IWH
Halle Institute for Economic Research
Member of the Leibniz Association

Discussion Papers

No. 18

August 2019

How Forecast Accuracy Depends on Conditioning Assumptions

Carola Engelke, Katja Heinisch, Christoph Schult

Authors

Carola Engelke
University of Cologne

Katja Heinisch
Corresponding author
Halle Institute for Economic Research (IWH) – Member of the Leibniz Association, Department of Macroeconomics
E-mail: katja.heinisch@iwh-halle.de
Tel +49 345 7753 836

Christoph Schult
Halle Institute for Economic Research (IWH) – Member of the Leibniz Association, Department of Macroeconomics
E-mail: christoph.schult@iwh-halle.de
Tel +49 345 7753 806

Editor

Halle Institute for Economic Research (IWH) – Member of the Leibniz Association

Address: Kleine Maerkerstrasse 8
D-06108 Halle (Saale), Germany
Postal Address: P.O. Box 11 03 61
D-06017 Halle (Saale), Germany

Tel +49 345 7753 60
Fax +49 345 7753 820

www.iwh-halle.de

ISSN 2194-2188

How Forecast Accuracy Depends on Conditioning Assumptions*

Abstract

This paper examines the extent to which errors in economic forecasts are driven by initial assumptions that prove to be incorrect ex post. Therefore, we construct a new data set comprising an unbalanced panel of annual forecasts from different institutions forecasting German GDP and the underlying assumptions. We explicitly control for different forecast horizons to proxy the information available at the release date. Over 75% of squared errors of the GDP forecast comove with the squared errors in their underlying assumptions. The root mean squared forecast error for GDP in our regression sample of 1.52% could be reduced to 1.13% by setting all assumption errors to zero. This implies that the accuracy of the assumptions is of great importance and that forecasters should reveal the framework of their assumptions in order to obtain useful policy recommendations based on economic forecasts.

Keywords: forecasts, accuracy, forecast errors, external assumptions, forecast efficiency, forecast horizon

JEL classification: C53, E02, E32

* We would like to thank Tara M. Sinclair, Ullrich Heilemann, and the participants at the International Symposium on Forecasting 2019, the IWH-CIREQ-GW Macroeconometrics Workshop 2018, and the Seminar on Statistics and Econometrics of Kiel University 2018, for fruitful comments and suggestions.

1 Motivation

The conduct of fiscal and monetary policy is based on the assessment of the current state of the economy and on future economic projections. For instance, governments have to estimate future tax revenues in order to plan their budgets and central banks need to set short-term interest rates to control the supply of credit, and therefore the amount of money, to meet inflation and output targets. Both variables depend on the current and expected states of the economy. Hence, the accuracy of the economic forecasts determines the effectiveness and efficiency of monetary and fiscal policy. Accordingly, both governments as well as central banks aim to minimize the errors in their economic forecasts so as to reduce errors in policy.

Forecasts for various economic aggregates are published regularly by national and international institutions. The most popular aggregate to summarize the state of an economy is the growth rate of the real gross domestic product (GDP). For Germany, the accuracy of the forecast of the GDP and major GDP aggregates has been intensively studied regarding both national (Döhrn and Schmidt, 2011; Heilemann and Stekler, 2003; Heilemann and Müller, 2018) and international forecasters (Pons, 2000; Timmermann, 2007; Júlio and Esperança, 2012). The rankings of the accuracy of the forecasts by different institutions based on multiple economic forecasts depend on the metric selected to assess this accuracy (see Sinclair, Stekler, and Müller-Droge, 2016).

An important research question regarding economic forecasts is whether the forecasters have used all relevant information available at the time the forecast was made (Kotchoni, Leroux, and Stevanovic, forthcoming), i.e. whether the forecasts are rational (Clark and McCracken, 2013). Using the definition of Muth (1961), Davies and Lahiri (1995) find evidence that economic forecasts for inflation and real GDP from the Blue Chip Survey of Professional Forecasters are not rational. Clements, Joutz, and Stekler (2007) extend this approach to test whether updates in forecasts over different horizons are predictable by updates in the data. If the forecasts were rational, forecast updates should not be predictable by data updates. Their results show that forecast updates are predictable by data updates.

A shortcoming of these studies is that they do not consider how the economic forecasts are constructed and how the information is used. None of the institutional forecasters provide explicit information on the process behind the creation of the economic forecast. It is easier to study the performance of economic forecasts which are solely based on econometric models. For instance, Heilemann (2002) conducts a case study using a medium-sized macroeconometric model to construct economic forecasts for Germany. In particular, the role of the assumptions fed into the model is investigated for one specific year. All in all, only a couple of studies analyse whether the underlying assumptions are a source of errors in the forecasts (Keereman, 2003; Takagi and Kucur, 2006; Fioramanti, Gonzalez Cabanillas, Roelstraete, and Ferrandis Valterra, 2016; Berge, Chang, and Sinha, 2019). Besides the well-known criticism that institutions' forecast models are not transparent (Heilemann, 2002), another reason might be that the assumptions are only sporadically published in the forecasting reports and the impact of those assumptions has just been ignored for analysing the evaluation (Döpke, Fritsche, and Waldhof, 2019). However, conditional forecasts are projections of future paths of some other variables and, hence, prior knowledge, albeit imperfect, of the future paths of some economic variables may contain information for the future paths of other variables.

The main contribution of this paper is an empirical analysis, using a novel data set for Germany, of the errors in the forecasts and assumptions made by institutions. Our data set comprises assumptions for the growth of world trade, price of oil, exchange rates, and interest rates, provided by national and international institutions. We analyse whether forecasters make systematic errors in their forecasts driven by systematic errors in the assumptions, and quantify the relative contribution of these errors to the errors in the forecast. Furthermore, we compare the accuracy of the forecasts of different institutions.

Forecasts of economic growth are based on (technical) assumptions, which might even constrain the forecasts.[1] For instance, an increase in the amount of world trade will affect exports and imports and, hence, will have an impact on the growth of the GDP. The mechanism for changes in exchange rates are similar. Interest rates directly influence the conditions of financing, and, hence, have effects on investments. Typically, assumptions can be set by (i) technical rules, e.g. the exchange rate is fixed at the current level or oil prices are assumed to remain constant at constant prices. In addition, (ii) external models could be employed for the assumptions of the forecast, such as world trade or (iii) expert-based assumptions, e.g. based on the announcements of central banks and markets, e.g. the price of futures. In the case of joint forecasts, like the *Gemeinschaftsdiagnose*, (iv) the set of assumptions is agreed upon by several institutions (Fritsche and Heilemann, 2010). Forecasts have been published for decades, but the details on the underlying assumptions are still very rarely published. Recently, international forecasters have provided more details on their underlying assumptions, in a technical appendix to the report of the forecast. National forecasters either place their assumptions in special sections, boxes, or just in the main text of the report, and hence make them hard to identify.

Our new data set on forecasts and the underlying assumptions for Germany published by national and international institutions is an unbalanced panel, since most institutions publish their forecast reports on different dates and do not always provide the same information regarding the underlying assumptions. We use an ordinary least squares (OLS) approach to regress the annual forecast errors at different horizons for German GDP on the errors in the underlying assumption. Our forecast target sample covers the period 1992 to 2018. Based on 12 institutions and 1–4 forecasts per institution and year, our sample includes in all 1390 GDP forecasts: 724 for the current year and 666 for the next year. For oil prices and the exchange rate, more than 1000 assumptions are provided. However, for world trade and interest rates, the number of assumptions provided is smaller.[2]

[1] Several models, such as AR and VAR models, do not require external assumptions.

[2] Since not all forecasters have provided full information on their assumptions, the sample is reduced to 620 observations for the baseline regression analysis.

The results of the regression show that about 75% of the variation of GDP squared forecast errors (SFE) are driven by the squared assumption errors (SAE). Squared assumption errors in world trade and oil prices have a strong positive statistically significant correlation. The results indicate that institutions make systematic forecast errors. Institution-fixed effects do not add explanatory power, as reflected by an unaffected adjusted R^2. This result implies that forecasters make similar forecast errors regarding the unobserved components. On the other hand, year-fixed effects increased the adjusted R^2 by 10 percentage points. In our regression sample for the baseline model, the observed root mean squared forecast error is roughly 1.52%. One could hypothetically reduce the root mean squared forecast error to 1.13% using our baseline regression estimates for the coefficients and residuals and setting all assumption errors to zero.

Including the errors in the assumptions about the interest rates of the main refinancing operations (MRO) set by the ECB restricts the evaluation sample to the period 1999 to 2018. Further, international institutions like the IMF do not report their assumptions on the refinancing rate, which reduces the number of observations to 340. Overall, squared assumption errors about world trade are found to have a robust positive correlation with squared forecast errors of the GDP.

This paper is structured as follows: Section 2 presents the data set for the forecasts, assumptions, and actual data. Section 3 provides details on the measurement of the forecast errors and assumption errors. Furthermore, the regression approach is presented. In Section 4, the regression results are presented for the squared forecast errors of GDP, including several extensions of our baseline specification.[3] Lastly, Section 5 summarizes the results.

[3] The results for various estimation specifications are available in the Online Appendix.

2 Data

We have collected the data for the forecasts and major external assumption and constructed a new data set for Germany.[4] We focus on annual GDP growth forecasts for the current and the next year. The forecasts of national and international forecasters are scrutinized. These comprise (i) the national economic research institutes – the German Institute for Economic Research (DIW), the Hamburg Institute of International Economics (HWWI, before 2006 known as the HWWA), the Kiel Institute for the World Economy (IfW), the Macroeconomic Policy Institute (IMK), the Leibniz Institute for Economic Research at the University of Munich (ifo), the Halle Institute for Economic Research (IWH), and the RWI–Leibniz Institute for Economic Research (RWI)[5] – (ii) the Joint Economic Forecast (GD) (iii) Deutsche Bundesbank (BBK) and (iv) international institutions – the European Commission (EC), the International Monetary Fund (IMF), and the Organisation for Economic Co-operation and Development (OECD).[6] National accounts data are regularly provided by the German Federal Statistical Office (Destatis) and first releases are taken from the Real-Time Database of the Deutsche Bundesbank. We manually extend the real-time data set for values before 2005, using printed editions by the German statistical office.[7]

Recently, institutions have been publishing the underlying assumptions for their forecasts. Commonly reported assumptions are oil prices, interest rates, growth rates for world trade, and exchange rates. Oil is a key factor in industrial production and its price is an important indicator of the development of costs. Kilian (2008) provides a comprehensive literature review on the effects of oil price increases on the US GDP. The growth in the GDP of countries which are net importers of oil, such as Germany, declines with unpredicted changes in the price of oil. Of course, countries that are net oil exporters benefit from higher oil prices (see Bergholt, Larsen, and Seneca, 2019). The price of oil is measured in US dollars per barrel, and the assumptions are usually based on market conditions and futures prices.

Germany depends heavily on exports, and the exchange rate is a common measure of the relative prices (competitiveness) of German goods. Forecasters commonly make assumptions about the bilateral exchange rate between Germany and the USA, because the USA is Germany's largest trading partner. The US dollar–euro (USD/EUR) exchange rate after 1999 and the US dollar–Deutsche Mark (USD/DM) exchange rate before 1999 are measures of the international competitiveness of German products. The other important trade partners of Germany are mostly part of the euro area. Forecasters commonly set the real exchange rate to be a constant and

[4] The data set is part of the German Research Foundation (DFG) Priority Programme 1859 "Experience and Expectation. Historical Foundations of Economic Behaviour".

[5] Further institutes, such as the IW and HWWI, have not provided any information on their assumptions and are therefore excluded from the analysis. Forecasts by the Macroeconomic Policy Institute (IMK) have been published since 2005, before that date, data from the Institute of Economic and Social Research (WSI) has been used. HWWA data is used until 2006.

[6] The complete list of institutions and their abbreviations are given in Table 3 in the Appendix.

[7] In the international context (IMF, OECD), the growth rates of annual calendar-adjusted series are published. Hence, to calculate the corresponding forecast errors, we compare those forecasts with the corresponding target.

adjust the nominal exchange rate according to the GDP deflator. Alternatively, they assume that the nominal exchange rate does not change.

As a net exporting nation, the business cycle in Germany is heavily affected by the world demand for goods and services. World trade is a measure of the state of the global economy and the share of German trade in world trade is about 7.5%.[8] World trade is a better indicator of the demand for German products and world supply of tradable products than world GDP. Since the growth in the world GDP is highly correlated with world trade, and more data on the assumptions about world trade are provided, we will use the latter series in this analysis. We refer to CPB data since it is more timely than IMF statistics on world trade and since it is used as a benchmark for most national forecasters.[9]

Interest rates set by the central bank define the refinancing costs of the financial sector, an important intermediary in modern economies to ensure sufficient investments in the capital stock of the economy. To reflect the objective of monetary policy with respect to price stability, the interest rate for main refinancing operations (MRO) set by the European Central Bank (ECB) is considered. Assumptions about interest rates often follow market expectations and announcements of the ECB. Unfortunately, assumptions about MRO rates are not provided by international institutions.

It is important to note that the details (precise description) of an institution's assumptions are often not reported. Furthermore, the target also varies across institutions and different forecast exercises, i.e. the assumptions can be provided either as end-of-year values or annual averages. To tackle this issue, we convert all assumptions to yearly averages, with the exception of the MRO rates. Further, we include interactions between institution-fixed effects and release-year fixed effects in our regression analysis to control for different potential assumption targets.[10]

Our evaluation sample covers the forecasting period 1992 to 2018, which includes 1390 GDP forecasts — 724 for the current year and 666 for the next year. Our data set consists of 1091 assumptions for oil, 1055 for the exchange rate, 704 for world trade, and 519 for the interest rate. We exclude interest rates as an explanatory variable in our baseline regression specification due to the limited number of observations. Since not all institutions have provided full information on the three remaining assumptions under review, the sample is reduced to 620 observations with full information.

[8] This is the average share of German exports in world exports from 1991 to 2018.

[9] Furthermore, IMF data and CBP data are highly correlated. Institutions did not specify the target of the world trade data, and the reference series might have changed over time as well.

[10] It is important to stress that one should not compare the errors in assumptions across institutions due to possible time-varying assumption targets. For example, institutions might change their assumption targets from average to end of year values from one forecast year to another.

3 Evaluation of Forecast and Assumption Errors

3.1 Descriptive statistics

We define four forecast rounds in each year to conduct a descriptive analysis conditional on the horizon and institution. This implies that we merge all forecasts conducted during a specific quarter. The longest forecast horizon is 8 quarters, i.e. the forecast for the next year conducted in the first quarter of the current year.[11] The final forecast is provided during October and December of the current year. However, this breakdown is only used for the descriptive part of our analysis. For the analytical part in Section 3.2, we make use of the exact date and, hence, distinguish between forecast horizons from 1 to 729, i.e. 365*2-1, days. This allows controlling for different sets of information, i.e. including updates of indicators and revision of ex-post data, available for each forecast conducted on different days.[12]

For each forecasting institution n, the forecast error for GDP y_t can be calculated as the difference between the forecast for period t at horizon h of y and the outcome (first release),[13]

$$e_{n,t+h|t} = \hat{y}_{n,t+h|t} - y_{t+h}. \tag{1}$$

The mean error (ME) is the average over all horizons and periods, but given that positive and negative errors can offset each other, the size of the error is distorted. Therefore, a criterion that takes into account the squared loss, i.e. the mean squared error (MSE) or the root mean squared error (RMSE), is recommended. In addition, the mean absolute percentage error (MAPE) and the normalized root mean squared error (NRMSE) based on the mean or standard deviation are shown.

We report different measures of the accuracy of the forecast of the GDP made by the different forecasting institutions.[14] To facilitate the understanding of our data set, consider, e.g. the Organization for Economic Cooperation and Development (OECD). The OECD has published more than 100 forecasts during the sample period. For graphical presentation, we summarize the forecasts for half-year intervals. The mean absolute error decreases with a decreasing horizon. The relation is not strictly monotonic, in contrast to theoretical predictions.[15] This result holds also for other forecasters. Figure 1 shows that the OECD gets closer to the first release of GDP as the horizon decreases. Longer horizon forecasts by the OECD exhibit a downward trend.

---

[11] Some institutions provide forecasts more than 8 quarters ahead. Others, in particular the international institutions, publish forecasts twice a year with some updates for major economic variables and selected countries in between.

[12] For the evaluation of forecast errors, other studies disentangle the projection for the next and current year, but usually take only one projection per year into account.

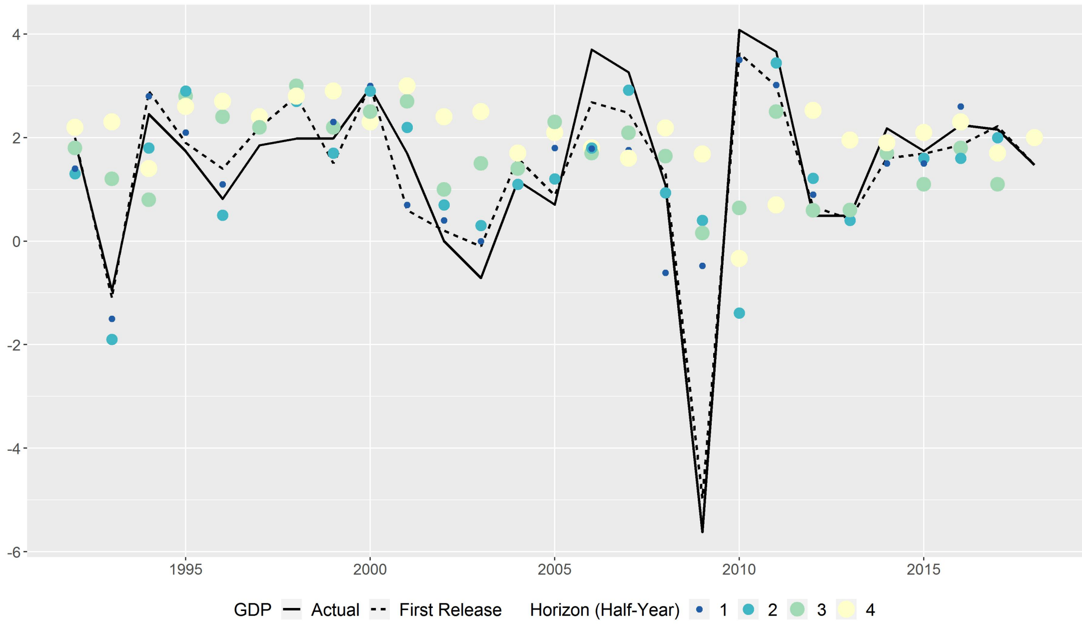
[13] Forecast errors can also be calculated with respect to the actual (and revised) values. The difference between first releases and revised data is shown in Figure 1. GDP forecasts and actual data are rounded to one decimal place.

[14] Table O.1 in the online Appendix.

[15] The monotonicity is distorted by forecasts conducted in forecast rounds two, four, six and eight, with only one observation.

This indicates that estimates for potential output growth decline over time. Longer horizon forecasts of GDP growth are successively revised, as new information is available, towards the true realization.

Figure 1: GDP growth and GDP forecasts by the OECD



Note: Growth rate and forecasts in percentages. Forecast errors for a longer horizon are depicted by larger points and lighter colours.

To estimate the bias in assumptions, we calculate the errors in the underlying assumptions for the following variables: oil price p_t^o, exchange rate s_t, interest rate r_t, and world trade w_t.[16] Since there are no revisions in the variables, the first releases are the same as the actual data. The only exception is world trade, where the first releases might be revised by new information on international trade flows. Descriptive statistics for the errors in the assumptions are reported (Tables O.2 to O.5 in the online Appendix). The mean absolute error for the prediction of the price of oil does not increase monotonically with the horizon. We observe similar characteristics as for the forecast errors for GDP.

Figure 2 shows scatterplots for the GDP forecast errors and the errors in the assumptions by the OECD. We find a strong correlation between the errors in the assumptions in world trade and in GDP. The same is true for the interest rate. The relation between the errors in the assumptions about the exchange rate and the errors in the forecast of the GDP is weak. The OECD systematically underpredicted exchange rates before 2005, leading to two clusters in the corresponding scatterplot. The errors in the assumptions about the price of oil are moderately correlated with the errors in the forecast of the GDP. The results are similar for the different forecasting institutions.

[16] The errors are calculated in the same way as the procedure described for GDP (Eq. 1), so that the assumptions are compared to the outcome. The assumptions are rounded to two decimal places.

Table 1: Summary Statistics

	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
			Realizations				
GDP	27	1.43	1.88	−5.62	0.76	2.21	4.08
Oil Price	27	50.60	33.28	12.85	19.95	71.60	111.63
Exchange Rate	27	1.21	0.15	0.90	1.12	1.32	1.47
Interest Rate	20	1.70	1.50	0.00	0.20	2.81	4.75
World Trade	27	4.95	5.12	−12.72	2.47	8.08	13.90
			Forecasts				
GDP	1,392	1.55	1.30	−6.50	1.20	2.20	4.00
Oil Price	1,091	63.62	32.48	12.00	33.30	95.18	135.00
Exchange Rate	1,055	1.23	0.21	0.43	1.12	1.35	1.85
Interest Rate	519	1.56	1.54	0.00	0.05	2.75	5.40
World Trade	704	5.03	3.60	−16.50	3.67	7.00	15.00
			Forecast and assumption errors				
GDP	1,392	0.18	1.29	−5.00	−0.30	0.50	6.80
Oil Price	1,091	−0.32	15.23	−47.73	−5.90	2.14	69.04
Exchange Rate	1,055	−0.002	0.17	−0.72	−0.04	0.03	0.74
Interest Rate	519	0.24	0.78	−2.25	0.00	0.34	3.50
World Trade	704	0.96	4.54	−25.90	−0.88	2.54	19.72

Note: GDP growth in percent, oil price in USD/Barrel, exchange rate in USD/EUR, interest rate in percent (MRO), world trade growth in percent (CPB).
The sample covers 1992–2018.

As an overview, Table 1 presents some characteristics of the GDP and the assumptions, both for the true data and the forecasts. Institutions tended to slightly overpredict annual GDP growth for Germany and the variation in GDP forecasts is lower than in the actual series, a finding that confirms previous studies, e.g. Heilemann and Müller (2018). The lower standard deviation in the forecasts for GDP growth than in the actual series might be driven by the larger number of observations. Further, we did not explicitly control for horizons and multiple forecasts for a particular year in the descriptive analysis.

The 50% confidence interval of the errors in the forecasts and assumptions include zero. Therefore, institutions do not systematically over- or underpredict the GDP and their assumptions. The average forecast for oil prices, the exchange rate, and world trade, is above the actual average for the sample. A higher standard deviation for the price of oil and world trade indicate that these variables are more difficult to forecast in general.

To assess the relation between GDP and the initial assumptions, we report unconditional correlation between GDP growth and the assumption variables for actual data at annual and quarterly frequency, for forecasts and assumptions, and for forecast errors and assumption errors (Table 4).[17]

The contemporaneous correlations between the forecast and actual GDP growth, interest rates, exchange rates, and world trade, have the same signs. The highest correlation is obtained for GDP with world trade (0.72) and between forecasts for GDP and world trade is similar. The

[17] Correlation estimates between forecasts and assumptions and the corresponding errors for different forecasters are provided in the online Appendix.

95% confidence interval for actual data and forecasts are overlapping. Furthermore, dynamic correlation between actual variables is analysed for lags up to four years, but magnitude and sign change already at lag order 1 for world trade (Table O.6).

Forecasters assume that assumptions and GDP growth comove with each other. It is straight forward to compare the correlation of actual data and forecasts for GDP growth and the underlying assumptions to judge whether forecasters make their forecasts consistent with empirical observations. However, the correlation estimates for annual data are only based on 20 to 27 observations. The uncertainty about these estimates is huge according to the confidence intervals reported in Table 4. Therefore, we additionally compute the correlation between actual variables using quarterly data (Table 4). The correlation estimate between GDP growth and world trade growth is 0.66 and the confidence interval overlaps with the confidence interval for the correlation between forecasts. Institutions tend to make no systematic error, indicating a good comprehension of the underlying relation between GDP growth and world trade growth. Of course, the correlation between world growth forecasts and GDP forecasts differs between institutions. Forecasts for GDP and world trade growth by the Deutsche Bundesbank comove even stronger compared to the actual data. Forecasts published by the Joint Economic Forecast for world trade and GDP growth have almost the same correlation as the unconditional correlation over all institutions. In order to estimate conditional correlations between forecast errors and assumption errors, a regression analysis is conducted in the following.

3.2 Regression Analysis

The aim of this section is to derive a simple econometric model for the forecast error in GDP growth. We show the strict hypotheses necessary to derive a linear relation between assumption errors and forecast errors. Based on this regression specification's including also unobserved components (κ), we conducted tests for some of the underlying strict hypotheses. The main hypotheses necessary to derive the baseline regression equation are:

H1: The data generating process for GDP growth is stationary and linear with respect to the underlying assumptions.

$$y_{t+h} = \alpha + \gamma^{p^o} p^o_{t+h} + \gamma^s \, s_{t+h} + \gamma^w \, w_{t+h} + \gamma^\kappa \, \kappa_{t+h}. \tag{2}$$

H2: Forecaster n uses a linear forecast rule for GDP growth.

$$\hat{y}_{n,t+h|t} = \hat{\alpha}_n + \hat{\gamma}^{p^o,n} p^o_{n,t+h|t} + \hat{\gamma}^{s,n} \, s_{n,t+h|t} + \hat{\gamma}^{w,n} \, w_{n,t+h|t} + \hat{\gamma}^{\kappa_n} \, \hat{\kappa}_{n,t+h|t}. \tag{3}$$

H3: The second moments of the data generating process for GDP are correctly estimated by the forecasters.

$$\mathrm{E}_t(\hat{y}_{t+h|t}\hat{x}_{j,t+h|t}) = \mathrm{E}_{t+h}(y_{t+h}x_{j,t+h})$$

H4: Assumption errors are uncorrelated.

$$\mathrm{E}_t(e_{j,n,t+h|t}\, e_{l,n,t+h|t}) = 0, \quad \forall l \neq j$$

H5: Unobserved components can be approximated as a function of institution-fixed effects γ_n, year-fixed effects γ_t, institution–year fixed effects $\gamma_{n,t}$, and the horizon h.

$$e^2_{\kappa,n,t+h|t} = f(\gamma_n, \gamma_t, \gamma_{n,t}, h_{n,t}) + \varepsilon_{n,t+h},\ \varepsilon_{n,t+h} \sim \mathrm{WN}(0, \Sigma)$$

It is convenient to assume that the true data generating process depends on the covariates in the regression analysis. We further show that the assumption H1 of a linear relation is not necessary, but reduces the algebra for the derivation. A derivation for a nonlinear data generating process is provided in Appendix A. The GDP growth process is usually assumed to be stationary. Tables O.9 and O.10 report the p-values for Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests for the null hypothesis that the process is stationary. The null hypothesis cannot be rejected for GDP growth at annual and quarterly frequencies. Since GDP growth itself is stationary and some of the underlying realizations of the assumptions are non-stationary, there is either a co-integration between the non-stationary assumptions or they are not true explanatory variables of GDP growth (see Baffes, 1997).

Under hypothesis H2, the reported assumptions are determinants of forecasts of economic growth. A forecast $\hat{y}_{n,t+h|t}$ of GDP growth by institution n for $t+h$ at t is therefore a function of the underlying assumptions $j \in \{p^o, s, w\}$ and an unobserved component $\hat{\kappa}_{n,t+h|t}$.[18] Given that our time series are not sampled at a constant frequency, our unbalanced panel does not allow us to use standard estimation methods, such as the augmented Dickey–Fuller test or the Phillips–Perron test. Therefore, the KPSS test is better suited for this specific data structure of the forecasts. Tables O.9 and O.10 also report the p-values of the KPSS test for the forecasts and assumptions for different institutions.

Hypothesis H3 assumes that the forecasters are rational, in the sense that they know in expectation the true dependency between the assumptions and the growth of GDP. Under H3, we know that $\hat{\alpha} = \alpha$ and $\hat{\gamma} = \gamma$. The forecast error or the assumption error of a variable is

[18] In our baseline specification we do not consider assumptions about the interest rate r, but the following equations can be easily extended.

the difference between the forecast rule Eq. (3) and the data generating process Eq. (2). In this setup, the forecast error is determined as follows:

$$e_{n,t+h|t} = \gamma^{p^o}\, e_{p^o,n,t+h|t} + \gamma^s\, e_{s,n,t+h|t} + \gamma^w\, e_{w,n,t+h|t} + \gamma^\kappa\, e_{\kappa,n,t+h|t}. \tag{4}$$

The forecast error by an institution is a linear function of the underlying errors in the assumptions and the unobserved component. A forecaster's objective is to be as close as possible to the true realization. Therefore, we square Eq. (4) to have a measure of the accuracy of the forecast. We derive a relation between the squared forecast errors of the GDP and the squared errors of the assumptions $j, l \in \{p^o,\ s,\ w\}$ as follows:

$$\begin{aligned} e^2_{n,t+h|t} = &\sum_j \underbrace{(\gamma^j)^2}_{\beta^j} e^2_{j,n,t+h|t} + \sum_j \sum_l \underbrace{(\gamma^j\, \gamma^l)}_{\beta^{j,l}}\, e_{j,n,t+h|t}\, e_{l,n,t+h|t} \\ &+ \sum_j \underbrace{(\gamma^j\, \gamma^\kappa)}_{\beta^{j,\kappa}}\, e_{j,n,t+h|t}\, e_{\kappa,n,t+h|t} + \underbrace{(\gamma^\kappa)^2\, e^2_{\kappa,n,t+h|t}}_{\gamma_n + \gamma_t + \gamma_{n,t} + \beta^h\, h_{n,t} + \beta^{h^2}\, h^2_{n,t} + \varepsilon_{n,t+h}} \end{aligned} \tag{5}$$

Under H4, the covariance between the errors in the assumptions is zero and, therefore, the coefficients for the interaction terms in Eq. (5) are zero. Tables O.11 and O.12 report the p-values for the KPSS tests applied to the errors in the forecasts and assumptions. We cannot reject the null hypothesis of a stationary time series at the 5% significance level for most of the institutions and errors. The results suggest controlling for spurious correlation in the regression analysis. Therefore, we will report the results of the KPSS test for residuals of the regression as well.

We have no observations for assumptions and realizations about $e^2_{\kappa_{n,t+h|t}}$. Under H5 we can approximate this term using institution-fixed effects γ^n, year-fixed effects γ^t, the horizon in days $h_{n,t}$, and $h^2_{n,t}$. Year-fixed effects capture specific effects for the corresponding release year of the forecast and the corresponding target year of the forecast. To control for different assumption targets over time, and by institution, we include an interaction term between the release year of the forecast and the institution considered. We further assume that the errors in the assumptions are independent of the unobserved components. Our regression residual $\epsilon_{n,t+h}$ is the variation in the unobserved components not explained by institution-fixed effects, year-fixed effects, and the horizon.

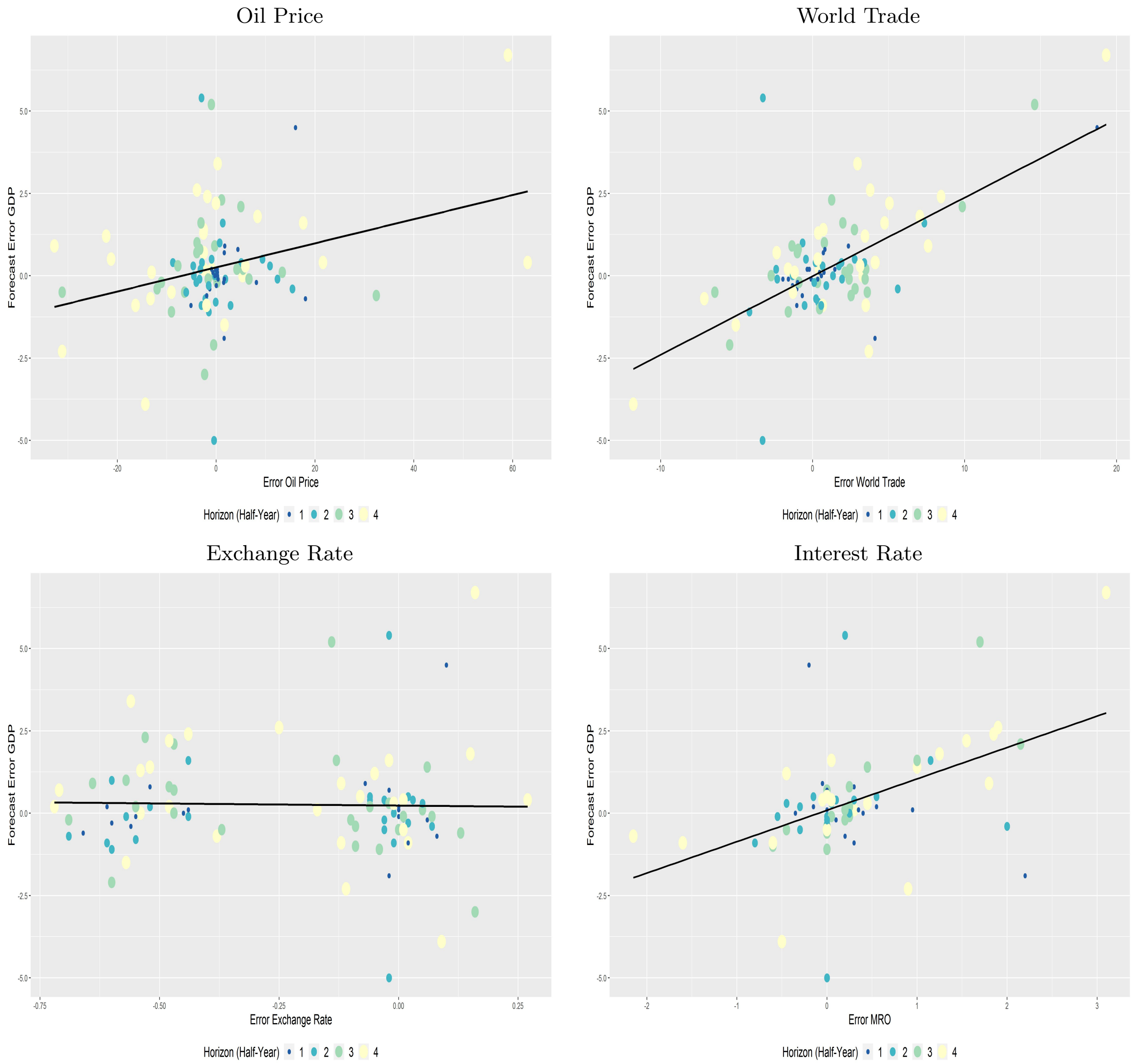
$$\begin{aligned} e^2_{n,t+h|t} = &\beta^{p^o}\, e^2_{p^o,n,t+h|t} + \beta^s\, e^2_{s,n,t+h|t} + \beta^w\, e^2_{w,n,t+h|t} + \gamma_n + \gamma_t + \gamma_{n,t} \\ &+ \beta^h\, h_{n,t} + \beta^{h^2}\, h^2_{n,t} + \epsilon_{n,t+h} \end{aligned} \tag{6}$$

In the first step, we estimate Eq. (6) applying OLS and do not estimate fixed-effects or random-effects models. The inclusion of institution-fixed effects allows investigating whether institutions make systematically different assumptions about the unreported variables. We control for the evolution in unobserved components by including year-fixed effects that comprise both the target and release years of the forecast. Including the horizon enables us to investigate whether the unobserved components are harder to predict with an increasing horizon. Our descriptive analysis of the forecast errors revealed a not strictly monotonic relation between them and the horizon. Therefore, we include a squared horizon term to control for a potential non-linearity.

The previous discussion illustrates the strict assumptions required to obtain the baseline regression specification. We will test the hypothesis of independent errors in the assumptions ($\beta^{j,l} = 0$) by including interaction terms between them. We also test whether errors in the forecast of the current year are drivers of errors in the next year's forecast. This test implicitly assumes that the growth of GDP depends partially on its past. Further, we include realizations of the assumptions to test whether forecasters are on average able to predict the comovement between the growth in GDP and the underlying assumptions ($\gamma = \hat{\gamma}$). Including squared actual assumptions as independent variables requires controlling for whether the estimated residuals are stationary, so as to apply standard inference procedures.

In addition, we conduct a set of robustness analyses: (i) We include the interest rate as an explanatory variable to check whether errors in the assumptions for the interest rate change our previous estimates. (ii) In order to evaluate the impact of the last recession in 2009 on our regression results, we neglect forecasts for the year 2009. (iii) To check for a time-varying relation between the squared errors in the assumptions and the squared errors in the forecast not captured by year-fixed effects, we split the sample into observations before 2008 and after 2008. (iv) Since it is likely that for a given institution its forecasts are based on the same models, the residuals of observations by the same institution might be correlated. We therefore calculate standard errors clustered at the institution level. We also cluster at the release year level since forecasters use the same information set. (v) While GDP forecasts are analysed with one digit, we assess the impact of the number of digits used for the assumptions. Lastly, (vi) we look at whether the regression results for international forecasters (OECD, IMF, EC) differ from those for national forecasters.

Figure 2: Scatterplots for errors in GDP forecast vs errors in assumption by the OECD



Note: Errors in the forecasts and the assumptions for a longer horizon are depicted by larger points and lighter colours.

4 Regression Results

Table 2 reports the regression results for the different specifications. The results of multiple robustness checks are provided in the tables of the online Appendix. The regression results for the GDP forecast errors show that about 75% of their variation is driven by errors in the associated assumptions (Table 2). The included explanatory variables are jointly significant. Squared assumption errors of world trade and oil prices have a strong and positive statistically significant association. Column (1) already shows that world trade has a positive and significant comovement with the squared forecast errors of the GDP, a result that is robust for all specifications without clustering of standard errors. The constant is insignificant when only squared assumption errors are considered.

Table 2: Regression Results for Squared Forecast Errors of GDP

	Dependent variable: SFE GDP					
	Biasedness	Horizon	Institutional FE	Year FE	Year and Inst. FE without interaction	Baseline
	(1)	(2)	(3)	(4)	(5)	(6)
SAE Oil Price	0.001	0.001	0.001	0.002**	0.001**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
SAE Exchange Rate	0.667	0.559	−1.251	0.568	−2.196	−4.139
	(0.813)	(0.785)	(1.475)	(1.076)	(1.652)	(5.203)
SAE World Trade	0.083***	0.083***	0.083***	0.037*	0.034*	0.031*
	(0.016)	(0.016)	(0.016)	(0.021)	(0.020)	(0.016)
Horizon (days)		−0.002			−0.0004	0.001
		(0.002)			(0.002)	(0.002)
Horizon squared (days)		0.00000			0.00001**	0.00001*
		(0.00000)			(0.00000)	(0.00000)
Constant	−0.090	−0.013	−0.162	−0.192	−1.603*	1.097
	(0.139)	(0.182)	(0.361)	(0.432)	(0.835)	(2.661)
Year FE	NO	NO	NO	YES	YES	YES
Institutional FE	NO	NO	YES	NO	YES	YES
Interaction Year and Inst. FE	NO	NO	NO	NO	NO	YES
Observations	622	622	622	622	622	622
R^2	0.754	0.755	0.758	0.864	0.872	0.909
Adjusted R^2	0.752	0.753	0.754	0.850	0.857	0.873
Residual Std. Error	3.370	3.367	3.361	2.619	2.561	2.414
F Statistic	629.911***	379.036***	159.269***	64.013***	56.548***	25.382***
Number of Parameters	4	6	13	57	68	338

Note: *p<0.1; **p<0.05; ***p<0.01.

Controlling for the horizon of the forecast measured in days does not significantly change the coefficients for squared assumption errors (column (2)). The estimated coefficient for the horizon is negative but insignificant and the estimated coefficient for the squared forecast horizon is positive and significant. This indicates that institutions tend to make better predictions as the horizon decreases. This is not surprising since institutions tend to predict that GDP returns to its trend growth for longer horizons.

Institution-fixed effects do not improve the adjusted R^2 (column (3)). This result indicates that institutions make similar forecast errors regarding the unobserved components. Year-fixed effects provide significant results for the price of oil (column (4)). Including an interaction term between release year-fixed effects and institution-fixed effects leads to an estimated coefficient for squared forecast errors in world trade that is lower compared to the regression without year-fixed effects. Controlling for institution- and year-fixed effects leads to an insignificant coefficient for errors in exchange rates in our baseline estimation (column (6)). Errors in world trade and oil prices are still significant and have a positive coefficient.

To provide an economic interpretation of our estimates, we compare the change in the forecast error by increasing one assumption error ceteris paribus by the respective standard deviation as tabulated in Table 1. Assumption errors for oil prices deviate on average from the sample mean by 15.23 USD/Barrel. An increase in the error in the assumption about the price of oil by one standard deviation comoves on average with an increase in the magnitude of the error in the GDP forecast by 0.68 percentage points. The standard deviation in the exchange rate is 0.17 USD/EURO. An increase in the assumption error about the exchange rate is, ceteris paribus, associated with a decrease in the magnitude of the GDP forecast error by, on average, 0.35 percentage points, according to the results in Table 2 column (6). The standard deviation in the assumption errors of world trade is 4.54%, and the associated increase in the GDP forecast errors is 0.80 percentage points. An increase of the forecast horizon by one year or 365 days will increase the magnitude of the forecast error by 1.75 percentage points. The root mean squared error for the growth of the GDP is, in the sample used for our regression analysis, 1.52%. If we use the estimated coefficients and residuals for (6) to predict the counterfactual squared forecast errors of the forecast of the GDP, without the squared errors of the assumptions ($e_{s,n,t+h|t}, e_{p^o,n,t+h|t}, e_{w,n,t+h|t} = 0, \forall n, t, h$), the predicted root mean squared forecast error of the growth of the GDP is 1.13%.[19] A reduction in the assumption errors by using instead 'correct' assumptions will not completely eliminate forecast errors of the GDP.

An increase in the squared forecast error for the current year's forecast for the same release date is associated with a higher squared error in the forecast for the next year forecasts of the GDP, as can be seen in Table 7 (column (2)). This implies that forecast errors of the growth of GDP by institutions for the next year comove with errors in the current year's forecast, if controlling for errors in the assumptions. Given our derivation in Section 3.2, we consider the possibility that the true data generating process or the forecasting rule includes assumptions for shorter horizons. The interaction term between the errors in the assumptions about the price of oil and world trade is significant and indicates that there is either a non-linear dependence of the assumption errors on the errors in the forecasts, or that the errors in the assumptions are not independent of each other (see Table 7 column (2)). When including the squared true realizations of the assumptions, only the coefficients for world trade and the price of oil are significant. Including the true realization of assumptions is an implicit test in a linear world of whether institutions on average know the true correlation between the assumptions and the forecast target. Our results show that institutions on average tend to predict the true

[19] Negative predictions for the squared forecast errors are set to zero.

comovement between growth in world trade and growth in GDP growth when the former is greater in magnitude.

The results of our robustness analyses show that including squared assumption errors about the interest rate in the regression analysis decreases the number of observations by almost 50%, to 340. Nevertheless, the coefficients for squared assumption errors about world trade only decreases slightly (see Table O.13 in the Appendix). The coefficient for squared assumption errors about the exchange rate is unaffected. The estimated coefficient for squared assumption errors about the price of oil is insignificant and negative. The squared assumption errors about the interest rate exhibit a positive and strong correlation with the squared forecast error of the GDP. An increase of the assumption error about the interest rate by 0.78 percentage points increases, ceteris paribus, the magnitude of the forecast errors of the GDP by 0.83 percentage points. This increase has a similar magnitude as the increase in the assumption error about world trade by one standard deviation. We have to keep in mind that including assumption errors about the interest rate restricts our analysis to forecasts published in the present millennium. However, we are not able to determine whether the significance of the estimate in the price of oil is caused by dropping the errors in the assumption about the interest rate or by extending the sample period to the early 1990s. Excluding observations before 1999 reveals that the coefficient for squared assumption errors about the price of oil is still significant without controlling for the interest rate. Therefore, the different estimate indicates an omitted variable bias between the assumption errors about the price of oil and about the interest rate. Including the interest rate as an explanatory variable yields an insignificant interaction term between errors in the forecast of world trade and errors in assumptions about the price of oil. As well, the interaction term between the rate of interest and world trade is significant at the five percent significance level (see Table O.13). Controlling for the interest rate leads to an insignificant coefficient for actual world trade squared. Our regression specification, including interaction terms, realizations of assumptions, and squared errors in the current year's forecast, has the highest adjusted R^2 of all the specifications. The finding that assumption errors about the price of oil do not have a robust correlation with forecast errors of the GDP is not surprising. In the past, higher oil prices led to investment in more energy efficient capital, a long-run consequence of high oil prices (see Elder and Serletis, 2010). Germany is one of the most energy efficient countries, according to Cohen, Jalles, Loungani, and Marto (2018). Blanchard and Gali (2007) state that the impact of oil price shocks on GDP growth for Germany even switched signs after 1984. An oil price shock will shift demand for imports from energy intensive countries to less energy intensive countries. German exports might benefit from higher oil prices, while German imports become more expensive.

The exclusion of forecasts for the year 2009 (see Table O.15 in the Appendix) reduces the coefficient for squared errors in the assumption about world trade from 0.03 to 0.01, but it remains significant. The coefficient for assumption errors about world trade depends on the inclusion of year-fixed effects. As reported in Table 2, the estimated coefficient drops from 0.083 in column (3) to 0.031 in column (4), that is, by more than 50%. In the sample, growth in world trade varies significantly from year to year. Our estimate for the regression coefficient of the exchange rate is still insignificant. The coefficient for assumption errors about the price of oil becomes insignificant and negative. It is obvious that in the last recession, the comovement between errors in the assumption and forecast errors was extraordinarily strong. However, excluding forecasts for the year 2009 reduces the number of observations for the regression by 6%. Separate analyses before and after 2008 did not affect our results.

It is not clear whether the residuals of the regression specification across institutions and years are correlated. We calculate standard errors for the estimated coefficients based on clustered residuals for both institutions as well as institutions and years (Tables O.25 and O.26). The analysis shows that the estimated coefficient for squared assumption errors about world trade is no longer significant. Nevertheless, the interaction terms for world trade and oil prices, and actual world trade squared are significant. Therefore, our finding, that squared assumption errors about world trade have a statistically significant impact, remains valid. The significant coefficient for actual world trade squared indicates that institutions tend to make forecast errors that are lower in magnitude in times of great changes in world trade. It also indicates that institutions tend to have a systematic bias in predicting the true relation between world trade and GDP growth.

The robustness analysis regarding the number of digits in the assumptions shows that the results do not differ for rounding procedures with more than two digits. Assumptions are reported in different units and exhibit different variations over time. The exchange rate varies mostly in the second digit after the decimal point. Hence, in order to get the variation in the assumption errors about the exchange rate, it is necessary to round appropriately. This is reflected by the tremendous differences between the coefficients for the squared errors in the assumption about the exchange rate.

Our final robustness analysis regarding systematic differences between international and national forecasters confirms our previous finding that institution-fixed effects do not add explanatory power to the regression model. The estimated coefficients for squared assumption errors about the price of oil and growth in world trade reported in Table O.30 are not very different for national and international forecasters. Only the coefficient for the exchange rate is negative, but insignificant for national forecasters, and positive and insignificant for international forecasters. Overall, the estimation results tend to be slightly better for national forecasters.

5 Conclusion

Economic forecasts are used to conduct fiscal and monetary policy. Good economic forecasts can improve the effectiveness of such policies. Macroeconomic forecasts are conditioned on ex-ante assumptions about the future development of exogenous factors linked to the global economy and financial markets. In this paper we have analysed the effects of errors in the assumptions on errors in the forecasts of the GDP of the German economy, and evaluated whether the forecasts are correctly conditioned. We find that errors in the assumptions about world trade and interest rates, in particular, have a large impact on the forecast errors of the GDP. However, our analysis only allows making statements about causality under the assumption that forecasters first make their forecasts of the assumptions and then make their forecast of the GDP. Otherwise, our analysis is only about correlation and not about causality.

Our results show that forcast errors for the current year comove with errors in the forecast for the next year even after controlling for errors in the assumptions. This implies that institutions take into account current year forecasts to determine the forecasts for the next year, either directly, or indirectly through the assumptions for the current year. The univariate as well as the multivariate correlation analysis results about the errors in the assumption about the exchange rate and the forecast errors of the growth of GDP indicate that the bilateral exchange rate between the US and the euro area might not be the relevant assumption to consider. The rise of the Chinese economy has led to a higher volume of exports and imports between China and Germany and, hence, the renminbi might be an assumption in the future. Furthermore, effective exchange rates are able to capture different decompositions of exports and imports by trade partners. Forecasters might already internally use assumptions on effective exchange rates, but they are not published yet.

To assess and improve the performance of economic forecasts, it is important to make the underlying assumptions transparent. In particular, publicly funded forecasters should publish these implicit assumptions. Our results indicate that the assumptions about world trade are the most important for the accuracy of forecasts of the German GDP. A more sophisticated and transparent way to make assumptions about world trade seems to be a necessary step, e.g. forecasters should state the source of the figures used and should publish their underlying assumptions about world trade.

Overall, our results are in line with those of Fioramanti, Gonzalez Cabanillas, Roelstraete, and Ferrandis Valterra (2016), who analyse the European Commission forecasts and underlying assumptions in a cross-country framework, finding that more than half of the variance in the errors in year-ahead forecasts appears to come from initial assumptions that prove to be incorrect ex post. However, they only find a limited impact on current-year forecasts of the growth of the GDP. Keereman (2003) finds that in the case of one-year ahead predictions, the initial assumptions together with the international economic environment outside the EU can account for up to about 60% of the error in the forecast of the GDP of the EU. Takagi and Kucur (2006) also find that the IMFs errors in its macroeconomic forecasts are largely correlated with errors in the underlying assumptions about oil prices and interest rates.

One major concern is that it is unclear whether and to what extent the set of assumptions are used in the institutions' forecast models. Another concern remains regarding last minute adjustments of the assumptions. During a forecasting exercise, which lasts up to two months, assumptions can be adapted to take into consideration the latest developments. However, they will not change the previous forecasting results if the changes are only minor. Publicly funded forecasters should consider publishing the protocols of their internal meetings, similarly to the publication of the accounts of the monetary policy meeting of the ECB or the minutes published by the Bank of England. As this analysis has only focused on Germany, further research might focus on differences between advanced and emerging economies, differences between oil importers and oil exporters, as well as floating/fixed exchange rate regimes.

References

Baffes, J. (1997): "Explaining stationary variables with non-stationary regressors," *Applied Economics Letters*, 4(1), 69–75.

Berge, T. J., A. Chang, and N. R. Sinha (2019): "Evaluating the conditionality of judgmental forecasts," Finance and Economics Discussion Series 002, Federal Reserve Board.

Bergholt, D., V. H. Larsen, and M. Seneca (2019): "Business cycles in an oil economy," *Journal of International Money and Finance*, 96, 283–303.

Blanchard, O. J., and J. Gali (2007): "The Macroeconomic Effects of Oil Shocks. Why are the 2000s so different from the 1970s?," *NBER Working Paper*, 13368.

Clark, T., and M. McCracken (2013): "Chapter 20 - Advances in Forecast Evaluation," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2, pp. 1107–1201. Elsevier.

Clements, M. P., F. Joutz, and H. O. Stekler (2007): "An evaluation of the forecasts of the Federal Reserve: a pooled approach," *Journal of Applied Econometrics*, 22(1), 121–136.

Cohen, G., J. T. Jalles, P. Loungani, and R. Marto (2018): "The long-run decoupling of emissions and output: Evidence from the largest emitters," *Energy Policy*, 118, 58–68.

Davies, A., and K. Lahiri (1995): "A new framework for analyzing survey forecasts using three-dimensional panel data," *Journal of Econometrics*, 68(1), 205–227.

Döhrn, R., and C. M. Schmidt (2011): "Information or Institution? – On the Determinants of Forecast Accuracy," *Jahrbücher für Nationalökonomie und Statistik*, 231(1), 9–27.

Döpke, J., U. Fritsche, and G. Waldhof (2019): "Theories, techniques and the formation of German business cycle forecasts," *Jahrbücher für Nationalökonomie und Statistik*, 239(2), 203–241.

Elder, J., and A. Serletis (2010): "Oil price uncertainty," *Journal of Money, Credit and Banking*, 42(6), 1137–1159.

Fioramanti, M., L. Gonzalez Cabanillas, B. Roelstraete, and S. Ferrandis Valterra (2016): "European Commission's Forecasts Accuracy Revisited: Statistical Properties and Possible Causes of Forecast Errors," European Economy Discussion Papers 027, European Commission.

Fritsche, U., and U. Heilemann (2010): "Too Many Cooks? The German Joint Diagnosis and Its Production," Discussion Paper 1, University of Hamburg, Department of Socioeconomics.

Heilemann, U. (2002): "Increasing the transparency of macroeconometric forecasts: a report from the trenches," *International Journal of Forecasting*, 18(1), 85–105.

Heilemann, U., and K. Müller (2018): "Wenig Unterschiede – Zur Treffsicherheit internationaler Prognosen und Prognostiker," *AStA Wirtschafts- und Sozialstatistisches Archiv*, 12(3–4), 195–233.

Heilemann, U., and H. O. Stekler (2003): "Has the accuracy of German macroeconomic forecasts improved?," Discussion Paper 31, University Dortmund.

Júlio, P., and P. Esperança (2012): "Evaluating the forecast quality of GDP components: An application to G7," CEE papers No 47, Gabinete de Estratégia e Estudos, Ministério da Economia e da Inovação.

Keereman, F. (2003): "External assumptions, the international environment and the track record of the Commission Forecast," European Economy Economic Papers 189, Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Kilian, L. (2008): "The economic effects of energy price shocks," *Journal of Economic Literature*, 46(4), 871–909.

Kotchoni, R., M. Leroux, and D. Stevanovic (forthcoming): "Macroeconomic forecast accuracy in a data-rich environment," *Journal of Applied Econometrics*.

Muth, J. F. (1961): "Rational expectations and the theory of price movements," *Econometrica: Journal of the Econometric Society*, 29(3), 315–335.

Pons, J. (2000): "The accuracy of IMF and OECD forecasts for G7 countries," *Journal of Forecasting*, 19(1), 53–63.

Sinclair, T., H. O. Stekler, and H. C. Müller-Droge (2016): "Evaluating forecasts of a vector of variables: A German forecasting competition," *Journal of Forecasting*, 35(6), 493–503.

Takagi, S., and H. Kucur (2006): "Testing the accuracy of IMF macroeconomic forecasts, 1994–2003," IEO Background Paper BP/06/1, International Monetary Fund IEO.

Timmermann, A. (2007): "An evaluation of the World Economic Outlook forecasts," *IMF Staff Papers*, 54(1), 1–33.

A General Derivation of the Baseline Regression Specification

A forecast $\hat{y}_n$ of GDP growth by institution n for $t+h$ at t is a function of the underlying assumptions $j \in \{p^o,\, s,\, w\}$ and an unobserved component $\hat{\kappa}_{n,t+h|t}$

$$\hat{y}_{n,t+h|t} = f(\hat{x}_{p^o,n,t+h|t}, \ldots, \hat{x}_{w,n,t+h|t}, \hat{\kappa}_{n,t+h|t}). \tag{7}$$

Furthermore, we assume that the data generating process of y also depends on the variables used to forecast y

$$y_{t+h} = g_(x_{p^o,t+h}, \ldots, x_{w,t+h}, \kappa_{t+h}). \tag{8}$$

In order to derive the baseline regression specification, we assume that the forecasters know the true data generating process and make forecasts accordingly: $f = g$. Both functions are assumed to be continuous and differentiable.

The error in the forecast of a variable or assumption $e_{t+h|t} = \hat{x}_{t+h|t} - x_{t+h}$ can be approximated by a Taylor approximation around the sample averages $\bar{x}_j, \bar{\kappa}$.

$$e_{n,t+h|t} \overset{1.T.A.}{=} \frac{\partial f}{\partial x_{p^o}}\, e_{p^o,n,t+h|t} + \frac{\partial f_i}{\partial x_s}\, e_{s,n,t+h|t} + \frac{\partial f}{\partial x_w}\, e_{w,n,t+h|t} + \frac{\partial f}{\partial \kappa}\, e_{\kappa,n,t+h|t} + \epsilon_{n,t} \tag{9}$$

The first order Taylor approximation states that the error in the forecast by an institution is a linear function of the errors in the underlying assumptions and the unobserved component and an approximation error $\epsilon_{n,t}$. A forecaster's objective is to be as close as possible to the true realization. Therefore, we square (9) to have a measure of the accuracy of the forecast. We derive a relation between the squared forecast errors and errors in the assumptions $j, l \in \{p^{no},\, s,\, r,\, w\}$ as follows:

$$\begin{aligned}
e^2_{n,t+h|t} =& \sum_j \underbrace{\left(\frac{\partial f}{\partial x_j}\right)^2}_{\beta^j} e^2_{j,n,t+h|t} + \sum_j \sum_l \underbrace{\left(\frac{\partial f}{\partial x_j}\,\frac{\partial f}{\partial x_l}\right)}_{\beta^{j,l}} e_{j,n,t+h|t}\, e_{l,n,t+h|t} \\
&+ \sum_j \underbrace{\left(\frac{\partial f}{\partial x_j}\,\frac{\partial f}{\partial \kappa}\right)}_{\beta^{j,\kappa}} e_{j,n,t+h|t}\, e_{\kappa,n,t+h|t} + \underbrace{\left(\frac{\partial f}{\partial \kappa}\right)^2 e^2_{\kappa,n,t+h|t}}_{\gamma_n+\gamma_t+\gamma_{n,t}+\beta^h\, h_{n,t}+\beta^{h^2}\, h^2_{n,t}} \\
&+ \underbrace{\sum_j \frac{\partial f}{\partial x_j}\, e_{j,n,t+h|t}\, \epsilon_{n,t} + \frac{\partial f}{\partial \kappa}\, \kappa_{j,n,t+h|t}\, \epsilon_{n,t} + \epsilon^2_{n,t}}_{\alpha+\varepsilon_{n,t+h}}
\end{aligned} \tag{10}$$

As described in the main text, we approximate $e_{\kappa,n,t+h|t}$ by including year-fixed effects γ_t, institution-fixed effects γ_n and the forecast horizon in days $h^{n,t}$. We further assume that the correlation between errors in the assumptions and the unobserved components is zero. Our regression residual is a composite of interactions between unobserved components and the errors due to the use of this Taylor approximation as well as the squared Taylor approximation residual. It is not clear whether these terms are zero in expectation; if not, one would have to include a constant. In our baseline specification, we assume that the interaction terms are zero and we derive the following regression equation.

$$e^2_{n,t+h|t} = \beta^{p^o}\, e^2_{p^o,n,t+h|t} + \beta^s\, e^2_{s,n,t+h|t} + \beta^w\, e^2_{w,n,t+h|t} + \gamma_n + \gamma_t + \gamma_{n,t} + \beta^h\, h_{n,t} + \beta^{h^2}\, h^2_{n,t} + \epsilon_{n,t+h} \tag{11}$$

B Tables

Table 3: List of Institutions and Forecast Publications

	institution	publication title	publication series
national			
BBK	Deutsche Bundesbank	Perspektiven der deutschen Wirtschaft	Monatsbericht
DIW	German Institute for Economic Research	Grundlinien der Wirtschaftsentwicklung	DIW Wochenbericht
GD	Joint Economic Forecast of German Economic Research Institutes (Gemeinschaftsdiagnose)	Gutachten	
HWWA, HWWI	Hamburg Institute of International Economics	HWWA–Konjunkturforum	Wirtschaftsdienst
IW	German Economic Institute	IW-Trends	
ifo	Leibniz Institute for Economic Research at the University of Munich	ifo Konjunkturprognose	ifo Schnelldienst: Daten und Prognosen
IfW	Kiel Institute for the World Economy	Deutsche Konjunktur	Kieler Konjunkturberichte; Kieler Diskussionsbeitrge
IMK	Macroeconomic Policy Institute	Die konjunkturelle Lage in Deutschland	IMK Report
IWH	Halle Institute for Economic Research	IWH Konjunktur aktuell	Konjunktur aktuell; Wirtschaft im Wandel
RWI	RWI - Leibniz Institute for Economic Research	Die wirtschaftliche Entwicklung	Konjunkturbericht
WSI	Institute of Economic and Social Research	WSI Konjunkturbericht	WSI Mitteilungen
Destatis	German Statistical Office	Fachserie 18	
international			
IMF	International Monetary Fund	World Economic Outlook	World Economic Outlook; World Economic Outlook Updates
OECD	Organisation for Economic Co-operation and Development	OECD Economic Outlook	OECD Economic Outlook
EC	European Commission	European Economic Forecast; Interim Forecast	European Economy Institutional Paper

B.1 Descriptive Statistics

Table 4: Correlation GDP growth and assumptions

Variables	Actual data (annual frequency)	Actual data (quarterly frequency)	Forecast and assumptions	Forecast and assumption errors
Oil Price	0.09	0.01	-0.01	0.38
	[-0.30,0.45]	[-0.19,0.21]	[-0.07,0.05]	[0.33,0.43]
Exchange Rate	-0.08	-0.06	-0.09	0.03
	[-0.45,0.31]	[-0.26,0.14]	[-0.15,-0.03]	[-0.03,0.09]
Interest Rate	0.18	-0.04	0.12	0.68
	[-0.29,0.57]	[-0.26,0.18]	[0.03,0.20]	[0.63,0.72]
World Trade	0.72	0.66	0.75	0.78
	[0.47,0.86]	[0.53,0.76]	[0.71,0.78]	[0.75,0.81]

Note: Pearson correlation coefficient and associated 95% confidence bands in parentheses.

Table 5: P-values of KPSS tests for baseline regression residuals

Institution	Biasedness	Horizon	Institutional FE	Year FE	Year and Inst. FE without interaction	Baseline
Bundesbank	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10
DIW	- / -	- / -	- / -	- / -	- / -	- / -
EC	- / -	- / -	- / -	- / -	- / -	- / -
GD	0.02 / 0.10	0.02 / 0.10	0.06 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10
HWWA	- / -	- / -	- / -	- / -	- / -	- / -
ifo	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10
IfW	- / -	- / -	- / -	- / -	- / -	- / -
IMF	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10
IMK	- / -	- / -	- / -	- / -	- / -	- / -
IWH	0.10 / 0.09	0.10 / 0.09	0.10 / 0.09	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10
OECD	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10
RWI	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10
WSI	- / -	- / -	- / -	- / -	- / -	- / -
IWH	0.10 / 0.09	0.10 / 0.09	0.10 / 0.09	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10

Note: P-values for the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. Under the null hypothesis the tested time series is stationary. The first p-value in a cell tests residuals for current year forecasts and the second for next year forecasts.

Table 6: P-values of KPSS tests for regression residuals with extensions

Baseline	Current Year Forecast	Interaction	Assumptions	All	
Bundesbank	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.02 / 0.10
DIW	- / -	- / -	- / -	- / -	- / -
EC	- / -	- / -	- / -	- / -	- / -
GD	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10
HWWA	- / -	- / -	- / -	- / -	- / -
ifo	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10
IfW	- / -	- / -	- / -	- / -	- / -
IMF	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10
IMK	- / -	- / -	- / -	- / -	- / -
IWH	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10
OECD	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10
RWI	0.10 / 0.10	0.10 / 0.10	0.09 / 0.10	0.10 / 0.10	0.10 / 0.10
WSI	- / -	- / -	- / -	- / -	- / -
IWH	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10	0.10 / 0.10

Note: P-values for the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. Under the null hypothesis the tested time series is stationary. The first p-value in a cell tests residuals for current year forecasts and the second for next year forecasts.

B.2 Regression results

Table 7: Regression Results for Squared Forecast Errors of GDP

	Dependent variable: SFE GDP				
	Baseline	Current Year Forecast	Interaction	Assumptions	All
	(1)	(2)	(3)	(4)	(5)
SFE Current Year		−0.379*			−0.401*
		(0.212)			(0.218)
AE Oil Price and Exchange Rate			0.097		0.156
			(0.177)		(0.172)
AE Oil Price and World Trade			0.016***		0.018***
			(0.005)		(0.005)
AE Exchange Rate and World Trade			0.659*		0.416
			(0.381)		(0.353)
Actual Oil Price squared				−0.001	−0.0002
				(0.001)	(0.001)
Actual Exchange Rate squared				−4.746	−14.421
				(20.617)	(17.019)
Actual World Trade squared				−0.006	−0.048***
				(0.016)	(0.017)
SAE Oil Price	0.002**	0.002**	−0.002	0.002**	−0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
SAE Exchange Rate	−4.139	−4.835	4.494	−4.139	2.971
	(5.203)	(5.399)	(4.574)	(5.203)	(4.529)
SAE World Trade	0.031*	0.028*	0.018	0.031*	0.014
	(0.016)	(0.015)	(0.013)	(0.016)	(0.011)
Horizon (days)	0.001	0.0001	0.002	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Horizon (days) squared	0.00001*	0.00001***	0.00001*	0.00001*	0.00001***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Constant	1.097	1.269	−3.825*	9.500	22.333
	(2.661)	(2.763)	(2.284)	(34.116)	(28.148)
Year FE	YES	YES	YES	YES	YES
Institutional FE	YES	YES	YES	YES	YES
Interaction Year and Inst. FE	YES	YES	YES	YES	YES
Observations	622	622	622	622	622
R^2	0.909	0.914	0.928	0.909	0.934
Adjusted R^2	0.873	0.880	0.899	0.873	0.907
Residual Std. Error	2.414	2.347	2.150	2.414	2.066
F Statistic	25.382***	26.841***	32.106***	25.382***	34.785***
Number of Parameters	338	339	341	341	345

Note: *p<0.1; **p<0.05; ***p<0.01.

Halle Institute for Economic Research –
Member of the Leibniz Association

Kleine Maerkerstrasse 8
D-06108 Halle (Saale), Germany

Postal Adress: P.O. Box 11 03 61
D-06017 Halle (Saale), Germany

Tel +49 345 7753 60
Fax +49 345 7753 820

www.iwh-halle.de

ISSN 2194-2188

Leibniz
Association