DIW Berlin

German Institute
for Economic Research

Discussion Papers

399

Jörg Döpke
Ulrich Fritsche

Growth and Inflation Forecasts in Germany – An Assessment of Accuracy and Dispersion

Berlin, February 2004

Opinions expressed in this paper are those of the author and do not necessarily reflect views of the Institute.

DIW Berlin

German Institute
for Economic Research

Königin-Luise-Str. 5
14195 Berlin,
Germany

Phone +49-30-897 89-0
Fax +49-30-897 89-200

www.diw.de

ISSN 1619-4535

Growth and Inflation Forecasts for Germany - An Assessment of Accuracy and Dispersion

by

JÖRG DÖPKE and ULRICH FRITSCHE*

January 2004

Abstract

Based on a panel of German professional forecasts for 1970 to 2003 we find that growth and inflation forecasts are unbiased and weakly, but not strongly efficient. Besides the effect of diverging forecasting dates, no other substantial differences in forecasting quality are found among forecasters. We argue that it is not always advisable to listen to the majority of forecasters. The dispersion of forecasts correlates positively with the volatility of macroeconomic variables. This suggests that forecasters do not behave predominately strategically, but rather share no common belief on the adequate model of the economy.

Keywords: Forecast error evaluation, consensus forecast, disagreement, uncertainty, Germany

JEL classification: E32, E37, C52, C53

* Corresponding author: Ulrich Fritsche, German Institute for Economic Research (DIW Berlin), Königin-Luise-Strasse 5, D-14195 Berlin, Germany. Phone: +49 30 8978 9315; Fax: +49 30 8978 9102; e-mail: ufritsche@diw.de. Jörg Döpke, Deutsche Bundesbank, Economics department, Wilhelm-Epstein-Strasse 14 Frankfurt, Germany. Phone: +49 69 9666 3051; Fax: +49 69 9566 4317; e-mail: joerg.doepke@bundesbank.de. The authors thank 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 helpful comments on an earlier draft of this paper. The usual disclaimer applies. The views presented in this paper are the author's, and do not necessarily reflect those of the DIW Berlin or the Deutsche Bundesbank.

I. Introduction

Expectations play a central role in modern macroeconomics. Assumptions about expectations are crucial for many important features of macroeconomic models. Consequently, a broad area of research analyses the properties of expectations of households and firms. Particular attention has also been given to forecasts of professional forecasters as a proxy for economic expectations. Though there are a lot of studies that refer to individual forecasts or very small groups of institutions[1] there is still no study, which takes into account a larger group of forecasting institutions for Germany. The present study tries to fill this gap by (to our knowledge) taking into account the prognoses of all forecasters that provide forecasts for Germany for a long time span, ie from (at least) the early seventies onwards.

This novel data set renders it possible to test whether there are any systematic differences in forecasting accuracy among the forecasters observed in this study. Moreover, the usual focus on individual forecasters or on the properties of a consensus forecast implies that an important stylised fact is, as Mankiw, Reis and Wolfers (2003:1) put it, "routinely ignored: not everyone has the same expectation." This lack of evidence comes as a surprise, as the dispersion of forecasts might provide useful information. For example, the degree of disagreement can be interpreted as a measure of forecast risk (Zarnowitz and Lambros 1987). Additionally, disagreement can help to understand the way in which the economy deals with information. For example, a large and persistent disagreement might point to the empirical relevance of "sticky information" models (Mankiw and Reis 2002, Carroll 2003). To shed light on some of the reasons for disagreement we link the measure of forecaster's disagreement to the volatility of some macroeconomic variables. We interpret the results in a way, that during turbulent times there is a large disagreement about the transmission of shocks between forecasters. This is partly in contrast to other studies, which stress the importance of forecaster's disagreement as a measure of uncertainty (Bomberger 1996), or as a result of strategic behaviour (Laster, Bennet and Geoum 1999). The level of disagreement concerning the effects of monetary policy shocks seems to be especially large among the institutions in our sample.

---

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

We use German annual data for growth and inflation forecasts from 1970 to 2003 stemming from up to 17 different forecasts to analyse the accuracy of the predictions and the disagreement among German professional forecasters. 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 rational expectations hypothesis is usually rejected. We document a substantial degree of disagreement among forecasters. Forecast dispersion moves with the stance of the business cycle, the level of inflation and the volatility of short-term interest rates, oil price movements and real exchange rate fluctuations.

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 discusses the properties of the consensus forecasts, measures of forecast dispersion and the underlying reasons for disagreements. Section 4 concludes.

II. Data and measures of forecast accuracy

II.1 Data

In the following section we evaluate forecasts of several institutions, which deliver forecasts regarding the German economy. In particular, the forecasts of the following institutions are under consideration:[2]

1. The Council of Economic advisors. The "five wise men" provide their report in the autumn of each year, normally in early November. The source for the forecasts is *Sachverständigenrat*, several years, Wiesbaden.
2. The joint forecast of the six (formerly five) "leading" German economic research institutes. The joint forecast is undertaken twice a year, in spring and in autumn. The concrete date of publication varies, but end of April and end of October are quite common publication dates. Numbers are taken from *Arbeitsgemeinschaft wirtschaftswissenschaftlicher Forschungsinstitute*, various issues, Berlin.

---

[2] Compare Antholz (2003) for a longer time series of forecasts for some institutions included in the study.

3. The German Institute of Economic Research (DIW Berlin), Berlin. Though the institute publishes its forecasts twice a year, we only refer to the forecast in January, which is based on the information available at the end of the previous year. The figures can be found in: German Institute of Economic Research (ed), Economic Bulletin, various issues, Berlin.
4. The Kiel Institute for World Economics (IfW), which mostly forecasts four times a year. We refer to the forecast published in December. The source for all figures is the institute's journal "*Die Weltwirtschaft*", various issues, Kiel.[3]
5. The HWWA Institute of Economic Research. The institute publishes twice a year. We refer to the autumn forecast. The source of the forecasts is the institute's journal "*Wirtschaftsdienst*", various issues, Hamburg.
6. The ifo Institute Munich publishes forecasts four times a year. We refer to the end-of year forecast published in the "*ifo Schnelldienst*", various issues, Munich.
7. The Institute for Economic Research Halle (IWH) was founded in 1993. It started publishing forecasts in 1995, mostly in the institute's journal "*Wirtschaft im Wandel*", various issues, Halle/Saale. We refer to the respective December forecasts.
8. The Rhine-Westphalia Institute for Economic Research (RWI), Essen. The main source for the forecasts is the institute's "Economic report", published on a biannual basis. We refer to the end-of-year forecast.
9. The OECD publishes forecasts for Germany twice a year in its "Economic Outlook" We refer to the forecast published in the autumn of each year.
10. The IMF provides two forecasts. We take account of both the spring forecast and the autumn forecast. The forecasts are published in the "World Economic Outlook".[4]
11. The European Commission provides two forecasts a year. We take into account both the spring (referring to the current year) and the autumn forecast (referring to the coming year) predictions. The forecasts have been published in the Commission's publication "*European Economy*", various issues, Brussels.
12. The Trade Union Institute (WSI) publishes its forecast at the end of each year in its journal "*WSI Mitteilungen*". [5]

---

[3] In two years, 1975 and 1976 the institute did not publish forecasts (see Stege (1989)) for a discussion of possible reasons). To ensure a complete time series we use the respective numbers from the joint forecast for these years, since the Kiel Institute was part of the joint forecast.

[4] However, some of the forecasts have not been published. Herewe use the data collected by Artis (1987).

13. The Institute of the German Economy (IW). The forecasts are published regularly in the journal "*IW-Trends*" at the end of each year.
14. The economic report of the German federal government ("*Jahreswirtschaftsbericht*") is regularly published by the responsible ministry at the beginning of each year.

The forecasters do not all publish at the same time. Figure 1 shows the average timing of the forecasts during the year.

— Insert figure 1 here —

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.

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).[6] In the case of interval forecasts the average is used. The numbers refer to West Germany up to 1992, and to the whole of Germany from 1993 to present. As a measure of the inflation forecast we use the predicted change of the deflator of private consumption when this figure was available. In some cases, however, no explicit reference was given 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. Table A1 and A2 in the appendix present all the forecasts we have used for our investigation. 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 ie we compare the forecasts made at the end of a certain year "x" or at the beginning of the following year "x+1" with the first published figure for the year "x+1".

---

[5] In 2001 the institute provided no growth forecast and in 2001 and 2002 no inflation forecasts. An inspection of the institute's comments makes clear, however, that it criticised the other forecasters as being too optimistic on the stance of the business cycle. Thus, we impute the minimum growth and inflation forecasts available at the end of the respective years.

[6] Sometimes the forecasts refer to "growth" rather than explicitly to either GDP or GNP. In this cases we have assumed that the forecasters made no distinction between the concepts and had the same forecast for both figures.

II.2 Forecast accuracy

To begin with, tables 1a and 1b present a couple of standard measures of forecast accuracy for each institution separately. In particular, we calculate the following statistics:[7]

- The mean error $ME = \frac{1}{T}\sum_{t=1}^{T} e_{t+1,t}$, 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) 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.
- The mean absolute error $MAE = \frac{1}{T}\sum_{t=1}^{T} |e|_{t+1,t}$.
- The root mean squared error $RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T} e_{t+1,t}^2}$.
- 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, ie running the following regression: $y_{t+1} = \beta_0 + \beta_1 y_{t+1,t} + u_t$ and test the hypothesis $H_0 : \begin{matrix}\beta_0 = 0 \\ \beta_1 = 1\end{matrix}$. We prefer regressing the forecast error on a constant: $e_{t+1} = \beta_0 + u_t$ and test the hypothesis $H_0 : \beta_0 = 0$. The respective column in the table reports the marginal significance level of this test.
- 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.
- The former test is a test for weak rationality; ie 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 it is emphasised by Tichy

---

[7] Unless otherwise stated our notation follows the textbook of Diebold (1998).

(1994), tests which 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 $y_{t+1} = \beta_0 + \beta_1 y_{t+1,t} + \beta_2 X_{t-1} + u_t$ where X_{t-1} represents the respective exogenous variable in the second quarter of the year in which the forecast is made. We use a standard Wald-Test to test the null hypothesis $H_0 : \begin{cases} \beta_0 = 0 \\ \beta_1 = 1 \\ \beta_2 = 0 \end{cases}$ (Holden and Peel 1990). If this hypothesis cannot be rejected, the forecast has to be considered as rational in a stronger sense.

— Insert tables 1a and 1b here —

In a nutshell, the results in table 1a and 1b confirm 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 suggesting on average a slight tendency of the forecasters to be too optimistic. However, the bias is not significant. A noteworthy 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 one percentage point.. Though the magnitude of these numbers sounds very familiar to economists in this field it is noteworthy that they still by far exceed the expectations of the public regarding the accuracy of forecasts (see, for example the statements surveyed in Döpke 2000). Moreover, the outcomes come close to what can 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, one should keep in mind that the information set represented by the

endogenous lagged variable is rather limited. Consequently, the picture given 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. One might argue that the rejection based on the information included in the past oil price change is not a particular reason for concern since oil price crises are infrequent events and, thus, it is hard for the forecasters to learn about them (see Keane and Runkle 1990). It is even possible to argue that the information contained in the change of the real exchange rate is difficult to exploit since this variable is extremely volatile. Moreover, domestic forecasters might be less informed over the business cycle development in other countries and thus, fail to fully take into account the information represented by the industrial production abroad.

But even if one accepts all these arguments, an astonishing puzzle still remaina: past short-term interest rates – which are definitely known to all forecasters – help to explain the forecast errors and, thus, violate the rationality of the forecasts (see Kirchgässner and Savioz 1999). Even a brief examination of the forecasters' publication 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 apparant 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, in contrast, less clear-cut. The often-stated hypothesis that the alleged poor quality of Germany's business cycle forecasts is merely due 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 respective number for the growth forecasts. Moreover, although the tests suggest unbiased

forecasts, the predictions turn out to be inefficient.[8] As regards the test for strong rationality the null is rejected less frequently than in case of growth forecasts. All in all, our results give support to doubts on the rationality of price forecasts raised recently by several authors (see Mankiw, Reis and Wolfers 2003, Caroll 2003).

II.3 Forecast accuracy – exploiting the panel dimension of the data

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 $y_{t+1,i} = \beta_0 + \beta_1 y_{t+1,i} + u_{t,i}$ with i denoting the number of forecasters. To our knowledge this is the first time that this dimension is used for German data. We only refer to forecasts, which have been made at the end of the year or in January to ensure a relatively common information base.[9] Furthermore, since some of the empirical techniques used require a balanced panel, we have left the Halle Institute out of the analysisand have restricted the sample to 1974 to 2003.

In a first step, we estimate the equation using OLS. This might be, however, incorrect as it 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). Moreover, we extend their approach and discuss also results that take into account both heteroscedasticity and possible correlation between the individual forecasts. 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.

— Insert table 2 here —

The results of these tasks are summarised in table 2. The most striking result is that, regardless of the estimation technique and the concrete test set-up, the null of rational exceptions is rejected strongly by the data. It is noteworthy that the rejection of the overall

---

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

[9] In particular we refer to the following forecasts in our panel analysis: the joint forecast, autumn, the Council of Economic Advisors, the Kiel Institute, the Berlin Institute, the Hamburg Institute, the Munich Institute, the Essen Institute, OECD, the Institute of the German Economy, the Trade Union Institute, the European Commission, autumn, and the IMF, autumn.

hypothesises stems apparently from different sources. The baseline regressions raise doubts on the unbiasedness and (weak) efficiency of the forecasts. In some cases, the additional variable does not make the results worse. However, the strong form of rationality is also rejected frequently, as it can be seen from the z-values of the exogenous variables.

The panel estimation results give thereby cause to serious doubts that the available information is processed by the bulk of forecasters in a way, which satisfies the postulate of a strictly rational expectation formation.

II.4 Tests on forecast equivalence and encompassing

An important question regarding the evaluation of forecast accuracy is how individual institutions perform in comparison to each other. To analyse this problem we use Diebold-Mariano and forecast encompassing tests.

In a first step, we test if the mean squared forecast error of one institution is significantly different from those of another institution. Formally, we apply a version of the Diebold-Mariano (1995) test and calculate the difference of the squared forecast errors ($D_t = (F_t^A)^2 - (F_t^B)^2$) and regress this term on a constant: $D_t = \alpha + \varepsilon_t$. Using again robust standard errors as proposed by Newey and West (1987) and the correction for small samples as proposed by Harvey, Leybourne and Newbold (1997) we test for $H_0 : \alpha = 0$.

— Insert table 3a and 3b here —

The results of these tests for growth and inflation forecasts are reported in tables 3a and b. The null hypothesis is stated in such a way, that the forecast in column is equal to the forecast in the row. We report the respective p-values of the test as explained in Harvey, Leybourne and Newbold (1997). According to standard criteria a p-value below 0.05 indicates a statistically significant difference of the mean squared errors. The findings suggest that most growth forecasts do not differ systematically. However, the IMF autumn forecast seems to have a significantly higher mean squared error as compared to most forecasts under investigation. This likely to be due to the fact that this forecast is the earliest annual forecast in our sample and, thus, suffers from the relatively small information basis available at this time. Also, the last forecasts – according to our ordering – seem to be significantly better than

the other forecasts. Again, the likely reason is obvious: these forecasts have a systematic advantage owing to more information.

Regarding inflation forecasts there is – according to the Diebold-Mariano-test – no evidence that the forecast quality is systematically different for most of the forecasters included in the study. Notable exceptions include the two institutions, which forecast late: namely the EU Commission spring forecast and the joint forecast in spring.

In a second step, we test the hypothesis that a specific forecast contains information from the other forecasts. Thus, we perform forecast encompassing tests as described in the literature (Diebold and Lopez, 1996, pp. 253) and estimate the following equation: $R_t = \beta_A P_t^A + \beta_B P_t^B + \varepsilon_t$ where R_t stands for the realised value of the relevant variable and $P_t^i, i = A, B$ represent the competing forecasts. The hypothesis that the forecast of institution A (in column) encompasses the forecast of institution B (in row) can be stated as $H_0 : \begin{cases} \beta_A = 1 \\ \beta_B = 0 \end{cases}$. If this hypothesis is not rejected, then the forecast A is systematically better than those of institution B. If we assume that forecasters are only interested in minimising the forecast loss, such a systematic and statistically significant outperformance of one institution compared to another one could be due either to more processed information or a better forecasting technology. To check this, all forecasting institutions were ordered according to the date of publication of their regular forecasts (see figure 1). As proposed by Harvey, Leybourne, and Newbold (1998) we performed the test in terms of forecast errors again and used the small-sample correction as in the case of the Diebold-Mariano-test. The table which contains the p-values of the above-mentioned data should be read as follows: The hypothesis is that forecast A encompasses B. Owing to the ordering we can expect the lower left area below the diagonal to be filled with zeros, ie the forecast published later should contain the antecedent forecasts. If there is no other reason than timing for varying performance, the upper right area should be filled with p-values clearly above the usual 0.05 significance criterion. Systematic deviations from that expected results have to be explained.

— Insert table 4a and 4b here —

The results of the encompassing tests are shown in tables 4a and b. All in all, the results appear to be not very surprising. We get the expected result that the in most cases, on average, the forecast published later contains the information of the forecast published earlier. With the

exception of the inflation forecasts made by the Kiel Institute and the Trade Union Institute there seems to be no systematic deviation from the results we would expect due to the ordering.

Taking the results from the table of forecast encompassing tests into account will give the impression that there is hardly any systematic difference in forecasting quality. The remaining differences between institutions can, at least to a large extent, be explained by the forecasting date.[10] This is illustrated by figure 2. We ordered the institutions according to the forecasting date as shown in figure 1 and regressed that on the mean absolute growth forecast error. The exhibit reveals that the absolute growth forecast error is declining as the forecasting date shifts more into the year for which the forecast is made.

— Insert figure 2 here —

II.5 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 (eg 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 seldom and irregular periods. Thus, since one has 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 5 (cf Diebold and Lopez 1996: 257).

— Insert table 5 here —

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

With the classification of table 5 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 $I \to 1$. If the forecast is perfect than $O_{ji} = O_{ij} = 0$ and $I = 2$. Therefore, any value of $1 < I \leq 2$ indicates a 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 $\hat{E}_{ij} = O_{i.} O_{.j} / O$. Finally, we construct the following measure $C = \sum_{i,j=1}^{2} (O_{ij} - \hat{E}_{ij})^2 / \hat{E}_{ij}$, 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.

—Insert table 6 about here—

The results of such tests are given in table 6. It turns out that all forecasters included in the test are at least as good as 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 in 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: OECD, the Institute of the German Economy and the *Sachverständigenrat* are early birds; it is therefore not all that astonishing that they miss the direction of changes more often than other institutions with access to more timely information.[11]

---

[10] For a recent analysis of the revision of forecasts see Kichgaessner and Müller (2003).

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

III. Analysing forecast dispersion

Up to this section, the analysis has focused on the question of forecast accuracy for individual forecasters and on the comparison of forecasts. Recently, the attention of the profession has also turned (again) to the cross-section dimension of the data. It has been argued that the dispersion of forecasts might provide interesting information and, furthermore, has implications for the relevance of theoretical models (see eg Mankiw, Reis and Wolfers 2003 and Caroll 2003). Since we are not aware of a study that analyses the dispersion of forecasts in Germany we aim in following at establishing a set of stylised facts regarding the cross-section variance of forecasts. In contrast to the vast majority of recent studies we do not exclusively rely on price expectation or inflation forecasts, but take also into account growth predictions.

To gain insights into the role of forecast dispersion we analyse in a first step whether the consensus forecast sufficiently covers all the information of the forecasts, or, if minority reports are worth looking at. In a second step, we try to figure out whether there are typical distributions of forecasts around pronounced business cycle turning points. For example, we study the pattern of growth and inflation forecasts before and during recessions. This leads immediately to the closely related question whether forecast dispersion is a measure of forecast risk. The last section is devoted to establishing under which circumstances forecasters disagree. Figuring out these circumstances might provide additional insights into the forecasting process, eg into the question as to what extent different theoretical standpoints or diverging incentives lead to diverging forecasts.

III.1 The consensus forecast and its properties: should we always listen to the majority?

Some authors have argued that the consensus forecast may be superior to individual forecasts (see Gavin 2003 for a discussion of related arguments). To shed light on the record of the consensus forecasts, table 7 reports the measures of forecast accuracy used in the previous chapters to gauge the quality of individual forecasts for the consensus forecasts. We have calculated the consensus forecast based on a restricted data set. In particular, we refer only to forecasts that have been published in January at the very latest. Moreover, we calculate two measures of a consensus: the mean and the median. The latter number puts less weight on outliers. We have also considered more sophisticated ways to calculate the consensus, ie we have tried to estimate optimal weights for each forecaster. However, since

there is a link between the optimal weight for each forecast and the result of the encompassing test mentioned above, we find only very few reliable weights because, in most cases, the forecasts published later encompass the earlier ones. Furthermore, the results obtained by this measure do not differ substantially from those for the mean and median.

— Insert table 7 here —

The results summarised in table 7 suggest that the consensus forecast has indeed reasonable properties. The measures of forecast accuracy are of more or less the same quality when compared to the individual forecasts. The tests on unbiasedness and weak efficiency also show results, which are at least not worse than the ones for individual forecasters – in spite of the fact that strong rationality is rejected frequently. Thus, relying on a consensus forecast seems to be nevertheless a natural choice for a person otherwise uninformed on the quality of forecasts.

Given the relative good forecasting record of the consensus forecasts in our data set as documented in the related literature, the question may arise 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 errors. To this end, we checked whether in a certain year all forecasters agreed on the direction of change. The classification of forecast errors is the same as above, ie relies on the acceleration or deceleration of growth rates. If at least one forecaster predicts an acceleration (deceleration) whereas the other forecasters forecast a deceleration (acceleration) we count this as a minority report.

— Insert table 8 here —

It appears that there are 12 (15) years in which a minority of forecasters disagrees on the question whether growth (inflation) will speed up or not. Table 8 gives the measures of information content introduced in the previous chapters. The results differ sharply between inflation and growth forecast. 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 however noteworthy 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

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.

III.2 The distribution of the forecasts at turning points

In this section, we analyse whether the cross-section distribution of forecasts displays any pattern, which can be seen as typical for certain spans of the business cycle. To shed light on this possibility, we analyse selected distribution of forecasts in periods of pronounced changes of the business cycle (see Mankiw, Reis and Wolfers 2003: 46 for a similar analysis of the Volcker disinflation period). For the purpose of the present analysis it is useful to keep in mind that the symmetry of the distribution of forecast may contain information. For example, Linden (2003) argues that the skewness of the distribution might be linked to the frequently used concepts of "downside" or "upside" risks in forecasting. If the distribution is skewed to the right (ie if the skewness is positive) the distribution has a long tail to the right and a sharp cut-off to the left. Thus, the majority of forecasters sees lower values as more likely than higher values, which can be interpreted as a downward risk. The opposite holds true in the case of a distribution skewed to the left.

— Insert figure 3 here —

First we examine the distribution of the growth forecasts before and during the recession of 1974/75. In figure 3 we plotted the Kernel density estimates using a normal distribution and the realisation as a dotted line. This economic downturn is usually seen as the most pronounced recession Germany has ever suffered. Most observers see the oil price shock which began in 1973 as the main cause of the recession. The oil price shock was the first one to ever hit the German economy. Thus, the forecasters had no experience to rely on. Under such circumstances one might expect a large degree of uncertainty resulting in a sharp disagreement between forecasters. However, the bulk of forecasters more or less forecasted the same growth rate for 1975 as they had forecasted one year before. Still the forecasters could have noticed the downward risk. But nothing consistent with such a behaviour is visible in the data. The distribution of the forecasts is more or less symmetric and the variance of the prediction turns out to be rather small.

The episode of 1981/82 reveals that the experience of the forecasters with the first oil price hike did not lead to greater forecasting success: the forecast errors are only slightly lower than those for the 1974/75 recession. However, the distribution of the forecasts is somewhat different. In particular, the relatively flat distribution of the forecast for 1981 is compatible with the notion of high uncertainty at that time. The 1981/82 recession differed in an important way from the 1974/75 recession: It was the first case of a so-called double dip in post-war Germany. In 1982 a very small peak left from the mass of the distribution reflects one outlier (here: the Kiel Institute) being more sceptical than the other forecasters. It turns out that the outsider was – whatever the reason might have been – the institute which forecasted the second dip with an impressive exactness.

The bulk of forecasters also missed the recession in 1993. However, the degree of uncertainty implied by the fat-tailed distribution seemed to have been high. Maybe influenced by their misfortune in 1982 some forecasters remained very sceptical regarding an upturn. All of them missed the dynamics of the upturn. All in all, the findings so far may point to the impression that expectations are rather adaptive – forecasters seem to learn from their misfortune but at the same time make other errors – and suggest that no business cycle is exactly the same as any other cycle.

— Insert figure 4 here —

The case for adaptive expectations is probably even more pronounced regarding inflation expectations. As an example, figure 4 depicts the inflation episode from 1978 to 1981 when inflation rates were exceptionally high by German historical standards. In contrast to figure 3, we plot the realisations of the previous year as a dotted line and the actual realisation as a thin line in figure 4. For 1978 most forecasts are concentrated around the previous year's realisation whereas a minority of forecasters believed in a lower rate. For 1979 and 1980 most forecasts are again concentrated around the previous year's realisation. Apparently, the learning curve must have been very flat at those times. The year 1981 is a curious example: it looks a bit like conspiracy. The majority of forecasters seemed to have opted for a moving average of the last year's realisations. Indeed, a sheer guess according to the last year's realisation would have produced better results. To be fair, the 1978 to 1981 episode clearly shows that in times of rapid changes in the conduct of economic policymakers – as is likely the case around the beginning of the 1980s (cf Clarida, Galí and Gertler 2000) – forecasting is a particularly difficult business.

III.3 Forecast disagreement over time

In the previous section, we asked whether the degree of disagreement is a measure of uncertainty or risk in a rather anecdotal manner. We now address this problem in a more systematic way. The question whether the cross-section variance of forecasts is a good measure for uncertainty has been frequently discussed in the literature, yet no consensus has been achieved. While Bomberger (1996) argues that disagreement is a useful proxy for uncertainty, this finding is challenged, eg by Rich and Tracy (2003). While Bomberger (1996) has established an empirical link between the conditional variance of a time series as a widely accepted measure of uncertainty and forecaster disagreement Rich and Tracy have not been able to confirm these results based on, as they claim, more robust econometric techniques.

In order to assess the usefulness of forecaster disagreement as a measure of uncertainty and risk, we first take a look at the development of the disagreement of forecasters over time (figure 5). We make use of two measures of forecast dispersion: the standard deviation of forecasts and, to exclude possible distorting outlier observations, the interquartile range.

Apparently, forecaster disagreement seems to decline over time. This finding holds true for both the growth and inflation forecasts. According to the interpretation of disagreement as a measure of uncertainty, this would imply that forecast uncertainty is recently lower than it was before. While this viewpoint is in line with results stating that the volatility of the German business cycle has declined over time (see eg Fritsche and Kouzine 2003) it is not quite clear whether this matches the forecasting success.[12] Taking the Council of Economic Advisors as an example, Dicke and Glissmann (2003) argue that the magnitude of forecast errors may have even increased over time.[13] Though "always wrong, but never in doubt" is a famous quip on business cycle forecasters, we would still prefer the view that constant errors in forecasting contradict the notion of decreasing forecasting uncertainty.

---

[12] Over the same time span average growth and inflation rates declined as well. This could have repercussions on the dispersion of forecasts as well.

[13] Note, however, the related paper of Weidmann (2003) who criticises the methodological approach of Dicke and Glissmann (2003).

Moreover, though there is no obvious theoretical link between the dispersion and the stance of the business cycle it can be seen from the exhibit that the variance of the forecast is relatively high during recessions and early upswings (eg 1975, 1982-84, 1994, 1996). This points to the possibility that the risk aversion of forecasters may differ, since it is commonly agreed among forecasters that predicting a business cycle turning point is particularly difficult.

— Insert figure 5 here —

To further analyse the hypothesis of disagreement as a measure of uncertainty, we plot in figure 6 the dispersion of the forecasts against the absolute forecast errors of the consensus forecasts. If disagreement were indeed a useful measure of uncertainty one would expect a positive correlation with the subsequent forecast error. Besides the standard deviation of the forecasts we also take into account the skewness of the forecasts. A positive skewness implies a long right tail of the distribution of the forecasts. In other words: the vast majority of the forecasters remains sceptical but a considerable minority has more optimistic forecasts. Using scatter graphs we tried to figure out if there is any systematic relationship between the forecast errors on the one hand and the dispersion respectively the skewness on the other hand.

— Insert figure 6 here —

The results regarding growth forecasts suggest that there is a positive, but relatively weak correlation between the dispersion of the forecasts and the subsequent forecast error. This finding is in line with an interpretation of forecast dispersion as risk. The evidence based on price forecast is much less clear. Though the slope of the respective regression line in figure 6 is still positive it is far from being significant. This raises doubts on the usefulness of forecast dispersion as a measure of inflation risks. There is now obvious relationship in the data between the forecast error and the degree of scepticism or optimism in the forecaster community.

A further problem regarding the usage of disagreement as a measure of uncertainty in forecasting is the low magnitude of the dispersion. Some studies have already noted (see eg Linden 2003) that the degree of disagreement is much lower than other measures of uncertainty. For practical purposes it might therefore be better to rely on conditional heteroscedasticity or other more direct measures of uncertainty. In the following section we

focus on another interesting aspect ie the association between dispersion and the volatility of macroeconomic variables.

III.4 Forecast dispersion and the volatility of macroeconomic variables

This section is devoted to studying the circumstances under which forecasters disagree. The question is interesting because it is a common assumption underlying a lot of economic analyses and models that economic agents share identical beliefs on the structure of the economy and common data sets. Some recent evidence based on US data makes clear that this assumption is almost surely violated for consumers or firms (Caroll 2003, Mankiw, Reis and Wolfers 2003). Thus, this evidence might point to the relevance of so-called "sticky information" models (Mankiw and Reis 2002). Our data set, however, stems from professional forecasters. It is quite plausible to assume that the incentives for professional forecasters not to ignore any valuable information are strong enough. However, they might be other reasons to differ from the consensus. The possibility of a strategic behaviour may be considered since professional forecasters may look at public attention as well as at forecast accuracy (see eg Laster, Bennet and Geoum 1999). Since the forecasters included in this study are all part of the public sector, we would assume that such an explanation is less appealing in our case. We prefer to presume that the forecasters do not share a common belief or theory or model of how the economy works. Therefore, the following section aims to establish a set of stylised facts representing the circumstances under which forecasters have disagreed.

For this purpose, we estimate a measure of forecasters' disagreement as a function of volatility for important macroeconomic variables. If the forecasters disagree on the grounds of the diverging theories underlying the forecasts, one would expect to find a positive impact of macroeconomic volatility on forecasters' disagreement. Consider, for example, the ongoing debate on the impact of monetary policy on business cycle fluctuations. If the forecasters have different opinions on the quantitative impact of short-term interest rate changes on the cycle, they should particularly disagree in times of rapid changes of these interest rates. Quite in contrast, if the forecasters disagree for purely tactical reasons – eg to gain attention from the media – such a correlation is less likely. Thus, we estimate equations of the type $\sigma_t^f = \beta_0 + \beta_1 \sigma_{t-1}^{macro} + u_t$ where σ_t^f is the standard deviation of forecasts regarding a certain year

and σ_t^{macro} is the volatility of macroeconomic variables in the year when the forecasts have been done. We calculate the latter based on two different concepts: first, we estimate a GARCH (1,1) model for each variable under investigation.[14] To this end we use quarterly, seasonally adjusted data. Second, we simply use the standard deviation of the variable during the year in which the forecast was made. The macroeconomic variables were choosen according to the criterion that we expected forecasters with different opinions about the sources of business cycles would disagree about their effects. We consider the following macroeconomic variables to have a possible impact on forecaster disagreement (see Mankiw, Reis and Wolfers 2003 and Giordani and Söderlind 2003):

- The volatility of short-term interest rates. We expect this variable to reflect changes in the course of monetary policy.
- The volatility of the overall stock market index. This variable can be rationalised as a measure of overall uncertainty, but it will also reflect the burst of a stock market bubble.
- The volatility of oil prices. Oil price changes are probably an uncontroversial type of supply shock that drives the cycle.
- The volatility of labour productivity. Some researchers argue that unpredictable changes in labour productivity are the underlying reasons for business cycles. While this may or may not be true, the German research institutions certainly do not agree upon this. Thus, in the face of high productivity volatility they might disagree on the forecast to be made.
- The volatility of real wages. Wage developments represent an area of strong disagreement between forecasters. Therefore, the forecasters might disagree when the volatility of real wage changes is high.
- The stance of the business cycle measured by the OECD's output gap. According to Mankiw, Reis, and Wolfers (2003) forecasters' disagreement varies with the current economic situation. Even if the forecasters are in agreement as to which shocks are driving the cycle or they rely on the same long-run model, it is still possible that they will hold differing opinions as regards the concrete dynamics of the economy.

The results of the analysis are given in the tables 9 and 10. In a nutshell, it turns out that the volatility of macroeconomic activity is positively correlated to forecast dispersion. In the

[14] Some information on the estimations are given in appendix table A3, further details are available upon

case of growth forecasts the volatility of short-term interest rates is most closely correlated to the degree of disagreement among forecasters. Obviously, the institutions share no common view on the impacts of interest rate changes on real activity. Surprisingly, the volatility of stock markets returns does not contribute to the understanding of the variance of forecasts. While the forecasters might still disagree on what impact stock market disturbances have on the real economy, they might consider this impact to be weak or not systematic.

— Insert table 9 and 10 here —

Some tentative conclusions might be drawn from the estimations. First, even professional forecasters do not share a common belief on the structure of the economy. Thus, models implying such homogeneity of information miss an important stylised fact. Second, the correlation of forecast dispersion with the volatility of macroeconomic variables is – taken for itself – not supportive of the idea of strategic behaviour as the main source of disagreement. And, third, the assessment of the effects of monetary policy seems to be the prime suspect for diverging beliefs regarding an appropriate model of the economy.

IV. Conclusions

Based on 33 years of forecasts of professional forecasters stemming from 14 institutions we analyse business cycle forecasts in Germany. We extend existing studies, since we take into account all forecasters, which have contributed regularly to the German forecast market over more than 30 years, rather than referring to just one or two institutions. 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 tests on forecasting quality equivalence and encompassing. The results of these tasks confirm results of earlier studies: German business cycle forecasts are, on average, unbiased and weakly rational. In 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, adding therefore an additional piece of evidence to the existing findings. The results based on a panel confirm the findings found for individual forecasters.

---

request from the authors.

As regards the relative accuracy, we do not find systematic differences between the forecasters besides the date when the forecast was made. 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.

Regarding the dispersion of forecasts we present evidence suggesting that even professional forecasters do not share a common belief on the structure of the economy. Thus, models implying homogeneity of information miss an important stylised fact. Moreover, the correlation of forecast dispersion with the volatility of macroeconomic variables is not supportive of the idea of strategic behaviour as the dominant source of disagreement. The assessment of monetary policy seems to be the prime suspect for diverging beliefs regarding an appropriate model of the economy.

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 dispersion.

As regards forecast dispersion, it would be interesting to evaluate whether the result of a pronounced disagreement among individuals does not only hold true for professional forecasters, but also to firms and households. This may give additional insights into the process of information processing within the economy. Specifically, it should be investigated how information might trickle down (see Mankiw, Reis, and Wolfers 2003).

Furthermore, the study makes clear that the investigation of monetary policy effects still warrants further efforts since it is obvious that no consensus seems to be achievable among professionals concerning the effects.

References

Antholz, B (2003). Genauigkeit deutscher Wirtschaftsprognosen 1950-2002. Mimeo, Hamburg.

Arbeitsgemeinschaft wirtschaftswissenschaftlicher Forschungsinstitute. Zur Lage der Weltwirtschaft und der deutschen Wirtschaft. Berlin.

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, D.C.: 1–36.

Bomberger, WA (1996). Disagreement as a Measure of Uncertainty. *Journal of Money, Credit, and Banking* 28: 318-392.

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

Carroll, CD (2003). Macroeconomic Expectations of Households and Professional Forecasters. *Quarterly Journal of Economics* 118: 269-298.

Clarida, R, J Galí and M Gertler (2000). Monetary Policy Rules and Macroeconomic Stability. *Quarterly Journal of Economics* 115: 147-180.

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

Diebold, FX (1998). *Elements of Forecasting*. Cincinatti.

Diebold, FX and RS Mariano (1995). Comparing Predictive Accuracy. *Journal of Business and Economics Statistics* 13: 3-63.

Diebold, FX and J Lopez (1996). Forecast Evaluation and Combination. In: Maddalla, G.S.GS and CR Rao (Eds.). Handbooks of Statistics. 241-268.

Döpke, J and E Langfeldt (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.

Fritsche U and V Kouzine (2003). Declining Output Volatility and Monetary Policy in Germany. Mimeo. Berlin.

Gavin, WT (2003). FOMC forecasts: Is all the information in the central tendency? Federal Reserve Bank of St. Louis Working PaperNo. 2003-002A.

Giordani, P and P Söderlind (2003). Inflation Forecast Uncertainty. *European Economic Review* 47: 1037-1059.

Harvey, DI, Leybourne, SJ and Newbold, P (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting* 13: 281-291

Harvey, DI, Leybourne, SJ and Newbold, P (1998). Tests for forecast encompassing. *Journal of Business and Economic Statistics* 16: 254-259.

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

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

Heilemann, U and HO Stekler (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 and DA Peel (1990). On Testing for Unbiasedness and Efficiency of Forecasts. The Manchester Scholl 58: 120-127.

Keane, MP and DE Runkle (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 and M Savioz (2001). Monetary Policy and Forecasts for Real GDP Growth: An Empirical Investigation for the Federal Republic of Germany. *German Economic Review* 2: 339-365.

Kirchgaessner, G and UK Müller (2003). Are forecasters recluctant to Revise Their Predictions? Mimeo, Swiss Institute of International Economics and Applied Economics analysis.

Laster, D, P Bennet and IS Geoum (1999). Rational bias in macroeconomic forecasts. *Quarterly Journal of Economics* 114: 114: 293-318.

Linden, S (2003). Assessment of GDP forecasts uncertainty. European Commission Economic Papers 184. Brussels.

Mankiw, NG, R Reis, and J Wolfers (2003). Disagreement about Inflation Expectations. NBER Working Paper No. 9796. Cambridge MA.

Mankiw, NG and R Reis (2002). Sticky Information versus Sticky Prices: a Proposal to Replace the New Keynesian Phillips Curve. *Quarterly Journal of Economics* 117: 1295-1328.

Neumann, MJM and HS Buscher (1980). Die Inflationsprognosen der Arbeitsgemeinschaft wirtschaftswissenschaftlicher Forschungsinstitute: Sind sie rational? *Weltwirtschaftliches Archiv* 116: 533–550.

Newey, W, and K West (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-Wirtschafts-Konjunktur* 50: 1–11.

Rich, R and J Tracy (2003). Modelling uncertainty: Predictive Accuracy as a proxy for Predictive Confidence. Federal Reserve Bank of New York Staff Reports No. 161. New York.

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

Zarnowitz, V and LA Lambros (1987). Consensus and Uncertainty in Economic Prediction. *Journal of Political Economy* 95: 591- 621.

Table 1a: Measures of forecast accuracy — Growth forecasts

Institution	Mean error	Mean absolute error	Root mean squared error	Test for unbiased-ness (p-value)	Test for weak rationality (p-value)	Test for strong rationality based on short-term interest rates (coef. / p-value)	Test for strong rationality based on the real external value of the currency (coef. / p-value)	Test for strong rationality based on the change of the oil price (coef. / p-value)	Test for strong rationality based on OECD industrial production (coef. / p-value)
IMF, autumn	-0.75	1.44	1.89	0.02	0.97	-0.30/ 0.00	-4.38 / 0.11	-0.48 / 0.13	-0.09 / 0.09
Joint forecast, autumn	-0.46	1.24	1.69	0.12	0.81	-0.24 / 0.03	-5.54 / 0.27	-0.53 / 0.36	-0.08 / 0.28
Europ. commission, autumn	-0.49	1.26	1.72	0.10	0.86	-0.28 / 0.01	-3.79 / 0.31	-0.53 / 0.36	-0.10 / 0.22
Council of economic advisors	-0.52	1.09	1.53	0.05	0.96	-0.21 / 0.02	-6.07 / 0.10	-0.49 / 0.20	-0.07 / 0.15
OECD	-0.31	1.11	1.63	0.27	0.69	-0.24 / 0.07	-3.32 / 0.61	-0.58 / 0.62	-0.11 / 0.37
Employer's Institute	-0.28	1.13	1.46	0.28	0.11	-0.16 / 0.37	1.07 / 0.76	0.16 / 0.76	-0.08 / 0.51
Kiel Institute	-0.28	1.12	1.59	0.32	0.87	-0.20 / 0.12	-5.48 / 0.29	-0.51 / 0.48	-0.07 / 0.41
Hamburg institute	-0.25	1.04	1.42	0.31	0.42	-0.17 / 0.24	-4.39 / 0.55	-0.14 / 0.79	-0.06 / 0.61
Trade union' institute	-0.19	1.18	1.62	0.49	0.57	-0.23 / 0.09	-5.99 / 0.47	-0.29 / 0.69	-0.11 / 0.37
Munich institute	-0.14	1.00	1.36	0.56	0.34	-0.17 / 0.33	-2.73 / 0.86	-0.26 / 0.93	-0.07 / 0.64
Berlin Institute	0.03	1.02	1.40	0.91	0.92	-0.20 / 0.13	-3.40 / 0.71	-0.80 / 0.60	-0.07 / 0.59
Halle Institute	-0.53	0.62	0.91	0.10	0.58	-0.62 / 0.21	7.72 / 0.10	0.21 / 0.46	-0.14 / 0.36
Essen institute	-0.30	1.11	1.41	0.22	0.94	-0.21 / 0.06	-2.12 / 0.57	-0.30 / 0.57	-0.06 / 0.46
Government's economic report	-0.28	0.98	1.43	0.26	0.30	-0.17 / 0.24	-3.96 / 0.57	0.00 / 0.74	-0.05 / 0.61
IMF, spring	-0.20	0.91	1.24	0.36	0.17	-0.15 / 0.15	-3.15 / 0.49	-0.36 / 0.56	-0.02 / 0.61
European commission, spring	-0.24	0.85	1.14	0.22	0.56	-0.15 / 0.12	-2.18 / 0.67	-0.27 / 0.59	-0.04 / 0.54
Joint forecast, spring	-0.26	0.88	1.21	0.21	0.76	-0.16 / 0.13	-1.73 / 0.64	-0.56 / 0.49	-0.05 / 0.47

See text for details. Shaded areas denote a violation of the hypothesis of rational expectation (at least at the 10 % level), ie of no bias or of weak or of strong efficiency, respectively.

Table 1b: Measures of forecast accuracy — Inflation forecasts

Institution	Mean error	Mean absolute error	Root mean squared error	Test for weak rationality (p-value)	Test for strong rationality based on short-term interest rates (coef. / p-value)	Test for strong rationality based on the real external value of the currency (coef. / p-value)	Test for strong rationality based on the change of the oil price (coef. / p-value)	Test for strong rationality based on OECD industrial production (coef. / p-value)
IMF, autumn	-0.15	0.70	0.95	0.40	-0.04 / 0.54	3.21 / 0.42	0.96 / 0.09	0.01 / 0.56
Joint forecast, autumn	0.02	0.84	1.09	0.90	0.04 / 0.31	2.74 / 0.20	0.94 / 0.08	0.08 / 0.09
Europ. Commission, autumn	0.08	0.74	0.98	0.66	0.03 / 0.25	1.87 / 0.21	0.90 / 0.05	0.0.7 / 0.07
Council of economic advisors	0.06	0.75	0.99	0.75	-0.01 / 0.15	5.25 / 0.04	0.96 / 0.02	0.04 / 0.09
OECD	-0.08	0.58	0.79	0.58	0.05 / 0.75	2.91 / 0.65	0.73 / 0.28	-0.00 / 0.86
Employer's Institute	-0.05	0.55	0.77	0.70	-0.07 / 0.19	2.90 / 0.20	0.09 / 0.31	-0.01 / 0.31
Kiel Institute	0.19	0.84	1.22	0.36	0.09 / 0.26	1.63 / 0.19	1.19 / 0.06	0.06 / 0.18
Hamburg institute	0.05	0.70	0.97	0.75	-0.05 / 0.43	4.90 / 0.21	0.52 / 0.33	0.03 / 0.41
Trade union' institute	0.00	0.85	1.11	1.00	0.04 / 0.94	4.75 / 0.67	1.08 / 0.29	0.05 / 0.78
Munich institute	-0.03	0.64	0.89	0.83	-0.04 / 0.75	3.81 / 0.46	1.12 / 0.77	0.01 / 0.79
Berlin Institute	0.20	0.73	0.96	0.24	0.03 / 0.27	2.50 / 0.26	0.98 / 0.03	0.06 / 0.09
Halle Institute	-0.29	0.36	0.52	0.12	0.32 / 0.09	1.71 / 0.29	0.26 / 0.27	-0.02 / 0.32
Essen institute	0.11	0.62	0.87	0.45	-0.02 / 0.02	5.01 / 0.00	1.11 / 0.00	0.04 / 0.01
Government's economic report	0.16	0.64	0.92	0.31	-0.03 / 0.62	3.39 / 0.50	0.16 / 0.63	0.02 / 0.61
IMF, spring	-0.19	0.61	0.86	0.22	-0.08 / 0.35	6.34 / 0.06	0.52 / 0.33	-0.03 / 0.46
European commission, spring	-0.06	0.38	0.48	0.48	-0.01 / 0.67	0.37 / 0.66	-0.43 / 0.25	-0.01 / 0.67
Joint forecast, spring	0.04	0.46	0.60	0.71	-0.03 / 0.59	3.36 / 0.29	-0.42 / 0.58	-0.02 / 0.48

See text for details. Shaded areas denote a violation of the hypothesis of rational expectation (at least at the 10 % level), ie of no bias or of weak or of strong efficiency, respectively.

Table 2: Test for forecast rationality - panel evidence

Institution	β_o	β_1	β_2	Test for rationality
		Growth forecasts		
OLS estimation	-0.116	0.877	(-)	25.41***
	(-0.68)	(13.12)***		
GLS estimation	-0.120	0.887	(-)	23.59***
	(-0.73)	(13.67)***		
GLS estimation with cross sectional correlation	0.129	1.572	(-)	383.54***
	(2.88)***	(6.96)***		
OLS, with short-term interest rates	1.512	0.673	-0.202	61.64***
	(4.71)***	(9.22)***	(-5.89)***	
OLS, with real exchange rate	1.874	0.899	-2.209	26.82***
	(1.37)	(13.23)***	(-1.46)	
OLS, with oil price changes	-0.100	0.889	-0.003	19.04***
	(1.30)	(12.96)***	(0.89)	
OLS, with industrial production abroad	-0.106	0.903	-0.063	32.52***
	(-0.63)	(13.58)***	(-2.73)***	
		Inflation forecasts		
OLS estimation	-0.439	1.099	(-)	21,.84***
	(-4.49)***	(38.28)***		
GLS estimation	-0.446	1.103	(-)	23.00***
	(-4..66)***	(39.12)***		
GLS estimation with cross sectional correlation	1.315	0.537	(-)	110.74***
	(8.01)***	(12.22)***		
OLS, with short-term interest rates	-0.474	1.084	0.013	21.83***
	(-4.06)***	(33.47)***	(0.69)	
OLS, with real external value	-4.183	1.035	4.277	49.95***
	(-5.73)***	(34.40)***	(5.19)***	
OLS, with oil price changes	-0.472	1.089	0.014	117.62***
	(-5.10)***	(37.91)***	(9.51)***	
OLS, with industrial production abroad	-0.459	1.099	0.008	21.76***
	(-4.24)***	(36.87)***	(0.64)	

z-values in brackets. *** (**,*) denotes rejection of the null hypothesis at the 1 (5, 10) percent level.

Table 3a: Diebold / Mariano tests, growth forecasts

	IMF, autumn	Joint forec. aut.	EC aut.	Council Eco. Advisors.	OECD	Kiel Inst.	Hamb. Institute	Trade Union Institute	Munich Institute	Berlin Institute	Institute of German Econ	Halle Institute	Essen Institute	Government's economic report	IMF, spring	European Commission,	Joint forecast, spring
IMF, autumn	X	0.10	0.23	0.02	0.10	0.07	0.00	0.12	0.01	0.05	0.04	0.03	0.01	0.00	0.01	0.00	0.00
Joint forecast, autumn	0.10	X	0.81	0.06	0.55	0.14	0.04	0.41	0.07	0.19	0.27	0.08	0.13	0.01	0.05	0.02	0.04
European Commission autumn	0.23	0.81	X	0.19	0.21	0.32	0.07	0.52	0.07	0.18	0.23	0.06	0.09	0.03	0.04	0.02	0.04
Council of Economic Advisors	0.02	0.06	0.19	X	0.43	0.56	0.21	0.23	0.22	0.41	0.62	0.48	0.43	0.15	0.09	0.04	0.06
OECD	0.10	0.55	0.21	0.43	X	0.66	0.18	0.92	0.16	0.32	0.38	0.06	0.22	0.11	0.09	0.04	0.07
Kiel Institute	0.07	0.14	0.32	0.56	0.66	X	0.24	0.78	0.23	0.42	0.57	0.96	0.37	0.16	0.14	0.07	0.11
Hamburg Institute	0.00	0.04	0.07	0.21	0.18	0.24	X	0.07	0.39	0.87	0.86	0.70	0.91	0.87	0.14	0.03	0.13
Trade Union Institute	0.12	0.41	0.52	0.23	0.92	0.78	0.07	X	0.09	0.22	0.33	0.80	0.26	0.05	0.04	0.02	0.04
Munich Institute	0.01	0.07	0.07	0.22	0.16	0.23	0.39	0.09	X	0.78	0.38	0.71	0.63	0.57	0.14	0.02	0.16
Berlin Institute	0.05	0.19	0.18	0.41	0.32	0.42	0.87	0.22	0.78	X	0.79	0.28	0.93	0.86	0.09	0.05	0.05
Institute of German Economy	0.04	0.27	0.23	0.62	0.38	0.57	0.86	0.33	0.38	0.79	X	0.09	0.82	0.98	0.04	0.00	0.05
Halle institute	0.03	0.08	0.06	0.48	0.06	0.96	0.70	0.80	0.71	0.28	0.09	X	0.33	0.47	0.26	0.35	0.29
Essen institute	0.01	0.13	0.09	0.43	0.22	0.37	0.91	0.26	0.63	0.93	0.82	0.33	X	0.88	0.08	0.00	0.06
Government's economic report	0.00	0.01	0.03	0.15	0.11	0.16	0.87	0.05	0.57	0.86	0.98	0.47	0.88	X	0.25	0.11	0.23
IMF, spring	0.01	0.05	0.04	0.09	0.09	0.14	0.14	0.04	0.14	0.09	0.04	0.26	0.08	0.25	X	0.09	0.80
European Commission, spring	0.00	0.02	0.02	0.04	0.04	0.07	0.03	0.02	0.02	0.05	0.00	0.35	0.00	0.11	0.09	X	0.34
Joint forecast, spring	0.00	0.04	0.04	0.06	0.07	0.11	0.13	0.04	0.16	0.05	0.05	0.29	0.06	0.23	0.80	0.34	X

Hypothesis: forecast in column is equal to the forecast in row.

Table 3b: Diebold / Mariano tests, inflation forecasts

	IMF, autumn	Joint forec. aut.	EC aut.	Counc. Eco. Advis. OECD	Kiel Institute	Hamb. Institute	Trade Union Institute	Munich Institute	Berlin Institute	Institute of German Econ.	Halle Institute	Essen Institute	Government's economic report	IMF, spring	European Commission,	Joint forecast, spring
IMF, autumn	X	0.53	0.58	0.76	0.40	0.49	0.46	0.41	0.28	0.46	0.02	0.76	0.29	0.14	0.47	0.03	0.05
Joint forecast, autumn	0.53	X	0.08	0.30	0.15	0.12	0.05	0.78	0.01	0.09	0.01	0.84	0.10	0.05	0.16	0.00	0.00
European Comm., autumn	0.58	0.08	X	0.96	0.32	0.04	0.77	0.06	0.17	0.59	0.13	0.85	0.23	0.28	0.51	0.00	0.01
Council of Econ. Advisors	0.76	0.30	0.96	X	0.13	0.11	0.78	0.15	0.20	0.68	0.11	0.23	0.14	0.39	0.48	0.00	0.01
OECD	0.40	0.15	0.32	0.13	X	0.27	0.51	0.05	0.72	0.39	0.54	0.49	0.81	0.49	0.67	0.05	0.09
Kiel Institute	0.49	0.12	0.04	0.11	0.27	X	0.05	0.37	0.02	0.04	0.03	0.54	0.04	0.02	0.14	0.00	0.01
Hamburg Institute	0.46	0.05	0.77	0.78	0.51	0.05	X	0.02	0.14	0.90	0.08	0.56	0.40	0.54	0.48	0.01	0.02
Trade Union Institute	0.41	0.78	0.06	0.15	0.05	0.37	0.02	X	0.01	0.07	0.01	0.55	0.04	0.03	0.10	0.00	0.00
Munich Institute	0.28	0.01	0.17	0.20	0.72	0.02	0.14	0.01	X	0.25	0.24	0.92	0.91	0.68	0.89	0.03	0.05
Berlin Institute	0.46	0.09	0.59	0.68	0.39	0.04	0.90	0.07	0.25	X	0.19	0.85	0.36	0.56	0.62	0.00	0.00
Institute of German Economy	0.02	0.01	0.13	0.11	0.54	0.03	0.08	0.01	0.24	0.19	X	0.54	0.81	0.51	0.43	0.09	0.22
Halle Institute	0.76	0.84	0.85	0.23	0.49	0.54	0.56	0.55	0.92	0.85	0.54	X	0.26	0.32	0.26	0.28	0.39
Essen Institute	0.29	0.10	0.23	0.14	0.81	0.04	0.40	0.04	0.91	0.36	0.81	0.26	X	0.51	1.00	0.00	0.03
Government's econ. report	0.14	0.05	0.28	0.39	0.49	0.02	0.54	0.03	0.68	0.56	0.51	0.32	0.51	X	0.76	0.01	0.02
IMF, spring	0.47	0.16	0.51	0.48	0.67	0.14	0.48	0.10	0.89	0.62	0.43	0.26	1.00	0.76	X	0.13	0.24
European Comm., spring	0.03	0.00	0.00	0.00	0.05	0.00	0.01	0.00	0.03	0.00	0.09	0.28	0.00	0.01	0.13	X	0.05
Joint forecast, spring	0.05	0.00	0.01	0.01	0.09	0.01	0.02	0.00	0.05	0.00	0.22	0.39	0.03	0.02	0.24	0.05	X

Hypothesis: forecast in column is equal to the forecast in row.

Table 4a: Tests on forecast encompassing, growth forecasts

	IMF, aut	Joint forec. Aut.	EC aut.	Counc. Eco. Advis. OECD	Kiel Inst.	Hamb Institute	Trade Union Institute	Munich Institute	Berlin Institute	Institute of German Econ.	Halle Institute	Essen Institute	Government's economic report	IMF, spring	European Commission	Joint forecast, spring
IMF, autumn	X	0.57	0.68	0.19	0.88	0.56	0.16	0.90	0.30	0.97	0.40	0.08	0.54	0.10	0.82	0.42	0.34
Joint forecast, autumn	0.03	X	0.22	0.55	0.58	0.86	0.10	0.55	0.20	0.73	0.62	0.13	0.51	0.06	0.60	0.35	0.24
European Comm. autumn	0.10	0.50	X	0.73	0.97	0.24	0.39	0.57	0.43	0.94	0.80	0.12	0.56	0.35	0.58	0.34	0.40
Council of Econ. Advisors	0.00	0.00	0.04	X	0.04	0.04	0.93	0.02	0.91	0.65	0.60	0.95	0.60	0.85	0.69	0.49	0.38
OECD	0.02	0.21	0.03	0.70	X	0.28	0.62	0.44	0.44	0.99	0.85	0.14	0.69	0.64	0.86	0.58	0.33
Kiel Institute	0.02	0.03	0.08	0.34	0.11	X	0.80	0.05	0.60	0.89	0.75	0.61	0.81	1.00	0.95	0.79	0.44
Hamburg Institute	0.00	0.02	0.02	0.03	0.05	0.05	X	0.00	0.66	0.44	0.39	0.42	0.27	0.44	0.92	0.71	0.76
Trade Union Institute	0.02	0.06	0.06	0.50	0.13	0.12	0.78	X	0.62	0.88	0.96	0.25	0.55	0.74	0.94	0.68	0.44
Munich Institute	0.00	0.03	0.02	0.04	0.06	0.09	0.21	0.01	X	0.29	0.10	0.42	0.07	0.36	0.49	0.85	0.95
Berlin Institute	0.00	0.05	0.03	0.07	0.09	0.14	0.22	0.02	0.40	X	0.04	0.10	0.04	0.30	0.92	0.94	0.65
Institute of German Econ.	0.02	0.15	0.09	0.21	0.21	0.23	0.66	0.10	0.94	0.21	X	0.18	0.29	0.57	0.41	0.93	0.82
Halle Institute	0.03	0.08	0.07	0.23	0.06	0.74	0.91	0.67	0.53	0.90	0.06	X	0.13	0.23	0.51	0.94	0.67
Essen Institute	0.00	0.05	0.01	0.10	0.07	0.07	0.23	0.04	0.25	0.26	0.10	0.90	X	0.14	0.73	0.46	0.60
Government's econ. report	0.00	0.01	0.01	0.01	0.03	0.01	0.69	0.00	0.88	0.49	0.50	0.87	0.25	X	0.97	0.74	0.78
IMF, spring	0.00	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.02	0.00	0.01	0.16	0.00	0.05	X	0.30	0.33
European com., spring	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.02	0.03	X	0.13
Joint forecast, spring	0.00	0.01	0.01	0.01	0.02	0.03	0.01	0.00	0.01	0.00	0.00	0.14	0.00	0.06	0.05	0.41	X

Hypothesis: forecast in column encompasses the forecast in row.

Table 4b: Tests on forecast encompassing, inflation forecasts

	IMF, aut	Joint forec. Aut.	EC aut.	Counc. Eco. Advis.	OECD	Kiel Institute	Hamb. Institute	Trade Union Institute	Munich Institute	Berlin Institute	Institute of German Econ.	Halle Institute	Essen Institute	Government's economic report	IMF, spring	European Commission	Joint forecast, spring
IMF, autumn	X	0.03	0.47	0.27	0.89	0.08	0.34	0.06	0.43	0.32	0.72	0.56	0.55	0.67	0.38	0.98	0.69
Joint forecast, autumn	0.33	X	0.24	0.79	0.78	0.03	0.63	0.12	0.79	0.64	0.55	0.14	0.64	0.72	0.26	0.45	0.66
European Comm. autumn	0.12	0.03	X	0.20	0.92	0.02	0.32	0.01	0.76	0.30	0.92	0.41	0.84	0.48	0.17	0.44	0.83
Council of Ec. Advisors	0.06	0.04	0.24	X	0.80	0.03	0.46	0.02	0.77	0.25	0.86	0.30	0.47	0.51	0.12	0.59	0.91
OECD	0.13	0.03	0.08	0.01	X	0.07	0.10	0.01	0.14	0.02	0.43	0.20	0.28	0.40	0.19	0.83	0.55
Kiel Institute	0.38	0.48	0.13	0.61	0.83	X	0.47	0.58	0.62	0.24	0.62	0.98	0.22	0.87	0.23	0.57	0.78
Hamburg Institute	0.03	0.01	0.10	0.25	0.51	0.02	X	0.00	0.66	0.32	0.53	0.27	0.69	0.72	0.18	0.14	0.24
Trade Union Institute	0.32	0.28	0.56	0.89	0.48	0.07	0.64	X	0.56	0.93	0.41	0.29	0.67	0.82	0.27	0.38	0.83
Munich Institute	0.01	0.01	0.01	0.02	0.23	0.01	0.02	0.00	X	0.07	0.83	0.46	0.22	0.28	0.17	0.08	0.37
Berlin Institute	0.04	0.01	0.03	0.09	0.12	0.01	0.09	0.01	0.48	X	0.57	0.15	0.62	0.27	0.06	0.70	0.96
Institute of German Economy	0.00	0.00	0.01	0.02	0.08	0.02	0.02	0.00	0.07	0.04	X	0.06	0.19	0.11	0.07	0.47	0.92
Halle Institute	0.41	0.56	0.58	0.19	0.91	0.26	0.98	0.96	0.36	0.35	0.79	X	0.17	0.85	0.71	0.85	0.54
Essen Institute	0.02	0.02	0.07	0.05	0.14	0.02	0.07	0.00	0.17	0.04	0.19	0.68	X	0.09	0.07	0.91	0.67
Government's econ. report	0.01	0.01	0.02	0.04	0.03	0.01	0.15	0.00	0.67	0.11	0.57	0.08	0.38	X	0.11	0.10	0.36
IMF, spring	0.08	0.02	0.06	0.04	0.36	0.03	0.07	0.01	0.10	0.04	0.34	0.06	0.08	0.14	X	0.65	0.44
European Comm., spring	0.01	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.02	0.00	0.06	0.06	0.00	0.00	0.06	X	0.01
Joint forecast, spring	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.06	0.12	0.01	0.00	0.07	0.41	X

Hypothesis: forecast in column encompasses the forecast in row.

Table 5: 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 6: Evaluating direction of changes forecasts

	Growth			Inflation		
	I	C	p-value	I	C	p-value
Joint forecast, autumn	1.36	5.32	0.021	1.43	9.29	0.002
Joint forecast, spring	1.57	12.27	0.000	1.74	20.10	0.000
Council of economic advisors	1.29	3.58	0.058	1.49	12.02	0.001
Kiel Institute	1.39	6.07	0.014	1.49	12.02	0.001
Berlin Institute	1.57	13.51	0.000	1.53	17.48	0.000
Hamburg Institute	1.34	4.58	0.032	1.49	12.02	0.001
Munich Institute	1.44	8.05	0.005	1.38	7.06	0.008
Essen Institute	1.39	6.07	0.014	1.43	11.03	0.001
Halle Institute	1.75	3.94	0.047	1.40	2.77	0.096
OECD	1.29	3.58	0.058	1.64	18.97	0.000
Employer's institute	1.29	3.58	0.058	1.37	8.30	0.004
Trade union's institute	1.39	6.07	0.014	1.43	11.03	0.001
Government economic report	1.39	6.07	0.014	1.50	14.25	0.000
European commission, autumn	1.36	5.32	0.021	1.56	15.25	0.000
European commission, spring	1.51	9.79	0.002	1.75	22.44	0.000
IMF, autumn	1.45	6.87	0.009	1.66	18.55	0.000
IMF, spring	1.30	6.02	0.014	1.55	18.24	0.000

Shaded areas denotes that the hypothesis of independence between forecast and realization cannot be rejected (at 5 % level).

Table 7: Measures of forecast accuracy — consensus forecasts

	Mean error	Mean absolute error	Root mean squared error	Test for weak rationality (p-value)	Test for strong rationality based on short-term interest rates (coef. / p-value)	Test for strong rationality based on the real external value of the currency (coef. / p-value)	Test for strong rationality based on the change of the oil price (coef. / p-value)	Test for strong rationality based on OECD industrial production (coef. / p-value)
				Growth forecasts				
Consensus forecast, mean	-0.32	1.06	1.47	0.21	-0.20 / 0.12	-4.45 / 0.47	-0.46 / 0.57	-0.09 / 0.30
Consensus forecast, median	-0.35	1.13	1.55	0.19	-0.23 / 0.06	-5.52 / 0.39	-0.63 / 0.50	-0.10 / 0.31
				Inflation forecasts				
Consensus forecast, mean	0.07	0.69	0.94	0.66	-0.02 / 0.16	2.15 / 0.11	0.70 / 0.05	0.04 / 0.09
Consensus forecast, median	0.08	0.68	0.96	0.64	-0.01 / 0.18	3.10 / 0.12	0.76 / 0.05	0.04 / 0.12

See text for details. Shaded areas denote a violation of the hypothesis of a rational expectation (at least at the 10 % level), ie of no bias or of weak or of strong efficiency, respectively.

Table 7: Directional forecast errors: majority vs minority

Growth forecasts (number of minority reports: 12)		
Error type	Majority	Minority
II	10	8
IJ	7	4
JI	12	15
JJ	4	6
II/(II+IJ) + JJ / (JJ+JI)	1.34	1.38
Inflation forecasts (number of minority reports: 15)		
II	9	3
IJ	3	7
JI	17	13
JJ	4	10
II/(II+IJ) + JJ / (JJ+JI)	1.56	0.86

Table 9: Forecast dispersion and macroeconomic volatility — Growth forecasts

Exogenous variable	Constant	Slope	R^2
Volatility of short-term interest rates, GARCH	0.37 (9.82)***	0.13 (4.14)**	0.47
Volatility of short-term interest rates, rate of change	0.33 (7.68)***	0.25 (6.07)***	0.34
Volatility of stock-market returns, GARCH	0.64 (7.26)***	-27.12 (-2.73)*	0.12
Volatility of stock-market returns, rate of change	0.62 (6.79)***	-2.56 (-2.29)	0.09
Volatility of oil price changes, GARCH	0.37 (9.82)***	0.13 (4.14)***	0.47
Volatility of price changes, rate of change	0.56 (8.24)***	-0.15 (-0.22)	0.00
Volatility of productivity, GARCH	0.34 (6.25)***	1671 (3.59)***	0.33
Volatility of productivity, rate of change	0.34 (4.82)***	24.71 (2.70)**	0.23
Volatility of real wages, GARCH	0.17 (0.73)	0.20 (1.39)	0.09
Volatility of real wages, rate of change	0.26 (2.73)**	22.26 (2.51)*	0.15
Volatility real exchange rates, GARCH	0.25 (3.11)***	746.97 (4.19)***	0.26
Volatility of real exchange rates, rate of change	0.27 (4.08)***	16.99 (5.40)***	0.23
Inflation rate, level	0.23 (4.27)***	0.08 (4.90)***	0.53
Output Gap	0.45 (9.51)***	0.02 (2.53)**	0.12

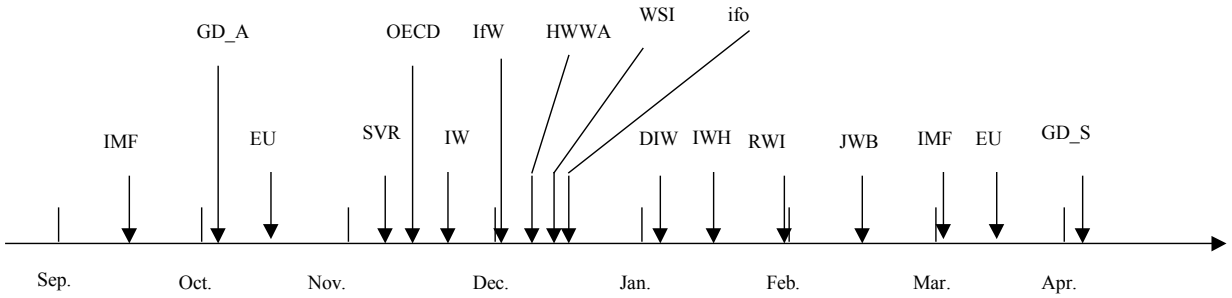
t-values in brackets. *** (**,*) denotes rejection of the null hypothesis at the 1 (5, 10) percent level.

Table 10: Forecast dispersion and macroeconomic volatility — Inflation forecasts

Exogenous variable	Constant	Slope	R^2
Volatility of short-term interest rates, GARCH	0.25 (11.70)***	0.09 (6.12)**	0.53
Volatility of short-term interest rates, rate of change	0.22 (6.27)***	0.19 (3.54)**	0.45
Volatility of stock-market returns, GARCH	0.39 (7.96)***	-9.38 (-1.99)*	0.04
Volatility of stock-market returns, rate of change	0.37 (6.99)***	-0.69 (-1.17)	0.02
Volatility of oil price changes, GARCH	0.25 (11.70)***	0.09 (6.12)***	0.53
Volatility of price changes, rate of change	0.27 (5.87)***	0.77 (1.21)	0.13
Volatility of productivity, GARCH	0.29 (7.02)***	559.45 (1.80)*	0.09
Volatility of productivity, rate of change	0.29 (7.99)***	7.80 (1.81)	0.05
Volatility of real wages, GARCH	0.28 (3.03)***	0.04 (0.73)	0.01
Volatility of real wages, rate of change	0.32 (4.31)***	1.82 (0.32)	0.00
Volatility real exchange rates, GARCH	0.18 (4.32)***	470.82 (4.39)***	0.24
Volatility of real exchange rates, rate of change	0.18 (6.01)***	11.59 (4.35)***	0.26
Inflation rate, level	0.20 (3.57)***	0.04 (2.41)**	0.33
Output Gap	0.29 (15.92)***	0.02 (2.25)**	0.27

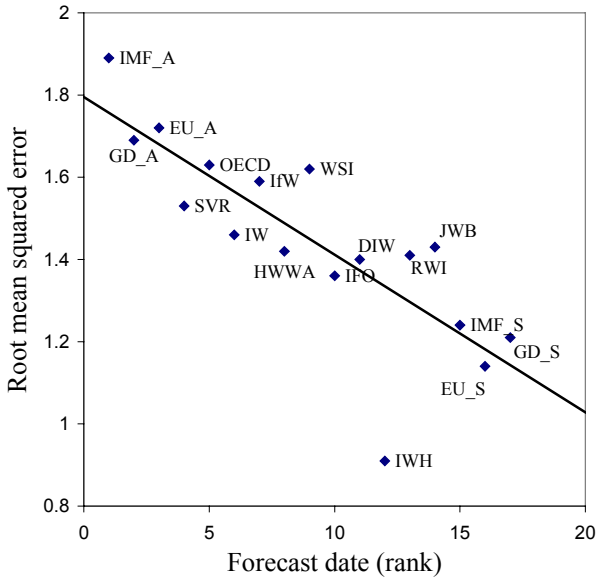
t-values in brackets. *** (**,*) denotes rejection of the null hypothesis at the 1 (5, 10) percent level.

Figure 1: Forecasting Institutions and Dates

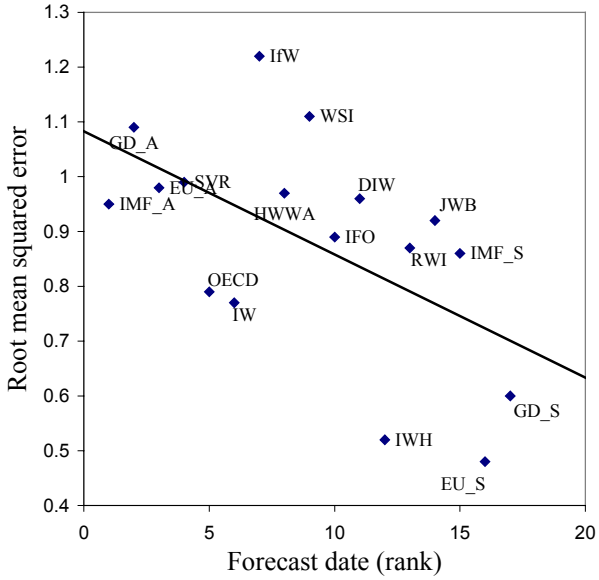


Notes: IMF: International Monetary Fund, GD_A: joint forecast, autumn, EU: European Commission, SVR: Council of Economic Advisors, IW: Institute of the German Economy (IW), OECD: Organization for Economic Co-operation and Development, Paris, IfW: Kiel Institute for World Economics, HWWA: HWWA Institute 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.

Figure 2: Date of forecast and forecast error



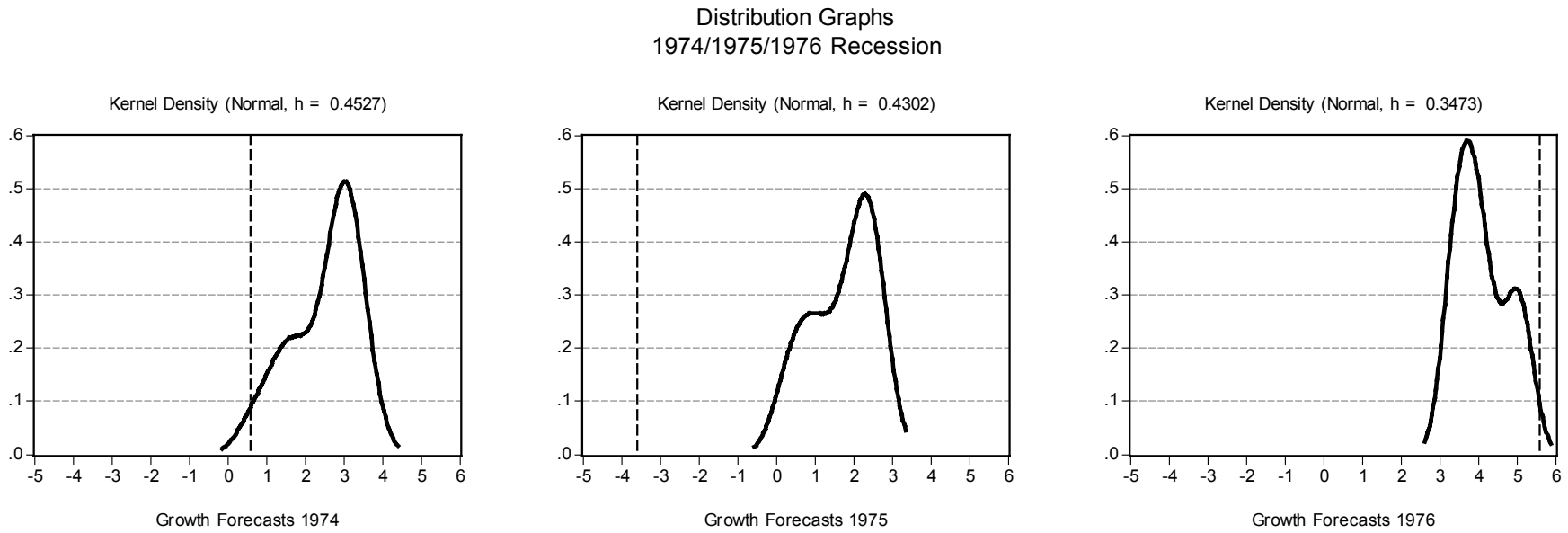
Panel (a): Growth forecasts



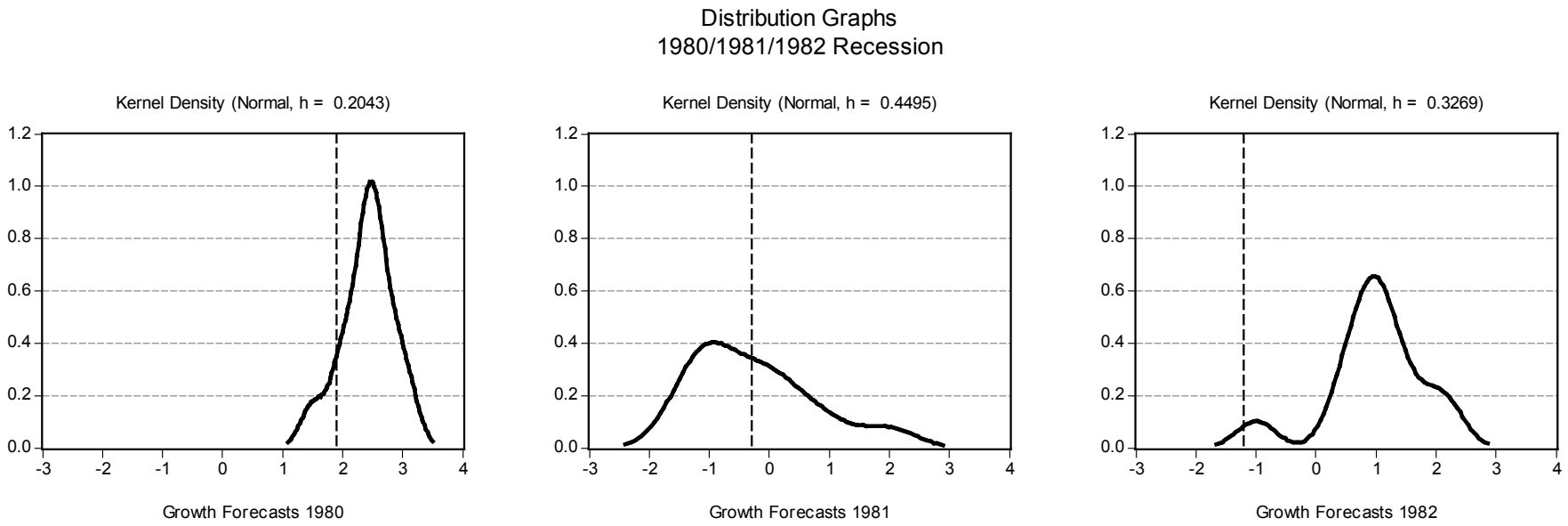
Panel (b): Inflation forecasts

Notes: 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.

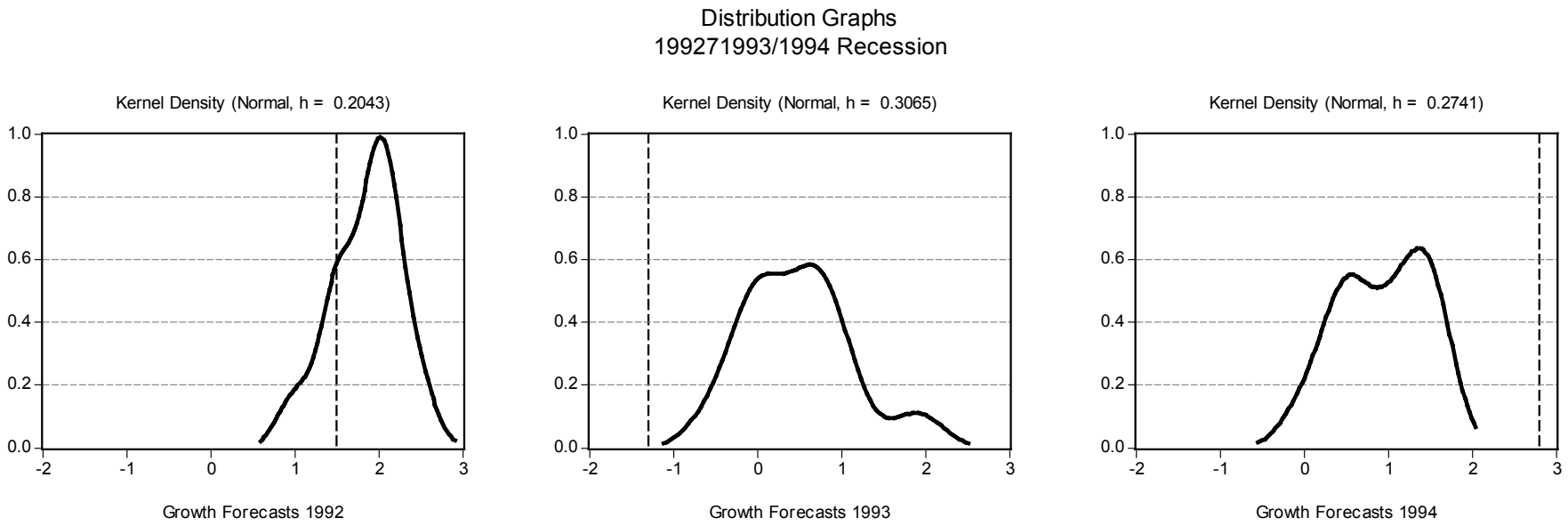
Figure 3: Distribution of forecasts in selected periods — Growth forecasts



Panel (a): The 1974/75 recession



Panel (b): The 1981/82 recession



Panel (c): The 1993 recession

Figure 4: Distribution of forecasts in selected periods — Inflation forecasts

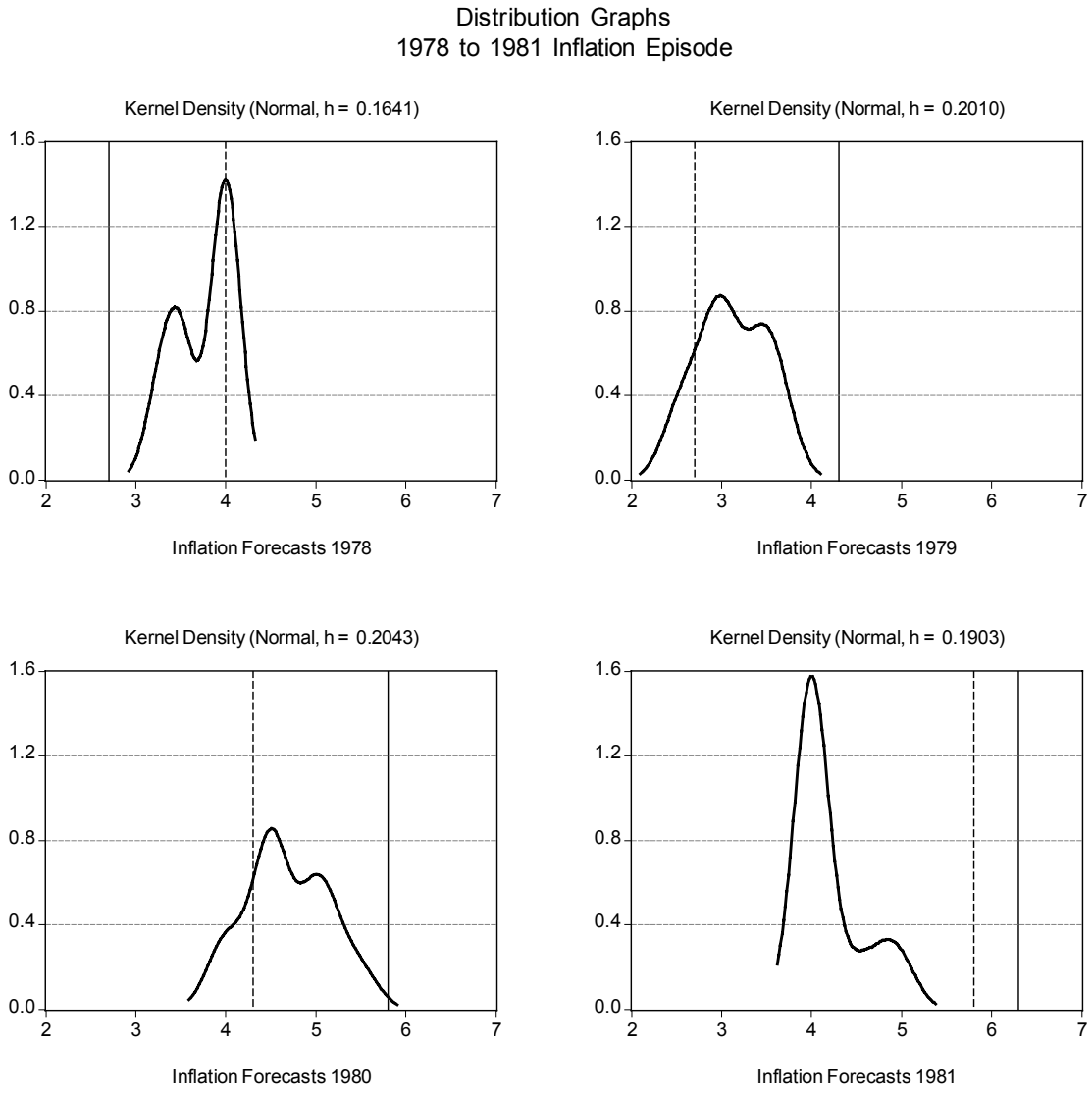


Figure 5: Forecast dispersion 1970 to 2003

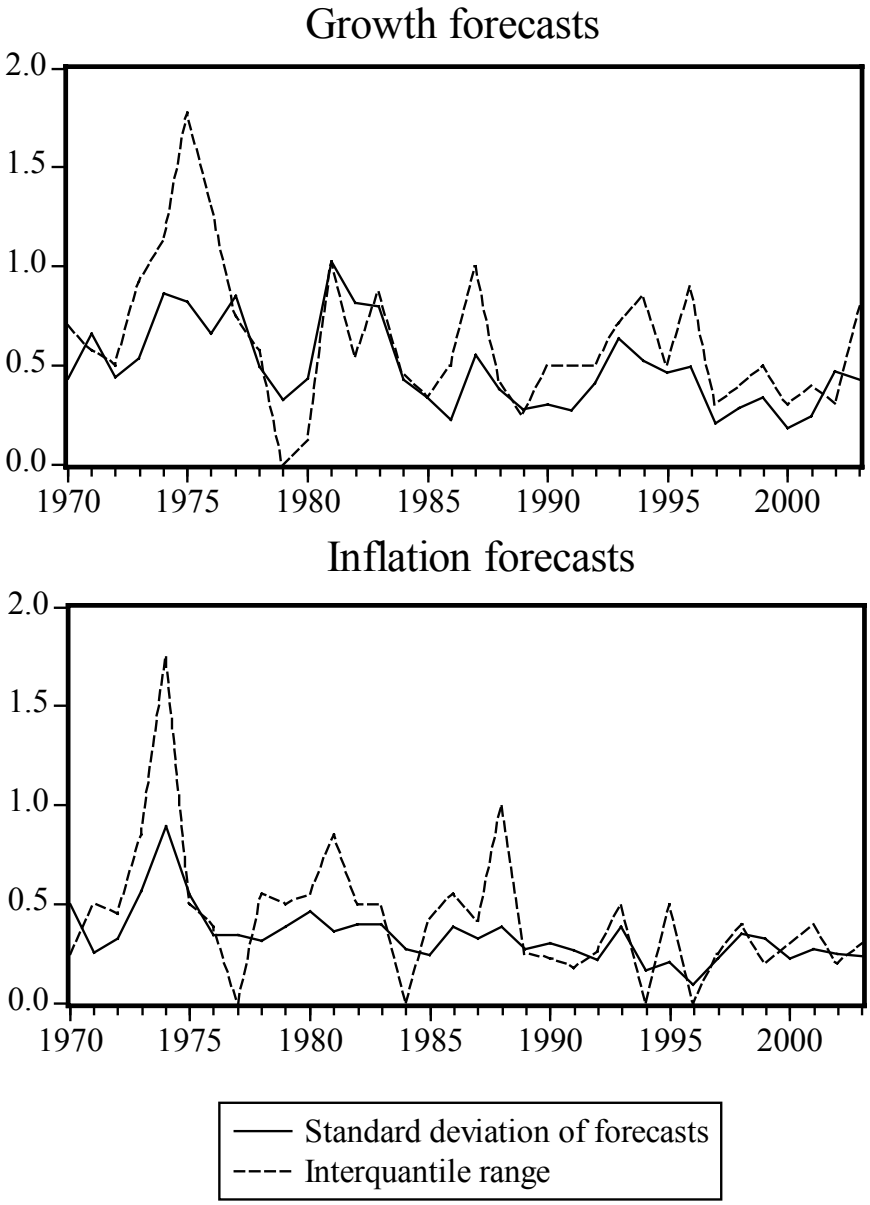
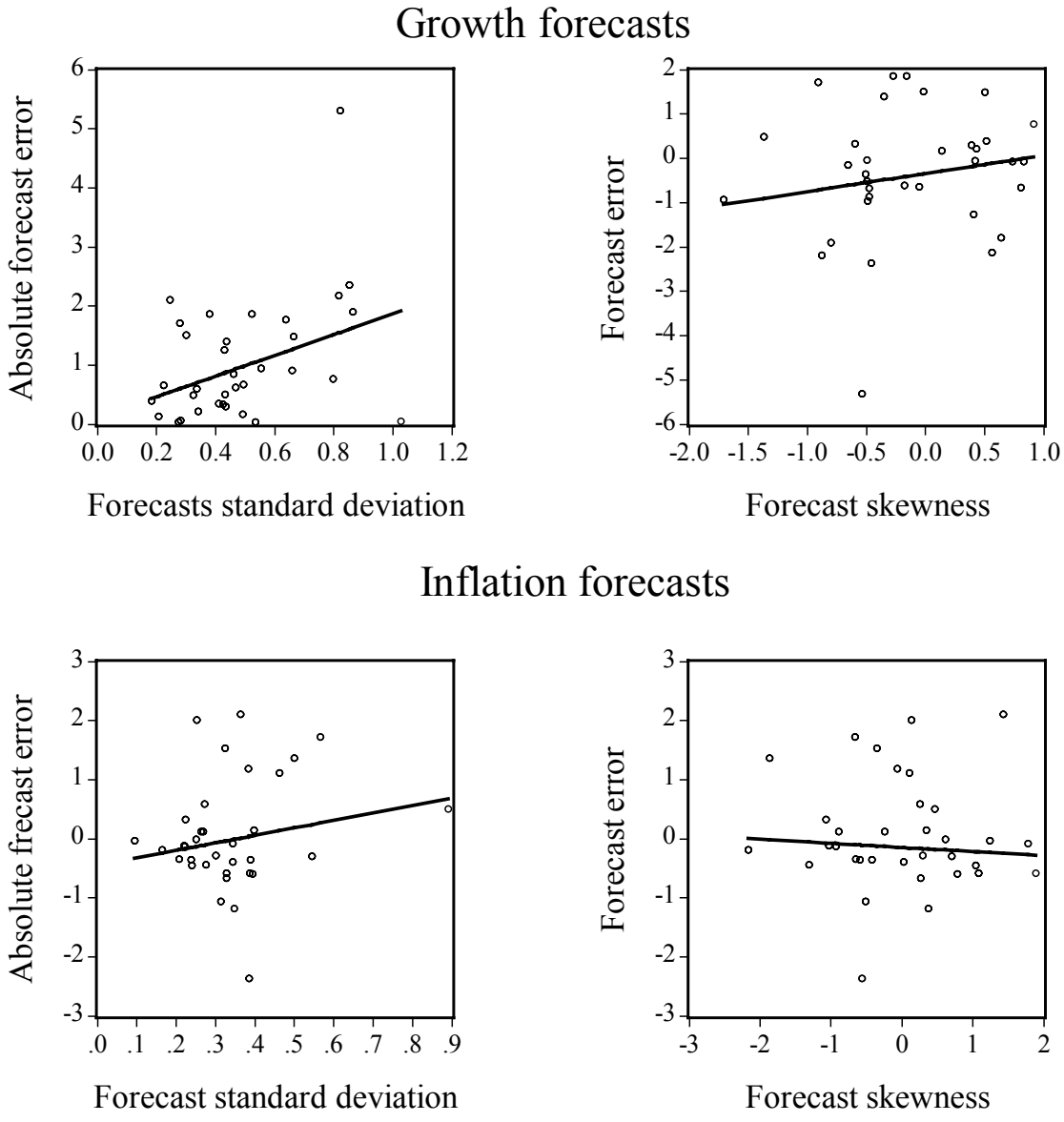


Figure 6: Forecasts' standard deviation, forecasts' skewness, and forecast errors



Appendix Table A1: Growth forecasts under investigation

	Actual outcome	Joint forecast, autumn	Joint forecast, spring	Council of economic advisors	Kiel Institute	Berlin Institute	Hamburg Institute	Munich Institute	Essen institute	Halle Institute	OECD	Employer's Institute	Trade union's Institute	Governme nt's economic	European commissio n, autumn	European commssio n, spring	IMF, autum	IMF, spring
1970	4.8	5	5	4.5	4	4.1	4	4.3	5		4.5		5.2	4.5	4.4	5.8		
1971	2.7	4	3.5	4	4	3.5	4	4	4 a)		3		2	3.5	3.7	3.4		3.3
1972	3	1	2.5	2	2	1.5	1.5	2	1 a)		2	1.5	2	2.5	1	1.8		2
1973	5.3	5	6	6	5.5	4.5	5	5	6.3		5.3	5.5	6	4.5	5	5.9	5	5.8
1974	0.6	3	2.5	2.5	3 a)	3	1.5	1.5	3 a)		3.3	0.75	2	1	3.5	2	3	2
1975	-3.6	2.5	0.5	2	2.5 a)	0.5	1.5	1	1		2.5	0.3	2	2	2.5	0.5	2.2	0.5
1976	5.6	4	5.5	5	4	5	4.5	4	5.2		3.3	3.8	3.5	4.5	3.5	6	3.5	6.8
1977	2.4	5.5	4.5	5.5	5	5.5	5	4.5	3.25		3.5	4.7	6	5	3.9	4	4.8	4.5
1978	3.4	3	2.5	4	3.5	3	3	3	2.5		3.3	2.5	3.5	3.5	4	2.7	3.4	3.1
1979	4.4	4	4	4.25	4	4	4	4	3.1		4	4.25	3.7	4	3.5	4	4	4
1980	1.9	2.5	2.5	2.75	1.5	2	2.5	2	2.5		2.3	2.5	2.4	2.5	2.9	2.5	3.1	2.5
1981	-0.3	0	-1.5	0.5	-1.5	-1	-0.5	-1	-1		-0.3	-1.2	0.1	-0.5	1	-0.7	2	-0.9
1982	-1.2	1	0.5	0.5	-1	0.5	1	1	1		1.3	1.5	0.8	1.25	2.2	1	2	1
1983	1	0	0.5	1	0	-0.5	-0.5	-0.5	-0.5		-0.3	0.5	0.5	0	1	0.6	2	0.5
1984	2.6	2	3	2.5	2.5	2	2.5	2.25	3		2.5	2.5	1.3	2.5	2.1	3	2	2.6
1985	2.6	2	2.5	3	2	2	2.5	2.25	2.25		2.8	2.75	2.1	2.5	2.5	2.5	2.4	2.7
1986	2.5	3	3.5	3	3	3	3	3	3.5		3.3	3.5	3	3	3.5	3.5	3.1	3.7
1987	1.6	3	2	2	3	1.5	2.5	2.25	2		3	3	2.2	2.5	3.2	1.5	3	1.9
1988	3.4	2	2	1.5	1.5	1	1.5	1	1.75		1.5	1	1.6	1.75	1.9	2.1	2.1	1.7
1989	4	2	3	2.5	2.5	2.5	2.25	2.25	2.25		2.5	2.5	1.7	2.5	2.5	3	1.9	2.4
1990	4.6	3	3.75	3	2.5	3.5	3	3	3.5		3.2	3	2.8	3	3.5	3.7	3	3.5
1991	3.2	3	3	3.5	3.5	3.5	3	3.25	3.5		3	3.5	2.7	3.25	3.1	2.7	3.3	2.8
1992	1.5	2	1.5	2.5	1.5	1	2	1.5	2		1.8	1.5	2.2	1.75	2.2	2	2	1.3
1993	-1.3	1	0.5	0.5	0.5	-0.5	0	0	0		0.7	0.75	0.9	0	0	-2	1.9	-2
1994	2.8	1.5	0.5	0.5	1.5	0	1.5	1.5	1		0.4	1	0.6	1.25	0.5	1.3	1.2	0.5
1995	1.9	3	2.5	3	3.5	2	3	3	2.5		2	3	2.3	3	3	3	2.8	3.2
1996	1.4	2.5	0.75	2	1.7	1	2	1.75	2	1.6	2.4	2.25	2.5	1.5	2.4	0.5	2.9	1
1997	2.2	2.5	2.25	2.5	2.5	2	2.5	2.5	2.5	2.1	2.2	2.5	2	2.5	2.2	2.3	2.4	2.3
1998	2.8	2.8	2.6	3	2.8	2.5	2.8	2.6	3.5	2.7	3	3	2.5	2.2	3.2	2.6	2.8	2.5
1999	1.5	2.3	1.7	2	2.3	1.4	2.1	1.7	2.8	2.2	2.2	2	1.9	2	2.2	1.7	2.3	1.5
2000	3	3	2.8	2.7	2.5	2.4	2.7	2.7	2.6	2.8	2.3	2.5	2.5	2.5	2.6	2.9	2.5	2.8
2001	0.6	2.7	2.1	2.8	2.4	2.5	2.7	2.7	2.8	2.7	2.7	2.9	2.3 b)	2.75	2.8	2.2	3.3	1.9
2002	0.2	1.3	0.9	0.7	1.2	0.6	0.7	0.6	0.8	0.6	1	1	-0.2	0.75	0.7	0.8	1.8	0.9
2003	-0.1	1.4	0.5	1	1	0.6	0.7	1.1	0.6	1.1	1.5	1.75	1	1	1.4	0.4	2	0.5

a) Taken from joint forecast, autumn of the respective year. b) Most pessimistic institute at the end of the year minus 0.1 percentage point.

Appendix Table A2: Inflation forecasts under investigation

	Actual outcome	Joint forecast, autumn	Joint forecast, spring	Council of economi	Kiel Institute	Berlin Institute	Hamburg Institute	Munich Institute	Essen institute	Halle Institute	OECD	Employer's Institute	Trade union's Institute	Government's economi	European commiss	European commssi	IMF, autum	IMF, spring
1970	4.1	2.5	4	3	1.5	2.7	3	3.1	3				3	3	2.8	3.8		
1971	5.5	3.5	4.5	3.5	3	3.5	3.5	4	3.5				3.5	3	3.4	4.5		6.3
1972	6	4.5	5	4.5	4	4.5	4.5	5	4.5			4.8	3.9	4.5	4.5	5		5.5
1973	7.4	5	6.5	6.2	4.5	6	5.75	6	6.5			6	5.5	5.75	5.5	7	5.5	6.5
1974	7.3	6	8.5	6.5	5.5	6	7.5	8.5	6		7	8	7	8.5	6.5	8.5	7	8
1975	6.1	6.5	5.5	5.7	6.5	6	6	6	6		7	6	7.5	6	6.5	6	7	6.3
1976	4.7	5.5	4.5	5	5	5.5	5	5	4.7		4.8	5	5.5	4.75	5.5	5	4.5	3.8
1977	4	4.5	4	4	4	4	4	4	4		4	3.8	5	4	4	4	3.7	3.7
1978	2.7	4	3	3.5	4	4	3.5	3.5	3.25		4	3.3	4	3.5	4	3.3	4	3.5
1979	4.3	3.5	3.5	2.5	3.5	3	3	3.5	2.6		2.8	3.3	3	3	3	4	3.7	4
1980	5.8	4.5	5.5	4.5	4.5	5	4.5	5	4		5	5.5	4	4.5	4.5	5.5	5.2	5
1981	6.3	4	5	4	4	4	4	4	5		4	4.8	4	4.5	4	5.8	4.5	5
1982	5	4.5	4.5	5.5	4.5	5	5	5	5		4.3	4.9	5.5	5	4.5	4.7	4.5	4.5
1983	3.2	3.5	3	4	3.5	3.5	4	3.8	3.5		3.8	3.3	4.5	4	3.6	3	4.5	4
1984	2.5	3	3	3	2.5	3	3	3	3		3.3	2.3	3	3	3.2	2.9	3	3
1985	1.7	2	2.5	2	2	2	2	2	2.5		2	2.5	2	2	2.2	2.4	2.6	2.2
1986	-0.6	2	0.5	1.5	2	1.5	2	2	1		1.3	2	2	1.5	1.5	0	2.3	2.6
1987	0.4	1.5	0.5	1	1	1	1	0.75	0.5		0.8	1	1.5	0.5	1.1	0.8	1.6	2.5
1988	1.3	2	1	1.5	2	1	1	1.5	1.5		1.8	1.5	2	1	1.8	1	2.2	2
1989	3	2.5	3	2	3	2.5	2.5	2.5	2.5		2.25	2	2.5	2.25	2.5	3	2.2	2.5
1990	2.7	3	3	3	3.5	3	3	3.5	3		2.6	3	3	2.5	2.7	2.8	2.5	2.9
1991	3.7	4	3.5	3.5	3.5	3	3.5	3.5	3.5		3.9	3.5	3.5	3.5	3.9	3.5	3.6	3.9
1992	3.8	4	3.75	4	4	3.5	4	4	4		4	3.5	4	3.5	4.2	3.9	3.7	4.3
1993	3.2	4.5	4.5	3.5	3.5	4	4.5	4	3.5		3.5	3.5	3.5	3.5	3.6	4	3.7	3.9
1994	2.8	3	3	3	3	3	3	3	3		3	3	2.5	3	3.2	3	3.1	3
1995	2	2.5	2	2.5	2.5	2.5	2.5	2.5	2		2.2	2.5	2	2	2.2	2.3	2.2	2
1996	2	2	1.5	2.25	2	2	2	2	2	2.2	2	2	2	2	2.1	1.6	1.9	1.7
1997	1.9	1.5	1.5	1.75	1.8	1	1.5	1.5	1.75	1.9	1.5	1.5	1.5	1.5	1.7	1.9	1.6	1.5
1998	0.9	2	1.3	2.75	2.5	1.8	1.8	2.2	2.5	2.1	1.9	2	1.8	1.9	2.2	1.7	1.5	1.6
1999	0.8	1.2	0.7	1.75	1.2	1.2	1.5	1.2	1.3	1.3	1.2	1.25	1.5	1	1.1	0.7	2.3	0.6
2000	1.2	1.6	1.5	1.5	1.6	1.1	1.4	1.2	1.3	1.3	1.4	1.5	1.4	1.25	1.2	1.5	0.8	1.2
2001	1.8	1.7	2.1	2	1.7	1	2	1.5	2	1.7	1.6	1.8	1.6	1.5	1.6	1.8	1.5	2
2002	1.4	1.5	1.5	1.5	1.1	1.4	1.3	1.5	2	1.2	1	1.5	1.4	1.5	1.6	1.7	1.3	1.5
2003	1	1.6	1.2	1.6	0.9	1	1.3	1.3	1.6	1.6	1.4	1.5	1.3	1.5	1.4	1.3	1.1	1

a) Taken from joint forecast, autumn of the respective year. b) most pessimistic institute at the end of the year minus 0.1 percentage point.

Appendix table A3 — Testing the GARCH Models for selected macro variables

Variable	Model specification of AR process	ARCH-term	GARCH-term
(1)	(2)	(5)	(6)
Short term interest rates	C, AR(1)	0.32 (2.81)***	0.66 (7.31)***
Stock market returns	C, AR(1)	0.17 (2.12)**	0.78 (8.05)***
Change of real exchange rates	C, AR(1)	0.19 (2.17)**	0.56 (2.88)***
Change of real wages	C	0.06 (1.49)	0.78 (4.27)***
Change of oil prices	C	1.70 (5.09)***	(-)
Change of industrial production in OECD countries	C, AR(1)	0.13 (3.00)***	0.86 (23.48)***

The number in brackets are z-statistics for a test whether the ARCH or GARCH coefficient is equal to zero. — ***(**,*) denotes rejection of the null hypothesis at the 1 (5, 10) percent level.