IWH
Halle Institute for Economic Research
Member of the Leibniz Association

Discussion Papers

No. 22
October 2024

Step by Step – A Quarterly Evaluation of EU Commission's GDP Forecasts

Katja Heinisch

Author

Katja Heinisch
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

The responsibility for discussion papers lies solely with the individual authors. The views expressed herein do not necessarily represent those of IWH. The papers represent preliminary work and are circulated to encourage discussion with the authors. Citation of the discussion papers should account for their provisional character; a revised version may be available directly from the authors.

Comments and suggestions on the methods and results presented are welcome.

IWH Discussion Papers are indexed in RePEc-EconPapers and in ECONIS.

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

Step by Step – A Quarterly Evaluation of EU Commisson‘s GDP Forecasts*

Abstract

The European Commission's growth forecasts play a crucial role in shaping policies and provide a benchmark for many (national) forecasters. The annual forecasts are built on quarterly estimates, which do not receive much attention and are hardly known. Therefore, this paper provides a comprehensive analysis of multi-period ahead quarterly GDP growth forecasts for the European Union (EU), euro area, and several EU member states with respect to first-release and current-release data. Forecast revisions and forecast errors are analyzed, and the results show that the forecasts are not systematically biased. However, GDP forecasts for several member states tend to be overestimated at short-time horizons. Furthermore, the final forecast revision in the current quarter is generally downward biased for almost all countries. Overall, the differences in mean forecast errors are minor when using real-time data or pseudo-real-time data and these differences do not significantly impact the overall assessment of the forecasts' quality. Additionally, the forecast performance varies across countries, with smaller countries and Central and Eastern European countries (CEECs) experiencing larger forecast errors. The paper provides evidence that there is still potential for improvement in forecasting techniques both for nowcasts but also forecasts up to eight quarters ahead. In the latter case, the performance of the mean forecast tends to be superior for many countries.

Keywords: consensus forecasts, data revision, forecast evaluation, forecast horizon, forecasting, nowcasting, professional forecasters

JEL classification: C32, C52, C53, E37

* I would like to thank Joan Paredes, Laura Bardone, John Duca, and participants of the International Symposium on Forecasting 2023, CFE 2023, ICMAIF 2024, IAAE 2024, and IWH macroeconomic seminar for helpful comments.

1 Introduction

Policymakers and central bankers focus on GDP estimates for the current quarter (nowcast) and short-term forecasts for the next quarter to assess the current state of the economy. While the assessment of the current state of the economy is certainly an important element in macroeconomic forecasting (Barhoumi *et al.*, 2008), economic policy decisions and financial projections are mainly based on annual projections.[1] This is particularly important in the European Union, where new fiscal regulations implemented in 2024 require member states to adhere to stringent guidelines for reducing deficits and debt levels relative to GDP. These regulations emphasize the importance of accurate GDP forecasts to ensure sustainable public finances and promote inclusive growth.

In this context, this paper makes several key contributions to the literature on GDP forecasting. First, it provides a comprehensive analysis of quarterly GDP forecasts published by the European Commission (EC) from 2000 to 2023, focusing on differences in the EC's forecast updating behaviour and forecast errors across various horizons and European countries, acknowledging the inherent uncertainty in these forecasts, especially during periods of economic instability. This analysis fills a significant gap in the existing research, which has largely focused on annual GDP forecasts rather than the more challenging quarterly predictions. Second, this paper compares the properties of quarterly forecasts by the EU Commission to those of mean forecasts and quarterly Consensus Economics forecasts. Finally, the paper examines the impact of data revisions on forecast accuracy, comparing EC forecasts with both first-release and current (final) data across countries, thereby contributing to the broader discussion on the use of real-time data in economic forecasting.

Forecasts by the European Commission provide a framework for policy decisions and are used as a benchmark for national forecasters as well. The Commission publishes a fully-fledged economic outlook on various macroeconomic indicators for the EU and its member countries twice a year (spring, and autumn) and in recent years updates key indicators in winter and summer as well. Since the initial investigation of the forecast performance of the EU Commission for annual GDP growth by Keereman (1999), several evaluation studies were published, e.g. by Melander *et al.* (2007); Cabanillas *et al.* (2012); Fioramanti *et al.* (2016); ICF, DIW, NIESR and OFCE (2017); Chabin *et al.* (2020), with the overall finding that the annual forecasts are unbiased and efficient. However, none of these studies covers the quarterly domain, despite that those are particularly important to determine the position in the business cycle. In addition, the consistency with available quarterly-frequency data is of utmost importance for the accuracy of the annual forecast. Furthermore, the high performance of annual GDP predictions does not necessarily imply that forecasters can perform well in the more difficult task of quarterly GDP predictions within the current year and even beyond. In general, if the forecast performance for the first two quarters of a year is high, then the overall performance of the annual growth rate tends to be high. However, predictions with a longer forecast horizon usually tend to be

[1] This holds even more so as the longer-term predictability of quarterly GDP growth has declined since the 1990s (D'Agostino, 2006).

less accurate than the forecast for the current and the next quarter. In general, analyses of quarterly GDP growth forecasts are sparse and have only been conducted for different forecast institutions and a different set of countries (e.g. Zarnowitz, 1979; Blunier and Hepenstrick, 2022; Fildes and Stekler, 2002; Öller and Barot, 2000). Due to the complex correlation structure of the multi-period quarterly forecasts, forecast accuracy tends to be difficult to interpret (Zarnowitz, 1979). Forecasts (and their errors) might be serially correlated within the sequence of multi-period forecasts estimated in a certain forecast period and across the successive sequences, which overlap and refer partially to the same target.

The delay in publishing quarterly national accounts poses a significant challenge to effective decision-making. Therefore, the existing literature has proposed several methods to receive earlier estimates based on (higher-frequency) indicators. These papers deal with large information sets and various methods to improve nowcasting accuracy through approaches such as bridge models, factor models, mixed-data sampling models (MIDAS), and mixed-frequency vector-autoregressive (MFVAR) models. Additionally, model averaging or forecast combination are nowadays standard tools for forecasters. Bańbura *et al.* (2013) provide an overview of nowcasting goals and methods, which is extended by a recent meta-study on nowcasting tools by Stundziene *et al.* (2023). A drawback of many empirical analyses is that they focus mainly on the euro area (Marcellino and Schumacher, 2010; Angelini *et al.*, 2011; Kuzin *et al.*, 2011; Drechsel and Maurin, 2011) or selected core countries such as Germany, France and Italy (Baffigi *et al.*, 2002; Rünstler and Sédillot, 2003; Schumacher and Breitung, 2008). Country comparisons are rather limited (Jansen *et al.*, 2016). Only, in the context of nowcasting, Cascaldi-Garcia *et al.* (2023) adopt a multi-country econometric framework to nowcast economic activity and show that simultaneously monitoring (nowcasting) the economic conditions of the euro-area aggregate and its three largest member countries (Germany, France, and Italy) increase country-specific forecast performance. Only a limited number of studies covers countries in Central and Eastern Europe, such as Poland, Hungary, and Lithuania (Rünstler *et al.*, 2009). In addition, the studies differ in the size of the information set, the method, and the sample period used. Some papers make use of Consensus forecasts published by Consensus Economics to compare pure model-based forecasts with private sector analysts' quarterly expectations for GDP (Jansen *et al.*, 2016; De Winter, 2011). These studies find stickiness in quarterly Consensus Economics forecasts and a tendency of herd behaviour, although forecasters adjust their forecasts quickly to new information about the state of the economy, they still fall behind the best model in most cases at the short horizon, except for Germany and Spain.[2]

Finally, a major issue in the assessment of forecast errors is the selection of its target, i.e. whether first-release data or current-release data are used for comparison of the respective quarters. Using different data vintages on activity in the euro area, Diron (2008) assesses the impact of using real-time data and data revisions on short-term forecasts of real GDP growth. She finds that while on average pseudo-real-time exercises seem to be reliable, the difference in

[2] Consensus forecasts are often used as a benchmark for annual GDP or inflation forecast comparisons. However, forecast data on quarterly forecasts is hardly noticed, due to the limited number of countries covered and the limited sample size.

specific quarters might even lead to a different assessment of economic activity. Therefore, this paper is also related to data revisions and compares the EC forecasts with both the first release and current data (final) across countries.

The analysis also sheds light on the important problem of how to make an accurate forecast in the presence of shocks and possible structural breaks. In addition, the degree of synchronization between business cycles in different countries plays a crucial role in macroeconomic forecasting. Understanding how closely aligned or divergent the business cycles of EU member states are can help explain variations in forecast accuracy and the challenges in predicting economic outcomes across regions. Business cycle fluctuations in the euro area were well synchronized before the Great Recession (see, e.g., Darvas and Szapáry, 2008), but synchronization of economic activity between core and the periphery decreased afterwards (Belke *et al.*, 2017). Beck (2020) finds evidence of two distinct business cycles in the Eurozone and Central and Eastern Europe. Over the last few years, European economies have experienced an unprecedented series of shocks. The Covid pandemic resulted in a partial shutdown of the economies. The Russian-Ukrainian war is leading to profound changes in energy markets and trade patterns. The directly visible impact of these shifts has been the return of high inflation. Hence, besides short-term impacts on the economy, these shifts could also have profound longer-term implications and come with fundamental changes like global economic interactions. The effects of accelerating climate change will likely lead to more frequent supply shocks in the future.

The paper is organised as follows. The next section describes the forecast dataset of the European Commission and Consensus Economics, as well as the target series of GDP growth by Eurostat. The empirical approach is presented in Section 3 and the results of the empirical analysis are discussed in Section 4. Finally, Section 5 provides conclusions on the European Commission's quarterly forecasting setup.

2 Data

2.1 GDP Forecasts

The European Commission publishes macroeconomic forecasts for the EU and its member countries up to four times a year. Recent forecasts cover both annual growth rates and quarterly growth rates for key macroeconomic variables. While annual GDP forecasts have been conducted since the 1960s, quarterly forecasts are only available from 2002 onwards. The fully-fledged forecasts are regularly published in spring and autumn. Furthermore, interim forecasts for key annual growth rates such as GDP have been provided since 2008. Since 2012, winter forecasts for quarterly data have been provided and since 2019, also quarterly forecasts have been published by the EC in summer. Therefore, the majority of the forecasts analysed were conducted in the spring and autumn. In each quarter, up to 8 forecast horizons have been covered, where $h = 0$ refers to the actual quarter (nowcast) and $h = 8$ refers to the quarter 8 periods ahead. Hence, two forecast definitions can be addressed within the analysis: the evaluation of fixed-horizon forecasts (GDP h-quarters ahead) and fixed-event forecasts (for a particular quarter). The sample covers 12 core countries (Austria, Belgium, Denmark, Finland, France, Germany, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden) and 10 Central and Eastern European countries (CEECs: Bulgaria, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovenia, Slovakia).[3] Furthermore, GDP growth forecasts for the EU aggregate and the euro area aggregate are analysed. Given their changing composition, the mean performance should be taken with care. For instance, although some CEECs joined in 2004 they were not included in the EU aggregate forecasts in 2004, mainly due to concerns about their data availability as well as stability and quality. Thus EU continues to refer to the EU15 for the years 2004 and 2005 (Melander *et al.*, 2007).

The main forecast evaluation period in this analysis covers data from 2000Q2–2019Q4, although data until 2023Q4 is available.[4] This is motivated first by the fact that so-called *final* quarterly GDP growth rates are published around 2-3 years after the respective quarter, i.e. when all annual revisions based on structural sources are incorporated. In addition, benchmark revisions take also place regularly. Furthermore, the shorter sample is not affected by the huge fluctuations in GDP growth during the COVID-19 period. While for most countries, the drop in GDP in 2020Q2 was well predicted by the EC in spring 2020, the size and speed of the recovery in the subsequent quarters were not expected, resulting in extreme forecast errors exceeding 25 percentage points for the euro area. Overall, more than 380 quarterly forecasts have been conducted for most of the member states, with up to 57 forecasts published for the nowcast, i.e. the current quarter, and almost 20 forecasts available for the period t+8 (Table A1). For the Central and Eastern European countries about 30 quarters can be evaluated for the nowcast (Table A9).[5]

The EU Commission's forecasts are not produced from a single global model. The procedure involves the interaction between the top-down (horizontal units providing guidance, assumptions, ex-ante and ex-post controls) and bottom-up approach (by country desks) in three iterations for the fully-fledged forecast (and two iterations for the interim forecasts). EU and euro area-wide forecast data are obtained by aggregation and incorporating inputs from a variety of models.[6] The national forecasts undergo thorough peer review processes to ensure their accuracy and reliability. Overall both intra-country consistency and cross-country consistency are important ingredients.

Besides comparing the EC's forecast performance across countries, I use the previous sample mean as a benchmark, given that for longer forecast horizons, mean GDP growth is moving in the direction of potential growth.Furthermore, quarterly Consensus Economics forecasts are evaluated as they represent the only source that offers quarterly forecasts for a wide range of EU countries. The private sector firm Consensus Economics collects and publishes economic

[3] Forecast data for Ireland, UK, Greece and Malta and Cyprus has not been considered due to the small number of forecasts.

[4] Results for the full sample are available upon request.

[5] Detailed tables and figures for all countries are provided in the Online Appendix (A).

[6] See ICF, DIW, NIESR and OFCE (2017) for a description of the EC forecasting procedure.

forecasts for a set of key macroeconomic variables on a monthly basis for a broad range of countries. While monthly Consensus Forecasts are best known for their expectations of annual GDP growth for the current and next year, they also provide quarterly forecasts for GDP once per quarter for the G7/Western Europe countries – compared to the previous quarter (in March, June, September and December). The panellists supply their forecasts for six consecutive quarters, starting with the first unpublished quarter. The countries covered are Germany, France, UK, Italy, Euro Area, the Netherlands, Spain and Sweden. Consensus quarterly forecasts for some CEECs countries (Poland, Czech Republic and Hungary) exist as well. However, for the latter, the quarterly sample starts very late, are published only twice a year and the forecasts are only depicted compared to the previous year. Due to the lack of a pure real-time time series in levels, it is not possible to convert those growth rates properly and, hence, those forecasts are not considered in the analysis below.

2.2 GDP Data

To evaluate the performance of the EC's forecasts, quarterly seasonally and working day-adjusted GDP growth rates provided by Eurostat published in March 2024 are used. GDP data, based on figures of GDP components, is published about t+65 days after the end of the corresponding quarter.[7] The absence of official real-time data for quarterly GDP has left the analysis of data revisions on forecast accuracy hardly addressed.[8]

The absence of official real-time data for quarterly GDP introduces uncertainty into the forecast evaluation process. However, data revisions could significantly impact quarterly short-term forecasts of GDP and their reliability (Aruoba, 2008; Croushore, 2011). Therefore, this study creates a unique *real-time data* set on first releases for almost all European countries. For German GDP data, the real-time dataset published by the Deutsche Bundesbank has been used to match the real-time values. For France, Italy, the Netherlands, Spain, and Sweden, the data have been extracted from Consensus Economics quarterly data, which includes both historical data and forecast data. The data for the other countries are extracted as a snapshot from the EC forecast publications, given that the forecast tables include historical data as well.[9]

The forecast performance can only be realistically judged based on the information set available when the actual forecast is produced. An advantage of using the first releases is that the forecaster's performance is not biased if the statistical agencies will revise the data in the future. Furthermore, the use of first available estimates in the assessment of current quarter forecast accuracy is motivated by the greater attention usually attracted by first available estimates,

[7] Flash estimates are available only for selected countries already 30 days after the end of the quarter since April 2016.

[8] Existing real-time datasets cover only the euro area (Diron, 2008) or a selection of countries (e.g. Germany, France, Italy, Spain (Asimakopoulos *et al.*, 2023)).

[9] For the periods where the EC has published forecast only biannually or three times a year, the real-time quarters for every second quarter refers to the first revision of the series. For instance, the forecast in April 2002, covers the real-time time data for the 2001Q4 (published in February) and the first revision for 2001Q3 (published in February). Given the publication schedule of the EC forecast, the real-time data does not cover flash estimates of GDP growth and, hence, refers to Eurostat's second estimate for GDP.

compared to later revisions. However, it could also be argued that it would be more appropriate to compare the forecast with the final data. Therefore, for completeness, I compare the forecasts to their final values, which were published in February/March 2024 (*current-release data*). This data release covers new information and several methodological changes in data revisions as well, e.g. the introduction of chain-linked volumes in 2005/06 or the ESA2010 concept in 2014. Therefore, the final data is based on a different information set than that available to the forecaster. Figure 1 illustrates the differences between real-time data and final data for the euro area.[10] Table 1 provides an overview of descriptive statistics for quarterly GDP growth in various EU countries from 2000–2023. The average quarterly growth rate ranges from 0.1% in Italy to 0.7% in Luxemburg. The euro area's average quarterly growth rate is about 0.3%. The measure of the GDP volatility (standard deviation) implies that the more volatile GDP growth is, the more difficult it usually is to predict, which is e.g. the case for Luxembourg. The GDP growth rates for countries in Central and Eastern Europe (CEECs) are higher (see Table 1, lower panel), with average growth rates compared to the previous quarter in the range of 0.6% up to almost 1%. The volatility of economic growth is also higher compared to the core countries. The COVID-19 period caused large outliers which is also evident in higher standard deviations compared to a sample until 2019Q4. However, the mean growth rates remain almost unaffected, both for core European countries as well as the CEECs.

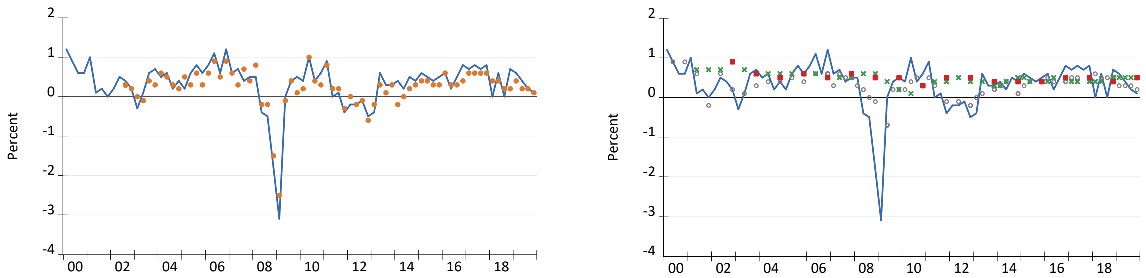
Data revisions in official statistics are a major issue in macroeconomic forecasting and incorporate new information that was not available at the time of the initial announcement. They are defined as the difference between the actual data (release in quarter v) compared to the first release y_t^0 for a certain quarter t, i.e. $\Delta y_t^{rev} = y_t^v - y_t^0$. In contrast to annual national accounts, quarterly national accounts revisions are more substantial and occur more frequently because of the characteristics of data sources and compilation methods. Eurostat (2013) points out that there is a trade-off between timeliness and accuracy, i.e. flash estimates are in general more prone to revision than traditional estimates. In addition, once annual information such as sectoral statistics on economic structures becomes available and the annual aggregate is revised, the corresponding quarterly figures are revised as well.

Data revision should have several properties (see Aruoba, 2008; Asimakopoulos *et al.*, 2023): i) the mean of data revision should be zero and, hence, not be biased, ii) the volatility of revision should be small in comparison to the volatility of the series themselves. iii) revision should not be predictable. Table 1 includes an overview of quarterly data revisions, covering mean revisions and mean absolute revisions. Further, the deviation from the mean revision and the correlation between first and current-release data are depicted. The mean revision of official statistics (until 2019) is zero for almost all countries (except Belgium with up to 0.1 percentage points and Sweden with –0.1 percentage points). A finding also confirmed by Asimakopoulos *et al.* (2023) who show that statistical offices rather revise other GDP components, on the expenditure side, and cancel each other out at an aggregate level. However, the mean absolute revision from the first publication to the current release (spring 2024) ranges from 0.2 percentage points in France to more than 1 percentage point in Luxembourg. Overall, the data revisions are larger for

[10] Comparisons for others countries are available in Figures A1 and A2 in the Online Appendix.

Figure 1: Euro area GDP growth

(a) Data **(b) Forecasts**



Note: Left figure: Actual data for GDP growth (line) compared to the initial releases (dots). Right figure: EC forecasts for each quarter with horizon h=0 (nowcast, circle), h=4 (one-year ago, cross) and h=8 (2 years ago, square) are depicted with the current-release data by Eurostat (blue line).

Table 1: Descriptive Statistics for Data Sample

	sample 2000Q1–2019Q4							sample 2000Q1–2023Q4						
	data			data revision				data			data revision			
	obs	mean	stdev	mean rev	mean abs rev	stdev rev	corr rev	obs	mean	stdev	mean rev	mean abs rev	stdev rev	corr rev
EU	80	0.366	0.578	0.046	0.152	0.188	0.955	96	0.353	1.752	0.024	0.188	0.389	0.975
EA	80	0.330	0.596	0.069	0.191	0.227	0.938	96	0.322	1.844	0.080	0.194	0.241	0.993
AT	80	0.386	0.645	0.046	0.336	0.464	0.702	96	0.350	1.896	0.017	0.371	0.539	0.961
DE	80	0.330	0.888	0.063	0.365	0.480	0.841	96	0.285	1.587	0.086	0.364	0.472	0.954
FR	80	0.335	0.486	0.048	0.198	0.232	0.884	96	0.326	2.372	0.040	0.206	0.272	0.995
ES	80	0.439	0.676	0.027	0.243	0.333	0.872	96	0.428	2.630	0.050	0.265	0.375	0.990
IT	80	0.075	0.686	0.026	0.271	0.352	0.862	96	0.124	2.132	0.042	0.308	0.442	0.982
PT	80	0.226	0.788	0.099	0.376	0.516	0.777	96	0.272	2.415	0.080	0.449	0.813	0.943
BE	80	0.416	0.507	0.120	0.263	0.312	0.793	96	0.416	1.801	0.089	0.291	0.467	0.966
LU	80	0.716	1.553	-0.057	0.958	1.320	0.503	96	0.650	1.865	-0.046	1.012	1.348	0.748
NL	80	0.361	0.651	0.048	0.323	0.541	0.700	96	0.371	1.306	0.057	0.341	0.558	0.922
DK	80	0.333	0.836	0.083	0.455	0.585	0.729	96	0.385	1.248	0.093	0.505	0.669	0.854
FI	80	0.394	1.196	0.039	0.639	0.857	0.705	96	0.327	1.409	0.004	0.617	0.844	0.801
SE	80	0.565	0.902	-0.111	0.566	0.794	0.571	96	0.530	1.447	-0.043	0.549	0.790	0.839
BG	80	0.961	1.414	-0.002	0.504	0.923	0.750	96	0.892	1.550	0.064	0.745	1.329	0.547
CZ	80	0.719	0.843	0.077	0.370	0.484	0.877	96	0.597	1.492	0.024	0.471	0.843	0.870
EE	80	0.944	1.969	-0.073	0.914	1.390	0.747	96	0.800	2.028	0.104	1.078	1.638	0.662
HR	80	0.514	1.320	0.024	0.673	0.879	0.555	96	0.604	2.298	0.085	1.341	2.713	0.412
LV	80	0.851	2.158	-0.173	1.292	1.910	0.598	96	0.777	2.277	-0.014	1.356	2.026	0.578
LT	80	1.049	1.875	0.184	0.687	1.321	0.803	96	0.953	1.961	0.211	0.839	1.505	0.746
HU	80	0.650	0.991	0.106	0.477	0.656	0.790	96	0.610	2.172	0.015	0.745	1.827	0.677
PL	80	0.929	0.802	0.042	0.465	0.741	0.468	96	0.880	1.483	0.018	0.677	1.347	0.579
RO	80	1.039	2.397	0.111	1.447	1.923	0.572	96	0.945	2.587	0.047	1.571	2.272	0.525
SI	80	0.611	1.124	0.061	0.464	0.620	0.887	96	0.604	1.960	0.131	0.605	1.052	0.884
SK	80	0.968	1.575	-0.033	0.480	0.859	0.880	96	0.856	1.957	0.016	0.551	0.965	0.899

Notes: Descriptive statistics for quarterly national GDP growth rates (number of observations, mean, standard deviation) and corresponding revisions from the first data release to the actual values (March 2024), i.e. mean revision, mean absolute revision, revisions' standard deviation, and correlation between first releases and current values.

individual EU member states than for the EU in total or the euro area. During the pandemic, data revisions have occurred to a particular extent (Table 1, left panel). In particular, in the second quarter of 2020 data revisions appear extremely large, e.g. with almost 8 percentage points for the euro area. Therefore, the mean absolute revision is higher in the full sample. The standard deviation of the data revision varies between 0.2 in the EU to 1.3 in Luxembourg. The correlation between the first release and actual data is quite high, only for Luxembourg, the correlation is only about 0.5. Another interesting finding is that the correlation between the first release and actual data is also weaker in the Scandinavian countries. The mean data revision in the CEECs is also zero, except in Lithuania, Latvia and Romania (Table 1, lower panel). To take into account the effects of data revisions, I make use of a real-time dataset described above and, hence, use both first-release data on the one hand, and latest-release data on the other hand as a benchmark for forecast evaluation.

3 Forecast Evaluation Measures

3.1 Forecast Revision

Forecasts for a particular target quarter are often revised to take into account new information or changes in the environment. This could include e.g. changes in economic conditions, changes in consumer behaviour, or other factors that could have an impact on the accuracy of the forecast. Furthermore, data revisions for previous quarters, as described above, may lead to forecast revisions for subsequent quarters. Forecast revisions at a particular horizon h are defined as the difference between two successive forecasts for the same target quarter t,

$$\Delta\hat{y}_{t,t-h} = \hat{y}_{t,t-h} - \hat{y}_{t,t-h+1}. \tag{1}$$

For example, in the actual quarter t, the forecast revision for GDP growth is defined as the difference between the current GDP forecast for that quarter ($h = 0$) and the previous prediction for the same quarter ($h = 1$), where the latter has been submitted in the previous forecast wave. Using EC forecast data with a forecast horizon of up to eight quarters allows for the assessment of seven forecast revisions for a particular quarter. The key question is whether the EC updates its nowcasts or also its forecasts up to 8 quarters ahead in a similar manner and whether the updating behaviour varies across countries. It is expected that the revisions of forecasts increase as more information becomes available, i.e., as h declines. Longer-term forecasts are usually based on the unconditional mean (Andrade and Le Bihan, 2013).

3.2 Forecast Errors

The assessment of forecast errors for point forecasts covers different forecast horizons (West, 2006; Clark and McCracken, 2013). The h-step ahead forecast error is defined as the difference between the forecast at a particular horizon h for the period t and the outcome,

$$e_{t,t-h} = \hat{y}_{t,t-h} - y_t. \tag{2}$$

Positive errors indicate overprediction, while negative errors represent cases of underprediction. In all forecast error analyses, it is crucial to consider whether forecasters are targeting the first or "actual" value (Croushore, 2012).

To determine forecast performance, several forecast evaluation measures are considered: i) the mean error (ME), ii) the mean absolute error (MAE), and iii) the root mean squared error (RMSE). The ME provides an initial indication of a possible bias in a forecast, but given that positive and negative forecast errors offset each other, the magnitude of this error measure is rather limited. Therefore, MAE is also considered when errors can no longer cancel each other out. To account for the relative size of the errors, the RMSE measure weights large forecast errors more heavily than small errors. It is expected that forecast errors decrease closer to the target (h=0). Finally, for comparing RMSEs across countries, the ratios of RMSE to the average magnitude (MAR) and to the standard deviation are considered. While the former is particularly relevant for variables of different magnitudes, the latter is used to account for variability.

Rather than limiting the comparison of methods to a specific forecast horizon, such as a one-step-ahead or eight-quarters-ahead forecast, it may be more informative to aggregate performance across multiple forecast horizons (e.g., by summarizing the error over the next h quarterly forecasts). One advantage of the cumulative-horizon error is its simplicity, and it is less influenced by outliers (Armstrong and Collopy, 1992). While using this single measure for calibration is preferable to examining the error measure for each forecast horizon, individual forecast horizon analyses are important for improving forecast methods.

3.3 Forecast Optimality, Unbiasedness and Efficiency

The properties of an optimal forecast are established under the assumption that the objective function is of the mean squared error (MSE) type, meaning the forecasts minimize a symmetric, quadratic loss function. Hence, a forecast is optimal if there is an absence of serial correlation in the forecast errors. This implies that subsequent errors are unrelated to the previous errors, and the mistake made earlier disappears. If a systematic relation between errors exists and forecasters repeat the same mistakes (or compensate for past mistakes by subsequent errors of the opposite sign), it could be exploited to improve the forecast. The null hypothesis of no serial correlation among the forecast errors can be tested with the Ljung-Box Q-statistics and their corresponding p-values.

Forecasts from public national or international institutions are often suspected of being too optimistic. To determine whether the forecasts are unbiased, Mincer and Zarnowitz (1969) type regressions are estimated, where the realized value of the predicted variable is regressed on an intercept and the forecast for a particular horizon $y_t = \alpha_h + \beta_h \hat{y}_{t,t-h} + \epsilon_{t,t-h}$. The forecasts are both unbiased and efficient if β is equal to one and α equals zero. A rejection of the joint hypothesis suggests that the forecasts are optimal.

To evaluate the multi-horizon forecast efficiency (Patton and Timmermann, 2012; Arai, 2014), the intermediate forecast revisions $d_{t|h_1,h_2}...d_{t|h_{H-1},h_H}$ are taken into account in the regression tests. If a sequence of forecasts is optimal, then the forecast revisions should themselves be unpredictable. To derive this test, a short-horizon forecast can be represented as a function of a long-horizon forecast ($\hat{y}_{t,t-h_H}$) and the intermediate forecast revisions:

$$\hat{y}_{t,t-h_0} \equiv \hat{y}_{t,t-h_H} + \sum_{j=0}^{H-1} d_{t|h_j,h_{j+1}}. \tag{3}$$

The implication of forecast efficiency in this context is that forecasts are the conditional mean and the subsequent revisions are orthogonal to the past forecasts. Using a Univariate Optimal Revision Regression, the test is undertaken by regressing the realized value on an intercept, the long-horizon forecast and the sequence of intermediate forecast revisions. A set of zero–one equality restrictions on the intercept and slope coefficients are then tested to evaluate the joint consistency of all forecasts at different horizons. The F-test is used in this regression to jointly test the null hypothesis.

$$y_t = \alpha + \beta_H \hat{y}_{t,t-H} + \sum_{j=0}^{H-1} \beta_j d_{t|h_j,h_{j+1}} + u_t \tag{4}$$

These univariate regression tests provide evidence that agents optimally and consistently revise their forecasts at the interim points between the longest and shortest forecast horizons and in addition that the long-run forecast is unbiased. Patton and Timmermann (2012) show that this multi-horizon forecast efficiency evaluation has a greater power to detect forecast inefficiency in finite samples. Given that the European Commission published quarterly forecasts only twice a year in the earlier years, I only make use of revisions among multiple horizons.

Furthermore, a correlation of forecast errors with regard to the forecast error for EU growth is conducted. Positively correlated forecast errors across the EU are an indication that the unexpected event that caused the prediction mistake has similar effects in the respective country and the EU (Keereman, 1999). It should be stressed that correlations do not say anything about the direction of causation.

Finally, the historical mean forecasts and the quarterly forecasts provided by Consensus Economics are compared to the EC's forecasts. This comparison allows to assess the relative accuracy and reliability of the EC's forecasts against well-established benchmarks in the field of economic forecasting. The historical mean forecasts represent a baseline that is often considered a simple yet powerful predictor in economic models, while Consensus Economics offers a composite view of expert forecasts from various sources, providing a broad perspective on expected economic trends. A relative RMSE value below 1 indicates that the EC's forecast is more accurate than the historical mean and the Consensus Economics quarterly forecasts, suggesting that the EC's methodologies might incorporate more relevant information or more sophisticated analytical techniques.

4 Results on Forecast Revision and Forecast Errors

In the following, I evaluate the revision pattern of forecasts and the forecast performance of the quarterly EC forecasts for the period 2000Q1–2019Q4 for all forecast horizons, both in real-time and pseudo-real-time. As an illustration, Figure 1 shows the forecasts for the euro area published for h=0, h=4 and h=8 quarters along with the final outcome.The figures as well as the results in Table A2, show that forecasts are revised regularly; the nowcast (from h=1 to h=0) is revised in about 60% of all forecasts. In contrast, forecasts at longer horizons are updated only in 25% of the cases.[11] The table highlights a distinct pattern in the revisions, showing that all the core countries have undergone forecast adjustments at a specific horizon. The size of the forecast revisions from one forecast round to another is negligible (Table A3). However, from the first forecast conducted for a particular quarter until the last forecast, the mean revision is up to 0.2 percentage points. There is no clear pattern on whether forecasts are revised upwards or downwards at a particular forecasting round. There is evidence, that for Germany, Austria and the euro area, forecasts were revised downwards in the last three forecasting rounds (h=0,1,2). For the CEECs, the share and the size of forecast revisions tend to be slightly larger (Tables A10 and A11). In particular, for Croatia and Poland, forecasts are updated in more than 70% of the cases in the last forecast round (nowcast). Forecasts are revised slightly downwards on average in all countries in the last revision round, except for Poland and Hungary, where forecasts are revised upwards in the current year forecast rounds. A correlation analysis between EU or euro area forecast revisions and country forecast revisions indicates that in particular for the nowcast revision, correlation exceeds 0.75 for all major EU countries – Germany, France, and Italy. This is in line with the same share of forecast revision across countries investigated above. The forecast updating behaviour increases when the sample extends to 2023Q4. In the core euro area countries, the updating share for the nowcast increases to 67%. For the Scandinavian countries, it increases to 50%.

The forecast performance varies across countries and depends on whether the results are compared to real-time data or current-release data. Table 2 shows that for most countries, the EC overestimates GDP growth at almost all horizons in real time; however, the majority

[11] The number of forecast revisions is lower than the total number, as the EC published forecasts only every second quarter in the earlier years.

of forecast errors is not significantly biased. When compared to the current-release data, the nowcast is more often underestimated. The results for Italy are the only ones that show a significant upward bias across all horizons, a trend that had been previously observed by Chabin *et al.* (2020) for annual EC forecasts.

Table 2: EC's Mean Forecast Error (ME) and Bias

	EU	EA	AT	DE	FR	ES	IT	PT	BE	LU	NL	DK	FI	SE
real-time data														
0	0.060	0.040	0.012	0.015	0.021	−0.060**	0.161***	−0.044	0.033	−0.121	−0.023	0.198*	0.204*	0.004
1	0.102	0.127*	0.148*	0.110	0.109**	−0.034	0.252***	0.157	0.077	0.057	0.107	0.154*	0.062	0.040
2	0.173***	0.185***	0.153**	0.232***	0.131**	0.035	0.341***	0.083	0.200**	−0.286	0.065	0.251**	0.298**	0.072
3	0.200**	0.228***	0.178*	0.207*	0.172***	0.116*	0.344***	0.318***	0.126**	0.200	0.212**	0.102	0.178	0.013
4	0.237***	0.238***	0.131**	0.241**	0.198***	0.111	0.340***	0.150	0.231***	0.043	0.124	0.160	0.278*	0.082
5	0.193**	0.217**	0.223**	0.198	0.220***	0.114	0.291***	0.328***	0.161**	0.171	0.239**	0.195*	0.136	0.098
6	0.271***	0.256***	0.216***	0.288***	0.242***	0.137*	0.374***	0.169*	0.237***	−0.136	0.184*	0.268**	0.260*	0.098
7	0.139**	0.175***	0.158*	0.158**	0.225***	0.067	0.229***	0.446**	0.108***	−0.178	0.162	0.191	0.058	0.046
8	0.371**	0.367***	0.272**	0.460***	0.256**	0.178	0.506***	0.544**	0.289*	−0.100	0.222	0.447**	0.278	0.200
current value data														
0	−0.005	−0.035	0.039	−0.086	0.012	−0.056	0.104	−0.083	−0.114	−0.350*	−0.067	0.060	0.153	0.043
1	0.089	0.098	0.071	0.075	0.041	0.014	0.282***	−0.002	−0.031	−0.300	0.057	0.126	0.102	−0.069
2	0.108	0.100	0.121	0.061	0.116	0.045	0.271***	0.078	0.055	−0.086	−0.006	0.091	0.168	0.153
3	0.196**	0.202**	0.098	0.216	0.116*	0.170*	0.398***	0.102	0.022	−0.021	0.146	0.040	0.243	−0.035
4	0.150*	0.127	0.173*	0.053	0.198**	0.127	0.289***	0.150	0.084	−0.129	0.071	0.048	0.216	0.149
5	0.209**	0.220**	0.159*	0.273*	0.145**	0.161	0.334***	0.144	0.061	0.014	0.207*	0.109	0.170	−0.018
6	0.174**	0.135	0.202*	0.091	0.219**	0.142	0.288***	0.169	0.098	0.036	0.119	0.105	0.181	0.163
7	0.112*	0.096	0.075	0.058	0.158**	0.062	0.229**	0.158	−0.025	−0.078	0.079	0.230	0.104	0.012
8	0.239	0.194	0.289	0.194	0.261*	0.228	0.394*	0.333	0.139	0.080	0.144	0.318	0.256	0.378

Notes: Mean forecast errors are provided for each country and each horizon (h=0 to 8) in the period 2000Q1–2019Q4 with regard to real-time first release data and current value data. A negative figure indicates an underestimation of GDP growth. ***, **, * indicates whether the null of unbiasedness is rejected at the 10%, 5% or 1% level.

Table A5 reveals that the absolute errors range from 0.2 to 0.7 percentage points, and with even higher errors for Luxembourg. In comparison to the current value data, the forecast errors also increase slightly. Furthermore, the forecast performance does not continuously increase from a forecast that had been published 8 quarters ago (h=8) to the nowcast (h=0). While the worst forecast performance is found for h=8, the best performance is not clearly identified for the current quarter (h=0) for all countries.

Implementing sectoral analysis (Figure 2) enhances the ability of forecasters to identify the specific sources of forecast errors and biases at both the national and EU-wide levels. By analyzing the contribution and variability of different sectors, forecasters can better understand the observed patterns of smaller GDP forecast misses at the EU level and the tendency to over-predict real GDP in real time. Sectoral analysis also provides a framework for addressing the challenges of parsing out value-added among member countries, tracking sectoral net exports and inventories, and dealing with the economic impact of non-reporting firms. Ultimately, this approach can lead to more accurate and reliable GDP forecasts by accounting for the unique characteristics of each sector within the broader economy. For example, the agriculture, forestry, and fishing sector contributes the least to the total GVA, while the service sector dominates across all countries, especially in Luxembourg (88.2%) and France (77.6%). On the other hand, the industrial sector has a higher share in Germany (30.6%) and Austria (29.3%) compared to countries like Luxembourg (11.4%) and Portugal (20.7%). For the Scandinavian

Table 3: EC's Mean Forecast Error (CEECs) and Bias

	EA	BG	CZ	EE	HR	LV	LT	HU	PL	RO	SI	SK
real-time data												
0	0.042	−0.317	−0.250	−0.207	−0.135	−0.084	0.069	−0.043	−0.249	0.077	−0.210	−0.292
1	0.116*	0.166	0.037	0.178	−0.244	0.253	0.065	0.071	−0.139	0.040	−0.113	0.270
2	0.204***	−0.180*	0.065	0.369	−0.173	0.195	0.355	0.057	−0.104	0.052	0.088	−0.134
3	0.210***	0.365	0.288	0.241	−0.067	0.522	0.521	0.122	−0.039	0.046	0.264	0.420
4	0.249***	0.009	−0.038	0.185	0.021	0.384	0.482	0.224	−0.027	0.157	0.094	0.109
5	0.195**	0.491	0.016	0.316	−0.262*	0.482	0.437	0.289*	−0.005	0.004	0.114	0.186***
6	0.265***	0.204	0.163	0.558	−0.038	0.345	0.615	0.410**	0.029	0.042	0.108	0.144*
7	0.150***	0.166	0.185	0.142	−0.112	0.319	0.072	0.253	−0.041	−0.133	−0.121	0.103
8	0.372**	0.290	0.317	0.000	0.175	0.564	−0.136	0.467	0.133	0.150	0.422	0.091
current value data												
0	−0.049	−0.099	−0.219*	−0.027	−0.194	0.243	−0.288*	−0.197	−0.229*	−0.128	−0.292*	−0.200
1	0.027	0.012	−0.097	−0.061	−0.365	0.206	0.006	−0.014	−0.238*	0.206	−0.203*	0.214
2	0.114	0.008	−0.032	0.624	−0.219	0.585**	−0.022	−0.051	−0.116	−0.062	0.050	−0.087
3	0.120	0.198	0.121	0.106	−0.119	0.594*	0.466	0.024	−0.053	0.166	0.046	0.397
4	0.131	0.242	−0.066	0.480	−0.047	0.524	0.171	0.082	−0.071	−0.138	0.063	0.124
5	0.175*	0.309	−0.032	0.019	−0.179	0.552	0.315	0.153	−0.025	0.148	−0.056	0.189*
6	0.151*	0.396*	0.077	0.856	0.031	0.584*	0.185	0.313	−0.038	−0.085	0.008	0.176*
7	0.108	0.128	0.009	0.197	−0.012	0.700	−0.059	−0.006	−0.065	−0.027	−0.179	0.166
8	0.217	0.770	0.283	0.946	−0.075	0.445***	−0.009	0.517	0.017	0.061	0.411	−0.155

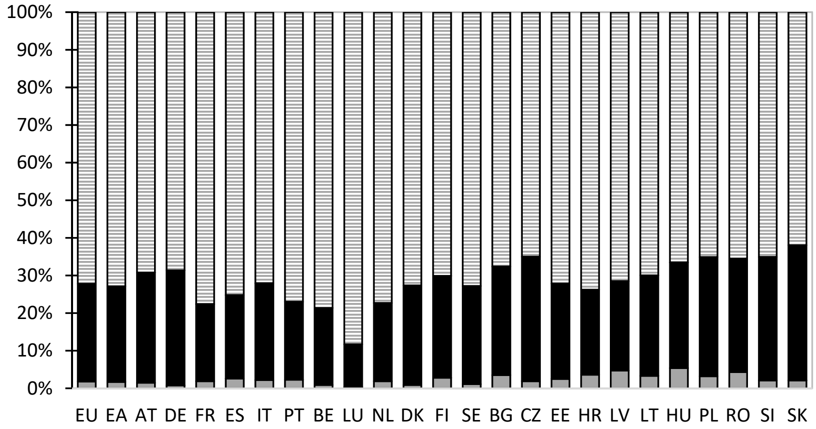
Notes: Mean Forecast Errors are provided for each country and each horizon (h=0 to 8) in the period 2000Q1–2019Q4. A negative figure indicates an underestimation of GDP growth. ***, **, * indicates whether the null of unbiasedness is rejected at the 10%, 5% or 1% level.

countries, the shares are comparable to the EU average. Countries with a higher manufacturing share show evidence of smaller forecast errors, while countries with a higher share of services tend to experience greater forecast errors due to the wide range of service activities, which are more sensitive to changes in consumer behaviour, external influences, and have more complex dynamics compared to manufacturing.

Overall, forecast errors for aggregates such as the EU or euro area tend to be smaller compared to errors of major countries, such as Germany, France, and Italy (Table A6); a result that also applies to the performance of annual forecasts, e.g. by Melander *et al.* (2007). Forecast errors for Spain are smaller than those for the euro area at almost all horizons. Luxembourg again stands out with the highest root mean squared forecast error for the nowcast (h=0). Furthermore, forecast errors for the Scandinavian countries are much larger at all forecast horizons, which might result from the high standard deviation of GDP compared to other EU countries, indicating that Swedish GDP growth is more difficult to forecast. When adjusting for the varying volatility of national GDP growth, the results based on the RMSE ratio using real-time data show smaller errors for the Scandinavian countries (Table A7). Although the differences between using real-time data and current vintage data are not substantial, the deterioration in error statistics at the country level is somewhat more pronounced than for the EU average, a fact also confirmed by Keereman (1999) in the context of annual EC forecasts.

The majority of individual country forecasts are not significantly biased at very short forecast horizons (Table 2), however, there is evidence of significant bias for the EU, the euro area, and many member states at medium horizons; for Italy, significant bias is present at all forecast horizons. A test for the persistence of forecast errors was carried out to investigate systematic

Figure 2: Sectoral shares



Note: The sectoral gross value added shares to the total national gross value added for the year 2023 are depicted as follows: for agriculture, forestry and fishing (A) in grey, industry (B-F) in black and services (G-U) with line pattern.
Source: Eurostat (nama10 a10).

correlation, and the Ljung-Box test statistic reveals that serial correlation among nowcast errors is absent or limited. Forecast errors are slightly persistent for horizon h=1 and h=3 for the EU, euro area and some countries (e.g. Germany, France, and Italy).[12] The test on multi-horizon forecast efficiency using univariate optimal revision regressions signals that the EC does not optimally and consistently revise their forecasts in addition to biased long-horizon forecasts. Analysing the errors over time indicates that the longer horizon errors have been reduced since 2000, however, the size of the nowcast error remained relatively stable.

A correlation analysis of forecast errors with regard to the forecast error for the EU reveals a significant disparity in the strength of these correlations across different member states (Table A8). The analysis demonstrates that the correlation of forecast errors is notably high for the core EU countries, indicating a strong alignment in the predictive inaccuracies among these nations. This suggests that the synchronization of economic activity across these regions plays a significant role in shaping forecast outcomes, particularly in response to common economic shocks. In contrast, the correlation is weaker for countries such as Austria – with strong ties to EU countries in Eastern Europe – and Spain. Furthermore, the Scandinavian countries exhibit even lower correlations with the EU forecast errors. In the latter case, economic connections with non-EU countries, particularly through trade and financial markets, might lead to forecast errors that reflect external economic influences not as strongly felt by core EU countries.

The analysis for the CEECs reveals that there is no systematic over- or underestimation in forecasts, with the exception of the EC's nowcasts, which tend to be overly pessimistic for the majority of these countries (Table 3). There is no significant bias in the forecasts for this group; however, forecast errors tend to be more persistent in Hungary. While forecast errors measured

[12] Melander *et al.* (2007) find also persistence in GDP forecast errors for annual GDP growth forecasts.

with real-time data are smaller, they are significantly higher for the CEECs compared to other EU countries. This disparity should be interpreted with caution, considering the differences in data availability and stability across countries. Controlling for the volatility of national GDP growth, the errors based on the RMSE ratio are lower. The forecast challenges may be also attributed to the distinct sectoral composition of CEEC economies (Figure 2). The share of agriculture in gross value added is significantly higher in several CEECs compared to the EU average of 1.8%. For instance, in Lithuania (4.7%), Poland (5.4%), Romania (3.2%), and Latvia (3.6%), agriculture plays a more prominent role in influencing GDP forecasts. Additionally, the industrial sector's share is generally higher in CEECs than the EU average of 26%, with up to 35.9% in Slovakia.

Finally, the performance of the European Commission's forecasts is evaluated in comparison to mean forecasts and Consensus Economics forecasts. The EC's forecasts generally outperform the mean forecasts for the nowcast (h=0), with Luxembourg being the only exception (Table 4). This trend extends to the next three quarters (up to h=3) for the EU, the euro area, and the Netherlands. For Spain, Portugal, and the Scandinavian countries, EC's forecasts consistently demonstrate superior accuracy across all forecast horizons. For other countries, the mean seems a suitable candidate to improve longer-horizon forecasts. However, the performance is almost the same, and only in a few cases is the mean forecast significantly better than the EC's forecast. The results show that the EC's forecasting performance for the CEEC is superior to the mean for the majority of countries and horizons (Table 5). However, for the Polish economy, the mean forecast consistently outperforms the EC's forecasts across all horizons, and similar superiority is observed for longer-term forecasts in the Czech Republic and Hungary.

Table 4: Relative Root Mean Squared Forecast Error (Mean)

	EU	EA	AT	DE	FR	ES	IT	PT	BE	LU	NL	DK	FI	SE
							real-time data							
0	0.620**	0.682**	0.890	0.857**	0.769**	0.341***	0.745**	0.813*	0.772*	1.140	0.746**	1.003	0.827**	0.736*
1	0.869	0.873	0.980	1.008	0.843	0.566**	0.945	0.873**	0.899	0.932	0.796***	0.932	0.855*	0.920
2	0.935	0.941	1.041	1.074	0.953	0.634***	1.032	0.820*	1.065	1.200	0.802**	0.941	0.923	0.910*
3	0.978	0.966	0.974	1.057	0.900	0.744**	1.023	0.907*	0.974	0.987	0.864***	1.001	0.892***	0.924
4	1.037	1.049	1.016	1.118	1.012	0.799**	1.073	0.870**	1.090	1.042	0.896**	0.930	0.940	0.964
5	1.023	1.011	1.032	1.070	0.965	0.783**	1.060	0.867***	0.971*	0.945	0.917***	1.096	0.893**	1.046
6	1.045	1.052	1.077	1.128	1.052	0.828**	1.086	0.827**	1.055	1.054	0.933**	0.963	0.937*	0.959
7	1.057	1.052	1.156	1.066	1.027	0.747***	1.216	0.901***	0.857***	1.097	0.943*	1.025	0.956	0.953
8	1.053	1.019	1.108	1.117	1.038	0.856***	1.059	0.883***	1.108	1.107	0.975	1.171	0.923*	0.957
							current value data							
0	0.727**	0.762**	0.853**	0.874**	0.761**	0.543***	0.802**	0.761**	0.903	0.935	0.888	0.913*	0.915	0.808*
1	0.883	0.898	1.029	0.974	0.882	0.664**	0.949	0.817**	1.039	1.022	0.824**	0.975	0.917	0.971
2	0.953	0.970	0.959	1.030	0.986	0.702***	0.993	0.842*	1.087	0.970	0.820**	0.907*	0.928	0.948
3	0.984	0.968	0.977	1.040	0.922	0.811**	1.016	0.896**	1.090	1.019	0.890**	0.980	0.970	0.976
4	1.029	1.040	0.999	1.043	1.013	0.835***	1.048	0.855**	1.098	1.027	0.892**	0.958	0.971	0.981
5	1.018	1.008	1.033	1.053	0.984	0.833**	1.043	0.880***	1.013	0.909*	0.935**	0.988	0.957*	1.034
6	1.025	1.041	1.025	1.050	1.020	0.845***	1.044	0.805***	1.087	0.985	0.930*	0.968	0.959	0.956*
7	1.027	0.983	1.082	0.963	1.064	0.829**	1.090	0.851***	1.013	1.075	0.993	1.023	0.984	0.962
8	1.047	1.016	0.988	1.059	1.037	0.881***	1.028	0.891**	1.084	0.984	1.004	1.103	0.966	0.991

Notes: Relative root mean squared forecast errors for the EC forecasts (quarter on quarter) in relation to mean forecasts are provided for each country and each horizon (h=0 to 8) for the period 2000Q1–2019Q4. For values lower than one, the EC forecast is superior to the mean forecast. ***, **, * indicates significance at the 1, 5 and 10% level.

For a subset of countries, the quarterly Consensus forecasts are evaluated against the EC's forecast (Table 6). The analysis reveals that the EC's nowcasts underperform relative to the

Table 5: EC's Relative Root Mean Squared Forecast Error CEECs (Mean)

	EA	BG	CZ	EE	HR	LV	LT	HU	PL	RO	SI	SK
						real-time data						
0	0.726**	1.326	0.965	0.832	1.271	0.610*	0.438	0.761	1.128	0.774*	0.964	1.908
1	0.902	0.939**	0.985	0.896	1.101	0.801	1.048	0.875	1.203	0.910	0.855	0.969
2	0.995	0.856	1.021	1.411	1.004	0.705*	0.873	0.882*	1.135	0.841	0.974	1.590
3	0.994	0.941**	1.011	0.946	0.974	0.875	0.915	0.943*	1.052	0.863*	0.949***	0.940
4	1.096	0.957	0.991	0.964	1.078	0.861	0.966	0.949	1.032	0.821*	0.961	0.815**
5	1.027	0.955**	1.053	0.949	0.934	0.957	0.958	1.036	1.019	0.902	0.900*	0.770**
6	1.091	0.972	1.048	1.566	1.007	0.961	1.029	1.169	1.145	0.929	0.970	0.817**
7	1.071	0.808	1.067	0.936	1.146	0.963	0.762	1.172	1.029	0.933	0.900**	0.952
8	1.050	1.106	1.039	0.968	1.148	0.889*	1.274	1.177	1.070	0.832	0.986	0.950
						current value data						
0	0.760**	0.762*	0.905*	0.876	1.045	0.752*	1.023	0.850	0.989	0.919	0.942	1.535
1	0.901	0.954	0.987	0.982	1.040	0.763*	1.000	0.988	1.070	0.999	0.937	0.978
2	0.941*	0.746*	0.987	1.181	1.047	0.798*	0.710	0.967	1.175	0.968	0.958	1.405
3	0.972	0.904**	1.009	0.887	0.986	0.820*	0.931	0.983	1.012	0.931*	0.949***	0.940
4	0.984	0.929*	1.040	0.965	1.025	0.914	1.002	0.960	1.110	1.022	0.980	0.956
5	0.987	0.924**	1.027	0.865	0.980	0.924*	0.997	1.006	1.037	0.945	0.904***	0.888
6	0.988	0.986	1.059	1.174	0.999	0.967	1.182	1.139	1.194	1.001	0.990	0.861**
7	0.950	0.885	1.091	0.884*	1.035	0.936	0.826	1.145	1.088	0.951	0.948*	0.981
8	1.014	0.962	1.029	0.953	1.132	0.786**	1.087	1.077	1.057	0.935	0.993	0.997

Notes: Relative root mean squared forecast errors for the EC forecasts (quarter on quarter) in relation to mean forecasts are provided for each country and each horizon (h=0 to 8) for the period 2000Q1–2019Q4. For values lower than one, the EC forecast is superior to the mean forecast. ***, **, * indicates significance at the 1, 5 and 10% level.

Consensus forecasts across all countries. This is not surprising, given that Consensus polls are conducted in the last month of the quarter, allowing professional forecasters to promptly incorporate the latest economic data into their predictions. My results confirm previous findings by Jansen *et al.* (2016) who demonstrated that Consensus forecasts, particularly for Germany, exhibit strong performance in stable periods. By contrast, during the COVID-19 pandemic, when GDP displayed extreme fluctuations, the results were less significant.

Table 6: Relative Root Mean Squared Forecast Error (Consensus Forecasts)

	DE	FR	IT	EA	NL	ES	SE
				real-time data			
0	1.217	1.126	1.094	1.234	1.003	0.950	0.600*
1	1.115	1.027	1.133	1.122	1.032	1.089	0.726
2	1.053	1.049	1.113	1.138	0.960	0.527	0.892*
3	1.016	0.968	0.974	1.020	0.976	1.011	0.786*
4	1.064	1.004	0.983	1.135	1.055	1.038	0.867**
5	1.027	1.006	1.045	1.019	1.016	0.994	0.981
6	1.047	1.039	0.997	1.07	1.018	0.979	0.843***
7	1.329	1.095	1.215	1.616	1.804	1.378	1.727
				current value data			
0	1.138	1.086	1.052	1.102	1.094	1.041	0.796*
1	1.087	1.013	1.090	1.064	1.040	1.054	0.852
2	1.047	1.041	1.068	1.072	0.929*	0.578	0.900***
3	1.014	0.976	1.006	0.985	1.013	0.993	0.883
4	1.027	0.977	1.002	1.053	1.013	1.021	0.910
5	1.016	1.024	1.017	0.994	1.007	1.006	0.939
6	1.007	1.007	1.005	1.023	0.990	0.958*	0.837***
7	1.337	0.852	1.233	1.133	1.266	1.108	1.156

Notes: Relative root mean squared forecast errors for the EC forecasts (quarter on quarter) in relation to Consensus forecasts are provided for each country and each horizon (h=0 to 7) for the period 2000Q1–2019Q4. For values lower than one, the EC forecast is superior to the Consensus forecast. ***, **, * indicates significance at the 1, 5 and 10% level.

5 Conclusion

This paper provides a comprehensive assessment of the quarterly GDP growth forecasts published by the European Commission, offering insights into the patterns of data revisions, forecast revisions, forecast errors, and varying performance across countries. The findings contribute to a better understanding of the complexities of quarterly macroeconomic forecasting and have implications for policy makers, businesses, and researchers who rely on such forecasts. Considering that European economic forecasts are a fundamental component of the economic surveillance framework established by the EC Treaty, they play a critical role in monitoring and guiding economic policies across member states. These forecasts are essential for assessing economic performance, ensuring compliance with fiscal regulations, and informing policy decisions aimed at promoting stability and growth within the European Union. My results show that the impact of data revisions on forecasting performance is notable, revealing higher data revision rates for member states compared to the EU as a whole. Scandinavian countries and Central and Eastern European countries are also affected by higher data revision by Eurostat from the initial release to the current release, including shifts from growth to slow down in particular quarters. Therefore, the results depend on the data target used.

Forecast revisions have a greater impact on short-term forecasts than on longer-term forecasts, which tend to remain more stable. Furthermore, the final forecast revision in the current quarter is generally downward biased for almost all countries. Surprisingly, forecast performance is not the best for the current quarter (h=0) in all countries. Forecast performance does not consistently improve from a forecast published eight quarters ago (h=8) to the nowcast (h=0). The analysis reveals varying levels of forecast accuracy across countries, with smaller and more volatile economies exhibiting greater uncertainty in their forecasts. This uncertainty is exacerbated by economic shocks and structural breaks, which can dramatically affect forecast performance, particularly for longer horizons. The analysis reveals two notable patterns: first, GDP forecast errors for the EU as a whole tend to be smaller compared to those for individual member states; second, there is a tendency to over-predict real GDP in real time across various countries. The findings underscore the importance of considering business cycle correlation when evaluating and improving GDP forecasts. Countries with synchronized business cycles tend to exhibit similar forecast performance, highlighting the need for forecasting models that can account for these correlations. Overall, there is no clear persistence in the Commission's forecast errors, and these findings are consistent with previous analyses of annual EC GDP forecasts (Fioramanti *et al.*, 2016; Keereman, 1999). The findings highlight the significant impact of both accurate (earlier) quarterly data by Eurostat and national statistics, as well as accurate forecasts, as crucial inputs for subsequent quarterly sequences and, hence, the annual forecasting performance derived from this. The results show that historical mean forecasts provide more accurate forecasts for longer-term horizons for many countries, and – given their later publication – Consensus forecasts are superior for many countries for the nowcast.

A drawback of the analysis is the small sample size. In particular, the shorter samples for some member states raise concerns about cross-country comparisons. Future research on forecast evaluation and improvement should not only extend the sample size for GDP forecasts by panel estimation (Chabin *et al.*, 2020) but also shed light on the underlying forecasts of GDP components. Additionally, investigating the potential impact of unforeseen events, such as the COVID-19 pandemic, on forecast accuracy could offer a forward-looking perspective on the resilience and adaptability of forecasting models in the face of unprecedented challenges. Further work might address multiple-horizon uncertainty as well, building upon earlier approaches by Knüppel (2014) or Clark *et al.* (2020), who suggest models that account for possible overlapping information in the multi-step forecasts (errors) and related forecast uncertainty.

Overall, the paper provides a comprehensive analysis of quarterly GDP growth forecasts for the EU and member states and, hence, provides insights how to improve annual forecasts. It reveals a tendency to overestimate short-term growth for some countries and highlights areas for potential improvement in forecasting accuracy by refining forecasting techniques, especially for short-term projections. The findings of this paper also highlight the need for further research into the cyclical behavior of different groups of economies within the European Union. Specifically, the potential emergence of a divergence between the more advanced EU economies and those of Central and Eastern European Countries (CEEC) warrants closer examination. Future research, with careful consideration, could explore whether differing economic structures, growth trajectories, and external vulnerabilities between these groups of countries are leading to distinct cyclical patterns. Such an inquiry could significantly contribute to the understanding of economic convergence and divergence within the EU, offering a nuanced perspective that could help policymakers tailor their strategies more effectively. Moreover, this research could enrich the broader discussion on EU economic convergence, supporting the long-term stability and integration goals of the European Union by ensuring that policy responses and economic coordination are appropriately adapted to these emerging dynamics.

References

Andrade, P. and Le Bihan, H. (2013). Inattentive professional forecasters. *Journal of Monetary Economics*, **60** (8), 967–982. DOI: https://doi.org/10.1016/j.jmoneco.2013.08.005.

Angelini, E., Camba-Mendez, G., Giannone, D., Reichlin, L. and Rünstler, G. (2011). Short-term forecasts of euro area GDP growth. *Econometrics Journal*, **14**, C25–C44. DOI: https://doi.org/10.1111/j.1368-423X.2010.00328.x.

Arai, N. (2014). Using forecast evaluation to improve the accuracy of the Greenbook forecast. *International Journal of Forecasting*, **30** (1), 12–19. DOI: https://doi.org/10.1016/j.ijforecast.2013.02.002.

Armstrong, J. S. and Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, **8** (1), 69–80. DOI: https://doi.org/10.1016/0169-2070(92)90008-W.

Aruoba, S. B. (2008). Data revisions are not well behaved. *Journal of Money, Credit and Banking*, **40** (2-3), 319–340. DOI: https://doi.org/10.1111/j.1538-4616.2008.00115.x.

Asimakopoulos, S., Lalik, M., Paredes, J. and García, J. S. (2023). *GDP revisions are not cool: the impact of statistical agencies' trade-off*. ECB Working Paper 2857, European Central Bank.

Bańbura, M., Giannone, D., Modugno, M. and Reichlin, L. (2013). Chapter 4 - nowcasting and the real-time data flow. In G. Elliott and A. Timmermann (eds.), *Handbook of Economic Forecasting*, vol. 2, Elsevier, pp. 195–237.

Baffigi, A., Golinelli, R. and Parigi, G. (2002). *Real-time GDP forecasting in the euro area*. 456.

Barhoumi, K., Benk, S., Cristadoro, R., den Reijer, A., Jakaitiene, A., Jelonek, P., Rua, A., Rünstler, G., Ruth, K. and Nieuwenhuyze, C. V. (2008). *Short-term forecasting of GDP using large monthly datasets-A pseudo real-time forecast evaluation exercise*. ECB Occasional Paper 84, European Central Bank.

Beck, K. (2020). Decoupling after the crisis: Western and eastern business cycles in the European Union. *Eastern European Economics*, **58** (1), 68–82. DOI: doi.org/10.1080/00128775.2019.1656086.

Belke, A., Domnick, C. and Gros, D. (2017). Business cycle synchronization in the EMU: Core vs. periphery. *Open Economies Review*, **28**, 863–892. DOI: https://doi.org/10.1007/s11079-017-9465-9.

Blunier, J. and Hepenstrick, C. (2022). *What were they thinking? Estimating the quarterly forecasts underlying annual growth projections*. SNB Working Paper 5, Swiss National Bank.

Cabanillas, L. G., Terzi, A. *et al.* (2012). *The accuracy of the European Commission's forecasts re-examined*. European Economy 476, Directorate General Economic and Financial Affairs (DG ECFIN).

Cascaldi-Garcia, D., Ferreira, T. R., Giannone, D. and Modugno, M. (2023). Back to the present: Learning about the euro area through a now-casting model. *International Journal of Forecasting*. DOI: https://doi.org/10.1016/j.ijforecast.2023.04.005.

Chabin, A., Lamproye, S. and Výškrabka, M. (2020). *Are we more accurate? Revisiting the European Commission's macroeconomic forecasts*. European Economy Discussion Paper 128, European Commission.

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

Clark, T. E., McCracken, M. W. and Mertens, E. (2020). Modeling time-varying uncertainty of multiple-horizon forecast errors. *Review of Economics and Statistics*, **102** (1), 17–33.

Croushore, D. (2011). Frontiers of real-time data analysis. *Journal of Economic Literature*, **49** (1), 72–100.

— (2012). *Forecast bias in two dimensions*. Working Papers 12-9, Federal Reserve Bank of Philadelphia.

D'Agostino, D. S. P., Antonello; Giannone (2006). *(Un)Predictability and macroeconomic stability*. ECB Working Paper 605, European Central Bank.

Darvas, Z. and Szapáry, G. (2008). Business cycle synchronization in the enlarged eu. *Open Economies Review*, **19**, 1–19. DOI: https://doi.org/10.1007/s11079-007-9027-7.

De Winter, J. (2011). *Forecasting GDP growth in times of crisis: private sector forecasts versus statistical models*. DNB Working Paper 320, DNB.

Diron, M. (2008). Short-term forecasts of euro area real GDP growth: an assessment of real-time performance based on vintage data. *Journal of Forecasting*, **27** (5), 371–390. DOI: https://doi.org/10.1002/for.1067.

Drechsel, K. and Maurin, L. (2011). Flow of conjunctural information and forecast of euro area economic activity. *Journal of Forecasting*, **30** (3), 336–354. DOI: https://doi.org/10.1002/for.1177.

Eurostat (2013). *European system of accounts - ESA 2010*. Tech. rep.

Fildes, R. and Stekler, H. (2002). The state of macroeconomic forecasting. *Journal of Macroeconomics*, **24** (4), 435–468. DOI: https://doi.org/10.1016/S0164-0704(02)00055-1.

Fioramanti, M., González Cabanillas, L., Roelstraete, B. and Ferrandis Valterra, S. (2016). *European Commission's forecasts accuracy revisited: Statistical properties and possible causes of forecast errors*. Tech. rep., European Commission.

ICF, DIW, NIESR and OFCE (2017). *Evaluation of DG ECFIN Forecasting Services ECFIN-108-2016/S12.738721*. Tech. rep.

Jansen, W. J., Jin, X. and de Winter, J. M. (2016). Forecasting and nowcasting real GDP: Comparing statistical models and subjective forecasts. *International Journal of Forecasting*, **32** (2), 411–436. DOI: https://doi.org/10.1016/j.ijforecast.2015.05.008.

Keereman, F. (1999). *The track record of the Commission forecasts*. European Economy - Economic Papers 2008 - 2015 137, Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Knüppel, M. (2014). Efficient estimation of forecast uncertainty based on recent forecast errors. *International Journal of Forecasting*, **30** (2), 257–267. DOI: https://doi.org/10.1016/j.ijforecast.2013.08.004.

Kuzin, V., Marcellino, M. and Schumacher, C. (2011). MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area. *International Journal of Forecasting*, **27** (2), 529–542. DOI: https://doi.org/10.1016/j.ijforecast.2010.02.006.

Marcellino, M. and Schumacher, C. (2010). Factor MIDAS for nowcasting and forecasting with ragged-edge data: A model comparison for German GDP. *Oxford Bulletin of Economics and Statistics*, **72** (4), 518–550. DOI: https://doi.org/10.1111/j.1468-0084.2010.00591.x.

Melander, A., Sismanidis, G., Grenouilleau, D. *et al.* (2007). *The track record of the Commission's forecasts – an update*. European Economy 291, Directorate General Economic and Financial Affairs (DG ECFIN).

Öller, L.-E. and Barot, B. (2000). The accuracy of European growth and inflation forecasts. *International Journal of Forecasting*, **16** (3), 293–315. DOI: https://doi.org/10.1016/S0169-2070(00)00044-3.

Patton, A. J. and Timmermann, A. (2012). Forecast rationality tests based on multi-horizon bounds. *Journal of Business & Economic Statistics*, **30** (1), 1–17. DOI: 10.1080/07350015.2012.634337.

Rünstler, G., Barhoumi, K., Benk, S., Cristadoro, R., Den Reijer, A., Jakaitiene, A., Jelonek, P., Rua, A., Ruth, K. and Van Nieuwenhuyze, C. (2009). Short-term forecasting of GDP using large datasets: a pseudo real-time forecast evaluation exercise. *Journal of Forecasting*, **28** (7), 595–611. DOI: https://doi.org/10.1002/for.1105.

— and Sédillot, F. (2003). *Short-term estimates of euro area real GDP by means of monthly data*. Ecb working paper, European Central Bank.

Schumacher, C. and Breitung, J. (2008). Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data. *International Journal of Forecasting*, **24** (3), 386–398. DOI: https://doi.org/10.1016/j.ijforecast.2008.03.008.

Stundziene, A., Pilinkiene, V., Bruneckiene, J., Grybauskas, A., Lukauskas, M. and Pekarskiene, I. (2023). Future directions in nowcasting economic activity: A systematic literature review. *Journal of Economic Surveys*, **n/a** (n/a). DOI: https://doi.org/10.1111/joes.12579.

West, K. D. (2006). Chapter 3 - Forecast Evaluation. In G. Elliott, C. Granger and A. Timmermann (eds.), *Handbook of Economic Forecasting*, vol. 1, Elsevier, pp. 99–134.

Zarnowitz, V. (1979). An analysis of annual and multiperiod quarterly forecasts of aggregate income, output, and the price level. *The Journal of Business*, **52** (1), 1–33. DOI: http://www.jstor.org/stable/2352661.

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

The IWH is funded by the federal government and the German federal states.