

Döhrn, Roland; Schmidt, Christoph M.

Working Paper

Information or Institution? – On the Determinants of Forecast Accuracy

Ruhr Economic Papers, No. 201

Provided in Cooperation with:

RWI – Leibniz-Institut für Wirtschaftsforschung, Essen

Suggested Citation: Döhrn, Roland; Schmidt, Christoph M. (2010) : Information or Institution? – On the Determinants of Forecast Accuracy, Ruhr Economic Papers, No. 201, ISBN 978-3-86788-228-6, Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI), Essen

This Version is available at:

<https://hdl.handle.net/10419/45323>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



RUHR

ECONOMIC PAPERS

Roland Döhrn
Christoph M. Schmidt

Information or Institution?

On the Determinants of Forecast Accuracy

Imprint

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Prof. Dr. Thomas K. Bauer
RUB, Department of Economics, Empirical Economics
Phone: +49 (0) 234/3 22 83 41, e-mail: thomas.bauer@rub.de

Prof. Dr. Wolfgang Leininger
Technische Universität Dortmund, Department of Economic and Social Sciences
Economics – Microeconomics
Phone: +49 (0) 231/7 55-3297, email: W.Leininger@wiso.uni-dortmund.de

Prof. Dr. Volker Clausen
University of Duisburg-Essen, Department of Economics
International Economics
Phone: +49 (0) 201/1 83-3655, e-mail: vclausen@vwl.uni-due.de

Prof. Dr. Christoph M. Schmidt
RWI, Phone: +49 (0) 201/81 49-227, e-mail: christoph.schmidt@rwi-essen.de

Editorial Office

Joachim Schmidt
RWI, Phone: +49 (0) 201/81 49-292, e-mail: joachim.schmidt@rwi-essen.de

Ruhr Economic Papers #201

Responsible Editor: Christoph M. Schmidt

All rights reserved. Bochum, Dortmund, Duisburg, Essen, Germany, 2010

ISSN 1864-4872 (online) – ISBN 978-3-86788-228-6

The working papers published in the Series constitute work in progress circulated to stimulate discussion and critical comments. Views expressed represent exclusively the authors' own opinions and do not necessarily reflect those of the editors.

Ruhr Economic Papers #201

Roland Döhrn and Christoph M. Schmidt

Information or Institution?

On the Determinants of Forecast Accuracy



Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über:
<http://dnb.d-nb.de> abrufbar.

ISSN 1864-4872 (online)
ISBN 978-3-86788-228-6

Roland Döhrn and Christoph M. Schmidt¹

Information or Institution? – On the Determinants of Forecast Accuracy

Abstract

The accuracy of macroeconomic forecast depends on various factors, most importantly the mix of analytical methods used by the individual forecasters, the way that their personal experience is shaping their identification strategies, but also their efficiency in translating new information into revised forecasts. In this paper we use a broad sample of forecasts of German GDP and its components to analyze the impact of institutions and information on forecast accuracy. We find that forecast errors are a linear function of the forecast horizon. This result is robust over a variety of different specifications. As better information seems to be the key to achieving better forecasts, approaches for acquiring reliable information early seem to be a good investment. By contrast, the institutional factors tend to be small and statistically insignificant. It has to remain open, whether this is the consequence of the efficiency-enhancing competition among German research institutions or rather the reflection of an abundance of forecast suppliers.

JEL Classification: C53, E27, E01

Keywords: Forecast accuracy, Forecast Revisions, Forecast Horizon, Economic Activity

September 2010

¹ Roland Döhrn, RWI and University of Duisburg-Essen; Christoph M. Schmidt, RWI, Ruhr-Universität Bochum, IZA Bonn, CEPR London. – We are grateful for comments by Thomas Bauer, Wim Kösters, and by two anonymous referees. – All correspondence to Roland Döhrn, RWI, Hohenzollernstr. 1-3, 45128 Essen, Germany, E-Mail: roland.doehrn@rwi-essen.de.

1. Introduction

Rankings of forecasters have gained popularity in Germany, at least since the *Financial Times Deutschland* has started to elect the “Forecaster of the Year” each December. However, a systematic analysis of the quality of forecasts and its determinants is still missing. Already the definition of forecast quality raises some important questions: Is it really sufficient to look only at *ex post*-forecast accuracy, i.e. the difference between outcome and prediction? More concretely, should we focus on GDP only or should we take a broader view and look also at the main components of GDP such as private consumption, investment etc.? Certainly, we would be hesitant to award a prize to an accurate forecast of GDP which combined large, yet offsetting forecast errors for its components.

And how should we appreciate the analytical and methodological foundations of a forecast? Is it sufficient to pick the right figure just by chance or should it be based on a solid background? A forecast that reflected the available information poorly at the time of its construction might be vindicated by surprises afterwards, but this does not make its originator a reliable forecaster. These questions demonstrate that one may not do justice to forecasters when reducing their work to one figure and evaluating them based on it.

Against this background this paper takes a broader view on forecast accuracy and its determinants. We do not only include GDP but also other relevant macroeconomic variables in our analysis, and focus on the role of the forecasting horizon and differences in institutional performance. Specifically, we use a large sample of real-time forecasts to find out whether some forecasting institutions tend to produce better forecasts than others. Our sample covers the years 1991 to 2008 which is a moderate time span, comprising two major recessions 1992/93 and 2001. Identifying institutional aspects is difficult, because forecast accuracy is heavily influenced by forecasting horizon, and forecasters systematically differ regarding the week when a forecast is completed.

It is obvious that new information about the state of the economy and its important components, such as interest rates, exchange rates or commodity prices, keeps arriving day after day, making forecasting easier for those who hold back their publication a while longer. Yet,

early forecasts are valuable for decision makers in business and for policy makers alike, and the different suppliers of forecasts therefore choose the date of publication in a trade-off between accuracy and novelty.

But timing is not the only aspect that sets different forecasting institutions apart. True, all professional forecasters use the same data and quite similar techniques. Therefore, the institutional factor can be expected to be small. At the same time, the personal experience of forecasters may enable them to combine the bits of information they collected more or less efficiently. And sometimes, also the bureaucratic procedures inside institutions may have an influence, e.g. when a forecast has to be approved by a governing council, as is typically the case for international institutions. Moreover, an important feature of all relevant professional forecasting institutions is that they do not rely simply on a single method but combine various techniques either in the sense that they employ different methods for different questions, or that the outcome of one method is assessed and potentially modified in the light of the results of another one¹.

The use of different methods is particularly evident for different time horizons. The forecast process mostly starts with an estimate, a “nowcast”, for the current quarter, and sometimes also for the quarter which just has ended, in the case official figures are not available yet. For these quarters, many indicators are already available which allow a more or less accurate estimate of GDP and its components. This part of the forecasting process which is sometimes extended to one quarter ahead provides the starting point for predicting the following quarters. For the remaining forecast horizon various techniques are combined, including time series as well as structural models, and also comparisons with reference cycles. All these different predictions are made consistent by linking them through the definitions of the system of national accounts.

As creating a forecast is a rather complex endeavour, it is difficult if not impossible to evaluate forecasts analytically, i.e. to assess the performance of the various techniques employed and the efficiency of information processing. To capture the empirical manifestation of the second aspect, we make use of the fact that forecasters receive a continuous stream

¹ An example for the interaction of different forecasting methods is provided by Nierhaus/Sturm 2003.

of data influencing their prediction. As the forecast horizon is shrinking, the information set is unequivocally growing. Therefore, we would expect a rather strong negative relation between forecast horizon and accuracy, which we, indeed, can find in our data. The institution-specific deviations from this trend line are small which on the one hand confirms that no forecaster commands a superior technique, and on the other hand allows some more general conclusions about the art of forecasting.

The paper is organized as follows. Section 2 offers a short review of the literature. To separate between institutional factors and the impact of information on forecast accuracy, our database contains, besides the forecasts of German GDP and its main demand side components, also the dates when the forecasts were constructed. A more detailed description of the database is given in section 3. Section 4 presents our analyses of forecast accuracy. After an inspection of the data, we at first assess the accuracy of GDP forecasts and, thereafter, we also investigate predictions of the main GDP components. In section 5, we draw some conclusions for improving forecast accuracy.

2. Review of the Literature

Rational forecasts are constructed by processing all available information about the state of the economy and all the time paths of its elements leading up to this state. While this general principle is universal, the concrete way how this information is processed is decisive for the forecast's accuracy. Quite obviously, it is undesirable to construct forecasts which could evidently be improved upon by using pieces of the available information in a different manner. One strand of the forecast evaluation literature addresses this question of information efficiency by devising concrete testing strategies (for an overview: Stekler 2002).

One test is provided by Isiklar et al. (2006) who use the autocorrelation of forecast revisions as a measure of "weak" information efficiency. If revisions are autocorrelated, they argue, it seems evident that forecasters have not used the full set of information available in their original forecast. Applying this approach to the Consensus Forecasts for 18 countries – each presenting an average of quite a large number of forecasts – they conclude that forecasts are not efficient, and that Germany belongs to the countries with the least efficient fore-

casts. Yet, the Consensus Forecasts collate predictions of quite different nature. Some forecasters adjust their predictions every month, whereas others only don't replace their forecasts for several months at a time. The German economic research institutes, e.g., which are also included in the Consensus, present new forecasts every three months, reflecting the publication rhythm of the quarterly national accounts. Hence, the inefficiency Isiklar et al. (2006) found may be in part an aggregation problem.

Moreover Isiklar et al. (2006) acknowledge that the published data that form the information set for the forecasts are typically revised thereafter and, thus, the inefficiency they describe may be a result of bad data quality. Incidentally, Öller/Teterukovsky (2007) even turn the perspective around and use the accuracy of forecasts as a measure of data quality. They argue that a statistical variable can hardly be forecasted accurately if the available information does not reflect previous values of the variable well. Similarly, Isiklar/Lahiri (2007) take the improvement of forecast accuracy as a measure of the information content of new data. By contrast, we shift our focus to the individual providers of the forecasts, recording in particular the dates of the finalization of individual forecasts.

The role of the forecaster on the accuracy of the forecast has only received little attention in the literature. Heilemann/Klinger (2005) analyzed in an international comparison the impact of the structure of the market for forecasts, measured by the number of forecasters, on the "market result" in terms of forecast accuracy. However, they could not find clear evidence that the forecast errors are smaller in countries with a more intense competition. Döpke/Fritsche (2006a, 2006b) analyzed the accuracy of growth and of inflation forecasts in a panel data set which comprised forecasts by some 14 German institutions. They illustrated a negative correlation of accuracy and the length of the period that had to be forecasted. However, they did not explicitly separate forecast horizon and institutional performance. Our analysis, by contrast, aims at gauging institutional performance net of any differences in forecast horizon.

Forecasts are constructed by researchers, not by lifeless automata. Consequently, they might be influenced by individual motives and personal or institutional characteristics. Most prominently, Lamont (2002) evaluated the accuracy of predictions provided by a large num-

ber of individual and institutional forecasters for the U.S., finding that the probability to make an outlier forecast increases with the age of the forecaster. he explained this behavior arguing that only experienced and well established forecasters dare presenting outlier forecasts which will increase their reputation, if they turn out to be right, but may do harm to the career if they are wrong. However, for the institutions included in his sample he did not find a bias in this direction. Hence, in our analysis one should expect the institutional factors to be small.

3. The Data Base

Our analysis comprises nearly two decades of data for the major professional institutions concerned with forecasting German economic activity. During a course of slightly more than two years all these institutions have published several forecasts for any given year, respectively. For the year 2008, e.g., the 'forecast season' was opened in the autumn of 2006, when the EU published its first forecast for this year. Subsequently the EU provided another two forecasts in the course of 2007, one in spring and one in autumn, and two further forecasts in 2008. Thus, the first forecast had a horizon of 10 quarters, i.e. the last quarter for which national accounts data were available at that time was the second quarter of 2006. Subsequent forecasts moved ever more closely to the end of 2008. Finally, the fifth and last forecast constructed in autumn 2008 had a forecast horizon of only two quarters.

Other forecasters provide a different number of forecasts. In particular RWI, Essen, and the Institute for World Economics, Kiel, synchronized their forecasts in the recent years with the releases of new GDP figures, resulting in four publications per year, which makes up eight forecasts for a given year. However, publication rhythms and forecast horizons have altered in the course of time, and they are sometimes changed occasionally².

The database only contains forecasts which were available for all years between 1991 and 2008. Due to this restriction, some forecasts had to be skipped. For instance, the spring forecasts of the Gemeinschaftsdiagnose and of RWI for the subsequent year were excluded from

² In such turbulent times as the recent recession, many international institutions publish interim forecasts to get a more up-to-date picture of the economic situation.

the analysis, as they are available only since the mid-1990s. For the same reason, institutions are missing which only entered the forecasting business during the 1990s, i.e. the Institut für Wirtschaftsforschung, Halle, as well as those who went out of business, i.e. HWWA. Furthermore, only institutions were included which provide information on all positions of the national accounts which are considered here. This is the reason why the International Monetary Fund forecasts have been skipped. In the end we analyze 32 forecasts from eight institutions, each providing 18 observations (Table 1) which consequently form a balanced sample.

=====

Table 1 about here

=====

In the database the forecasts of the year-over-year growth of GDP are collected as well as those of its demand side components: private consumption, government consumption, gross fixed capital formation, investment in equipment and machinery, investment in buildings, exports, and imports. Some forecasts also contain data on the contributions of net-exports to growth. As far as these data were not provided directly in the publications, they were derived from the export and import forecasts³.

As we focus in our paper on the efficiency of information processing, a very important part of the database is precise information on the dates when the forecasts were finalized. In quite the majority of forecasts in our database, the day when the forecast was completed is mentioned explicitly. In others, the editorials of the forecast are dated, and it seems logical that they were finalized close to this date. As far as no explicit reference is made, the publication date of the volume which contains the forecast provides some indication. If only the month of publication is known, which is the fact in rare cases, the middle of this month is regarded as the publication date.

³ The contribution of real net exports (ne) to growth was calculated as

$$ne_t = EXP_{t-1}/GDP_{t-1} * g(exp_t) - IMP_{t-1}/GDP_{t-1} * g(imp_t).$$

with EXP, IMP and GDP being *nominal* exports, imports and GDP (in bn Euro resp. DM) and $g(exp)$ resp. $g(imp)$ being the growth rates of *real* exports and imports. The nominal values were taken from the *Gemeinschaftsdiagnose* published in spring. This procedure may lead to wrong results for the second year of a two-step forecast, as in this case not the actual values of EXP, IMP and GDP should be taken, but the forecasted values. However, as these data are not at hand for many forecasts, we have to rely on these calculated figures. Our experience from cases in which figures on net exports were published shows, that the error implied by our procedure tends not to be very large.

During the period under study, German unification took place. This created a considerable challenge to all forecasters, given the tremendous uncertainty about the true value of East German GDP and its growth potential. Until 1994, most institutions provided separate projections for West and for East Germany. To compare the comparable, our database contains West German figures for the years 1991 to 1994 and data for unified Germany thereafter. However, due to unification, the data on export and import forecasts may be contaminated to some extent for the years leading up to 1994, as West German figures include intra-trade with Eastern Germany to be consistent with the System of National Accounts. The intensity of this trade was difficult to measure because of the lack of any statistics on that issue.

The results of forecast evaluations, of course, depend a lot on the choice of the figures that are considered as realizations. National accounts data are revised several times, and the revisions may be considerable, in particular when they reflect conceptual changes in the System of National Accounts. If these changes had already been known when the predictions were made, they would have altered the relevant information sets and, thus, have led to different forecasts. For the sake of a fair comparison, only those national accounts figures should be used as references which were released as close as possible to the forecast, to make sure that prediction and realization follow the same conceptions⁴. Therefore, we will always compare the forecasts under real time conditions using the figures as a yardstick that were published by the German Statistical Office in February of the year which followed the year the forecast was made for. These data were taken from the *Gemeinschaftsdiagnose* published in the spring of each year.

4. Accuracy of GDP Forecasts Over Time

4.1. Inspection of the Data

We start our analyses by an inspection of our data. In Chart 1 the errors of the forecasts included are plotted against the length of the forecast horizon expressed in calendar years. It becomes evident that there is a strong correlation between production date and forecast

⁴ In Germany, e.g., revisions tend to make recessions look less severe with every revision of the system of national accounts (Räth 2009).

accuracy in some years, whereas no clear trend can be discovered in many others. The relationship is particularly strong in years that mark turning points in the business cycle (1993, 1994, 2001). But also in 2002 and 2003, when an upswing was forecasted that did not occur, and in 2006 and 2007, when the upswing turned out to be stronger than expected, forecast errors become significantly smaller over time. For the remaining years, forecast errors often are generally low, and additional information apparently confirmed the forecasts which were made earlier.

=====

Chart 1 about here

=====

Retaining our focus on forecast horizons and individual institutions, Table 2 presents summary statistics on forecast accuracy. Across all institutions, mean absolute errors (MAE), mean squared errors (MSE) and root mean squared errors (RMSE) all decline visibly as the forecasting horizon, here expressed in average days, becomes smaller. For instance, the MAE tends to fall from approximately 1.2 percentage points for a forecast horizon of 7 quarters to 0.5 percentage points or less for a horizon of just 3 quarters. The comparison of MAE and RMSE which can serve as a rough measure for the contribution of outliers to the forecast errors, exhibits large differences in particular for long forecast horizons. This makes evident that a good deal of the errors results from large deviations in a small number of years. For shorter forecast horizons there are small differences between MAE and RMSE.

=====

Table 2 about here

=====

Finally, the mean errors (ME) suggest that, if at all, then most forecasts tend to be biased upwards. Numerically, this bias is more substantial for large horizons, but in most cases, the ME is not significantly different from zero. To illustrate the intimate relation between forecast performance and horizon further, Chart 2 plots MAE against the forecast horizon, generating an almost perfect straight line. The deviations of individual institutions from this straight line are small, with two exceptions: The two early forecasts by the European Commission and the OECD with forecast horizons above two years. They exhibit errors that are nearly the same as those published about 200 days later by the same institutions. The adhe-

rence of all other average forecasts to this line indicates that institutional factors – if there are any – have to be considered as being rather small.

=====
 Chart 2 about here
 =====

All three forecasts of the Joint Forecast (Gemeinschaftsdiagnose) can be found below the average. This seems to confirm that pooling the knowledge of various forecasters may lead to better predictions. By contrast, the average errors of all four OECD forecasts lie above trend. This indicates that the publication of OECD forecasts is delayed by internal coordination requirements. But it remains to be shown whether these differences are statistically significant.

4.2. Accuracy of GDP Forecasts

Therefore, we now proceed from the analysis of summary measures by institution to the investigation of individual forecasts. To this end, we consider the variation of absolute forecast errors (AE) across institutions and individual years, and relate them to their precise forecast horizon (h). This variable is measured in days and it is defined as the distance between the day the forecast was released and the 31st December of the year the forecast was made for⁵. Here, we utilize 30 forecasts for 18 years starting from 1991 and ending 2008, generating a sample of 540 observations⁶. Thus, the basic equation of our analyses can be written as follows:

$$(1) \quad AE_{i,q,t} = a_0 + a_1 h_{i,q,t} + \varepsilon_{i,q,t}$$

In eq. (1), i denotes the institution, q the forecast horizon (in quarters, as defined in Table 1) and t the year the forecast is constructed for. The variable ε is the disturbance term. The parameter a_0 captures that forecast error which remains at the end of the forecast year, whereas the parameter a_1 reflects the (negative of the) accuracy gain per day.

⁵ Negative values indicate that the forecast for year t was published in year $t+1$, but before the official figures were released.

⁶ In our regressions the two early forecasts of OECD and EU are omitted, since they are three-step forecasts.

To control for institutional factors, we additionally include a set of indicator variables D_i^j . These indicators take the value of 1 whenever a forecast was published by institution j , and 0 in all other cases. As these variables allow for institution-specific intercepts, their variation captures any permanent performance differences across institutions⁷:

$$(2) \quad AE_{i,q,t} = \sum_j a_{0,j} D_i^j + a_1 h_{i,q,t} + \varepsilon_{i,q,t}$$

In the pooled regression according to equation (1), on average, AE are reduced by 0,002 percentage points with every day the forecast horizon gets shorter (column 1 in table 3), or 0.06 percentage points per month. As the constant shows, there still is a small error of 0.07 percentage points on average when the forecast horizon is 0. This reflects the fact that even at the end of the year the data are not fully known to the forecaster. The results for equation (2) (column 2 of table 3) do not differ very much from the pooled results. As was already suggested by the inspection of Chart 2, the influence of institutional factors is estimated to be insignificant. In a joint test, the hypothesis that the intercepts are identical is not rejected at the 1% level of significance.

=====
 Table 3 about here
 =====

Chart 1 demonstrates that forecast errors vary considerably across years. In our experience, years differ systematically in ways which are known a priori to the forecaster, and one should perhaps account for such differences in the regressions. As a first factor, we consider a working day effect: In Germany, different from most other countries in Europe, annual GDP figures are not adjusted for the number of working days. As the number of working days is known in advance, and the influence of working days on GDP can be gauged from past experience, this is a piece of information which the forecaster can take into account in the forecast. A second factor which might display a serious impact on forecast accuracy is

⁷ In our concrete application, the RWI forecasts were used as a numeraire, retaining an overall constant. This leads to numerically identical results for the corresponding sum of squared residuals and the slope coefficients. Expressing institutional intercepts as deviations from a hypothetical mean (Haisken-DeNew/Schmidt 1997) has not become an issue, due to the apparent irrelevance of performance differences across institutions.

the (absolute value of the) carry-over from the previous year, which marks the starting point of the forecast⁸.

Hence, we extend expressions (1) (and analogously equation (2)) by including the variables wd_t which is the (mean-adjusted) absolute value of the working day effect on GDP in year t compared to year $t-1$, as it is published by the German Statistical office, and co_t which is the (mean-adjusted) absolute value of the carry-over relevant for year t . Both variables enter the regression directly and as interaction terms with the forecast horizon:

$$(3) \quad AE_{i,q,t} = a_0 + a_1 h_{i,q,t} + a_2 wd_t + \alpha_3 wd_t \cdot h_{i,q,t} + a_4 co_t + \alpha_5 co_t \cdot h_{i,q,t} + \varepsilon_{i,q,t}$$

Including these two factors into our regressions does not influence the coefficient of the forecast horizon. But the coefficients for the additional variables are statistically significant and their inclusion reduces the unexplained variance of the equation considerably (columns 3 and 4 of table 3). Taking α_2 and α_3 together, the forecast errors are as expected generally smaller in years with an important working-day effect. However, the impact gets smaller with a decreasing forecast horizon. Furthermore, forecast errors tend to be smaller, whenever there is a substantial carry-over, negative or positive, from the previous year. The interaction of these two factors might explain, e.g., why all forecasts for 2004 exhibited a rather small error, although the year should have been difficult to predict given its position in the business cycle. But the working day effect as well as the carry-over was quite high in this year. Surprisingly, the signs of α_2 and α_3 are positive indicating that forecast errors might be higher in years with a large working day effect or carry-over when the forecast horizon becomes very short. However, this result might reflect a non-linearity which becomes evident when the sample is split into one- and two-step forecasts.

In a final set of specifications we refrain from accounting for the complete set of institutional indicators and rather include a dummy variable indicating whether the forecast is conducted by an international organization. This restricted version of institutional heterogeneity accounts for the different conditions under which the forecasts of international organizations are typically conducted: A lengthy process of internal approval might cost valuable time. In

⁸ The carry-over is defined as deviation of the last quarters' seasonally and working-day adjusted GDP from the annual average in per cent of the annual average.

our regressions (columns 5 and 6 of table 3) we find no evidence that the performance of international forecasting institutions is significantly weaker, taking into account the horizon alone and also together with other systematic factors.

Up to this point all forecasts have been treated the same way, irrespective of whether they are looking one or two years ahead. However, the predictions involving long horizons – i.e. those in Table 1 with an index number larger than four – are two-step forecasts. Rather than being observations for the previous year, their starting points are themselves forecasted values. Therefore, it would not be too surprising, if the forecasts with index numbers five and over were performing significantly worse. To yield a catchy interpretation of the coefficients we define forecast horizon here as the difference between the publication date and the end of the year the forecast was published. For the one-step forecasts this implies no change from the practice hitherto. For the two step forecasts the constant can be interpreted as the forecast error remaining after the first year of the forecast season. Yet, splitting the sample accordingly makes evident that the relation between forecast horizon and accuracy does not differ strongly between the groups (Table 4). However, the working-day effect and the effect of the carry-over turn out to be significantly stronger in the case of the two-step forecasts. For the one step forecasts, all coefficients are smaller and they are, if at all, only significant on a 90% level, which holds in particular for α_2 and α_3 .

=====

Table 4 about here

=====

Given the simultaneous fashion by which the two groups of forecasts are constructed, the errors of the two-step and the one-step forecasts might be interrelated within institutions. The error of the one-step forecast will be realized a year ahead of that of the two-step forecast which is published by the forecaster at the same time. Specifically, if it were typical that forecasters are correctly predicting tendencies but are wrong concerning the timing of events, we would observe a negative correlation of these two forecast errors. If it is correctly anticipated that taxes will be cut soon, for instance, but contrary to the presumption underlying the one-step forecast the political decision is delayed to the second year, then growth might be overestimated in the first and underestimated in the second year.

To gauge the relevance of such a pattern, we check whether the sign of the forecast error corresponding to the one-step forecast is a significant predictor for the two-step error, estimating the following regression:

$$(4) \quad e_{i,q+4,t+1} = b_0 + b_1 e_{i,q,t} + b_2 h_{i,q,t} \cdot \text{sign}(e_{i,q,t}) + u_{i,q,t}.$$

Here $e_{i,q,t}$ is the error of the forecast published by institution i for the current year t , looking q quarters ahead, and $e_{i,q+4,t+1}$ is the error of the coincidentally published forecast for year $t+1$, looking $q+4$ quarters ahead. As the sign of the forecast error matters in this calculation, controlling for the influence of the length of the forecast horizon h requires taking it into account in the specification, via an interaction term. Table 5 documents that the estimate of b_1 is positive, i.e. that errors of coincidentally produced forecasts tend to display the same sign. Hence it does not seem to be the case that a bad timing of events is a major culprit for forecast errors. To the contrary, it seems that forecasters frequently err on the general tendencies instead.

=====
 Here about Table 5
 =====

4.3 Demand Side Components of GDP

While evaluating the forecasts of the demand side components of GDP we restrict ourselves to estimating equation (1). According to our results (Table 6), the correlation between forecast accuracy and forecast horizon is weaker in most cases compared to the results found for GDP (which are replicated from Table 3 for convenience). They are especially low for government consumption and net exports. The information-gain per day is particularly high for investment, exports and imports. This is consistent with the fact that these components of GDP are more volatile than the other components, so that each additional piece of information may reduce uncertainty more than for less volatile categories. Furthermore, the constants of the regressions which indicate the forecast errors remaining at the end of each year, are higher than those for GDP for all individual components. Partly this indicates that component errors tend to offset one another at the aggregate level.

=====

Here about Table 6

=====

But the result is also consistent with the fact that the first national accounts data published for the GDP components are not very reliable in some cases and subject to major revisions. This particularly holds for the data in foreign trade. As exports and imports are often revised in the same direction new data often provide little information on net exports. Hence, the accuracy of projections of net exports improves little over time. Here, the correlation between accuracy and forecast horizon is very low, and consequently the information gain per day is quite modest.

Also for the GDP components we additionally included dummy variables to control for institution-specific effects, as it is described in equation (2). Again, the estimated coefficients of the institutional variables were small and statistically insignificant. A joint test did not reject the null hypothesis that there are no institutional effects in the forecast errors.

5. Conclusions

Although the accuracy of forecasts provided by various institutions differs markedly in concrete situations, the differences become small when a longer-term perspective is taken. The variation can be explained to a large extent by differences in the date when the forecasts were completed. Taking into account the information that can be collected by delaying the finalization of the forecast by one day, the absolute error of the GDP forecast can be reduced, on average, by 0.002 percentage points. This amounts to an average reduction 0.06 percentage points per month, with a remaining absolute error of 0.07 percentage points as the horizon approaches 0. This result is quite robust with respect to different specifications of the model as well as to the inclusion of control variables capturing time-specific effects such as the working-day effect and the carry-over effect.

The institutional factors, captured in our analyses by institution-specific intercepts of the regressions exploring the correlates of absolute forecast errors, are extremely small and sta-

tistically insignificant. Similar conclusions hold for a more restrictive specification which merely distinguishes national and international institutions.

As forecast accuracy varies substantially more across time than across institutions, it is tempting to draw the conclusion that one could do without one or the other of the forecast suppliers. Yet, our results may also indicate that it is precisely the presence of these competent competitors which keeps providers on their toes, exploiting the available information to the fullest of their abilities. Of course, it is completely unclear, what would happen to overall forecast performance, if the number of competitors were reduced. In the extreme, it seems very probable, that a forecast monopoly would perform worse. But we cannot really say anything about interim cases. What we can conclude confidently, though, is that our results did not single out an obvious underperformer which could definitely be kicked out of the market without a loss in overall performance.

Moreover, our results demonstrate that it would be very difficult to find evidence for the prejudice that many forecasters are engaged in herding behavior. This popular view has its origin in the observation that many forecasters publish predictions at almost the same time and that their forecasts very much look alike. However, in the light of our results a more convincing explanation for this effect is the – actually quite comforting – fact that forecasters base their publications on similar data, methods and, most importantly, on similar information sets. When important new information arrives about the state of the economy, competent forecasters will revise their forecasts accordingly and, consequently, in a similar fashion. Of course, it cannot be ruled out that herding plays some role either. But this would require more than just an eyeballing of correlations. Until shown otherwise, the null hypothesis of no herding has not been rejected yet in a careful quantitative analysis.

Finally, our analyses demonstrate that even delaying forecasts for a few days may pay off in terms of a somewhat more accurate result, albeit progress is small. To governments demanding accurate macroeconomic forecasts as a basis for their decisions, this insight suggests a promising avenue for improving forecast performance: As timely information seems to be the key to achieving accurate forecasts, investing into approaches which provide information more timely might be a very valuable option. By contrast, one might not gain too

much from attempts to improve further the way in which the currently available information is translated quickly and accurately into forecast revisions: In general, macroeconomic forecasts provided by the major professional suppliers tend to be information efficient. This notion is not only confirmed by our results, but also holds for more sophisticated tests regarding strong information efficiency (Döhrn 2006).

Simply acquiring the data from official sources earlier, however, would threaten to reduce the quality of the data which in turn would depress forecast quality. Another approach which seems to be quite promising, though, could be the exploitation of new data sources from the internet. They may provide information on the decision making e.g. of consumers in advance to the market outcome (Schmidt, Vosen 2009). However, even the progress made by using these new data sources will be limited. If they were able to “buy” half a month of information gain via an earlier revelation of relevant information, according to our estimates this would be worth, on average, a reduction of the absolute forecast error by 0.03 percentage points. Only a careful assessment of the costs involved will be able to show whether this investment is really worthwhile.

References

- Döhrn, R. (2006), Improving Business Cycle Forecasts' Accuracy – What Can We Learn from Past Errors? RWI: Discussion Paper 51. RWI, Essen.
- Döpke, J. and U. Fritsche (2006a), Growth and inflation forecasts for Germany: A panel-based assessment of accuracy and efficiency. *Empirical Economics* 31: 777-798
- Döpke, J. and U. Fritsche (2006b), When do forecasters disagree? An assessment of German growth and inflation forecast dispersion. *International Journal of Forecasting* 22: 125-135
- Haisken-DeNew, J. and C. M. Schmidt (1997), Inter-Industry and Inter-Region Differentials: Mechanics and Interpretation, *The Review of Economics and Statistics* 79: 516–521.
- Heilemann, U. und S. Klinger (2005), Zu wenig Wettbewerb? Zu Stand und Entwicklung der Genauigkeit von makroökonomischen Prognosen. Technical Report / Universität Dortmund, SFB 475 Komplexitätsreduktion in Multivariaten Datenstrukturen 2005,16
- Isiklar, G. and K. Lahiri (2007), How far ahead can we forecast? Evidence from cross-country surveys. *International Journal of Forecasting* 23: 167-187.
- Isiklar, G., K. Lahiri and P. Loungani (2006), How quickly do forecasters incorporate News? Evidence from Cross-country Surveys. *Journal of Applied Econometrics* 21: 703-725.
- Lamont, O. A. (2002), Macroeconomic forecasts and microeconomic forecasters. *Journal of Economic Behavior and Organization* 48: 265-280.
- Nierhaus, W. und J.-E. Sturm (2003), Methoden der Konjunkturprognose. *Ifo-schnelldienst* 56 (4): 7-23.
- Öller, L.-E, and A. Teterukowsky (2007), Quantifying the quality of macroeconomic variables. *International Journal of Forecasting* 23: 204-217.
- Räth, N. (2009), Rezessionen in historischer Betrachtung. *Wirtschaft und Statistik* 2009: 203-208.
- Schmidt, T. and S. Vosen (2009), Forecasting Private Consumption – Survey-based Indicators vs. Google Trends. *Ruhr Economic Paper* 155.
- Stekler, H. O. (2002), The Rationality and Efficiency of Individuals Forecasts. In: Clements, M. P. and D. Hendry (2002), *A Companion to Economic Forecasting*. Blackwell: 222-240.

Table 1

Forecasts Included in the Analysis

Abbreviation	Forecast (first quarter to be predicted)
German Institute for Economic Research (Deutsches Institut für Wirtschaftsforschung)	
DIW-7	Summer forecast year t-1 (q2,t-1)
DIW-5	January forecast year t (q4, t-1)
DIW-3	Summer forecast year t (q2, t)
DIW-1	January forecast year t+1 (q4, t)
European Commission	
EU-10	Autumn forecast year t-2 (q3, t-2)
EU-8	Spring forecast year t-1 (q1, t-1)
EU-6	Autumn forecast year t-1 (q3, t-1)
EU-4	Spring forecast year t (q1, t)
EU-2	Autumn forecast year t (q3, t)
Joint forecast of the German Institutes (Gemeinschaftsdiagnose)	
GD-6	Autumn forecast year t-1 (q3, t-1)
GD-4	Spring forecast year t (q1, t)
GD-2	Autumn forecast year t (q3, t)
ifo Institute for Economic Research (Ifo Institut für Wirtschaftsforschung)	
Ifo-7	June Forecast year t-1 (q2, t-1)
Ifo-5	December forecast year t-1 (q4, t-1)
Ifo-3	June Forecast year t (q2, t)
Ifo-1	December forecast year t (q4, t)
Kiel Institute of World Economics (Institut für Weltwirtschaft)	
IfW-7	June Forecast year t-1 (q2, t-1)
IfW-5	December forecast year t-1 (q4, t-1)
IfW-3	June Forecast year t (q2, t)
IfW-1	December forecast year t (q4, t)
Organization for Economic Co-operation and Development	
OECD-9	Economic Outlook December year t-2 (q4, t-2)
OECD-7	Economic Outlook June year t-1 (q2, t-1)
OECD-5	Economic Outlook December year t-1 (q4, t-1)
OECD-3	Economic Outlook June year t (q2, t)
OECD-1	Economic Outlook December year t (q4, t)
Rheinisch-Westfälisches Institut für Wirtschaftsforschung	
RWI-7	Summer forecast year t-1 (q2, t-1)
RWI-5	December forecast year t-1 (q4, t-1)
RWI-4	February/March forecast year t (q1, t)
RWI-3	Summer forecast year t (q2, t)
RWI-1	December forecast year t (q4, t)
The German Council of Economic Experts (Sachverständigenrat)	
SVR-6	Annual Report year t-1 (q3, t-1)
SVR-2	Annual Report year t (q3, t)

The index t marks the year the forecast was made for. The forecasts are named after the acronym of the forecasting institution and the forecast horizon expressed in quarters.

Table 2
Accuracy of GDP Forecasts Across Institutions and Horizons
 1991 - 2008

Forecast ¹	Horizon (avg. days)	MAE	MSE	RMSE	ME
DIW-7	543	1.23	2.54	1,59	-0.36
DIW-5	360	0.72	0.95	0,97	0.08
DIW-3	175	0.43	0.38	0,62	0.13
DIW-1	-6	0.12	0.03	0,17	-0.03
EU-10	787	1.34	2.78	1,67	-0.88**
EU-8	622	1.30	2.73	1,65	-0.64*
EU-6	422	0.93	1.29	1,13	-0.29
EU-4	258	0.57	0.55	0,74	0.00
EU-2	59	0.16	0.05	0,33	-0.04
GD-6	439	0.93	1.27	1,13	-0.42
GD-4	257	0.47	0.40	0,63	-0.03
GD-2	74	0.18	0.05	0,22	-0.05
ifo-7	529	1.13	2.31	1,52	-0.63*
ifo-5	376	0.62	0.67	0,82	-0.23
ifo-3	166	0.41	0.26	0,51	-0.01
ifo-1	11	0.11	0.03	0,17	-0.01
IfW-7	565	1.16	2.13	1,46	-0.65*
IfW-5	384	0.74	0.91	0,96	-0.35
IfW-3	201	0.52	0.39	0,63	-0.09
IfW-1	18	0.09	0.01	0,26	-0.05*
OECD-9	768	1.27	2.82	1,68	-1.02***
OECD-7	591	1.24	2.61	1,62	-0.79**
OECD-5	403	0.96	1.45	1,20	-0.41
OECD-3	226	0.54	0.47	0,69	0.02
OECD-1	38	0.19	0.07	0,26	0.01
RWI-7	538	1.21	2.49	1,58	-0.80**
RWI-5	382	0.84	1.01	1,01	-0.41*
RWI-4	325	0.72	0.88	0,94	-0.36
RWI-3	173	0.44	0.35	0,59	-0.15
RWI-1	19	0.10	0.02	0,14	-0.06**
SVR-6	417	0.86	1.20	1,10	-0.26
SVR-2	52	0.14	0.03	0,18	-0.03

Authors' calculations:

$$MAE_{i,q} = \frac{1}{n} \sum_{t=1}^n |R_t - F_{i,q,t}| \quad MSE_{i,q} = \frac{1}{n} \sum_{t=1}^n (R_t - F_{i,q,t})^2$$

$$ME_{i,q} = \frac{1}{n} \sum_{t=1}^n (R_t - F_{i,q,t}) \quad RMSE_{i,q} = \sqrt{MSE_{i,q}}$$

with F denoting the forecast and R the realization, subscript t indicating the year for which the prediction was constructed, and subscripts i and q indicating the institution and the number of quarters to be predicted, respectively. */**/** denote significance at 90/95/99% level respectively. – ¹see Table 1.

Table 3
Absolute Forecast Errors of GDP and Forecast Horizon
 1991-2008

Regression	(1)	(2)	(3)	(4)	(5)	(6)
Explanatory variable						
Constant	0.069 (2.6)	0.097 (1.5)	0.068 (3.0)	0.095 (1.8)	0.063 (7.3)	0.062 (2.5)
Forecast horizon (days)	0.00197 (13.3)	0.00196 (13.2)	0.00198 (18.2)	0.00196 (17.9)	0.00196 (13.2)	0.00197 (18.1)
Working-day effect (absolute)			0.661 (4.1)	0.661 (4.1)		0.662 (4.1)
Working-day*horizon			-0.00784 (10.2)	-0.00784 (10.2)		-0.00785 (10.3)
Carry-over ¹ (absolute)			0.138 (4.1)	0.138 (4.1)		0.138 (4.1)
Carry-over * horizon			-0.00136 (10.0)	-0.00136 (9.9)		-0.00136 (10.0)
Forecast of International Institution					0.032 (0.5)	0.032 (0.7)
Institutional effects	NO	YES	NO	YES	NO	NO
Significance of institu- tional effects ²		0.21 (98.3%)		0.37 (92.0%)		
Number of observations	540	540	540	540	540	540
R ² (adj.)	0.289	0.281	0.548	0.543	0.288	0.547

Authors' calculations. In parentheses: t-statistics. – ¹White heteroskedasticity-consistent standard errors. – ²End of previous year. – ²F-Statistic, in parentheses: probability in %.

Table 4
Absolute Forecast Errors:
One- and Two-Step Forecasts
 1991-2008

Regression	One-Step Forecasts		Two-step Forecasts	
	(1)	(2)	(3)	(4)
Constant	0.083 (4.4)	0.084 (4.6)	0.753 (11.8)	0.753 (14.4)
Forecast horizon (days) ¹	0.00188 (8.2)	0.00187 (8.6)	0.00227 (3.7)	0.00228 (5.0)
Working-day effect (absolute)		0.035 (1.6)		1.793 (1.4)
Working-day*horizon		-0.00057 (2.1)		-0.010 (4.6)
Carry-over ² (absolute)		0.209 (1.6)		0.684 (2.8)
Carry-over * horizon		-0.00237 (1.4)		-0.00254 (4.6)
Number of observations	288	288	252	252
R ² (adj.)	0.268	0.277	0.053	0.454

Authors' calculations. In parentheses: t-statistics. – White heteroskedasticity-consistent standard errors. – ¹End of the year the forecast was published = 0 – ²End of previous year.

Table 5

Two-Step-Forecast Errors: One-Step-Forecast Errors as Predictors

1992 – 2008

Constant	0.080 (1.4)
Error of the one step forecast	0.323 (2.9)
Forecast horizon (days) * sign of the forecast error	0.002 (17.4)
Number of observations	238
R ² (adj.)	0.579
F-Statistic	164.0

Authors' calculations. In parentheses: t-statistics. – White heteroskedasticity-consistent standard errors

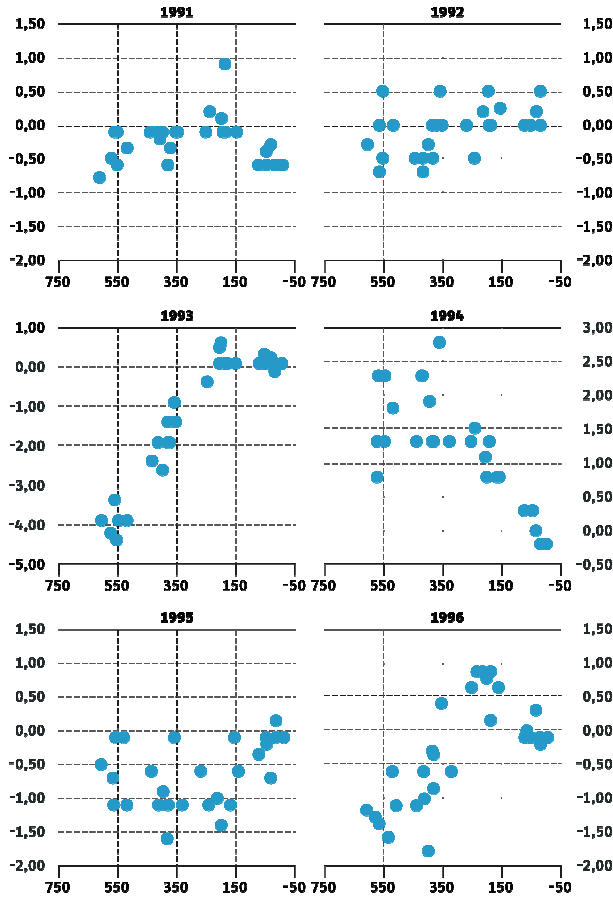
Table 6
MAE and Forecast Horizon for GDP Components
 Regressions 1991 – 2008

	Constant	Forecast horizon (days)	Institutional effects	Test on Institutional Effects ¹	R ² (adj)
GDP	0.069 (2.6)	0.00197 (13.3)	NO		0.289
	0.097 (1.5)	0.00196 (13.1)	YES	0.21 (98.3)	0.281
Private Consumption	0.251 (8.9)	0.00177 (14.1)	NO		0.293
	0.285 (4.9)	0.00177 (13.9)	YES	0.29 (95.8)	0.286
Government Consumption	0.491 (15.1)	0.00086 (7.4)	NO		0.086
	0.440 (7.6)	0.00085 (7.2)	YES	0.58 (78.8)	0.079
Gross Investment	0.672 (6.6)	0.00507 (10.6)	NO		0.187
	0.575 (2.6)	0.00502 (10.3)	YES	0.32 (94.7)	0.180
Exports	1.140 (9.4)	0.00442 (10.5)	NO		0.188
	1.083 (5.2)	0.00445 (10.6)	YES	0.46 (86.2)	0.183
Imports	1.377 (9.5)	0.00402 (8.4)	NO		0.127
	1.206 (5.2)	0.00403 (8.4)	YES	0.62 (73.9)	0.123
Net-exports ²	0.476 (10.4)	0.00041 (3.4)	NO		0.022
	0.506 (7.4)	0.00041 (3.4)	YES	0.20 (98.5)	0.011

Authors' calculations. 540 observations. In brackets: t-statistics- – White heteroskedasticity-consistent standard errors. – ¹F-Statistics, in parentheses: probability in %. ²Contribution to growth in percentage points.

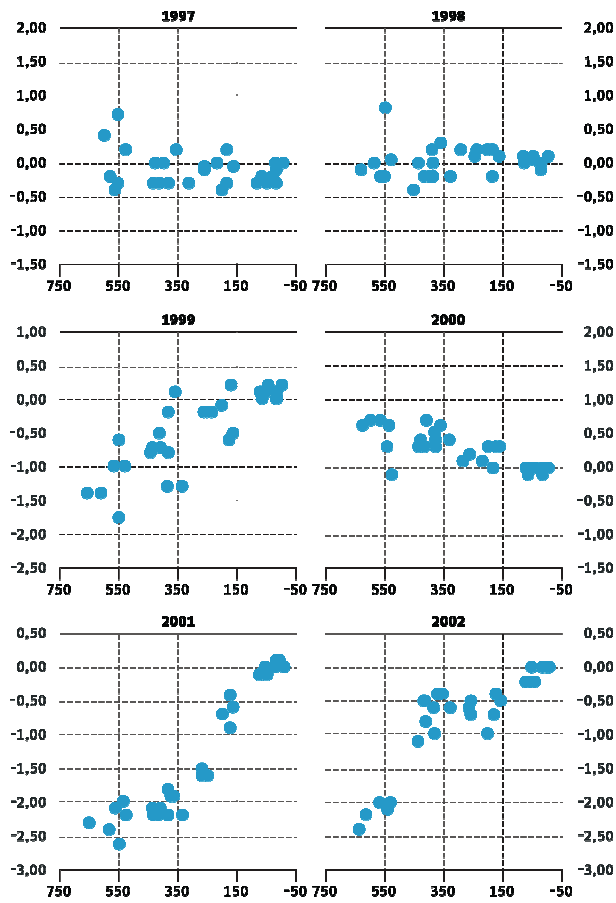
Chart 1

Forecast Error versus Forecast Horizon¹ for Individual Years



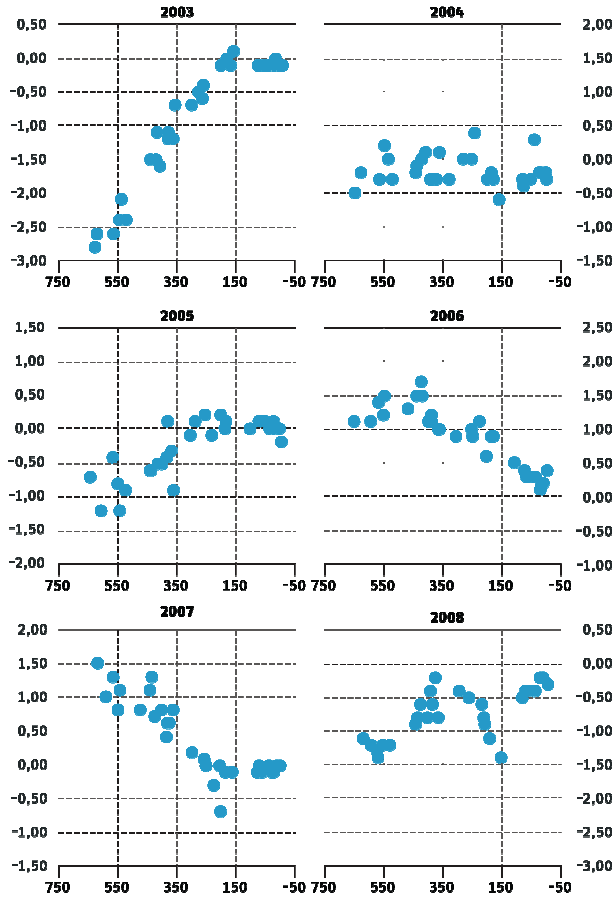
Authors' calculation - ¹Difference between the publication date of the forecast and December, 31rd of the year, the forecast was made for. Values on the horizontal axis.

Chart 1 (continued)

Forecast Error versus Forecast Horizon^a for Individual Years

Authors' calculation - ^aDifference between the publication date of the forecast and December, 31st of the year, the forecast was made for. Values on the horizontal axis.

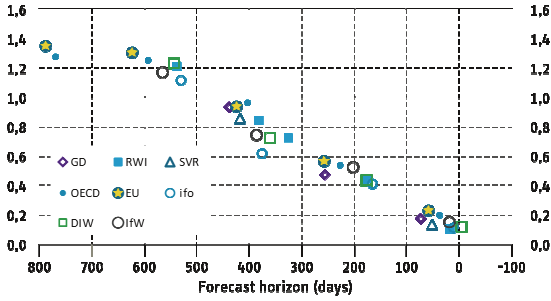
Chart 1 (continued)

Forecast Error versus Forecast Horizon¹ for Individual Years

Authors' calculation - ¹Difference between the publication date of the forecast and December, 31rd of the year, the forecast was made for. Values on the horizontal axis.

Chart 2

**Mean Absolute Forecast Error¹ and Forecast Horizon²
GDP forecasts for Germany by eight institutions**



Authors' calculations – ¹1991 to 2008. – ²Difference between the publication date of the forecast and December, 31st of the year, the forecast was made for. Average 1991 to 2008.