Evaluation and accuracy of economic forecasts

Groemling, Michael

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

gesis
Leibniz-Institut
für Sozialwissenschaften

Mitglied der

Leibniz-Gemeinschaft

Evaluation and Accuracy of Economic Forecasts

Michael Groemling[*]

Abstract: The economic forecasts for 2001 grossly missed reality. While forecasters estimated a growth rate of 3 per cent, real GDP actually grew by only 0.6 per cent. The analysis of forecasts in the years 1995 to 2001 shows nevertheless that they were fairly accurate. In addition, the article gives several arguments which may explain forecast errors: data revisions, unpredictable events, behavioural and political feedback and imitation behaviour of forecasters.

Justifying economic forecasts for the year 2001 brought a great deal of trouble with it. At the beginning of 2001 economists down the line expected real Gross Domestic Product (GDP) in Germany to grow by around 3 per cent. In the course of the year all forecasts were revised downwards. Finally, real GDP grew only by 0.6 per cent.

Such forecast errors can have far-reaching consequences for entrepreneurial decisions, economic policy and wage bargaining. For example, expected economic growth is a component of productivity forecasts. Productivity growth in turn determines the bargaining margin. Thus, forecast errors can be misleading for wage settlements, and this again may have severe consequences for future labour demand. Table 1 shows to what extent the expected productivity growth for Germany in 2000 and 2001 diverged from the actual one according to the forecasts by the German Council of Economic Experts (Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung, SVR), the Association of Economic Research Institutes (Arbeitsgemeinschaft deutscher wirtschaftswissenschaftlicher Forschungs-institute, Institute) and the Cologne Institute for Business Research (Institut der deutschen Wirtschaft Köln, IW).

---

[*] Address all communications to Michael Groemling, Institut der deutschen Wirtschaft Köln, Gustav-Heinemann-Ufer 84-88, D-50968 Köln. E-Mail: groemling@iwkoeln.de.

Table 1: Productivity Forecasts for Germany

Real GDP per employee
Percentage change from previous period

	Forecast			Actual Value
	SVR	Institute	IW	
2000	2,4	2,4	2,0	1,4
2001	1,8	1,7	1,7	0,4
Annual Average	2,1	2,1	1,9	0,9

SVR: German Council of Economic Experts
Institute: Association of Economic Research Institutes
IW: Cologne Institute for Business Research

Source: Groemling, 2002

The shortcomings of productivity forecasts for the last two years can be explained to a large degree by errors in forecasting real GDP growth (Groemling, 2002). For the years 2000 and 2001, growth predictions ran up to an annual average of 2.8 per cent. Finally, the actual growth rate was 1 percentage point lower.

This article will show the purpose of economic forecasts, the accuracy of the forecasts by the IW and, finally, what justifiable reasons there are for forecast errors.

The meaning and purpose of economic forecasting is to gain information about the future development of an economy. Forecasts generally cover the development in the current and following year.

- On the one hand, predictions are expected to gauge the magnitude of the future growth rate, in particular the growth rate of real GDP.
- On the other hand, forecasts are expected to state whether there will be a cyclical turning point. That means that it should foresee a recession or a recovery.

In evaluating the predictive performance of forecasts one has to keep in mind that economic predictions are as-if statements. They are based on certain assumptions about future economic and political conditions and data. Such forecasts are labelled conditional predictions. The predicted path of development is determined by the level of information at the time of forecasting. Forecasts are efficient or rational, provided that all relevant and available information are used (Nordhaus, 1987). Contrary to scenarios, economic forecasts only

point out that upcoming growth path with the highest probability. Again one has to keep in mind that the purpose of a forecast is to reduce uncertainty, not to eliminate it.

Modern forecasting techniques differ from ancient prophecies in so far as they are based on scientific methods and that they can be traced and can be evaluated by other forecasters or, in particular, the users. Everyone who knows the relevant forecasting techniques and who is provided with the necessary data should be able to evaluate a certain prediction. The purpose of a forecast, the available data, the computing capacities and, of course, the preferred time consumption to do the forecast determines the kind of method that will be chosen. But to choose one method does not exclude the application of others – for example, to check results or to support parts of the forecast. Aside from econometric methods, such as simple regression models, structural models and time-series models, most forecasters still use so-called iterative forecasting techniques.

The annual forecasts by the IW are produced in September for the following year and there are occasional revisions in the course of the following year. The Association of Economic Research Institutes (Institute) publishes ist semi-annual forecasts in April and October, and the German Council of Economic Experts (SVR) publishes its annual forecast in November. The underlying method of the IW-forecasts is the iterative forecasting technique, or the formal SNA forecasting technique, where SNA stands for System of National Accounts. It is based on the three approaches to analyse GDP: the expenditure, the output and the income approach. The expenditure approach reflects GDP by type of expenditure (e.g. private consumption, gross fixed capital formation, exports, imports). The output approach views the total gross value added by industries (e.g. manufacturing, construction, services), and the income approach covers compensation of employees, property and entrepreneurial income. The first step is a forecast for each of these approaches. The second step is to bring the single forecasts into line. In order to get a consistent forecast of GDP, it is necessary to adjust the single forecasts. By means of the iterative forecasting technique the final prediction will be attained by moving step by step towards a consistent as well as the most probable version.

One advantage of this forecasting method is its flexibility and relevance for reality. New information and special events (e.g. terror attacks, election outcomes, regional crises in Asia or Russia) can be taken into account immediately. Special forecasts by experts – forecasts for certain types of expenditures or industries – and the results of business surveys or consultations can be included. But this way of forecasting is said to be intransparent for others and prone to manipulations. This disadvantage may be diminished by explaining the underlying assumptions, conjectures and conditions, what makes the forecasted values and conclusions more comprehensive. Other forecasting techniques allow an econometric control at any time.

244

After this short glance at the forecasting technique of the IW, we will now take a look at the forecasting accuracy since 1995. Figure 1 shows the predicted and actual growth rates of real GDP for Germany. The actual values are final rates (except for 2001). From 1999 onwards they are based on the revised European System of Accounts (ESA 1995). The forecasts for 1999 were based on the unrevised SNA, but actual values are only available from the revised version.

Figure 1 : IW-Forecasts and Actual Values

real GDP, percentage change from previous period,
Germany 1995 to 2001
forecasts (square, with deviations of +/- 0.25 percentage points)
and actual values (rhombus)
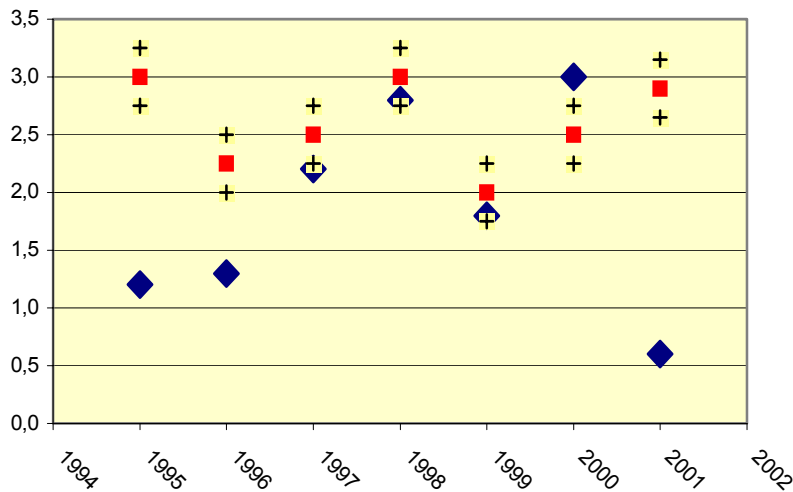
According to figure 1, there were only three major discrepancies (1995, 1996, 2001) between the predicted and the actual values in the seven covered years. In three years (1997, 1998, 1999) the forecasts are well within the scope of tolerance. Normally a discrepancy of a quarter of a percentage point in both directions is in accordance with an accurate forecast. Figure 2 also shows the forecasts by the Institute. The IW-forecasts compare well with those of the Institute. In 1995, 1996 and 2001 both institutions missed reality. Even in the successful forecasting years there are no pronounced differences.

245

Figure 2 : Forecasts by IW and Institute and Actual Values

real GDP, percentage change from previous period, Germany
1995 to 2001
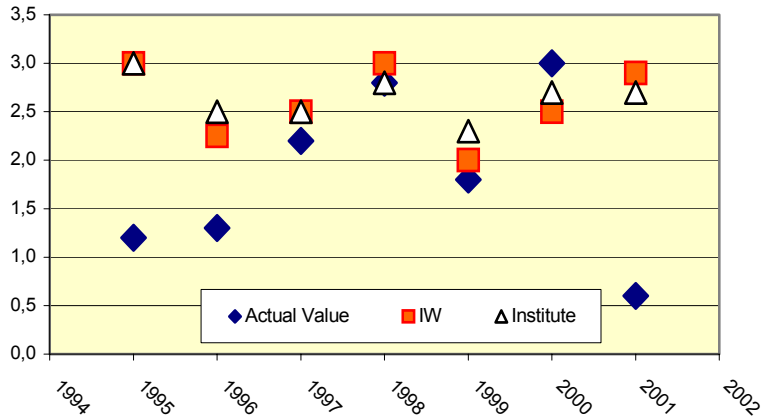
The accuracy of the IW-forecasts is also illustrated in the scatter diagram relating predictions to actual values (Mincer/Zarnowitz, 1969). Points on the 45°-line, the line of perfect forecasts, mean that forecasted and realized values are identical. A regression $R = a + b*P$, relating expected values (P) to realized

Figure 3 : Prediction-Realization Diagram

IW-forecasts for real GDP, percentage change from previous period, Germany 1995 to 2001
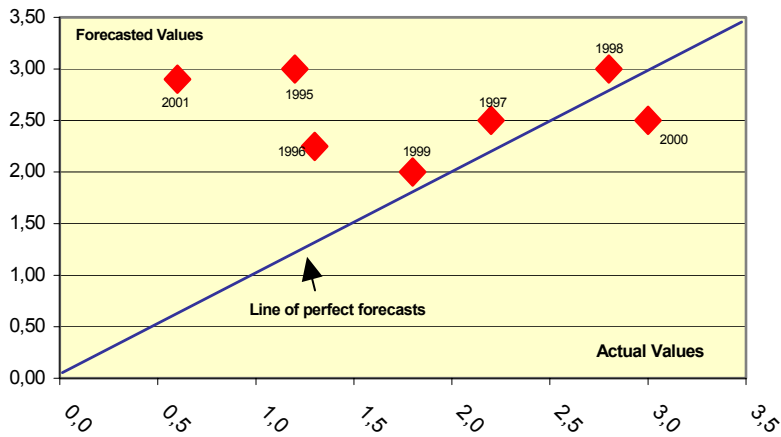
values (R), would have the following parameters in the case of a perfect forecast: a = 0 and b = 1. Such a regression line would be equivalent with the line of perfect forecasts (45°-line). Looking at the forecasts for Germany in 1995 up to 2001 there is a pronounced gap to those values: a = 2.4 and b = -0.2.

The prediction-realization diagram shows once again, which forecasts are more or less accurate. Deviations above (below) the 45°-line mean that the growth rate of GDP was overestimated (underestimated). Thus the IW-forecasts show optimistic tendencies. With one exception the forecasted values were always higher than the actual ones. Moreover, according to figure 1, 2 and 3, the predicted values fluctuate less than the actual outcomes. In other words, the forecasts were smoother than reality. In the period 1995 to 2001 the actual values ranged from 0.6 to 3 per cent, whereas the predicted values covered a range from 2 to 3 per cent only. A value for the parameter a in the regression of 2.4 suggests, that the forecasts for these years expected an average value of this magnitude.

In addition to this graphical analysis there are several measures to evaluate the accuracy of forecasts (Wallis, 1989; Stekler, 1991; Doepke/Langfeldt, 1995):

1) Mean error: $ME = (1/n) \Sigma(P-R)$, where n is the number of observations (years), P is the predicted and R is the realized value. The mean error takes into account the sign of the annual forecasting error. It can be used to evaluate annual outcomes according to figure 1, 2 and 3. But this measure is not useful in determining the forecasting accuracy for more than one year, because positive and negative deviations compensate each other. Therefore, the sum of annual deviations allows only limited conclusions.

2) Mean absolute error: $MAE = (1/n) \Sigma|P-R|$. The mean absolute error solves the aggregation problem of the mean error. Positive and negative deviations do not compensate each other. Therefore the mean absolute error can be used to evaluate the sum of single deviations.

3) Mean square error: $MSE = (1/n) \Sigma(P-R)^2$. This also applies to the mean square error. This measure for the predictive performance attributes a higher weight to large forecasting errors.

4) Standardized error: $SE = 100*(MAE/SD)$, where SD is the standard deviation of R. The standardised forecasting error takes into account the fluctuations of the realized values (real GDP). Pronounced fluctuations of the actual outcomes bring about higher forecasting risks and therefore a higher probability of forecasting errors.

Table 2 shows how well the IW did in comparison with other forecasting institutions in the period 1995 to 2001. Moreover, there is additional information for other time periods and for West Germany.

247

Table 2: Measures of Forecasting Accuracy

Forecasting Institutions	Source	Area	Time Period	ME	MAE	MSE	SE
IW	Groemling, 2002	Germany	1995-2001	0,8	0,9	1,2	102
Institute	Groemling, 2002	Germany	1995-2001	0,8	0,9	1,2	101
SVR	Groemling, 2002	Germany	1995-2001	0,7	0,8	1,1	93
For comparison:							
Institute	Hinze, 1996	Germany	1992-1995	0,6	1,3	2,1	73
Institute	Hinze, 1996	West-Germany	1980-1990	-0,1	1,1	1,2	69
SVR	Hinze, 1996	West-Germany	1980-1990	0,1	0,9	1,0	56
SVR	Rinne, 1997	West-Germany	1975-1994	0,1	1,3	2,4	n.a.
SVR	Doepke/Lang-feldt, 1995	West-Germany	1976-1994	-0,1	1,3	n.a.	64
Institute	Doepke/Lang-feldt, 1995	West-Germany	1976-1994	-0,2	1,4	n.a.	70

SVR: German Council of Economic Experts

Institute: Association of Economic Research Institutes

IW: Cologne Institute for Business Research
ME: Mean Error; MAE: Mean Absolute Error; MSE: Mean Square Error; SE: Standardized Error

Source: Groemling, 2002

The mean error, a measure of limited usefulness, indicates no differences between the forecasts for the period 1995 to 2001 in Germany. However, it is higher than in earlier periods in all three forecasts. It must be kept in mind, that the latest period only covers seven years. The mean absolute error also shows no difference for the recent forecasts. It is even lower than for the forecasts for West Germany. Likewise the mean square error shows no pronounced differences. This error is today not larger than in former times. However, the standardised error is larger for the latest forecasts. According to this measure the forecasts by the SVR were more accurate than the forecasts by the Institute and the IW.

248

The quality of a forecast depends not only on how accurately it predicts GDP growth but also on whether it foresees turning points. Figure 4 shows to what extent the IW-forecasts predicted an acceleration or deceleration of the growth rate in the time period 1995 to 2001.

Figure 4 : Forecasted and Actual Changes

changes of growth rates of real GDP from previous period, in percentage points,
IW-forecasts for Germany, 1995 to 2001
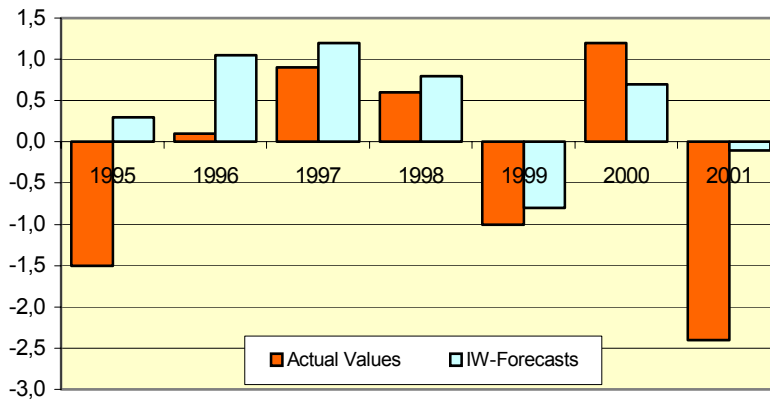
The columns in figure 4 show the change of the growth rate – not the growth rate itself – from the previous period in percentage points. It plots the change of the actual and predicted value from the actual outcome of the previous year. For example, the IW expected that the predicted growth rate of 3 per cent for 1995 would exceed the actual growth rate for 1994 (2.7 per cent) by 0.3 per-centage points. In fact, the actual growth rate for 1995 (1.2 per cent) fell below the actual growth rate of the previous year by 1.5 percentage points. Finally, figure 4 also shows that in six out of seven years the IW-forecasts did well in predicting an acceleration (1996, 1997, 1998, 2000) or a deceleration (1999 and 2001). An error in forecasting a change happened in 1995 only. The extent of the change was predicted fairly accurately for 1997, 1998, 1999 and more or less for the year 2000. Although the direction of change was predicted correctly for 1996 and 2001, the forecasts failed to foresee the extent of the change.

An additional measure to assess the informative quality of forecasts is given in table 3. The contingency table indicates whether the forecasts predicted the change of the growth rate correctly or not (Diebold/Lopez, 1996). The columns show the actual accelerations and decelerations, and the lines the predicted ones. Cell AA and BB mean that actual accelerations or decelerations were predicted correctly, whereas AB and BA indicate faulty forecasts.

249

Table 3: Contingency Table for Actual and Predicted Changes

	Actual acceleration	Actual deceleration
Predicted acceleration	AA	AB
Predicted deceleration	BA	BB

In combination with table 3 there is a measure for the informative substance (IS) of a forecast (Doepke, 2000): IS = {(AA) / (AA + AB)} + {(BB) / (BB + BA)}. The values of this measure range from 0 to 2, with 0 meaning that no acceleration or deceleration was predicted correctly at all, and 2 meaning that all accelerations or decelerations were predicted correctly. If the measure is 1, it means that all cells of the contingency table are filled with the same number of observations. In this case the number of correct forecasts equals the number of erroneous ones.

Figure 4 has already shown that four accelerations and two decelerations were predicted correctly. Only in 1995 the actual deceleration was not forecasted properly. Therefore cell AA contains four, cell BB two, and cell AB one observation. According to this result the measure IS takes a value of 1.8, which is close to the ideal grade of 2. For West Germany in the time period 1976 to 1999 Doepke (2000) found out IS-values of 1.63 (forecasts by the German Institute for Economic Research – Deutsches Institut für Wirtschaftsforschung, DIW) and 1.39 (forecasts by the Kiel Institute of World Economics – Kieler Institut für Weltwirtschaft). However, it must be kept in mind that this analysis covers a longer time period, so that a comparison with the latest results is possible only to a limited degree.

After assessing the quality of economic forecasts several arguments are presented which may explain forecast errors.

1. Data revisions make forecasts prone to errors

Forecasting errors are sometimes due to revisions of the underlying data. Subsequent revisions of official data change the statistical fundamentals of forecasts and make them prone to errors.

The revision of the data for Germany´s GDP growth in the third quarter of 2000, for example, largely diminished the statistical carry-over for the year 2001. When forecasting for 2001 in autumn 2000 the preliminary growth rate for the third quarter 2000 was according to the official data substantially higher than later on. In the course of time it was revised down several times by the Federal Statistical Office (Statistisches Bundesamt) – from 0.6 per cent in November 2000 to 0.3 per cent in March 2001, and finally to 0.1 per cent in August 2001. For that reason the statistical carry-over was halved to merely 0.5 per cent. This revision of data cannot explain the entire

forecasting error for 2001 (2.3 percentage points), but at least for half of a percentage point. For similar results see Nierhaus (2002).

2. Unpredictable events make forecasts obsolete

The extent of the forecasting error for Germany in 2001 can be explained to a large degree by the business slump of the U.S. economy and the loss of purchasing power due to inflation. Unforeseeable shocks, like excessive wage settlements (Germany 1995), overshooting exchange rates (Germany 1995), natural disasters (e.g. earthquake in Japan 1995), country-specific crises (e.g. Mexiko 1995; Russia and Asia 1997), oil price shocks (2000/2001), animal epidemics (Europe 2000/2001), the hard landing of the U.S. economy (2001) and political shocks (terror attacks in the United States 2001) contribute to invalidate forecasts.

In addition to spreading more rapidly international shocks overthrow some forecasts faster nowadays.The velocity with which global events affect economies around the world has accelerated and has obviously been under-estimated by forecasters. Apart from traditional trade relations, financial markets and multinational corporations are today important routes for transmitting global impulses. Business cycles in certain countries affect in-vestment decisions of multinational firms directly and thus have a higher in-fluence on business activities in other countries (Dalsgaard/Elmeskov/Park, 2002).

However, not all European economies have recently been affected in the same way by external shocks, although they are equally tied up in the international division of work (Beyfuß/Groemling, 2001). Germany was hit harder by external disruptions than other European countries. Its lacking flexibility to adjust to external and internal shocks is obviously the main reason for its higher susceptibility.

3. Conditional forecasts ignore feed-back

Most of the forecasts are conditional. They are based upon certain assump-tions about the political situation and the economic behaviour of households and firms. Political conditions – for example, taxes and transfers – that pre-vail at the time of forecasting are taken into account, of course. However, unforeseeable policy changes cannot be included. Political changes and their effects on growth may also be predicted but are more suitable for scenarios.

Forecasting errors happen, in particular, if the forecast itself induces changes of political conditions or economic behaviour. Because economic agents react to forecasts, they influence whether the forecasted situation will happen or not:

- On the one hand, predictions may unfold a momentum which fosters the predicted development (self-fulfilling prophecy). For example, a forecasted recession may aggravate the economic situation because in-vestors or consumers cut down on expenditure.

251

- On the other hand, forecasts may set off reactions that destroy the forecast. For example, a forecasted recession may trigger political programmes, that help to improve the economic situation and to avert the forecasted development.

Obviously, the optimism to steer the economy through stabilization politics in the sixties and seventies of the last century coincided with the heyday of economic forecasting.

4. Forecasters are human beings
Forecasters often succeed in predicting an acceleration or deceleration of the growth rate. But the predictive power is often not sufficient enough to forecast an absolute decrease of real GDP: „Recessions generally arrive before the forecast." (Loungani, 2001). There are at least two reasons to explain this:

- Forecast smoothing
Forecasters adjust their predictions slowly and in small steps only. As human beings generally adhere to the status quo, forecasters hold on to their present forecast. They may fear that large and, in particular, abrupt corrections confuse the users (Nordhaus, 1987). Fast and abrupt changes may be interpreted as inconsistency of the forecasters point of view or an exaggerated reaction on events of the day.

- Forecast clustering
In particular, in times of increased uncertainty forecast clusters seem to emerge. The forecasts gather around a consensus value. Imitation or herd behaviour of the forecasters leads to a convergence towards a common value (Loungani, 2001; Batchelor/Dua, 1992; Gallo/Granger/ Jeon, 2002). Particularly in the case when forecasters copy the latest forecasts of others, this imitation behaviour may result in a consensus that increasingly departs from reality. New information on the business situation will have less and less influence on the forecasts. Under certain circumstances forecasters may have a closer look on other forecasts than on economic fundamentals (Gallo/Granger/Jeon, 2002).

Additionally, it is not easy for single forecasters to deviate from the „common sense" of the forecasting community. According to the motto „Great minds think alike" (Loungani, 2001) there may be pressure to converge or stick to the consensus. At best, well established forecasters may risk to abandon the herd (Lamont, 1995). In sum, a high conformity of forecasts is not necessarily proof of high clarity of the forecasts.

Figure 5-a : Highest and Lowest Forecasts for Germany for 2001
predicted growth of real GDP for 2001 according to Consensus
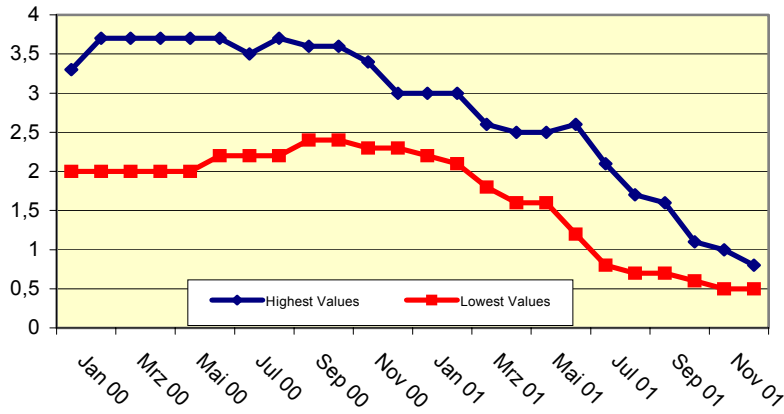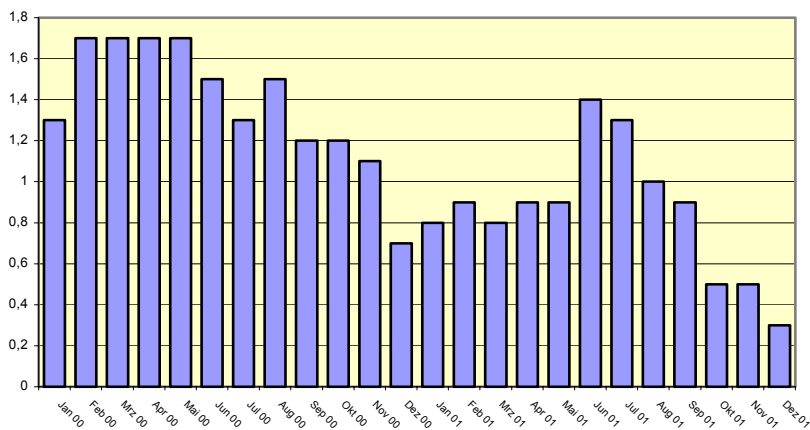Forecasts, in percent

Figure 5-b : Differences between Highest and Lowest Forecasts
in percentage points

Finally, figure 5 indicates that there is no clearcut evidence for imitation behaviour in the forecasts for Germany for 2001. Figure 5-a shows how the monthly growth predictions for 2001 have changed between January 2000 and December 2001. Not only the highest but also the lowest single forecasts, published monthly by Consensus Forecasts, are taken into account. According to figure 5-a the forecasted growth rates for real GDP were reduced more and more in the

253

course of time. This, of course, might be attributed to the fact that the level of information improved as time passed by. Therefore, both, the highest and the lowest forecasts, were close to the actual outcome by the end of 2001, and there was no pronounced difference between them anymore. However figure 5-b shows that the difference between the highest and lowest monthly forecast did not decline continually. It was stable to a large degree during the first six months of 2001, but increased markedly in the middle of the year 2001.

References

Batchelor, Roy; Dua, Pami, 1992, Conservatism and consensus-seeking among economic forecasters, in: Journal of Forecasting, Vol. 11, pp. 169 - 81.

Beyfuß, Joerg; Groemling, Michael, 2001, Deutsche Wirtschaftsperspektiven im Schatten hoher konjunktureller Störanfälligkeit, in: iw-trends, 28. Jg., No. 2, pp. 5 - 20.

Dalsgaard, Thomas; Elmeskov, Jorgen; Park, Cyn-Young, 2002, Ongoing Changes in the Business Cycle – Evidence and Causes, OECD Economics Department Working Papers, Nr. 315, Paris.

Diebold, Francis X.; Lopez, José A., 1996, Forecast Evaluation and Combination, in: G. S. Maddala; C. R. Rao (Ed.), Handbook of Statistics, Vol. 14, pp. 241 - 268.

Doepke, Joerg, 2000, Haben Konjunkturprognosen in Deutschland einen politischen Bias?, in: Schmollers Jahrbuch, Vol. 120, pp. 587 - 620.

Doepke, Joerg; Langfeldt, Enno, 1995, Zur Qualität von Konjunkturprognosen für Westdeutschland 1976 - 1994, Kieler Diskussionsbeiträge, No. 247, Kiel.

Gallo, Giampiero M.; Granger Clive W. J.; Jeon, Yongil, 2000, Copycats and Common Swings: The impact of the use of forecasts in information sets, in: IMF Staff Papers, Vol. 49, No. 1, pp. 4 – 21.

Groemling, Michael, 2002, Konjunkturprognosen: Methoden, Risiken und Treffsicherheiten, in: iw-trends, Vol. 29, No. 2.

Hinze, Joerg, 1996, Konjunkturprognosen unter veränderten Rahmenbedingungen – Auswirkungen der deutschen und europäischen Einigung, in: Hamburger Jahrbuch für Wirtschafts- und Gesellschaftspolitik, Vol. 41, Tübingen, pp. 59 - 84.

Lamont, Owen, 1995, Macroeconomic Forecasts and Microeconomic Forecasters, NBER Working Paper Series, No. 5284, Cambridge, MA.

Loungani, Prakash, 2001, How Clear is the Crystal Ball?, Reflections on the accuracy of growth forecasts, in: World Economics, Vol. 2, No. 1, pp. 1 - 8.

Mincer, Jacob; Zarnowitz, Victor, 1969, The Evaluation of Economic Forecasts, in: Economic Forecasts and Expectations: Analyses of Forecasting Behavior, NBER Studies in Business Cycles No. 19, New York.

Nierhaus, Wolfgang, 2002, Deutsche Konjunktur 2001: Prognose und Wirklichkeit, in: ifo Schnelldienst, Vol. 55, pp. 32 - 34.

Nordhaus, William D., 1987, Forecasting Efficiency: Concepts and Applications, in: The Review of Economics and Statistics, Vol. 69, No. 4, pp. 667 - 674.

Rinne, Horst, 1997, Bemerkungen zur Qualität der Konjunkturprognosen des Sachverständigenrates zur Begutachtung der Gesamtwirtschaftlichen Entwicklung, in: Peter von der Lippe (Ed.), Wirtschafts- und Sozialstatistik heute: Theorie und Praxis, Ludwigsburg, Berlin, pp. 319 - 340.

Stekler, H. O., 1991, Macroeconomic forecast evaluation techniques, in: International Journal of Forecasting, Vol. 7, pp. 375 – 384.

Wallis, Kenneth F., 1989, Macroeconomic Forecasting: A Survey, in: Economic Journal, Vol. 99, pp. 48 - 81.