to be consistent, the efficiency of forecasts differs, however, owing to the differences in horizons.

Turning again to the panel tests, first, it is noteworthy that the panel estimation with a fixed effects approach—which allows for a forecaster-specific constant and, thus, for some variation due to diverging forecasting dates—does not alter the results. To further control for an influence of the diverging forecasting dates, we split the forecasts into four groups: "early" forecasts, forecasts as of November and December and "late" forecasts. The relevant groups are represented by dummy variables and we control for the dummy variables as well as for interaction terms of the dummies with the slope coefficients. In both cases, the results regarding the rationality tests are not affected.

The data set at hand allows us also to discuss in outline a question that has been debated in Germany recently, namely, whether forecast quality has improved over time. Dicke and Glissmann (2003) argue in a provocative manner that forecast accuracy has not increased, despite the fact that the forecasters' information set has improved. There are several reasons why the information set has improved: for example, data are available with a shorter time lag, the quality of first data releases may also have become better and, perhaps, even the quality of the economic and statistical forecasting models may have improved since the 1970s. On the contrary, Weidmann (2002) defends, in particular, the forecasting record of the Council of Economic Experts, arguing that the forecast quality has, at least, not worsened over time. Heilemann and Stekler (2003) conclude that there is some evidence for improved absolute forecast accuracy in Germany, although there are no clear signs of an improved accuracy in comparison with other countries. To test whether our data set supports the notion of more accurate forecasts, we make use of two rather simple approaches and regress the absolute and the root of the squared forecast error, respectively, on a constant and a set of dummies representing the 1970s, the 1980s and the period form the 1990s onwards. The results—which are available in detail upon request—suggest that the forecast error in the 1970s was indeed significantly higher (by about 1/2% point) than in the subsequent decades. We also regress the absolute and the root of the squared forecast errors on a time trend, which yield a significant negative coefficient and, thus, an improvement of accuracy over time. However, although we find some evidence for declining forecast error, we hesitate to interpret this as a systematic improvement in forecast quality, since results are strongly affected by the large forecast errors 1975 and the first oil price shock.

All in all, the panel estimation results give cause to doubt that the available information is processed by the bulk of forecasters in a way that satisfies the postulate of a strictly rational expectation formation.

5 Tests on forecast equivalence based on model confidence sets

To assess the respective forecasting performance of the different forecasts, we did make use of the relatively new concept of a model confidence set (MCS, see Hansen et al. 2004) in order to select the best model(s) from a collection of models according to a given loss function.[8] Using the in-sample information on forecasting performance of a given model population, the relevant procedure constructs a

[8] We are grateful to Peter Reinhard Hansen and Asgar Lunde, who helped to calculate the respective MCS and p-values.

Table 3 Model confidence set

	Growth forecasts		Inflation forecasts	
	p-value	MCS *p*-value	*p*-value	MCS *p*-value
IMF, autumn	0.00	0.00	0.07	0.07
Joint forecast, autumn	0.01	0.01	0.02	0.02
Europ. Commission, autumn	0.01	0.01	0.09	0.09
Council of Economic Advisors	0.07	0.07	0.06	0.06
OECD	0.04	0.04	**0.12**	**0.12**
Employer's Institute	0.05	0.05	**0.13**	**0.16**
Kiel Institute	0.01	0.01	**0.11**	**0.11**
Hamburg institute	**0.18**	**0.18**	**0.15**	**0.16**
Trade union's institute	0.02	0.02	0.05	0.05
Munich institute	**0.22**	**0.22**	**0.13**	**0.13**
Berlin institute	0.07	0.07	0.06	0.07
Essen Institute	0.10	0.10	**0.14**	**0.16**
Government's economic report	**0.12**	**0.12**	**0.13**	**0.16**
IMF, spring	**0.19**	**0.29**	**0.16**	**0.16**
European Commission, spring	**1.00**	**1.00**	**1.00**	**1.00**
Joint forecast, spring	**0.29**	**0.29**	0.08	0.16

Loss function: mean absolute error
Bold numbers denote membership to level 0.1 model confidence set

confidence set which contains those models which perform best and delivers a confidence bound. As the authors state, "Thus, the MCS is a random data-dependent set of models that contains the best forecasting model(s), as a standard confidence interval covers the population parameter." (Hansen et al. (2004), p. 2).

There are two attractive features of the MCS approach: first, it acknowledges the limitations of the data (or competing objects) and from the results we can evaluate the information content. Very informative data will result in a confidence set which contains only a few optimum models or even only one "best" model. Less informative data will contain possibly all or most of the models under consideration. Second, we can make statements about significance in the traditional sense, which is not possible in the approach of reporting *p*-values from multiple pairwise comparisons and we can therefore have more than model to be chosen as "best performing".

The MCS procedure is based on an equivalence test and an elimination rule. The equivalence test (*p*-value in table) is applied to the models under consideration. If the models are not equally 'good' in the statistical sense, objects with poor performance are eliminated sequentially. Then, the MCS *p*-value (in Table 3) is calculated using a block bootstrap procedure with 5,000 replications. From the proposed test statistics, the deviation from the common average was used. We applied α=0.1 for the confidence level. All tests were calculated with the mean absolute error in the loss function (see Hansen et al. 2004).

The results indicate that, for the growth forecasts, the model confidence set contains more of the "late" forecasts than "early" forecasts. This is consistent with our hypothesis that information content matters in terms of forecast accuracy. However, the Hamburg institute (HWWA) as well as the Munich-based Ifo institute—

both publish the forecast in the middle of the respective sample—seem to be in the same boat as some "late" forecasters.

For the inflation forecasts, we have the case that the forecasts do not appear to be very informative about the realised inflation rates. The procedures pick up the majority of forecasts. However, as in the case of the growth forecasts, there is a clear tendency for "late" forecasts to outperform "early" forecasts and the above-mentioned institutes may be seen as superior to other institutes.

6 Direction-of-changes analysis

Business cycle analysis deals with different regimes such as booms and recessions. Therefore, the turning points between regimes are extremely important for the user of forecasts. It is frequently suspected that professional forecasters tend to miss turning points (e.g. Zarnowitz 1999). In particular, it is a widely accepted stylised fact that forecasters are often unable to forecast recessions (see also Heilemann 1998). This statement, however, relies to a large extent on the NBER tradition of dating business cycles. According to this specific tradition, recessions are rare and irregular periods. Thus, since there is less experience to draw on, these periods are likely to be more difficult to forecast than more regular growth cycles. We rely on a milder version of turning points analysis, namely direction-of-change tests of growth rates.*

To analyse the information content of the forecasts regarding turning points, the forecast errors are classified in a contingency table as shown in Table 4 (cf Diebold and Lopez 1996: 257).

For example, the symbol "ii" stands as an acronym for a forecasted growth acceleration, which at the end has taken place in reality.

With this classification at hand, the information content of a forecast can be summarised using the measure $I = \frac{O_{ii}}{O_{ii}+O_{ji}} + \frac{O_{jj}}{O_{jj}+O_{ij}}$. In a pure coin flip case, where the forecast is made using a magic cube, we have $O_{ii} \approx O_{ji}$ and $O_{jj} \approx O_{ij}$ and therefore $I \rightarrow 1$. If the forecast is perfect then $O_{ji} \approx O_{ij} = 0$ and $I = 2$. Therefore, any value of $1 < I \leq 2$ indicates positive information content. Furthermore, it can be tested whether the information content is significantly different from 1 (cf Diebold and Lopez 1996: 257). To this end, we estimate the expected cell counts under the null hypothesis of no information content. The consistent estimator for the cell counts is given by $\widehat{E}_{ij} = O_{i.}O_{.j}/O$. Finally, we construct the following measure

Table 4 Classification of directional forecast errors

		Actual outcome		Sum
		Upswing	Downswing	
Predicted	Upswing	O_{ii}	O_{ij}	$O_{i.}$
	Downswing	O_{ji}	O_{jj}	$O_{j.}$
	Sum	$O_{.i}$	$O_{.j}$	O

Source: Diebold and Lopez (1996, page 257)

Table 5 Evaluating direction of changes forecasts 1970–2004[a)]

	Growth			Inflation		
	I	C	p value	I	C	p value
Joint forecast, autumn	1.34	5.38	0.02	1.46	9.19	0.00
Joint forecast, spring	1.52	10.62	0.00	1.76	22.32	0.00
Council of Economic Advisors	1.31	3.94	**0.05**	1.51	11.82	0.00
Kiel Institute	1.41	6.33	0.01	1.51	11.82	0.00
Berlin Institute	1.47	8.24	0.00	1.50	16.35	0.00
Hamburg Institute	1.36	4.90	0.03	1.51	11.82	0.00
Munich Institute	1.47	8.24	0.00	1.40	7.05	0.01
Essen Institute	1.36	4.90	0.03	1.39	9.91	0.00
Halle Institute	1.83	8.92	0.00	1.43	4.81	0.03
OECD	1.65	14.24	0.00	1.72	20.59	0.00
Employer's institute	1.78	20.35	0.00	1.33	7.18	0.01
Trade union's institute	1.41	6.33	0.01	1.45	10.62	0.00
Government economic report	1.41	6.33	0.01	1.52	13.73	0.00
European Commission, autumn	1.39	5.85	0.02	1.58	14.92	0.00
European Commission, spring	1.53	10.15	0.00	1.76	22.32	0.00
IMF, autumn	1.47	7.43	0.01	1.62	16.47	0.00
IMF, spring	1.39	6.87	0.01	1.47	14.74	0.00

Bold numbers denote that the hypothesis of independence between forecast and realization cannot be rejected (at 5% level)
[a)] Or longest time span available. See Appendix table for details

$C = \sum_{i,j=1}^{2} \left(O_{ij} - \widehat{E}_{ij}\right)^2 \Big/ \widehat{E}_{ij}$, which is χ^2 distributed with one degree of freedom under the null hypothesis of no information content. This measure is also known as Pearsons χ^2.

The results of such tests are given in Table 5. It turns out that all forecasters included in the test are better than a pure coin flip. This result holds true for growth as well as for inflation forecasts. Indeed, it is comforting that all forecasts contain information content. The argument—sometimes brought into the discussion by the yellow press or some frustrated politicians—that all these forecasts are nothing more than a coin flip is therefore simply not true. On the contrary, many institutions seem to forecast the direction of changes quite well. Again, the impression is that timing matters: for example, the "joint diagnosis" in autumn and the *Sachverständigenrat* (German Council of Economic Experts) are early birds; it is therefore by no means surprising that they miss the direction of changes more often than other institutions with access to more timely information.[9]*

Given the results for individual forecasts, the question may arise as to whether one should rely on the consensus only and ignore possible "minority reports". To shed light on this question, we take a look at the direction-of-change forecast

[9] Again, the result for the IWH has to be interpreted with caution since we only have 12 exploitable observations.

Table 6 Directional forecast errors: majority vs minority 1970 to 2004

Growth forecasts (number of minority reports: 12)		
Error type predicted/outcome	Majority	Minority [a)]
Acceleration/acceleration (II)	11	9
Acceleration/deceleration (IJ)	6	5
Deceleration/acceleration (JI)	12	13
Deceleration/deceleration (JJ)	5	7
Information content: II/(II+IJ)+JJ/(JJ+JI)	1.35	1.29
Inflation forecasts (number of minority reports: 15)		
Acceleration/acceleration (II)	10	8
Acceleration/deceleration (IJ)	1	7
Deceleration/acceleration (JI)	19	13
Deceleration/deceleration (JJ)	4	6
Information content: II/(II+IJ)+JJ/(JJ+JI)	1.74	1.22

[a)] We count as minority report, if at least two forecasts out of all forecasts disagree with respect to the predicted acceleration/deceleration

errors. To this end, we checked whether, in a given year, all forecasters agreed on the direction-of-change. The classification of forecast errors is the same as above, i.e., relies on the acceleration or deceleration of growth rates. If at least two forecasters predict acceleration (deceleration) whereas the other forecasters forecast a deceleration (acceleration), we count this as a minority report.

It appears that there are 12 (15) years in which a minority of forecasters disagrees on the question of whether growth (inflation) will speed up or not. Table 6 gives the measures of information content introduced in the previous chapters. The results differ sharply between inflation and growth forecasts. In the case of the growth forecasts, the minority provides as much information as the majority. Whereas the majority has more often predicted an upswing correctly, the opposite holds true for downswings. On average and according to this criterion, it does not matter to whom the user of the forecasts listens. It is notable, however, that the minority might do somewhat better in predicting recessions. Unfortunately, we cannot test the idea rigorously, since there are only three major recessions in our sample (if we count a negative growth rate as a recession). In two cases, however, the majority has even predicted an acceleration of growth rather than a deceleration, not to mention a negative growth rate. In the case of the inflation forecasts, however, the rational consumer of forecasts should rely on the opinion of the majority. Despite the fact that minority reports are more frequent as regards inflation, the minority provides substantially less information than the majority and even less information than a coin flip.

7 Conclusions

Based on 35 years of forecasts of professional forecasters stemming from 14 institutions and up to 17 different regularly published forecasts, we analyse business cycle forecasts in Germany. We extend existing studies, since we take into account all forecasters who have contributed regularly to the German forecast

market over a period of more than 30 years. This novel data set renders it possible to exploit the cross-section dimension of the data, too. For example, besides usual measures of forecast accuracy for individual forecasts we are able to test on the relative performance of the forecasting institutions by means of model confidence sets. The results of these tasks confirm the findings of earlier studies: German business cycle forecasts are, on average, unbiased and weakly rational. By contrast, the hypothesis of strong rationality has to be rejected. Since we have more than just one series of forecasts at hand, we can test these hypotheses in a panel framework, thereby adding an additional piece of evidence to the existing findings. The results based on a panel confirm the findings found for individual forecasters. As regards the relative accuracy, we do not find systematic differences between the forecasters besides the date when the forecast was made. Exceptions are the Hamburg-based HWWA institute and the Munich-based Ifo institute. However, if we refer to the criterion of direction-of-change tests, we show that—as far as growth forecasts are concerned—it is not always advisable to listen to a majority of the forecasters. This is not the case for inflation forecasts.

The results of this study may and should be extended in various directions. Concerning forecast accuracy, a data set with a higher frequency might be helpful, since it renders it possible to analyse the changes of forecasts over time and the dynamic interaction of forecasters with the consensus. Moreover, in further studies, the reasons for forecast errors should be addressed more extensively. In particular, the violation of the hypothesis of strong rationality deserves a sufficient explanation. Also, a data set covering privately financed forecasting institutions would allow us to check whether different incentive structures due to the institutional position of the forecaster have an impact on accuracy and efficiency.

Acknowledgements The authors wish to thank two anonymous referees as well as Imke Brüggemann, Andreas Cors, Roland Döhrn, Felix Hüffner, Hans-Martin Krolzig, Thomas Knetsch, Udo Ludwig, Francis Piotrowski, Jürgen Wolters, Rudolf Zwiener and seminar participants at the Institute for Statistics and Econometrics at the Free University Berlin for their helpful comments on an earlier draft of this paper. We also wish to thank Birger Antholz, Hamburg, for pointing us to errors in our data and for providing us with additional data. We are particularly grateful to Peter Reinhard Hansen and Asgar Lunde, who used their procedures to calculate the relevant MCS and *p*-values for us. The usual disclaimer applies. The views presented in this paper are the authors', and do not necessarily reflect those of the DIW Berlin or the Deutsche Bundesbank.

Appendix

Institutions and forecasts under investigation

Institution	Source of forecast	Start/ending dates of forecasts	Average month until end of forecast horizon	Additional notes
Council of Economic Advisors	*Sachverständigenrat*, annual reports, several years, Wiesbaden	1970–2004	13	

Institution	Source of forecast	Start/ending dates of forecasts	Average month until end of forecast horizon	Additional notes
Joint forecast of the six (formerly five) "leading" German economic research institutes, autumn forecast	*Arbeitsgemeinschaft wirtschaftswissenschaftlicher Forschungsinstitute*, various issues, Berlin	1970–2004	14	
Joint forecast of the six (formerly five) "leading" German economic research institutes, autumn forecast	*Arbeitsgemeinschaft wirtschaftswissenschaftlicher Forschungsinstitute*, various issues, Berlin	1970–2004	8	
German Institute of Economic Research (DIW Berlin)	German Institute of Economic Research (ed), Economic Bulletin, various issues, Berlin	1970–2004	11	Two forecasts per year; we refer to the one in January only
Kiel Institute for World Economics (IfW)	"*Die Weltwirtschaft*", various issues, Kiel	1970–2004	12	No forecasts for 1975 and 1976 (see Stege (1989)) for a discussion of possible reasons). We use numbers from the joint forecast for these years, since the Kiel Institute was part of the joint forecast. For forecasts, we use the December forecast.
HWWA Institute of Economic Research, Hamburg	"*Wirtschaftsdienst*", various issues, Hamburg	1970–2004	12	Two forecasts per year, we refer to the autumn forecast.
Ifo Institute, Munich	"*ifo Schnelldienst*", various issues	1970–2004	12	
Institute for Economic Research Halle (IWH)	"*Wirtschaft im Wandel*", various issues, Halle/Saale	1993–2004	11	Institute founded in 1992
Rhine-Westphalia Institute for Economic Research (RWI), Essen	Institute's "Economic report"	1970–2004	11	
OECD	Economic Outlook, various issues, Paris	1970–2004	13	Two forecasts, we use the autumn forecast.
IMF, spring forecast	World Economic Outlook, various issues, Washington	1974–2004	9	Some forecasts have not been published. Here we use the data collected by Artis (1987).

Growth and inflation forecasts for Germany

Institution	Source of forecast	Start/ending dates of forecasts	Average month until end of forecast horizon	Additional notes
IMF, autumn forecast	World Economic Outlook, various issues, Washington	1974–2004	15	Some forecasts have not been published. In this cases we use the data collected by Artis (1987).
European Commission, spring forecast	European Economy, various issues, Brussels	1974–2004	9	
European Commission, autumn forecast	European Economy, various issues, Brussels	1974–2004	14	
Trade Union Institute (WSI)	*WSI Mitteilungen, various issues, Düsseldorf*	1974–2004	12	In 2001 no growth forecast and in 2001 and 2002 no inflation forecasts.
Institute of the German Economy (IW)	*IW-Trends; various issues, Cologne*	1970–2004	13	
Economic report of the German federal government	*Jahreswirtschaftsbericht*, various issues, Bonn/Berlin	1970–2004	10	

References

Artis MJ (1987). How accurate are the IMF's short term forecasts? Another examination of the world economic outlook. Staff Studies for the World Economic Outlook. Washington, District of Columbia: 1–36

Artis, MJ, Marcellino M (2001) Fiscal forecasting: the track record of the IMF, OECD and EC. Econometrics Journal 4:20–36

Büttner T, Horn GA (1993) Wie gut sind Prognosen? Eine Überprüfung der Vorhersagen der Gemeinschaftsdiagnose und des DIW am Beispiel des BSP. DIW Diskussionspapier Nr. 73. Berlin

Christofferson PF, Diebold FX (1996) Further results on forecasting and model selection under asymmetric loss. Journal of Applied Econometrics 11:561–571

Dicke H, Glissmann HG (2002) Konjunkturprognosen und wissenschaftlich-technischer Fortschritt. Wirtschaftsdienst 82:167–169

Diebold FX (1998) Elements of Forecasting. Cincinnati

Diebold FX, Lopez J (1996) Forecast evaluation and combination. In: Maddalla GS, Rao CR (eds.) Handbooks of statistics 241–268

Döpke J, Langfeldt E (1995) Zur Qualität von Konjunkturprognosen für Westdeutschland 1976–1994. Kieler Diskussionsbeitrag Nr. 274. Kiel

Döpke J (2000) Haben Konjunkturprognosen einen politischen Bias? Schmollers Jahrbuch-Zeitschrift für Wirtschafts-und Sozialwissenschaften 120:587–620

Elliot G, Kommunjer I, Timmermann A (2004) Biases in macroeconomic forecasts: irrationality or asymmetric loss? http://repec.org/esNASM04/up.6440.1075609439.pdf

Hansen PR, Lunde A, Nason JM (2004) Model confidence sets for forecasting models. http://ssrn.com/abstract=522382

Heilemann U (1998) Zu den Projektionen des Jahreswirtschaftsberichts der Bundesregierung. In: Heilemann U, Kath D, Kloten N (Hrsg.) Entgrenzung als Erkenntnis-und Gestaltungsaufgabe: Festschrift für Reimut Jochimsen zum 65 Geburtstag: 79–100 Berlin

Heilemann U (2004) Besser geht es nicht-Genauigkeitsgrenzen von Konjunkturprognosen. Jahrbücher für Nationalökonomie und Statistik. Forthcoming

Heilemann U, Stekler HO (2003) Has the accuracy of German macroeconomic forecasts improved? Discussion Paper of the German Research Council's research centre 475. 31/03
Hinze J (1996) Konjunkturprognosen unter veränderten Rahmenbedingungen: Auswirkungen der deutschen und europäischen Einigung. Hamburger Jahrbuch für Wirtschafts-und Gesellschaftspolitik 41:59–84
Holden K, Peel DA (1990) On testing for unbiasedness and efficiency of forecasts. The Manchester School 58:120–127
Keane MP, Runkle DE (1990) Testing the rationality of price forecasts: new evidence from panel data. American Economic Review 80:714–735
Kirchgässner G (1984) Sind die Erwartungen der Wirtschaftssubjekte rational? Eine empirische Untersuchung für die Bundesrepublik Deutschland. Weltwirtschaftliches Archiv 120:279–300
Kirchgässner G, Savioz M (2001) Monetary policy and forecasts for real GDP growth: an empirical investigation for the Federal Republic of Germany. German Economic Review 2:339–365
Kirchgässner G, Müller UK (2003) Are forecasters reluctant to revise their predictions? Mimeo, Swiss Institute of International Economics and Applied Economics analysis
Neumann MJM, Buscher HS (1980) Die Inflationsprognosen der Arbeitsgemeinschaft wirtschaftswissenschaftlicher Forschungsinstitute: sind sie rational? Weltwirtschaftliches Archiv 116:533–550
Newey W, West K (1987) A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55:703–708
Nierhaus W (1998) Ifo-Konjunkturprognosen und Wirklichkeit-eine Ergebnisanalyse für die Bundesrepublik im Zeitraum 1991–1997. ifo-Wirtschaftskonjunktur 50:1–11
Stege C (1989) Zur Beurteilung der Prognosetätigkeit der "führenden Wirtschaftswissenschaftlichen Forschungsinstitute" in der Bundesrepublik Deutschland. Hamburg
Tichy G (1994) Konjunktur. Stilisierte Fakten, Theorie, Prognose. Berlin
Weidmann J (2002) Hat sich die Prognosequalität des Sachverständigenrates systematisch verschlechtert? Wirtschaftsdienst 82:741–747
Zarnowitz V (1999) Has the business cycle been abolished? NBER Working Paper No. 6367. Cambridge MA