Research Program on Forecasting

HAS THE ACCURACY OF MACROECONOMIC FORECASTS FOR GERMANY IMPROVED?

Ullrich Heilemann
Institut für Empirische Wirtschaftsforschung
Universität Leipzig
Grimmaische Straße 12
04105 Leipzig, Germany

Herman O. Stekler
Department of Economics
George Washington University
Washington, DC 20052 USA

RPF Working Paper No. 2010-001
http://www.gwu.edu/~forcpgm/2010-001.pdf

February 28, 2012

RESEARCH PROGRAM ON FORECASTING
Center of Economic Research
Department of Economics
The George Washington University
Washington, DC 20052
http://www.gwu.edu/~forcpgm

HAS THE ACCURACY OF MACROECONOMIC FORECASTS FOR GERMANY IMPROVED? [1]

By Ullrich Heilemann, Leipzig, Germany, Herman Stekler, Washington, D.C., USA[2]

JEL E37

Forecast evaluations, macroeconomic forecasting, accuracy limits

Summary

This paper asks whether the accuracy of German macroeconomic forecasts has improved over time. We examine one-year-ahead forecasts of rates of real GDP growth and inflation for the years 1967 to 2010 by three major German forecasters and the OECD. We find that overall error levels are high but not much different from those for the U.S. and U.K. In the 1980s and 1990s accuracy improved somewhat, but has now returned to its 1970s level, indicating that it reflects the variance of growth and inflation. Benchmark comparisons of these predictions with ex post forecasts of a macroeconometric model indicate that accuracy can be improved but it will be difficult to achieve.

---

1 Dedicated to the memory of *Victor Zarnowitz* (November 3, 1919 – February 21, 2009) a great economist, a sceptic on economic forecasting, and a good friend to both authors.

2 For excellent research assistance the authors are indebted to Sissy Issleb. For critical comments we thank Roland Döhrn, Heinz Josef Münch, Essen, Germany, and two anonymous referees. The financial support of Deutsche Forschungsgemeinschaft (SFB 475, “Reduction of complexity in multivariate data structures”) is gratefully acknowledged.

1. Introduction

Quantitative forecasting in Germany began in earnest in the early 1960s when the "Gemeinschaftsdiagnose" (Joint Diagnosis (JD)) of five (now four) large economic research institutes started to publish forecasts fully consistent with the National Accounts (NA) framework (Antholtz 2005, pp. 31ff.). These forecasts were soon followed by the "Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung" (Council of Economic Experts) and the Federal Government's "Jahreswirschaftsbericht der Bundesregierung" (Annual Economic Report of the Federal Government). In the 1970s, an increasing number of private forecasters, most of them from the banking sector, also started to issue macroeconomic forecasts. Including the IMF, the OECD, the World Bank, and the EU-Commission, there are now more than 35 institutions that regularly publish macroeconomic forecasts for Germany. Only a small number of them are as detailed as the JD, CEE or the GAER, and only about a dozen of them receive public attention.

The general state of the art in such macroeconomic forecasting is surveyed in Fildes/Stekler (2002) and Stekler (2009). The present paper presents an in-depth examination of the accuracy of the rate of growth forecasts for real GDP and inflation (GDP deflator) for Germany from 1967 to 2010 to examine (1) whether accuracy has improved over time and (2) whether the characteristics of these forecasts are similar to those of the U.S. and U.K. In addition, we use ex post econometric model results to reflect on the increase in forecast accuracy that can be expected. Of course, there are other criteria besides accuracy that determine the quality of forecasts such as their informational content or their theoretical foundation (Wild, 1974, p. 138). But there are few users for whom accuracy is not the most important criterion, not least because there are few differences between major forecasters on the other criteria.

Related papers include Öller and Barot, 2000; Pons, 2000; Kreinin, 2000; Blix et al., 2001; Döpke and Fritsche, 2006; and Kholodilin and Siliverstovs, 2009. These studies report the usual statistics for absolute and relative accuracy or other forecast characteristics over a specific span of time. Depending on the forecasters and the forecasting period, the mean absolute errors (MAE) of the forecasts of the rate of real GDP growth (inflation) vary between 1.2 (0.6) and 1.6 (0.8). Most studies concluded that there is no

forecaster (or method) that is uniformly superior (Döpke and Fritsche, 2006). These findings are similar to results for the U.S. (e.g. Zarnowitz, 1992).

However, these studies lack an explicit systematic analysis of the way that accuracy has changed over time. Dicke/Glismann (2002) offer a brief analysis of forecast accuracy over time for one of the research institutions, and Döpke/Fritsche (2006) suggest that accuracy may have improved. Other studies make implicit references to this question, e.g. Döpke and Langfeldt, 1995; Heilemann, 1998; and Döpke, 2000. The findings are contradictory, and there are no definitive conclusions (e.g. Burns, 1986; McNees, 1988; Melliss, 1997; Öller and Barot, 2000; Vogel, 2007; and Timmerman, 2007). A remarkable but much overlooked result was Zarnowitz' (1992) finding that the accuracy of U.S. real GDP forecasts for the period 1953 to 1989 had not improved. It is therefore appropriate to revisit the question of whether forecasts have improved over time. The results are taken as an occasion to inquire after the limits of accuracy that can currently be expected from macroeconomic forecasts. Though an explanation of our main findings is beyond the scope of this paper, they represent a first step in this direction.

The next section will discuss our sample of forecasters, the time periods that will be examined, and the methods of analysis. Section 3 has the full sample results; section 4 presents the results over time. The final section summarizes the findings and offers conclusions and suggestions.

2. Forecasters, samples, data, methods of analysis

Among the dozen major macroeconomic forecasts for Germany, only four sets of forecasts are examined here. The criteria used in selecting these forecasting organizations were: they should play an important role in the public discussion of economic policies and they must have produced a sufficient number of forecasts to determine whether their accuracy has improved over time. Private sector forecasters with a published long enough record do not exist. In addition, we wanted to include forecasts from both a government institution and an international organization. Finally, the forecasts had to be comparable in terms of the set of variables forecast, the forecast horizon, and the date of their publication. This led to the selection of the forecasts produced by (1) the JD (Ar-

beitsgemeinschaft, 1967ff.),[3] (2) the CEE (Sachverständigenrat, 1967ff.), (3) the GAER (Bundesregierung, 1967ff.), and (4) the OECD (1967ff.).

We consider two crucial variables: the rates of change of real GDP and of the GDP deflator. "Growth" and "no inflation" are the most important macroeconomic goals for German economic policy. Given the strong dependence of employment, government deficits, etc. on GDP, the accuracy of these two variables are good indicators of what might be expected for the other variables.

In order to have a common base, the analysis starts with 1967, the first year for which the GAER published a forecast. The sample ends in 2010. Despite German unification in 1990, for the period 1991 to 1994 we still use forecasts for West Germany. To examine the evolution of forecast accuracy, the sample is divided into four sub periods: 1970-1979, 1980-1989, 1990-1999, and 2000-2010. These decades are frequently used in analyses of the type done here. Each decade experienced at least one recession. Other events affecting forecast accuracy such as the oil-shocks in the 1970s and 1980s, German unification, the first fiscal consequences of the Maastricht treaty, the Asia/Russia crisis in 1997/98, and the beginning of the Great Recession of 2008ff. are also included. As a consequence, the size and tendencies of the error measure values calculated for one or two cycles hardly differ from the decade-based results reported in the following pages. (This and all other results not reported here, data and sources are available as supplementary material on the GER website.)

The forecasts offer predictions for the remaining part of the current year as well as for the following year, but we analyse only the year-ahead predictions. Forecasts are published over a stretch of four months: October (JD), November (CEE), December (OECD), and January (GAER), but the information on which they are based are not too different (on this point see Döhrn and Schmidt, 2011). The JD, CEE and, given its three months of reconciliation, also the OECD forecasts, have to start from NA data ending

---

3 The JD is a group of four to six major German research institutes who produce macroeconomic forecasts twice a year under a contract with the Federal Government. The composition of the JD has changed several times; currently main contractors are: Ifo-Institut für Wirtschaftswirtschaftsforschung, München; Institut für Weltwirtschaft (IfW), Kiel; Institut für Wirtschaftsforschung Halle (IWH), Halle; and Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI) Essen. Until recently the group included also Deutsches Institut für Wirtschaftsforschung (DIW), Berlin, and Hanseatisches Weltwirtschaftsarchiv (HWWA), Hamburg.

with the second quarter. The GAER, however, can start from data for the third quarter and can probably also use the Federal Statistical Office's first estimate of GDP for the past year, which is issued in mid-January (of the following year). In the period studied here, there were only a few cases in which macroeconomic developments and events of essential importance occurred between October and January. Although the GAER forecasts use more information, notably more recent data, and thus should be more accurate, this is hardly the case.

The JD's and the CEE's forecasts have been published for some time with rates of change rounded to 0.5 percentage points; if forecasts have been presented as ranges (e.g. the GAER), the mean value has been taken. As actuals we used, as is now common practice, the NA-data released by the German Statistical Office as close as possible to the forecast, usually in February of the following year. In order to provide a fair basis of comparison, all forecasts and actual data were rounded. A comparison of the results of rounded with decimal point OECD forecasts for Germany showed hardly any differences, and if there were differences they rarely exceeded the value of 0.1 (Table 4). Of course, rounded values offer less information, but the quality of forecasts as determined by their degree of disaggregation, theoretical foundation, timeliness, past accuracy etc. (Wild, 1973, pp. 134ff.) is of no interest here. In 1993, the Federal Statistics Office changed its NA concepts and replaced GNP with GDP as its measure of output. Hence, until 1993, "growth" is associated with real GNP, and thereafter with real GDP; the inflation indicator was changed correspondingly. In 2004, the Federal Statistical Office started a new NA system and switched to chain weighted deflation. As forecasts followed this standard, this switch should not affect the present analysis.

Our measures of forecast accuracy include descriptive statistics as well as parameter tests, tests for directional accuracy, and rationality tests. To measure forecast accuracy we focus on mean absolute error (MAE, error: $p_t - a_t$ with p: forecast value, a: actual), bias, and root-mean-square error (RMSE). As a benchmark, comparative accuracy is measured by Theil's U coefficient (based on extrapolating the previous rate of change $p_t = a_{t-1}$). Given the general decline of both growth and inflation rates, the test may be seen as biased against an extrapolation of the previous year's rates of change. Since the decline extended over more than 40 years, this distortion should be small. The forecast performance associated with the difficulty of the task is measured by the relationship of

RMSE/σ (McNees, 1988). Similarly, the forecast performance was measured by RMSE/average growth rates of real GDP and inflation. The results were similar and not shown here.

In analysing directional accuracy, we first describe the type of errors that were observed, e.g. the failure to predict turning points and the number of over- and underestimates that occurred. Then, we determine whether accelerations and decelerations in growth and inflation rates were correctly predicted. We use the concept of "informational content" (IC), which compares the number of forecasted accelerations (or decelerations) of changes to the number of those actually realized (see e.g. Diebold and Lopez, 1996):

$$\mathrm{IC} = \frac{\mathrm{AC}}{\mathrm{AC} + \mathrm{AW}} + \frac{\mathrm{DC}}{\mathrm{DC} + \mathrm{DW}}$$

with AC: increase forecast and realized; AW: increase forecast, decrease realized; DC: decrease forecast and realized; and DW: decrease forecast, increased realized. For a forecast to have "informational content", IC has to be > 1 (Merton, 1981).

The rationality of forecasts, based on unbiasedness and efficiency, is tested using the Mincer/Zarnowitz equation (Mincer, Zarnowitz 1969). A sufficient condition for forecasts to be unbiased is that the joint null hypothesis $\alpha_1 = 0$ and $\beta_1 = 1$ in regression (1) cannot be rejected.

$$a_t = \alpha_1 + \beta_1 \cdot p_t + u_t, \qquad (1)$$

with a_t:= realized; p_t:= forecast and u_t the error term for which the assumptions of the classical regression model hold. The forecasts are efficient if $\beta_2 = 0$ in (2)

$$e_t = \alpha_2 + \beta_2 \cdot p_t + u_t, \qquad (2)$$

and $\rho = 0$ in (3).

$$e_t = \alpha_3 + \rho \cdot e_{t-1} + u_t. \qquad (3)$$

Like Theil's inequality coefficient, this test is based on the assumption that the previous year's *actual* data are known, which the case is at best for the GAER.

3. Results: The complete sample—a summary

Forecasts and actual rates of growth and inflation for the full range of years are shown in Figure 1 and Table 5 (Appendix). The results of the accuracy analysis may be seen in Table 1. The MAE of the growth forecasts was about 1.5, about 40% of the mean growth rate. The bias is comparatively large. The MAE of the inflation forecasts is about 0.7, only 20% of the mean rate of inflation. Inflation forecasts are free from bias and all in all more accurate than growth predictions, contrary to findings for the U.S. and the U.K. [4]

(*Figure 1 about here*)

Theil's U coefficient indicates that the forecasts were superior to simple extrapolations of the previous actual rates of change. The average errors of all four groups were similar for both variables, with the JD's higher growth errors being a possible exception. Although the forecasts were highly correlated, we tested whether there was a statistically significant difference in the accuracy of the four groups. The forecasts for each year were ranked based on their accuracy. The average ranking test (also called analysis of variance by ranks) was also used (Batchelor, 1990). There was no significant difference among the four groups' predictions of inflation but for growth. The values of χ^2 were 9.46 and 0.62 for the growth and inflation forecasts, respectively. The critical 5% value of the statistic with three degrees of freedom is 7.81. The tests of the growth forecasts were based on all 44 observations, but we only used the last 39 observations for the inflation predictions because the OECD did not forecast inflation in either 1967 or 1971.

(*Table 1 about here*)

Considered in light of a four-phase business cycle classification, forecasts are more accurate during upswings and upper turning point phases than in lower turning point phases (Figure 1). The JD and CEE failed to predict any of the five recessions that occurred in this period. The OECD and GAER were more successful, but mostly failed to predict the gravity of the recessions. This result is similar to U.S. and U.K. forecasts.

---

4 Although the forecast periods are not the same, it is possible to compare these results with those that Fildes/Stekler (2002, pp.443ff.) reported for the U.S. and U.K. In the U.S. the errors were about 25% of the mean absolute changes of both variables, while in the U.K. they averaged about 60%. It should also be noted that between 1968 and 1999, the MAE between the first and the final actual data had been 0.4 percentage points for growth and 0.3 percentage points for inflation.

German forecasts also displayed most of the systematic errors that have been long observed in many countries. Fildes/Stekler (2002) note that the U.S. and U.K. forecasters underestimate GDP when it is growing and overestimate when it is declining; similar errors were observed when inflation was accelerating and decelerating. The forecasts for Germany exhibit the same patterns. About 25% of the growth forecasts are "overestimates" and 50% of them are "underestimates", while about 10% accurately predict the development and about 10% miss the turning points (Table 2). A similar pattern holds for the inflation forecasts, though with more incidences of agreement and fewer "turning point errors".

(*Table 2 about here*)

All of the forecasters produced forecasts with sufficient informational content, yielding IC values > 1, indicating that most of the periods of acceleration and deceleration were anticipated correctly (Table 3). Differences in directional errors of growth and inflation forecasts are about 0.7 on average for both accelerations and decelerations, varying little between forecasters.

(*Table 3 about here*)

Finally, the regression rationality test did not reject the null hypothesis that forecasts are unbiased. However, the hypotheses of the efficiency of both GDP growth and inflation forecasts were rejected. The β-test (Equation 2) indicated that the forecast errors were positively related to the forecasts, and the ρ-test (Equation 3) revealed that most forecast errors were autocorrelated.

4. Results: accuracy over time

Starting with directional errors, for GDP growth forecasts the number of coincidences as well as turning point errors remain rather stable until the most recent period; the relationship between overestimates and underestimates fluctuates (Table 2). Inflation forecasts were generally underestimated when inflation was increasing (in the 1970s) and overestimated when it was decreasing (in recent years).

The small number of observations in each sub period precludes formal statistical tests, but descriptive results can be obtained from the information content (IC) statistics. If there had been an increase in accuracy over time, these statistics should have increased

from the 1970s to the present. However, this has not generally been the case (Table 3). The information content of the growth forecasts deteriorates in the 1980s, generally improves in the 1990s, and declines again thereafter, suggesting that there is no tendency towards a monotonic improvement in directional accuracy. A similar result can be observed for inflation.

The interpretation of the time trend of the quantitative forecast errors depends on which error measure is used. For the growth forecasts simple trend results over the complete sample suggest for all four institutions stable average errors (AE). Looking at the decades, while MAEs declined considerably in the 1980s and 1990s but in the most recent decade returned to 1970s levels (Table 1); pairwise t-tests show no differences for the MAE (90%). The bias developed similarly. Theil's U is rather constant, but in the most recent decade forecasts by both the JD and CEE were considerably higher than that of the GAER – a result mainly due to the errors of the former in 2009. The picture for inflation is much better: the trend shows a clear but small (-0.02) significant decline. The large errors in the 1970s, results of the wage explosion in the early 1970s and the oil shock, both seen only after the fact, are followed by a decline of the MAE to about 0.5, with accompanying small values for bias and Theil's U. The pairwise tests of the MAE, suggest significant improvements in the 1980s vs. the 1970s and the 1990s vs. the 1980s for the JD and the CEE, and for the 1990s vs. the 1990s vs. the 1980s for the OECD.

We have not examined the results from a comparative perspective. In particular, we did not adjust for the difficulties involved in forecasting, though Theil's U might be seen as doing this implicitly. One possible adjustment is to divide the RMSE by the standard deviation of the actual changes that occurred in each period. The last entry in each panel of Table 1 presents this measure. The results indicate that the forecast errors, adjusted for this variability, improved by a small degree for the growth forecasts and considerably for the inflation forecasts in the 1980s and 1990s, but exceeded or returned to their 70s' levels in the 2000s. Overall, there is some evidence of minor improvements in absolute forecasting accuracy, particularly if the oil and wage shocks of the 1970s are taken into account; however, relative stability has been rather constant (though the 2009f. experience is a serious challenge to this finding). These findings may be further qualified by pointing out the falling trend of the rates of growth and inflation over the last 45

years. Indeed, standardizing MAEs by the average rates of growth and of inflation also shows statistics move steadily.

The data reveal some factors that reduce accuracy and suggest areas where a forecaster should invest more effort in order to improve accuracy. The effects of the errors made in predicting the recessions and downswings of 1974, 1980/81, 2001, and 2009/10 can be identified even in the recursive accuracy of growth forecasts mentioned above (not shown here). While this finding suggests that greater emphasis should be placed on timely predictions of recessions (not a new demand), it must be remembered that forecasters in other countries also fail to predict the onset of recessions (and recoveries), as the recent crisis has demonstrated once more.

Similarly, the impacts of wage inflation and oil shocks on inflation can be observed in the first half of the 1970s, but the error statistics steadily return towards their previous "limits" after these events. The most plausible explanation is that exogenous inflation impulses and internal inflation behaviour simply normalized (see Figure 1), and forecasters were able to forecast more accurately in this environment. However, the error statistics show a declining trend up to the Great Recession of 2009f. that seems to be approaching a limit, i.e. a level beyond which accuracy cannot go, at least not with the current state of knowledge, forecasting methods, and data. This result is corroborated when examining the OECD's forecasts for the G7-countries (Table 4). The anticipated major improvements cited above do not seem likely to materialize in the near future. While Fildes/Stekler (2002) did not discuss the limits of accuracy, their results are compatible with this view.

(*Table 4 about here*)

Further improvement is, of course, an important area of investigation, which can both help to judge the present state of forecasting and also help to identify some areas in which improvement is possible. We employ a macroeconometric model to do so, with similar aggregative detail as the forecasts analysed above. Static ex post-simulations within the sample period, that is successively solving the model for each with actual values of the exogenous and lagged endogenous ("predetermined") variables, produce the most accurate results possible with this model and thus supply a benchmark. The errors produced by this model are the sum of the single equation errors and the one-period "model errors". They are lower than that of ex ante model forecasts because they

are free from errors caused by incorrect predetermined variables, the model's limited capacity to capture the dynamics of multi-period forecasts, or the model's limited stability outside the sample period, i. e. in ex ante-forecasts.

Suppose that the ex ante forecast errors produced by a model are comparable to those of the institutions' forecasts for the same period. In this case a comparison of the ex-ante and the static ex post-forecasts would indicate the limits of forecast accuracy, i.e. the accuracy that could be attained with perfect forecasts of the predetermined variables, perfect capture of the model dynamics, and stable relationships in the forecast period. Increased accuracy would be possible only by increasing the accuracy of the model's single equations and their static interactions.

For this exercise, we use the RWI business cycle model, a medium sized (quarterly) macroeconometric model employed since the late 1970s for short term ex ante forecasting and simulations. The model was constructed to complement the JD; it has the same aggregative details and uses the same set of assumptions. For details of the model and its ex post and ex ante performance, see Heilemann (2004). The model produces higher forecasting accuracy than that produced by time series models, which are often used as benchmarks (Döhrn, Kitlinski and Münch, 2009). Of course, other models and other types of models could also be used. It would be interesting whether they come to different results. As to applied structural models with a similar closure (the set of predetermined and explained variables) as in the RWI-model we speculate that the outcome would not differ too much from the present results.

We used this model to produce static ex post forecasts for each of the years from 1980 to 1989. Each forecast was based on the actual values of all predetermined variables. As an example, based on the sample period 1971-I to 1980-IV (the model estimation uses a 40-quarter-moving sample) data up to the first half of 1979 were used to forecast the second half of that year and all of 1980. Hence the year ahead forecast was the result of six consecutive static forecasts for the current year's third and fourth quarters and for each quarter of the next full year.

The model's MAE for the growth forecasts for the period 1980 to 1989 was 0.6 percentage points, though the results seem to have benefitted from some small aggregations

gains on the demand side.[5] For inflation, the MAE was 0.4 percentage points, which is very similar to the ex ante errors of the four forecasting institutions analysed here. This suggests that the inflation forecasts for this period achieved the highest possible level of accuracy. However, like all other accuracy statements in an interdependent environment (which carry the risk that the figures are "right for the wrong reason"), this result should be taken with a grain of salt. Unsurprisingly, the model was substantially more accurate in predicting the rate of growth. It made no turning point errors, and its errors were only about 60% as large as those made by the four organizations. If the model's errors are seen to represent some kind of maximum accuracy attainable given the state of macroeconomics, modelling, and forecasting in the 1980s, which still holds now, we provide the following interpretation: The accuracy of ex ante growth forecasts can still be considerably improved. However, such improvement requires a perfect forecast of world trade, interest rates, government investment outlays, and the price of oil (to name some of the macroeconomic assumptions used in the forecasts for Germany), as well as stable economic reactions.

5. Summary and conclusions

Quantitative macroeconomic forecasting for Germany is only modestly accurate. For the last 45 years the absolute sizes of the errors in predicting both growth and inflation have been about 1.3 and 0.7, respectively. These results are similar to previous findings for Germany as well as the U.S. and the U.K.

More disappointingly, better theories, data, and methods have not appeared to offer substantial improvement. While the accuracy of growth forecasts improved somewhat in the 1980s and 1990s, it deteriorated in the past decade, returning to the levels of the 1970s. This trend can also be observed in OECD forecasts for the G7 countries. Even worse, temporary improvements seem to be mainly due to the decline of the variance of

---

5 Though the present paper does not aim to compare different methods, it should be noted that the model's ex ante performance for this period was unusually impressive. The MAE of its forecasts for growth (inflation) was 0.7 (0.4), while with actual figures for the exogenous variables this figure was 0.2 (0.1). For 1980/1990 results were 0.9 (0.5), and for 1990/1999 1.2 (0.6). With actual figures for the exogenous variables the corresponding results were 0.3 (0.2) and 0.3 (0.2). These results included considerable add factoring, dynamic errors, etc., and therefore cannot serve as benchmarks for the model's accuracy.

growth and inflation (and their levels). Improvements in directional accuracy have been even slighter. With few exceptions, the recessions in 1975, 1981/82, 1993, 2001 and most of the 2008ff. crisis were seen only after the fact. The upswings in the late 1960s, early 1990s, and 2010 were similarly not predicted. Consequently, we should hold only modest expectations of sizeable or continuous improvements in forecasting accuracy in the near future.

As the macroeconometric model exercises suggest, there is room for improvement, though it would be interesting to see how much this differs according to model time. But these exercises also demonstrate that this margin is much smaller than often assumed, and improvement will be hard to achieve. To realize improvement, forecasters should start with detailed forecast evaluations determining, tracing, categorizing, and publicizing forecast errors. Are errors the result of faulty assumptions, misleading theories, empirical irregularities, insufficient data, etc.? As little is known about the forecasting techniques and their operation by the four institutions examined here, forecasters themselves must carry out these evaluations. Third parties, who have carried out the bulk of forecast evaluation so far, are not in a position to accomplish such investigations. Macroeconomic forecasts such as those we have examined here are conjoined, multivariate, and multi-period forecasts – examining only forecasts of highly aggregated variables such as GDP and inflation would be insufficient.

References

Antholz, B. (2005), Zur Treffsicherheit von Wachstumsprognosen. Dissertation, EHA-Verlag, Münster.

Arbeitsgemeinschaft deutscher wirtschaftswissenschaftlicher Forschungsinstitute e.V., Essen (1967ff), Die Lage der Weltwirtschaft und der deutschen Wirtschaft im Herbst 1966 ff. Arbeitsgemeinschaft, Essen.

Batchelor, R. (1990), All forecasters are equal. Journal of Business and Economics Statistics 8, 143-144.

Blix, M. and K. Wadefjord, U. Wernicke and M. Adahl (2001), How good is the forecasting performance of major institutions? Sveriges Riksbank Economic Review 3, 38-68.

Bundesregierung (1967ff.), Jahreswirtschaftsbericht der Bundesregierung, Hegner, Bonn, after 1994: Bundesanzeiger-Verlag, Köln.

Burns, T. (1986), The interpretation and use of economic predictions, Proceedings of the Royal Society, Series A 407, 103-125.

Dicke, H. and H. G. Glisman (2002), Konjunkturprognosen und wissenschaftlich technischer Fortschritt, Wirtschaftsdienst 83, 167-169.

Diebold, F. X. and J. A. Lopez (1996), Forecast evaluation and combination, NBER Technical Working Paper 192, NBER, Cambridge, MA.

Döhrn, R., T. Kitlinski and H. J. Münch, H. J. (2009), Forecasting accuracy of the RWI short term model compared to time series models, in: A. Wagner (Hrsg.), Empirische Wirtschaftsforschung heute, Poeschel, Stuttgart, 171-181.

Döhrn, R. and C. M. Schmidt (2011), Information or institution – on the determinants of forecast accuracy, Jahrbücher für Nationalökonomie und Statistik 231, 9-27.

Döpke, J. and U. Fritsche (2006), Growth and inflation forecasts in Germany: a panel-based assessment of accuracy and dispersion, Empirical Economics 31, 777-798.

Döpke, J., U. Fritsche and B. Siliverstovs (2009): Evaluating German Business Cycle Forecasts under an asymmetric loss function, KOF Working Paper. KOF, ETH Zürich, Zürich.

Döpke, J. and E. Langfeldt (1995), Zur Qualität von Konjunkturprognosen für Westdeutschland 1976-1994, Kieler Diskussionsbeiträge, 247, IfW, IfW.

Fildes, R. and H. O. Stekler (2002), The state of macroeconomic forecasting, Journal of Macroeconomics 24, 435-468.

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

Heilemann, U. (2004), Besser geht's nicht – Genauigkeitsgrenzen von Konjunkturprognosen, Jahrbücher für Nationalökonomie und Statistik 224, 51–64.

Heilemann, U. and R. Schuhr (2008), Zur Evolution des deutschen Konjunkturzyklus 1958 bis 2004 – Ergebnisse einer dynamischen Diskriminanzanalyse, Jahrbücher für Nationalökonomie und Statistik 228, 84-109.

Kholodilin, K. A. and B. Siliverstovs (2009), Do forecasters inform or reassure? Evaluation of the German real-time data, KOF working paper 215, KOF ETH Zürich, Zürich.

Kreinin, M. (2000), Accuracy of OECD and IMF Projections, Journal of Policy Modeling 22, 161-79.

McNees, S. K. (1988), How accurate are macroeconomic forecasts? New England Economic Review, July/August, 15-26.

Melliss, C. (1997), The Treasury forecasting record: An evaluation, University of Warwick, Warwick.

Merton, R. C. (1981), On market timing and investment performance I: an equilibrium theory of value for market forecasts, Journal of Business 54, 363-406.

Mincer, J. and V. Zarnowitz (1969), The evaluation of economic forecasts, in: J. Mincer (ed.), Economic forecasts and expectations: analysis of forecasting behavior and performance, Columbia University Press for NBER, New York, 14-20.

Öller, L.-E. and B. Barot (2000), The accuracy of European growth and inflation forecasts, International Journal of Forecasting 16, 293-315.

OECD (1967 ff.), Economic outlook, 1ff., OECD, Paris.

Pons, J. (2000), The accuracy of IMF and OECD forecasts for G 7 countries, Journal of Forecasting 19, 53-63.

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (1967ff.), Jahresgutachten 1967/68 ff. Kohlhammer, Mainz, after 1989: Metzler-Poeschel, Stuttgart.

Stekler, H.O. (1991), Macroeconomic forecast evaluation techniques, International Journal of Forecasting 7, 375-384.

Stekler, H. O. (2009), What do we know about G-7 macro forecasts? in: A. Wagner (Hrsg.), Empirische Wirtschaftsforschung heute, Poeschel, Stuttgart, 207- 213.

Timmerman, A. (2007), An evaluation of the World Economic Outlook, IMF Staff Papers 54, 1–33.

Vogel, L., (2007). How do the OECD growth projections for the G7 economies perform? OECD Economics Department Working Papers No. 573, OECD, Paris.

Wild, J. (1974), Grundlagen der Unternehmensplanung, Rowohlt, Reinbek b. Hamburg.

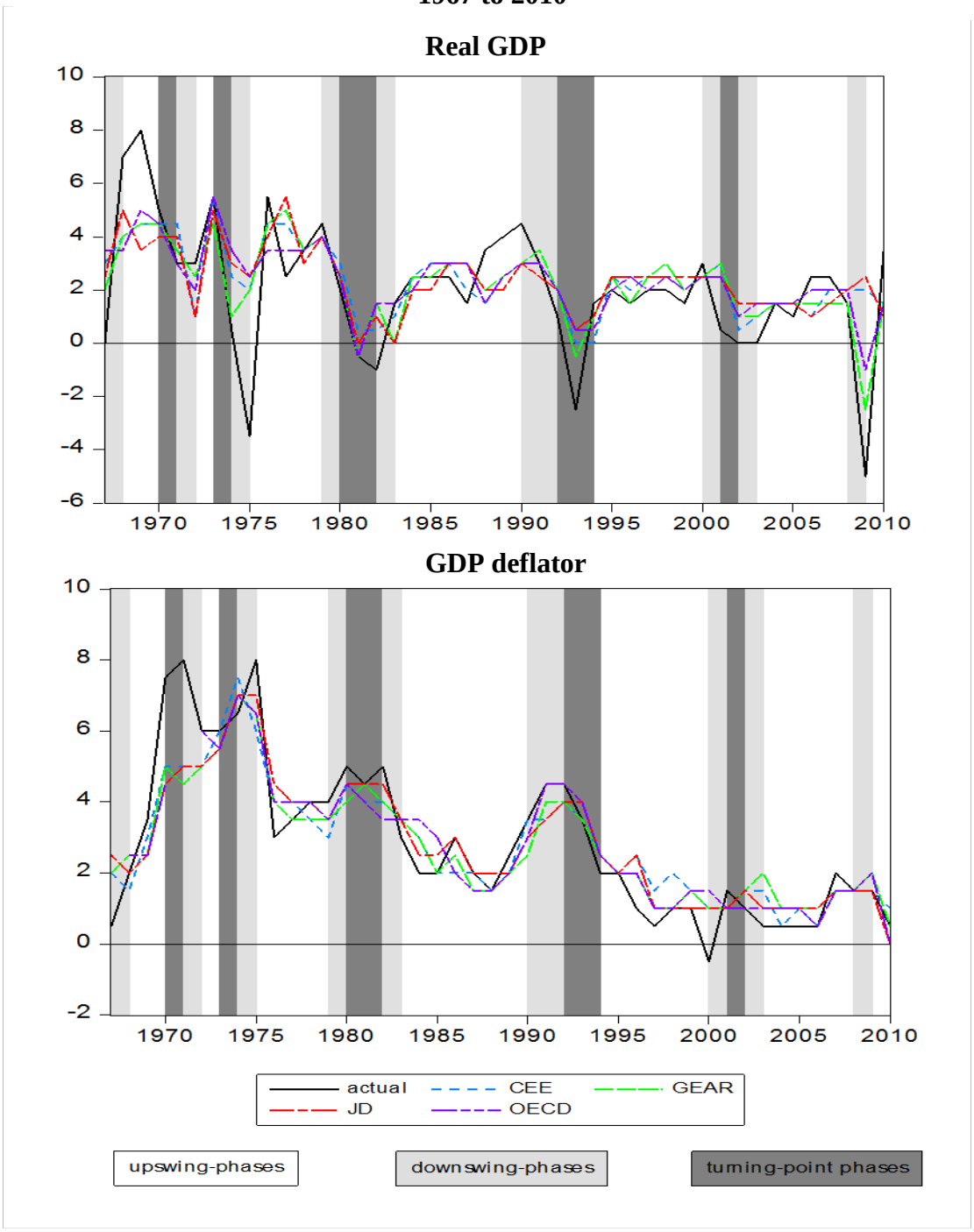
Zarnowitz, V. (1992), Has macroeconomics failed? Cato Journal 12, 129-160.

Ullrich Heilemann, Institut für Empirische Wirtschaftsforschung, Universität Leipzig, Grimmaische Straße 12, 04105 Leipzig, Germany. Phone: +49173 255 1403. E-Mail: heilemann@wifa.uni-leipzig.de. .

Herman .O. Stekler, Department of Economics, George Washington University, Washington DC 20052 USA. Phone + 202-994-1261. E-mail: hstekler@gwu.edu.

Figure 1

Accuracy of forecasts of real GDP and of GDP price deflator for Germany 1967 to 2010



Sources: Federal Statistical Office, JD, CEE, OECD, GEAR and own Computations.
For details of the cyclical classification see Heilemann and Schuhr, 2008.

Table 1

Annual forecasts of percentage changes of real GDP and of GDP price deflator: summary measures of error, 1967 to 2010

	real GDP				GDP price deflator			
	JD	CEE	OECD	GAER	JD	CEE	OECD	GAER
				1967 to 2010				
MAE	1.5	1.3	1.3	1.1	0.6	0.7	0.6	0.7
Bias	0.3	0.3	0.3	0.2	0.0	-0.1	0.0	-0.1
U	0.7	0.6	0.6	0.5	0.6	0.6	0.6	0.6
RMSE/ σ	0.9	0.8	0.7	0.6	0.4	0.4	0.4	0.4
				1970 to 1979				
MAE	1.9	1.5	1.4	1.3	1.2	1.3	0.8	1.2
Bias	0.7	0.7	0.6	0.6	-0.7	-0.8	-0.4	-0.9
U	0.6	0.5	0.6	0.5	0.7	0.7	0.5	0.7
RMSE/ σ	0.9	0.8	0.8	0.7	0.8	0.8	0.7	0.8
				1980 to 1989				
MAE	1.1	0.9	1.0	1.0	0.4	0.5	0.8	0.5
Bias	-0.1	0.1	0.2	0.1	0.1	-0.2	-0.2	-0.2
U	0.8	0.7	0.8	0.8	0.4	0.6	0.9	0.6
RMSE/ σ	0.8	0.7	0.8	0.8	0.3	0.5	0.7	0.5
				1990 to 1999				
MAE	1.0	0.9	0.9	0.8	0.5	0.6	0.4	0.5
Bias	0.5	0.3	0.4	0.4	0.1	0.3	0.3	0.1
U	0.7	0.6	0.6	0.5	0.8	0.9	0.6	0.7
RMSE/ σ	0.7	0.6	0.7	0.6	0.4	0.5	0.3	0.4
				2000 to 2010				
MAE	1.7	1.5	1.2	1.1	0.5	0.5	0.5	0.6
Bias	0.7	0.7	0.5	0.3	0.2	0.3	0.2	0.4
U	0.8	0.7	0.5	0.4	0.6	0.7	0.7	0.8
RMSE/ σ	1.1	1.0	0.7	0.6	0.9	0.9	1.0	1.1

Authors´ computations. – For abbreviations and computation of the error measures, see text.

Table 2

Annual forecasts of percentage changes of real GDP and of GDP price deflator: summary measures of directional errors, 1968 to 2010

	real GDP				GDP price deflator			
	JD	CEE	OECD	GAER	JD	CEE	OECD	GAER
				1968 to 2010				
Overestimates	11	9	7	10	11	10	11	11
Underestimates	21	22	19	18	17	17	16	18
Turning p. errors	7	4	5	5	2	3	3	2
Coincidences	0	4	8	6	10	10	9	9
Not defined	4	4	4	4	3	3	4	3
				1970 to 1979				
Overestimates	3	3	2	3	3	3	3	2
Underestimates	6	4	4	5	6	6	4	7
Turning. p. errors	1	1	1	1	0	0	0	0
Coincidences	0	2	3	1	1	1	2	1
Not defined	0	0	0	0	0	0	1	0
				1980 to 1989				
Overestimates	2	1	1	1	3	2	4	2
Underestimates	5	6	4	3	4	4	4	4
Turning p. errors	2	1	2	2	0	1	1	1
Coincidences	0	1	2	3	3	3	1	3
Not defined	1	1	1	1	0	0	0	0
				1990 to 1999				
Overestimates	5	4	2	4	3	3	1	3
Underestimates	3	5	4	4	2	2	4	3
Turning pt. errors	2	0	1	1	1	1	0	0
Coincidences	0	1	3	1	3	3	4	3
Not defined	0	0	0	0	1	1	1	1
				2000 to 2010				
Overestimates	1	1	2	2	2	2	2	3
Underestimates	5	5	5	4	4	3	3	3
Turning pt. errors	2	2	1	1	1	1	2	1
Coincidences	0	0	0	1	2	3	2	2
Not defined	3	3	3	3	2	2	2	2

Authors' computations. For sources and computation see text. – [1] $Q_t = (p_t - a_{t-1}) / (a_t - a_{t-1})$, p_t a_t: forecast, actual, rates of change, with: $Q_t > 1$: overestimate; $0 \leq Q_t < 1$: underestimate; $Q_t < 0$: turning point estimate; $Q_t = 1$: coincidence, for $p_t = a_{t-1}$ the measure is not defined (for details see Heilemann, 1998).

Table 3

Accuracy of forecasts of directional change of real GDP growth and of GDP price deflator for Germany 1968 to 2010

	JD					CEE					OECD					GAER				
	IC	AC	AW	DC	DW	IC	AC	AW	DC	DW	IC	AC	AW	DC	DW	IC	AC	AW	DC	DW
										Real GDP										
1968 to 2010	1,2	9	7	18	9	1,6	14	4	21	4	1,4	11	5	19	8	1,5	14	7	18	4
1970 to 1979	1,4	3	2	4	1	1,9	3	0	6	1	1,9	3	0	6	1	1,4	3	2	4	1
1980 to 1989	0,9	1	2	4	3	1,6	3	1	5	1	1,2	2	2	4	2	1,2	2	2	4	2
1990 to1999	1,9	3	0	6	1	1,9	3	0	6	1	1,2	2	2	4	2	1,7	4	2	4	0
2000 to 2010	0,8	1	3	4	3	1,3	3	3	4	1	1,5	3	1	5	2	1,6	3	1	6	1
										GDP price deflator										
1968 to 2010	1,4	10	4	21	8	1,3	8	5	22	8	1,5	7	2	23	8	1,3	9	7	20	7
1970 to 1979	1,3	3	1	3	3	1,2	2	1	4	3	1,5	2	0	3	3	1,0	2	2	3	3
1980 to 1989	1,8	2	0	6	2	1,8	2	0	6	2	1,8	2	0	6	2	1,6	3	1	5	1
1990 to 1999	1,3	2	2	5	1	1,3	2	2	5	1	1,5	2	1	6	1	1,5	2	1	6	1
2000 to 2010	1,5	2	1	7	1	1,2	1	2	7	1	1,4	1	1	8	1	1,1	1	3	6	1

Authors' computations. For sources and computations, see text. – AC (AW): acceleration correctly (wrongly) forecast, DC (DW): deceleration correctly (wrongly) forecast, IC: informational content.

Table 4

OECD annual growth and inflation forecasts: summary measures of error, 1967 to 2010

		Canada	France	Germany	Italy	Japan	UK	US	G7	OECD
					1967 to 2010					
MAE	a	1.3	0.9	1.4	1.4	1.7	1.0	1.0	0.8	0.9
	b	1.3	0.7	0.6	1.3	1.5	1.4	0.8	0.6	0.7
Bias	a	0.3	0.2	0.2	0.4	0.1	0.1	0.0	-0.1	0.1
	b	-0.1	-0.1	-0.1	-0.8	0.3	-0.5	-0.1	-0.2	-0.3
U	a	0.6	0.6	0.6	0.7	0.8	0.6	0.5	0.5	0.5
	b	0.8	0.6	0.6	0.9	0.9	0.6	0.8	0.5	0.9
RMSE/σ	a	0.8	0.6	0.8	0.8	0.6	0.7	0.6	0.5	0.7
	b	0.5	0.2	0.4	0.3	0.6	0.4	0.4	0.5	0.4
					1970 to 1979					
MAE	a	1.2	1.1	1.5	2.7	2.3	1.4	1.0	0.4	1.1
	b	1.8	1.2	0.9	3.4	3.6	3.4	1.5	0.4	1.4
Bias	a	0.6	0.2	0.5	0.4	0.8	0.8	0.7	-0.1	0.5
	b	-1.4	-0.8	-0.6	-2.9	-1.0	-2.7	-0.8	-0.1	-1.1
U	a	0.6	0.6	0.6	0.7	0.8	0.7	0.4	0.2	0.5
	b	0.9	0.6	0.6	1.0	0.8	0.6	0.7	0.2	0.9
RMSE/σ	a	0.8	0.8	0.8	1.1	0.9	1.0	0.5	0.2	0.7
	b	0.8	0.5	0.7	0.8	1.0	0.7	0.8	0.2	0.9
					1980 to 1989					
MAE	a	1.5	0.7	1.1	0.9	1.1	1.0	1.0	1.0	0.7
	b	1.0	0.5	0.6	1.4	1.1	1.1	0.8	0.8	0.6
Bias	a	-0.2	-0.2	-0.1	-0.2	-0.1	-0.8	-0.6	-0.7	-0.4
	b	0.2	0.0	-0.3	-0.9	0.8	-0.2	0.8	-0.5	0.3
U	a	0.6	0.8	0.9	0.6	0.8	0.6	0.4	0.6	0.6
	b	0.7	0.4	0.7	0.6	1.5	0.5	0.8	0.5	0.8
RMSE/σ	a	0.8	0.9	0.8	0.6	1.0	0.5	0.5	0.7	0.7
	b	0.4	0.2	0.6	0.3	1.3	0.3	0.4	0.6	0.3
					1990 to 1999					
MAE	a	1.5	0.9	1.0	0.9	1.4	1.1	1.3	0.6	0.7
	b	1.0	0.5	0.5	0.8	0.5	0.3	0.3	0.6	0.5
Bias	a	0.8	0.4	0.1	0.7	0.7	0.4	-0.4	-0.2	0.1
	b	1.0	0.5	0.2	-0.2	0.4	0.0	0.1	-0.2	-0.4
U	a	0.8	0.7	0.7	0.8	0.8	0.9	0.9	0.7	0.8
	b	1.2	1.2	0.6	0.8	0.9	0.7	0.9	0.7	1.5
RMSE/σ	a	0.8	0.9	0.8	1.1	0.8	0.8	0.9	0.9	1.0
	b	1.2	0.6	0.3	0.5	0.6	0.2	0.4	0.9	1.1
					2000 to 2010					
MAE	a	1.1	0.7	1.3	1.2	1.6	0.8	1.0	1.1	1.0
	b	1.0	0.4	0.4	0.4	0.8	0.6	0.5	0.6	0.4
Bias	a	0.3	0.6	0.4	0.9	0.3	0.4	0.4	0.4	0.4

	b	-0.5	0.0	0.1	-0.2	0.7	-0.1	-0.3	0.0	-0.2
U	a	0.6	0.6	0.5	0.6	0.6	0.5	0.6	0.6	0.5
	b	0.4	0.7	0.7	0.7	1.4	0.9	1.0	0.9	1.0
RMSE/σ	a	0.7	0.7	0.7	0.8	0.8	0.6	0.7	0.7	0.7
	b	0.7	0.8	0.9	0.6	1.7	1.2	0.9	1.1	1.0

Authors´ computations. For sources, abbreviations and computation of the error measures, see text. – a) real GDP; b) GDP deflator.

Appendix

Table 5
Forecasts and actual data, 1967 to 2010

	real GDP					GDP-Deflator				
	JD	CEE	OECD	GAER	Actual	JD	CEE	OECD[1]	GAER	Actual
1967	2.5	3.0	3.5	2.0	0.0	2.5	2.0	.	2.0	0.5
1968	5.0	4.0	3.5	4.0	7.0	2.0	1.5	2.5	2.5	2.0
1969	3.5	4.5	5.0	4.5	8.0	2.5	3.0	2.5	2.5	3.5
1970	4.0	4.5	4.5	4.5	5.0	4.5	5.0	4.5	5.0	7.5
1971	4.0	4.5	3.0	3.5	3.0	5.0	5.0	.	4.5	8.0
1972	1.0	1.0	2.0	2.5	3.0	5.0	5.0	6.0	5.0	6.0
1973	5.0	5.5	5.5	4.5	5.5	5.5	6.0	5.5	5.5	6.0
1974	3.0	2.5	3.5	1.0	0.5	7.0	7.5	7.0	7.0	6.5
1975	2.5	2.0	2.5	2.0	-3.5	7.0	6.0	6.5	6.5	8.0
1976	4.0	4.5	3.5	4.5	5.5	4.5	4.0	4.0	4.0	3.0
1977	5.5	4.5	3.5	5.0	2.5	4.0	4.0	4.0	3.5	3.5
1978	3.0	3.5	3.5	3.5	3.5	4.0	3.5	4.0	3.5	4.0
1979	4.0	4.0	4.0	4.0	4.5	3.5	3.0	3.5	3.5	4.0
1980	2.5	3.0	2.5	2.5	2.0	4.5	4.5	4.5	4.0	5.0
1981	0.0	0.5	-0.5	-0.5	-0.5	4.5	4.0	4.0	4.5	4.5
1982	1.0	0.5	1.5	1.5	-1.0	4.5	4.0	3.5	4.0	5.0
1983	0.0	1.0	1.5	0.0	1.5	3.5	3.5	3.5	3.5	3.0
1984	2.0	2.5	2.0	2.5	2.5	2.5	3.0	3.5	3.0	2.0
1985	2.0	3.0	3.0	2.5	2.5	2.5	2.0	3.0	2.0	2.0
1986	3.0	3.0	3.0	3.0	2.5	3.0	2.0	2.0	2.5	3.0
1987	3.0	2.0	3.0	3.0	1.5	2.0	2.0	1.5	1.5	2.0
1988	2.0	1.5	1.5	2.0	3.5	2.0	1.5	1.5	1.5	1.5
1989	2.0	2.5	2.5	2.5	4.0	2.0	2.0	2.0	2.0	2.5
1990	3.0	3.0	3.0	3.0	4.5	3.0	3.5	3.0	2.5	3.5
1991	2.5	3.0	3.0	3.5	3.0	3.5	3.5	4.5	4.0	4.5
1992	2.0	2.0	2.0	2.0	1.0	4.0	4.0	4.5	4.0	4.5
1993	0.5	0.0	0.5	-0.5	-2.5	4.0	3.5	4.0	3.5	3.5
1994	1.0	0.0	0.5	1.0	1.5	2.5	2.5	2.5	2.5	2.0
1995	2.5	2.5	2.0	2.5	2.0	2.0	2.0	2.0	2.0	2.0
1996	2.5	2.0	2.5	1.5	1.5	2.5	2.5	2.0	2.0	1.0
1997	2.5	2.5	2.0	2.5	2.0	1.0	1.5	1.0	1.0	0.5
1998	2.5	2.5	2.5	3.0	2.0	1.0	2.0	1.0	1.0	1.0
1999	2.5	2.0	2.0	2.0	1.5	1.0	1.5	1.5	1.5	1.0
2000	2.5	2.5	2.5	2.5	3.0	1.0	1.0	1.5	1.0	-0.5
2001	2.5	3.0	2.5	3.0	0.5	1.0	1.0	1.0	1.0	1.5
2002	1.5	0.5	1.0	1.0	0.0	1.5	1.5	1.0	1.5	1.0
2003	1.5	1.0	1.5	1.0	0.0	1.0	1.5	1.0	2.0	0.5
2004	1.5	1.5	1.5	1.5	1.5	1.0	0.5	1.0	1.0	0.5

Table 4, continued

	real GDP					GDP-Deflator				
	JD	CEE	OECD	GAER	actual	JD	CEE	OECD[1]	GAER	actual
2005	1.5	1.5	1.5	1.5	1.0	1.0	1.0	1.0	1.0	0.5
2006	1.0	1.0	2.0	1.5	2.5	1.0	0.5	0.5	1.0	0.5
2007	1.5	2.0	2.0	1.5	2.5	1.5	1.5	1.5	1.5	2.0
2008	2.0	2.0	2.0	1.5	1.5	1.5	1.5	1.5	1.5	1.5
2009	2.5	2.0	-1.0	-2.5	-5.0	1.5	1.5	2.0	2.0	1.5
2010	1.0	1.5	1.5	1.5	3.5	0.0	1.0	0.0	0.5	0.5

Sources: Arbeitsgemeinschaft, 1966ff.; Sachverständigenrat, 1966ff., OECD, 1966ff.; Bundesregierung 1967ff.,. All numbers rounded.

Supplementary material on the GER website

Table 1, GER Supplement
Correlation of major institutions' forecasts for Germany, 1967 to 2010

		JD	CEE	OECD	GAER
JD	1967 to 2010	-	0.96	0.96	0.97
	1970 to 1979	-	0.92	0.94	0.96
	1980 to 1989	-	0.92	0.83	0.95
	1990 to 1999	-	0.95	0.96	0.95
	2000 to 2010	-	0.50	0.81	0.72
CEE	1967 to 2010	0.90	-	0.95	0.96
	1970 to 1979	0.92	-	0.92	0.96
	1980 to 1989	0.80	-	0.92	0.95
	1990 to 1999	0.95	-	0.93	0.91
	2000 to 2010	0.75	-	0.53	073
OECD	1967 to 2010	0.76	0.84	-	0.97
	1970 to 1979	0.72	0.80	-	0.95
	1980 to 1989	0.83	0.81	-	0.90
	1990 to 1999	0.92	0.95	-	0.99
	2000 to 2010	-0.11	0.28	-	0.69
GAER	1967 to 2010	0.76	0.80	0.87	-
	1970 to 1979	0.74	0.80	0.55	-
	1980 to 1989	0.96	0.80	0.89	-
	1990 to 1999	0.87	0.92	0.86	-
	2000 to 2010	-0.12	0.25	0.97	-

Author's computations, for sources see text.

Table 2, GER Supplement

Annual forecasts of percentage changes of real GDP and of GDP price deflator: summary measures of error cycles, 1967 to 2010

	real GDP				GDP price deflator			
	JD	CEE	OECD	GAER	JD	CEE	OECD[1]	GAER
				1967 to 2010				
MAE	1.5	1.3	1.3	1.2	0.7	0.7	0.6	0.7
BIAS	0.3	0.3	0.3	0.2	-0.0	-0.1	-0.0	-0.1
U	0.7	0.6	0.6	0.5	0.6	0.6	0.6	0.6
RMSE/ σ	0.9	0.8	0.7	0.7	0.4	0.4	0.4	0.4
				1967 – 1974				
MAE	2	1.9	1.8	1.4	1.4	1.3	0.9	1.4
BIAS	-0.5	-0.3	-0.2	-0.7	-0.8	-0.6	-0.6	-0.8
U	0.5	0.5	0.4	0.3	0.3	0.3	0.2	0.3
RMSE/ σ	0.8	0.8	0.8	0.6	0.7	0.6	0.7	0.6
				1975 – 1982				
MAE	1.8	1.6	1.6	1.6	0.6	0.9	0.8	0.7
BIAS	1.2	1.2	0.9	1.2	-0.1	-0.5	-0.4	-0.4
U	0.8	0.7	0.7	0.7	0.2	0.2	0.2	0.2
RMSE/ σ	0.8	0.7	0.8	0.8	0.5	0.7	0.6	0.6
				1983 – 1994				
MAE	1.2	1	1.1	1	0.5	0.4	0.5	0.5
BIAS	-0.2	-0.2	0	-0.1	0	-0.1	0.1	-0.1
U	0.5	0.5	0.5	0.4	0.2	0.2	0.2	0.2
RMSE/ σ	0.8	0.7	0.8	0.7	0.5	0.6	0.7	0.6
				1995 – 2001				
MAE	0.9	0.8	0.6	0.8	0.6	0.9	0.6	0.6
BIAS	0.7	0.6	0.5	0.6	0.4	0.7	0.5	0.4
U	0.5	0.5	0.5	0.6	0.7	0.9	0.8	0.6
RMSE/ σ	1.2	1.4	1.2	1.4	1.1	1.3	1.2	1
				2002 – 2010				
MAE	1.8	1.5	1.2	1	0.4	0.3	0.3	0.5
BIAS	0.7	0.6	0.5	0.1	0.2	0.2	0.1	0.4
U	1.1	1	0.7	0.5	0.4	0.4	0.4	0.6
RMSE/ σ	1.1	1	0.7	0.5	0.8	0.8	0.7	1.1

Authors´ computations. For cycle demarcation see Heilemann, Schuhr 2008 (updated), for abbreviations and computation of the error measures, see text. – [1] Missing value for 1967 and 1971

Table 3, GER supplement

Annual forecasts of percentage changes of real GDP and of GDP price deflator: summary measures of error for two cycles, 1967 to 2010

	real GDP				GDP price deflator			
	JD	CEE	OECD	GAER	JD	CEE	OECD[1]	GAER
				1967 to 2010				
MAE	1.5	1.3	1.3	1.2	0.7	0.7	0.6	0.7
BIAS	0.3	0.3	0.3	0.2	-0.0	-0.1	-0.0	-0.1
U	0.7	0.6	0.6	0.5	0.6	0.6	0.6	0.6
RMSE/ σ	0.9	0.8	0.7	0.7	0.4	0.4	0.4	0.4
				1967 – 1974				
MAE	2	1.9	1.8	1.4	1.4	1.3	0.9	1.4
BIAS	-0.5	-0.3	-0.2	-0.7	-0.8	-0.6	-0.6	-0.8
U	0.5	0.5	0.4	0.3	0.3	0.3	0.2	0.3
RMSE/ σ	0.8	0.8	0.8	0.6	0.7	0.6	0.7	0.6
				1975 – 1982				
MAE	1.8	1.6	1.6	1.6	0.6	0.9	0.8	0.7
BIAS	1.2	1.2	0.9	1.2	-0.1	-0.5	-0.4	-0.4
U	0.8	0.7	0.7	0.7	0.2	0.2	0.2	0.2
RMSE/ σ	0.8	0.7	0.8	0.8	0.5	0.7	0.6	0.6
				1983 – 1994				
MAE	1.2	1	1.1	1	0.5	0.4	0.5	0.5
BIAS	-0.2	-0.2	0	-0.1	0	-0.1	0.1	-0.1
U	0.5	0.5	0.5	0.4	0.2	0.2	0.2	0.2
RMSE/ σ	0.8	0.7	0.8	0.7	0.5	0.6	0.7	0.6
				1995 – 2001				
MAE	0.9	0.8	0.6	0.8	0.6	0.9	0.6	0.6
BIAS	0.7	0.6	0.5	0.6	0.4	0.7	0.5	0.4
U	0.5	0.5	0.5	0.6	0.7	0.9	0.8	0.6
RMSE/ σ	1.2	1.4	1.2	1.4	1.1	1.3	1.2	1
				2002 – 2010				
MAE	1.8	1.5	1.2	1	0.4	0.3	0.3	0.5
BIAS	0.7	0.6	0.5	0.1	0.2	0.2	0.1	0.4
U	1.1	1	0.7	0.5	0.4	0.4	0.4	0.6
RMSE/ σ	1.1	1	0.7	0.5	0.8	0.8	0.7	1.1

Authors´ computations. For cycle demarcation see Heilemann and Schuhr, 2008 (updated), for abbreviations and computation of the error measures, see text. – [1] Missing value for 1967 and 1971