Ex-post evaluation of the official short-term fiscal forecasts: Evidence from Finland

WORKING PAPER

Working paper

Ex-post evaluation of the official short-term fiscal forecasts: Evidence from Finland

Suvi Kangasrääsiö

June 2024

National Audit Office of Finland
Fiscal Policy Monitoring Function
University of Oulu, Finland

Abstract

D/97/04.04.01/2022

This study presents an ex-post evaluation of the short-term fiscal forecasts prepared for the central government budgeting in Finland, a Member State of the EU and the Economic and Monetary Union (EMU). The analysis relies on a unique dataset including forecasts of the general government balance-to-GDP ratio, debt-to-GDP ratio, and expenditure-to-GDP ratio. The official forecasts by the Ministry of Finance are compared with those of other key forecasters of Finnish public finances. The dataset covers the period from 1997 to 2021, which is the longest period analyzed in the Nordic fiscal forecast literature so far. Also, the set of variables analyzed is expanded compared to the previous literature. I find that, on average, there is bias towards pessimism in the general government balance-to-GDP ratio forecasts by the Ministry of Finance and other key forecasters if the year-ahead forecasts for non-crisis years are considered. In addition, similar bias is detected in the current-year forecasts for fiscal balance-to-GDP ratio by the Ministry of Finance and the OECD. No statistically significant bias is detected in debt-to-GDP or expenditure-to-GDP forecasts. The mean absolute errors of the Ministry of Finance and the other forecasters were large on average and statistically significant in most cases. The Ministry of Finance's forecasts do not contain more bias than those of the other forecasters. In addition, the Ministry of Finance's forecasts are as accurate or inaccurate as those of the other forecasters, except for the fact that the naïve forecasts and the OECD debt forecast are more inaccurate in some cases.

Helsinki, 18 June 2024

JEL codes: H68, E62

Keywords: Forecasting, forecast errors, fiscal forecasts, fiscal policy

Acknowledgements:

I would like to thank Juha-Pekka Junttila, Pentti Pikkarainen, Arto Kokkinen, Matthias Strifler, Mika Sainio, Matti Okko, Antti Ripatti, and participants in several different seminars at the National Audit Office, Ministry of Finance, Bank of Finland, Etla economic research, and University of Oulu, and the participants in a conference of the Baltic–Nordic cooperation network of the Independent Fiscal Institutions (IFI) in Tallinn on 9 June 2023 for their valuable comments.

The National Audit Office of Finland audits the management of central government finances and assets, monitors fiscal policy, and oversees political party and election campaign funding. The fiscal policy monitoring function is responsible for monitoring compliance with fiscal rules as well as the setting and implementation of fiscal objectives. The aim of the monitoring is to promote the transparency and intelligibility of fiscal rules as well as stable and sustainable general government finances. The fiscal policy monitoring function reports its assessments twice a year. All publications are available online. The opinions expressed in this paper do not necessarily reflect the official views of the National Audit Office of Finland.

Contents

1 Introduction

Fiscal forecasts play a significant role in policymaking. In Finland, the short-term fiscal forecasts produced by the Ministry of Finance (MoF) are used by the government in state budget planning. The official forecasts also play a key role when fiscal policy reactions to various macroeconomic shocks are planned. Systematic biases and large inaccuracies in the official forecasts make it more difficult to exercise appropriate fiscal policy. This weakens the policy credibility. Therefore, it is essential to evaluate the government's fiscal forecast errors. In recent years, the global economy has encountered a large shock caused by Covid-19. The fiscal policy response to the crisis has been substantial worldwide as well as in Finland. A better understanding of the quality of the official fiscal forecasts would help the Independent Fiscal Institutions (IFI) to give sharper advice to policymakers in crisis situations. Also, the results of ex-post analysis may help the MoF and forecasters in general to develop their forecasts towards better accuracy and less bias. In the public discussion, the forecasts of the MoF are compared to those of international organizations as well as the Finnish key forecasters. Sometimes the forecasts by various institutions differ from each other by several percentage points and the outlook of the MoF is questioned. Therefore, it is important to know whether any forecaster outperforms another in ex-post forecast comparison.

Assessing the reliability of macroeconomic and fiscal forecasts has become an important topic in international literature since the forecasts play a key role in the EU fiscal policy rules (see e. g. Frankel and Schreger, 2013; Merola and Perez, 2013; Cronin & McQuinn, 2023; Gilbert and de Jong, 2017). The member states have agreed to revise the rules in 2024[1]. In the old rules, the forecasts are essential for the European Commission's supervision since the Excessive Deficit Procedure (EDP) can be activated if the fiscal balance-to-GDP ratio is forecast to exceed the 3% limit. In the new rules, there are no major changes in this respect, which means that fiscal forecasts remain an important topic.

This paper evaluates the official Finnish fiscal forecasts from two important perspectives. First, are there systematic overestimations or underestimations (bias) in fiscal forecasts? Second, how inaccurate are the fiscal forecasts? The above questions are first assessed regarding the Ministry of Finance, after which comparisons are made to determine whether the Ministry of Finance performs significantly differently from the other forecasters.

To address these questions, a novel dataset is constructed. I use data from the autumn forecasts by the Ministry of Finance and other key forecasters from 1997 to 2021. The research period is relatively long compared with the previous literature and includes different macroeconomic shocks that the forecasters have confronted. During the research period, there were the change of exchange rate regime from the Finnish markka to the euro, the dotcom bubble, the global financial crisis, the debt crisis in the euro zone, and Covid-19. I compare the MoF to institutes that are well-established as economic forecasters and whose forecasts are available for the research period. They are the European Commission, the Bank of Finland, the ETLA Economic Research (Etla), and the Organization for Economic Co-operation and Development (OECD). The data used in this report covers the key variables of general government finances: the GDP ratio of general government expenditure, fiscal balance (net lending or borrowing), and debt. In the analysis I focus on the data on forecast errors**,** which are calculated as a difference between the forecast for the reference period and the outturn data collected 6 to 8 months after the reference period, depending on the publication schedules of Statistics Finland. This is because the outturn data is subject to revisions.

The measures used in the analysis are common in the literature evaluating the goodness of out-of-sample forecasts (see e. g. Jiao et al., 2022; Fortin et al., 2020; Boot & Pick, 2020; Reifschneider & Tulip, 2019; Duncan & Martinez-Garcia, 2019).

The main contributions of this paper are the following: (1) I expand the set of variables to be assessed and utilize the longest dataset in the Nordics so far for the purpose of exploring the official fiscal forecasts' quality; (2) I carry out the first long-term comparison with other key forecasters producing economic forecasts for Finland.

I find that there is bias towards pessimism in the general government balance forecast by the Ministry of Finance and the other key forecasters if I omit the large outliers of the crisis years of 2009 and 2020. In addition, bias is detected in the current-year forecasts for the fiscal balance-to-GDP ratio by the Ministry of Finance and the OECD. No statistically significant bias is detected in debt-to-GDP or expenditure-to-GDP forecasts. The MoF's forecasts do not contain more bias compared with the others. The Ministry of Finance's forecasts are as inaccurate as those of the other forecasters, except for the fact that the naïve forecasts and the OECD debt forecast are more inaccurate. Some of the forecast errors are very large and related to the errors made in the short-term GDP growth forecasts. The findings show that there is still room for improvement in the forecasting methods.

[1] See the communication of the European Commission of 23 April 2023 https://ec.europa.eu/commission/presscorner/detail/en/qanda_23_2394 and of 10 February 2024 https://www.consilium.europa.eu/en/press/press-releases/2024/02/10/economic-governance-review-council-and-parliament-strike-deal-on-reform-of-fiscal-rules/.

The remainder of this paper is organized as follows. The next section presents the previous literature. The third section presents the data used in this study. The fourth section presents the ex-post evaluation of the forecast accuracy and bias. The fifth section summarizes the conclusions. The Appendixes contain supplementary material.

2 Previous literature

2.1 Fiscal forecasting errors

There are some ex-post evaluations available of the Finnish official budgetary forecasts from the past few decades. These forecasts have been assessed in international comparisons where the role of Finnish fiscal forecasts is often impossible to identify in a group of country-level forecasts (see e. g. Cronin & McInerney, 2023; Beetsma et al., 2023; Frankel, 2011). In some reports, Finnish fiscal forecast data are the main topic, but the sample is often short (see e. g. National Audit Office, 2016; IMF, 2015). A larger strand of literature evaluates the fiscal forecasts of large international institutions such as the International Monetary Fund (IMF), the OECD, the World Bank, the Economic Intelligence Unit (EIU), and the European Commission (see e. g. Beaudry & Willems, 2022; Rybacki, 2021; Gatti et al., 2024; Gilbert & de Jong, 2017; Estefania-Flores et al., 2023; Cronin & McInerney, 2023).

In the existing literature, the length of the forecasting horizon varies, depending on the setting of the research questions. Forecast horizons are typically divided into short-term (up to one year ahead), medium-term (two to five years ahead), and long-term (beyond five years) as Leal et al. (2008) point out. In studies concerning short-term fiscal forecasts, the focus is on the annual budgeting (see e. g. Beetsma et al., 2023; Cronin & McQuinn, 2021; Fioramanti et al., 2016). Assessments covering medium-term forecasting horizons are linked to fiscal strategies for the medium term and fiscal sustainability (see e. g. Estefania-Flores et al., 2023), but long-term projection evaluations are also linked to fiscal sustainability (see e. g. Morikawa, 2022).

In the existing literature, ex-post evaluations have been examined from three different perspectives: (un)biasedness, accuracy, and effectiveness (Musso & Philips, 2002). The existing reports discussing Finnish fiscal forecasts have mainly examined unbiasedness and accuracy but with short samples (National Audit Office, 2016; IMF, 2015; Lahtinen et al., 2009). The information efficiency of Finnish fiscal forecasts has not been examined previously since the analysis requires parametric methods and longer samples (see e.g. National Audit Office, 2018; Keereman, 1999; Auerbach, 1999).

As a general conclusion, the empirical results of biasedness and accuracy of fiscal forecast evaluations depend on the sample of countries and the time period. However, a quite typical and universal finding is over-optimism of forecasts during times of economic crisis. Another typical and universal finding is that forecasts become more inaccurate the longer the forecast horizon, which also applies to other forecasts than fiscal forecasts (see e. g. Frankel, 2011; Giovanelli & Pericoli, 2020).

In the previous literature, the scale of fiscal forecast errors has been documented in several research studies. Jalles et al. (2015) studied the accuracy and unbiasedness of fiscal balance forecasts by private forecasters in advanced and emerging economies during the years 1993–2009. The mean absolute error, which reflects the accuracy, was 1.33 (0.68) percentage points in advanced economies for the year ahead (current year) and higher, 2.38 (1.75), in emerging economies. They detected that the bias toward optimism was smaller in advanced economies than in emerging economies (0.15 vs. 0.62) for the full sample. They also found that the optimistic bias also remained in contemporaneous *t+0* forecasts in emerging economies (-0.50 percentage points) but not in advanced economies (0.01 percentage points). Fioramanti et al. (2016) studied the European Commission's forecasts for fiscal balance. The mean absolute error for the year ahead (current year) was 1.40 (0.93) for Finland during 1995–2014 and lower, 0.90 (0.51), for the eurozone during 1998–2014. This indicates that it could be more difficult for the European Commission to forecast the public finances of Finland than those of the rest of the eurozone countries. As the Finnish economy is a small open economy, the fluctuation in the business cycles can be higher than in larger economies. Giuriato et al. (2016) found national governments' forecasts in the Stability and Convergence Programmes (SCP) for fiscal balance to be biased towards optimism in the EU during 1999–2013 by 0.31 and 0.24 percentage points for the *t+1* and *t+0* horizons, respectively. However, according to Giuriato et al. (2016), Finland, Denmark, and Germany are the only countries which display prudence in their fiscal forecasts while drafting the budget.

Cronin and McQuinn (2023) studied the debt-to-GDP forecasts in the SCPs of the EU member states and found an underestimation of an average of 1.029 percentage points in the year-ahead forecasts and an overestimation of 0.592 percentage points in the current-year forecasts during 2012–2019. The expenditure-to-GDP forecast errors are rarely studied. Beetsma et al. (2023) found that the EU countries had underestimated the expenditure-to-GDP ratios in the

year-ahead forecasts on average by 0.14 percentage points. Giuriato et al. (2016) found that the expenditure forecasts presented by 13 national governments in the EU during 1999–2013 were overestimated on average by 0.37 and 0.62 percentage points on the *t+1* horizon and the *t+0* horizon, respectively.

2.2 Causes of fiscal forecast errors

A large strand of literature investigates the causes of fiscal forecast errors. Auerbach (1995) divided the causes of errors into three different categories: 1) political, 2) economic, and 3) technical errors. Political errors arise when the impacts of policy changes are not incorporated into a forecast, or a policy change already assumed in the forecast is cancelled. Economic errors arise from errors in the forecasts on macroeconomic variables, such as errors in cyclical forecasts (see e.g. Beetsma et al., 2023) or errors in inflation forecasts (see e.g. Hadzi–Vaskov et al., 2023). According to Auerbach, the rest of the errors are of technical nature. Some factors are not compatible with the framework determined by Auerbach. For example, the poor quality of economic institutions is a factor studied by von Hagen (2010).

Some of the existing research investigates the possibility of government agency bias, i.e. the phenomenon where the forecasts prepared by the government agencies are more biased than those of the private sector or other independent institutions (see e. g. Jonung & Larch, 2006; Krol, 2014). Keereman (1999) suggested that the bias in budgetary forecasting could come from political motivations: there could be an incentive to build a safety margin to meet budgetary targets, also in case of slippage in government revenues or expenditures. According to Leal et al. (2008), the existing research has found that this kind of behaviour is more typical in the EU than in the US.

One possible source of bias in the EU is the fiscal policy rules agreed as part of the Maastricht Treaty and the Stability and Growth Pact (SGP). Assessing the official fiscal forecasts of the EU member states and the European Commission in this context has become an important topic in research literature. The fiscal policy rules set out in the SGP impose restrictions on the general government deficit-to-GDP and debt-to-GDP ratios.[2] There is evidence that some of the EMU member states have presented over-optimistic forecasts when there is the risk of breaking the rules (see e. g. Frankel and Schreger, 2013; Merola and Perez, 2013) or when the Maastricht Treaty threshold values have already been exceeded in the previous period (Cronin & McQuinn, 2023)[3]. Also, Gilbert and de Jong (2017) showed evidence that the European Commission's forecasts[4] are biased upwards when national governments expect European fiscal rules to bind (for fiscal balance and other related variables). The rules have failed to prevent the eurozone countries from running deficits and accumulating debt, and hence, the member states have recently agreed to amend the fiscal rules. In the new fiscal rules agreed at the beginning of 2024[5], the reference values of debt and deficit remain the same, but the focus is on debt sustainability and net expenditure, which is a concept quite similar to the expenditure benchmark in the old rules. Hence, the forecasts of these variables (GDP ratios of debt, fiscal balance, and expenditure) are still important variables in the context of the new fiscal rules.

Beetsma et al. (2019) analyzed the fiscal forecasts in the Stability and Convergence Programmes (SCP) of various countries, including Finland, and found that the presence of a fiscal council is associated with more accurate and possibly less optimistic fiscal forecasts, as well as greater compliance with fiscal rules.

The other domestic forecasters and large international organizations such as the OECD and the European Commission are seen as reference forecasters to the official forecasts in the public discussion. In the forecast evaluation literature, they are used as benchmarks. Auerbach (1999) pointed out that there is a risk that the possible bias of government agencies' forecasts could spread indirectly to other private forecasters as well, if they had incentives not to deviate from the "consensus" estimates too much, or private forecasters had little independent information. Merola and Perez (2013) found evidence of political biases spreading into the forecasts of independent forecasters in their empirical study of the EU countries during 1999–2007.

Despite the interest towards reference forecasters, there is only a handful of forecast comparisons between the Finnish MoF and reference forecasters available. The IMF (2015) compares the general government balance forecast of the Ministry of Finance with the forecasts of some other forecasters[6] over a period of only two years (2012 and 2013),

[2] An EU Member State with excessive deficits and debts may be subject to the disciplinary procedure set out in the Maastricht Treaty. It should be added, however, that the sanctions set out in the EU law have never been applied.

[3] Cronin and McQuinn show this with the SCP forecast data on the debt-to-GDP ratio.

[4] The national officials supply information to the European Commission while the Commission is preparing its forecast.

[5] Economic governance review: Council and Parliament strike deal on reform of fiscal rules, 10 February 2024 https://www.consilium.europa.eu/en/press/press-releases/2024/02/10/economic-governance-review-council-and-parliament-strike-deal-on-reform-of-fiscal-rules/

[6] The Bank of Finland, Etla, the Labour Institute for Economic Research LABORE (Labore), the OECD, and the IMF.

but the sample is very short, and the results cannot be generalized. There are some studies available that compare domestic short-term macroeconomic forecasts (see e. g. National Audit Office, 2018).

For other countries, there are some recent studies of this subject available. Based on Japanese data, Morikawa (2022) compared the business cycle forecasts of academic researchers, professional forecasters in private institutes, and government institutes. The study shows that the economic growth forecasts by academic researchers are less upward biased than those by the government.

Some recent empirical research studies have examined economic errors and found that fiscal forecast errors are at least partly explained by or associated with errors in business cycle forecasts. Some find relevance of forecast errors for real GDP growth or output gap (Cronin & McQuinn, 2023; Cronin & McInerney, 2023; Estefania-Flores et al., 2023). In addition, other factors have been found to be important: the unemployment rate (Croushore & van Norden, 2018), the terms of trade, and the GDP deflator reflecting inflation (Hadzi–Vaskov et al., 2023). The fiscal variables that are typically studied are partly the same variables that my study investigates. The GDP ratios of expenditure, fiscal balance and debt have been investigated by Estefania–Flores et al. (2023), Hadzi–Vaskov et al. (2023), Cronin & McQuinn (2023), Cronin & McInerney (2023), and Croushore & van Norden (2018). Also, variables that are not included in my research have been studied: the structural budget balance and revenues have been evaluated by Cronin and McQuinn (2023), Cronin and McInerney (2023), Hadzi–Vaskov et al. (2023), and Croushore and van Norden (2018).

3 Data and measures

3.1 Data

Each autumn the Ministry of Finance publishes the official economic forecasts which are used for planning the state budget of the upcoming year. I focus on the forecasts of general government balance, i.e. the net lending / net borrowing[7], expenditure[8], and debt[9] in relation to nominal GDP for the current and next year. My sample covers the fiscal balance-to-GDP and debt-to-GDP forecasts for 1997–2021 and the expenditure-to-GDP forecasts for 2000–2021. Hence, for fiscal balance and debt, the current-year (*t+0*) and year-ahead (*t+1*) forecast time series are 25 and 24 years, respectively. For expenditure, the corresponding time spans are 22 and 21 years. The concepts used are in line with the European System of Accounts (ESA) accounting standards ESA 1995 and ESA 2010 (see European Commission, 2000; European Commission, Eurostat, 2014). The changes in the accounting standards have been taken into account in the analysis so that the data set is consistent.

In addition to the Ministry of Finance, there are four other forecasters represented in the sample. They are well-established as forecasters, and their forecasts are adequately available. These are the European Commission (see European Commission, 2022a & 2022b), the ETLA Economic Research (See Etla, 2021), the Organization for Economic Co-operation and Development (OECD) (see OECD, 2022), and the Bank of Finland (see Bank of Finland, 2021)[10]. The forecast data of these forecasters were collected from the forecast publications so that the publication dates were as close as possible to the publication date of the Ministry of Finance's forecast.[11] This was to ensure that the forecasters had approximately the same amount of outturn data of the current year available when formulating their forecasts.

[7] The general government net lending / net borrowing is calculated as a net of general government total revenue and expenditure.
[8] The general government expenditure refers to the consolidated general government expenditure calculated according to the European System of Accounts ESA 1995 and ESA 2010 accounting standards (see European Commission, Eurostat, 1996; European Commission, 2000; European Commission, Eurostat, 2014).
[9] The debt concept used in EDP reporting.
[10] Several other institutions produce forecasts for Finnish general government debt and deficit, but the forecast time series are shorter.
[11] The publication dates of the autumn forecasts have varied over the decades. The Ministry of Finance has published its forecast in September or October. In 1990–1998, 2002–2007, and in 2009–2010, the Bank of Finland published its comprehensive economic forecasts in September or October, while in 1999–2001, 2008, and 2011–2021, the forecast was published in December. Etla has published all its forecasts in September. In 1996–2012 and 2020–2021, the OECD published its forecast in December, while in 2013–2019, the forecast was published in November. The European Commission has published its forecast in November. Not all forecasters had provided the exact publication date of each of their forecasts. In most cases, the information was available at an accuracy of one month. The forecast publication dates differ from each other by a maximum of three months, which probably has some impact on the findings.

In the analysis, I focus on forecast errors, which are calculated as a difference between the forecast for the reference period and the outturn[12] data collected 6 to 8 months after the reference period, depending on the publication schedules of Statistics Finland. This is done since the outturn data are subject to revisions. Therefore, the latest vintage, i.e. the ex-post figures, cannot be directly downloaded from the Statistics Finland database. Hence, I collect the real-time data vintages from Statistics Finland. To create the outturn time series on realized values of the variables focused on, I pick the observation of each outturn data vintage of June or July after the end of the statistical year and, for debt, the autumn data (August–October)[13]. For example, the realized value of the government expenditure-to-GDP ratio of the year 2020 is collected from the data vintage of 14 July 2021.

The data collected for this study are unique. Several studies have investigated the forecast errors of Stability and Convergence Programmes (SCP), which are produced in April[14] for the European Commission (see e. g. Cronin & McInerney, 2023; Cronin & McQuinn, 2021; Strauch, Hallerberg & von Hagen, 2004; Beetsma et al., 2023). However, in practice, the Finnish state budget for the upcoming year is planned based on the autumn forecasts by the Ministry of Finance. This makes the dataset important from the short-term fiscal policy point of view. Also, the comparison in this study is not cross-country but country specific.

3.2 Measures for the goodness of forecasts

A good forecast is on average correct and accurate. To evaluate these characteristics, a forecast error (FE) is calculated as a difference e_t^i between the value forecast $\hat{y}_t^i$ for the variable i and the actual value y_t^i. In the current year *t* (*t+0*), it can be defined[15] as in equation (1):

$$FE_t^i = e_t^i = \hat{y}_t^i - y_t^i. \tag{1}$$

In this presentation, t is always the period when the forecast is produced. In the rest of the descriptions of the measures that I have focused on, the superscript i is removed to ensure simpler presentation. The forecast error for the year ahead (*t+1*) is also shown as the difference between the forecast for the year ahead (*t+1*) and the actual value for that year:

$$e_{t+1} = \hat{y}_{t+1} - y_{t+1}. \tag{2}$$

The mean forecast error (ME) from year *t* to year *T* is obtained by calculating the arithmetic mean of the forecast errors:

$$ME = \bar{e} = \frac{1}{T}\sum_{t=1}^{T} e_t. \tag{3}$$

In this paper, the naïve accuracy of forecasts is evaluated using various indicators. The first of them is the mean absolute error (MAE), which is written as

$$MAE = \frac{1}{T}\sum_{t=1}^{T} |e_t|. \tag{4}$$

The second accuracy indicator is the mean absolute percentage error (MAPE), which is defined as

$$MAPE = \frac{1}{T}\sum_{t=1}^{T} \left|\frac{e_t}{y_t}\right|. \tag{5}$$

[12] Also called realized or actual values.

[13] An exception to this is the debt observation for 2021, which was corrected by 5.9 percentage points downwards, as the exceptional revision carried out by Statistics Finland on 28 June 2022 increased the debt level by about EUR 14 billion. The revision was due to a methodological change in the processing of interest subsidy loans granted to rental housing and right-of-occupancy housing approved by the Housing Finance and Development Centre of Finland (Statistics Finland 2022, 12 December).

[14] Before 2011, the SCPs had to be submitted in the autumn. This makes the SCP data subject to structural change in the middle of the time series.

[15] Traditionally, the forecast error is defined through an equation $y = \hat{y} + e$ in which the actual variable y is obtained as the sum of the forecast $\hat{y}$ and the forecast error e. Solving the forecast error results in $e = y - \hat{y}$. The definition used in this publication is $e = \hat{y} - y$, which is consistent with the assessment produced by the National Audit Office (2018). The method does not affect the findings but only the interpretation of negative and positive errors.

The third way to describe the accuracy of the forecast is the mean squared error (MSE), where the errors are squared before the mean is calculated. In that case, the unit of measurement is also squared. As a result, large errors cause a greater loss than in the above-mentioned indicators. The MSE is written as

$$MSE = \frac{1}{T}\sum_{t=1}^{T} e_t^2. \quad (6)$$

Taking a square root of the MSE gives us the root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T} e_t^2} \quad (7)$$

The accuracy and quality of the forecast is examined by comparing the forecast with a simple naïve forecast. Here[16] the naïve forecast for the current year (*t+0*) is computed as:

$$\hat{y}_t = y_{t-1}. \quad (8)$$

and for the year ahead (*t+1*) as:

$$\hat{y}_{t+1} = y_{t-1}. \quad (9)$$

In the previous forecast evaluation literature, the naïve no-change forecast has been used as a benchmark for example in Fortin et al. (2020) and in Eicher et al. (2019). Simple forecasts are inaccurate: a naïve forecast is not suited for forecasting turning points of a time series. Comparing the forecast with the naïve forecast gives an idea of how much additional benefit the forecaster can achieve by producing the forecast.

The accuracy of the forecasts can be compared using the Theil U_2l statistics (Theil 1966), which is written as

$$U_2 = \sqrt{\frac{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\hat{y}_t - y_t}{y_{t-1}}\right)^2}{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{y_{t-1} - y_t}{y_{t-1}}\right)^2}}. \quad (10)$$

In the numerator, the forecast errors of the compared model have been divided by the previous observation, and in the denominator, the errors of the naïve forecast have been divided by the previous observation.

By definition, this measure is positive (≥ 0). The naïve forecast is assigned the value 1, and the perfect forecast (forecast = outturn) is assigned the value zero. If Theil's U2 is less than one, the forecaster's projections are more accurate than the naïve forecast. If the value of Theil's U2 exceeds one, the naïve forecast is more accurate than that of the other forecaster. The advantage of Theil's U_2 is that variables of different sizes can be scaled to the same scale.

4 Empirical results on fiscal forecast analysis for Finland

Optimistic and pessimistic forecasts vary over time in line with the findings of Giuriato et al. (2016) (Figures 1, 2, and 3). As the forecast error (FE) is defined according to Equation (1), a positive forecast error indicates overestimation and a negative forecast error underestimation. To judge whether fiscal forecast errors are associated with errors in macroeconomic business cycle forecasts, I produce scatter plots between the forecast errors of real GDP growth[17] and fiscal variables as in Hadzi–Vaskov et al. (2023). As seen in Figures 5–7 in Appendix 1, the *t+1* errors are associated with errors made in the business cycle forecasts, but in the *t+0* horizon, the correlation is smaller[18]. The results are not surprising since many recent studies have shown with sophisticated methods that errors in fiscal forecasts are at least partly

[16] A naïve forecast can also be created in other ways, such as by using the autoregressive AR(1) model.

[17] I use the real GDP growth forecast error data by the MoF.

[18] The exception is the expenditure-to GDP forecasts for which the correlation is large also in *t+0* horizon. The reason could be that the automatic fiscal stabilizers are large in Finland (Kellokumpu et al., 2023).

explained by errors made in business cycle forecasting (see e.g. Cronin & McQuinn, 2023; Cronin & McInerney, 2023; Estefania-Flores et al., 2023). These results highlight the importance of high-quality business cycle forecasts.

Some of the forecast errors are very large especially in year-ahead forecasts during times of economic recession, such as the financial crisis in 2009 and the Covid-19 crisis in 2020. The MoF's largest *t+1* forecast errors occur in 2009: in the fiscal balance-to-GDP ratio 6 percentage points, in the debt-to-GDP ratio 13 percentage points, and in the expenditure-to-GDP ratio 9 percentage points. In current-year forecasts, the difference between the errors of crisis years and non-crisis-years is smaller. Also, the fiscal forecasts for the current year (*t+0*) are more accurate than the forecasts for the year ahead (*t+1*). These findings are intuitive, as it is easier to produce forecasts of the current year: some outturn data (1st and 2nd quarter) on the year have already been collected by the time the autumn forecasts are prepared.

In short samples, mean forecast errors are sensitive to outliers. Random factors may be overemphasized, and for this reason, I formed two samples for the analysis: a full sample and a reduced sample, where I excluded the years 2009 and 2020 from the analysis. Omitting the crisis years as a robustness check has also been done in some previous studies, but the selected years vary (see e. g. Giuriato et al., 2016; Beetsma et al., 2023).

The forecasters make errors typically in the same direction with each other, particularly as regards general government fiscal balance in the *t+1* forecasts (see Figures 1–3 and Appendix 2). The reason could be that the forecasters have access to the same information and/or use similar methods when preparing their forecasts. The largest differences in fiscal forecasts between the MoF and the rest of the forecasters arise during the years of economic downturns 2000, 2009–2012, and 2020–2021. During these years, there was a lot of uncertainty of the economic development of Finland. The biggest difference arises in 2021 when the debt-to-GDP forecast error of the OECD is 6,2 percentage points higher than that of the MoF. The finding that there is greater forecaster disagreement in downturns than in upswings has been documented in the literature concerning forecasts of GDP growth (see e. g. Fortin et al., 2020). Already here, the OECD stands out from the rest of the forecasters with its inaccurate fiscal forecasts. Especially the forecasts for the debt-to-GDP ratio for both forecast horizons are clearly more inaccurate than the rest of the forecasters, but the balance-to-GDP forecasts for the current year (*t+0*) are also more inaccurate. In the previous literature, international organizations, such as the OECD, have been found to be more inaccurate and biased than forecasters based in the country they forecast (see Batchelor, 2001). According to Leal et al. (2008), these kinds of findings are also documented in the literature more widely.

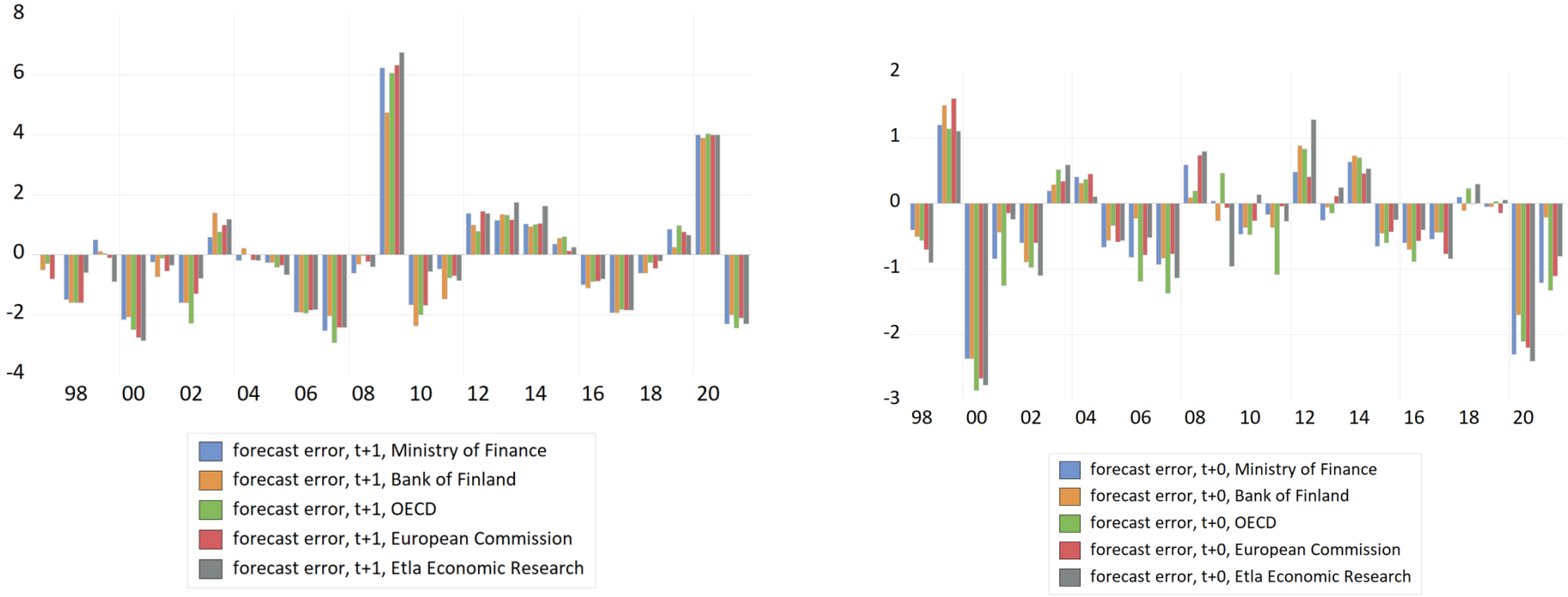

Figure 1: Errors in general government fiscal balance-to-GDP forecasts, percentage points. Forecasts for the year ahead (left panel) and for the current year (right panel).

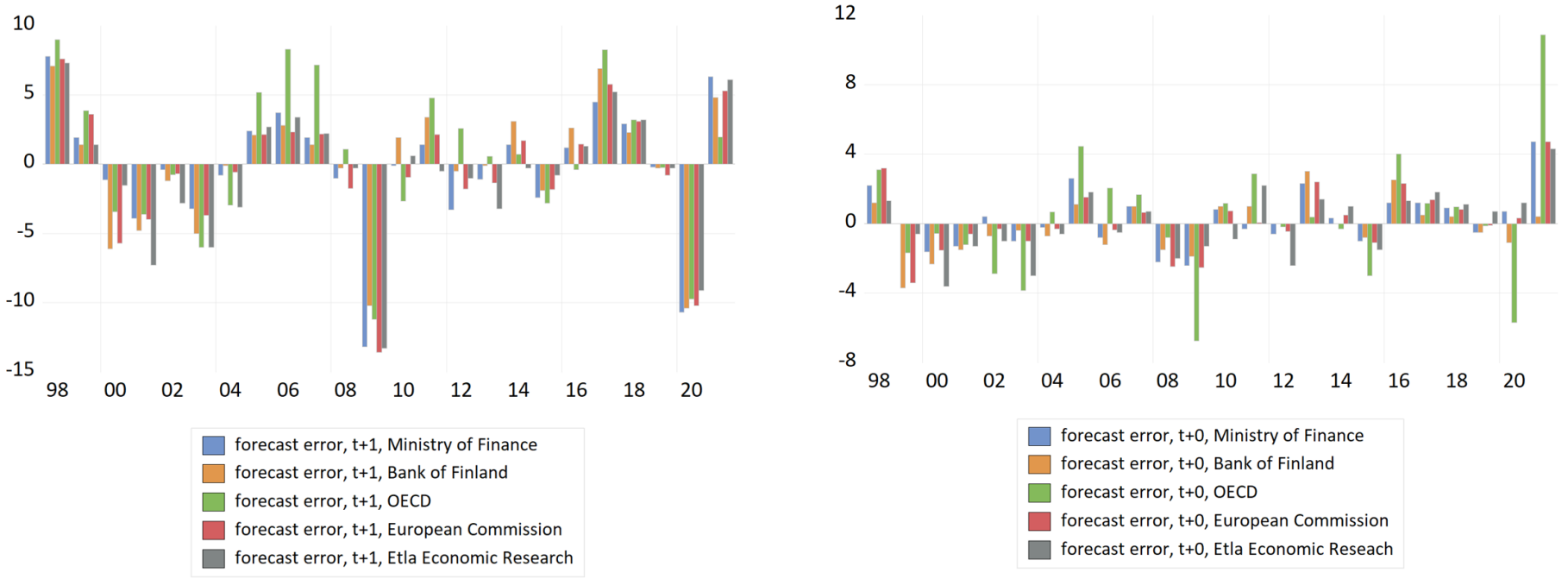

Figure 2: Errors in general government debt-to-GDP forecasts, percentage points. Forecasts for the year ahead (left panel) and for the current year (right panel).

Figure 3: Errors in general government expenditure-to-GDP forecasts, percentage points. Forecasts for the year ahead (left panel) and for the current year (right panel).

4.1 Bias

First, I investigated whether there is bias (systematic overestimation or underestimation) in the MoF's fiscal forecasts. This was examined using a zero-mean forecast error test with Newey–West heteroskedasticity and autocorrelation robust standard errors, when necessary[19], in a somewhat similar way as in Croushore and van Norden (2018) due to the risk of small-sample bias[20]. The findings are presented in Tables 1–2.

General government balance-to-GDP ratio

On average, the general government balance-to-GDP ratio is underestimated, i.e. the forecast is more negative than the outturn, by some decimals (from -0,121 to -0.440) by the MoF and the other forecasters[21] on both forecast horizons (*t+1* and *t+0*) (Table 1). However, the underestimation is statistically significant only for the MoF and the OECD on the current-year (*t+0*) forecast horizon. The sample includes large overestimations in the years 2009 and 2020, and hence, as a robustness check, I calculated similar measures for the reduced sample. After omitting the outliers, the underestimations become clearer in the budgeting horizon (*t+1*): all forecasters underestimated the general government balance-to-GDP ratio by -0.487 to -0.658 percentage points (Table 1) in their autumn forecasts. The underestimations of the *t+1* forecasts are statistically significant in all cases except for Etla and the OECD. In the *t+0* forecasting horizon, omitting the outliers decreases the mean error in all cases except for the naïve forecast, and no statistical significance is documented. This means that the current-year forecasts are less biased than the forecasts for budgeting horizon (*t+1*).

The negative bias is also seen in the error distributions of the Ministry of Finance and the other forecasters (Appendix 3, Figures 11–12) and even more clearly in the forecast error medians (Table 2). The forecast error distributions of the MoF (and the other forecasters) are typically centred between 0 and -1 for both the year-ahead and the current-year forecast, and the large outliers of the years 2009 and 2020 are positive. This gives further indication of a small but systematic underestimation (negative bias) in the fiscal balance forecasts. A recurring underestimation of the fiscal balance was also noted in the forecast error diagrams (Figure 1). Also, the medians are negative in all cases, except for the naïve *t+0* forecast. (Table 2).

Similar results of underestimations by the MoF have already been reported in the previous literature using Stability and Convergence Programme (SCP) data (forecasts made in the spring) and shorter samples. Using forecasts of 1998–2001, Strauch et al. (2004) found signs of pessimism in fiscal balance forecasts for Finland. Beetsma et al. (2023) documented underestimation of fiscal balance in SCP spring forecasts during 1999–2021. In both studies, pessimism is also found in the official balance forecasts of some other countries, such as Sweden and Luxemburg.

[19] The number of observations (21–25) contained in the data is relatively small for statistical testing, and for this reason, the decrease in the power of the test and the increase in the risk of rejection or acceptance error concerning the tested hypothesis have been considered. In the case of a small sample, it cannot be assumed that the distribution of the sample mean follows the conventional distribution of the *t*-test statistic even if the forecast errors of the entire population (for example, during Finland's independence) were not distributed normally. The t-test also requires independent forecast errors over time (no autocorrelation) and homoscedasticity (constant variance of the forecast error over time). Therefore, these characteristics of the forecast error series were tested before the t-tests were applied. In cases where the t-test assumptions (normality of distribution, no autocorrelation, no heteroscedasticity) were met, the normal t-test was used. When the assumptions were not met, the sample mean standard errors were calculated using HAC standard errors that consider heteroscedasticity and autocorrelation (Newey & West 1987). When the HAC standard errors were calculated, the confidence intervals became extreme large (Tables 4–9). The result reflects the uncertainty associated with the parametric testing of short time series when the underlying assumptions of the t-test are not met. Due to the small number of observations in the sample, the results of the parametric tests should be interpreted with caution.

[20] The Mincer–Zarnowitz test, where the realized value of a variable is regressed on a constant and forecasts, is typically used for testing bias. This is however not recommendable with short samples. (Croushore & van Norden, 2018.)

[21] An exception is the Etla Economic Research, whose mean error is exactly zero in the full sample. However, the median is negative (Table 2).

Table 1: Mean errors (ME) of the fiscal forecasts.

Full sample	balance-to-GDP	debt-to-GDP	expenditure-to-GDP
t+1	***t+1***	***t+1***	***t+1***
Ministry of Finance	-0.121	-0.251	-0.729
Etla Economic Research	0.000	-0.672	-
Bank of Finland	-0.234	-0.047	-0.238
European Commission	-0.130	-0.402	-0.681
OECD	-0.184	0.534	-
Naïve forecast	0.146	-0.954	-0.686
t+0	***t+0***	***t+0***	***t+0***
Ministry of Finance	-0.360**	0.415	0.200
Etla Economic Research	-0.192	0.159	-
Bank of Finland	-0.264	-0.025	0.473
European Commission	-0.308	0.330	0.114
OECD	-0.440**	0.425	-
Naïve forecast	-0.020	-0.340	-0.264
Reduced sample	balance-to-GDP	debt-to-GDP	expenditure-to-GDP
t+1	***t+1***	***t+1***	***t+1***
Ministry of Finance	-0.596**	0.551	-0.116
Etla Economic Research	-0.487	0.285	-
Bank of Finland	-0.646**	0.885	0.268
European Commission	-0.609**	0.642	-0.116
OECD	-0.658**	1.536	-
Naïve forecast	-0.414	0.167	-0.079
t+0	***t+0***	***t+0***	***t+0***
Ministry of Finance	-0.292	0.525	0.165
Etla Economic Research	-0.061	0.220	-
Bank of Finland	-0.200	0.103	0.505
European Commission	-0.235	0.455	0.095
OECD	-0.405	1.004	-
Naïve forecast	-0.522	0.496	0.260

*Notes: All values are in percentage points. Negative (positive) value refers to underestimation (overestimation). Reduced sample refers to a sample where the years 2009 and 2020 have been omitted as outliers. Balance-to-GDP refers to the net lending / net borrowing of general government calculated as a net of total revenues and total expenditures. Debt-to-GDP refers to the general government debt concept used in the Excessive Deficit Procedure (EDP) reporting. Expenditure-to-GDP refers to the consolidated total expenditure concept used in European System of Accounts (ESA) accounting standards. ** denotes significant at 5%.*

Table 2: Median errors of the fiscal forecasts.

Full sample	balance-to-GDP	debt-to-GDP	expenditure-to-GDP
t+1	***t+1***	***t+1***	***t+1***
Ministry of Finance	-0.366	-0.150	-0.200
Etla Economic Research	-0.487	-0.300	-
Bank of Finland	-0.450	0.650	-0.100
European Commission	-0.413	-0.642	-0.300
OECD	-0.164	0.642	-
Naïve forecast	-0.500	-0.489	-0.200
t+0	***t+0***	***t+0***	***t+0***
Ministry of Finance	-0.400	0.300	0.000
Etla Economic Research	-0.250	0.700	-
Bank of Finland	-0.262	0.000	0.000
European Commission	-0.148	0.062	-0.150
OECD	-0.441	0.372	-
Naïve forecast	0.096	0.600	0.250
Reduced sample	balance-to-GDP	debt-to-GDP	expenditure-to-GDP
t+1	***t+1***	***t+1***	***t+1***
Ministry of Finance	-0.536	0.550	0.000
Etla Economic Research	-0.583	-0.300	-
Bank of Finland	-0.673	1.400	0.000
European Commission	-0.494	0.420	-0.300
OECD	-0.338	0.894	-
Naïve forecast	-0.670	1.661	0.100
t+0	***t+0***	***t+0***	***t+0***
Ministry of Finance	-0.400	0.150	0.000
Etla Economic Research	-0.241	0.100	-
Bank of Finland	-0.222	0.000	0.000
European Commission	-0.148	-0.019	-0.150
OECD	-0.441	0.512	-
Naïve forecast	-0.202	0.661	0.550

Notes: All values are in percentage points. Negative (positive) value refers to underestimation (overestimation). Reduced sample refers to a sample where the years 2009 and 2020 have been omitted as outliers. Balance-to-GDP refers to the net lending / net borrowing of general government calculated as a net of total revenues and total expenditures. Debt-to-GDP refers to the general government debt concept used in the Excessive Deficit Procedure (EDP) reporting. Expenditure-to-GDP refers to the consolidated total expenditure concept used in European System of Accounts (ESA) accounting standards. No statistical testing has been done for the median errors.

In addition, Giuriato et al. (2016) found signs of prudency in the forecasts of Finland, Denmark, and Germany during the years 1999–2013 while on average the group of 13 EU countries were overly optimistic by 0.31 (0.24) percentage points at the *t+1* (*t+0*) forecast horizons, respectively. My sample is longer, and the forecasts have been made in the autumn instead of spring, but for Finland, the same result holds.

Furthermore, my results are in line with Cronin and McInerney (2023), who showed that biases of a pessimistic form are found in the forecasts of budget balance and output variables in the data of the Stability and Convergence Programmes (SCP) of the 28 EU member states as well as in the European Commission's forecasts between 2013 and 2019. According to their findings, these errors were explained by the errors in the output variables and by EU fiscal rule requirements. Also, older forecast data of the European Commission has been previously evaluated. Fioramanti et al. (2016) showed that the European Commission's forecasts for Finland during 1995–2014 were underestimating the fiscal balance by 0.21 (0.26) percentage points for the year ahead (current year), respectively. However, the underestimations were not statistically significant. Significant underestimations were found in the forecasts for Sweden, Luxemburg, Austria, and Germany. My results show that using a longer time period and controlling for

the years of economic crisis reveals a more pessimistic forecasting behaviour of the European Commission since the mean errors are larger and more negative in this study.

The Ministry of Finance's mean error in the forecasts of the fiscal balance-to-GDP ratio is compared with the mean errors of the reference forecasters. The Ministry's forecast error differs from that of the other forecasters by a few decimal points, but the differences are not statistically significant. The Ministry of Finance's point estimate falls within the confidence intervals of the other forecasters. Thus, the mean forecast errors of the Ministry of Finance are not statistically significantly higher than the mean forecast errors of the reference forecasters.

It seems that the other forecasters than the MoF have also been prudent in their fiscal balance forecasts even though there is no requirement for that except for the BoF[22]. This could be because the forecasters seek to produce forecasts that do not deviate too much from the MoF's forecasts or consensus forecasts as hypothesized by Auerbach (1999) in his study of US data. This would not be surprising, since similar results have been obtained before. In an empirical study of 15 EU countries[23], Merola and Perez (2013) found evidence of political biases spreading into the forecasts of independent forecasters during 1999–2007.

General government debt-to-GDP ratio

Concerning the full sample, the MoF and the other forecasters underestimated the general government debt-to-GDP ratio on the *t+1* forecast horizon on average by -0.047 to -0.672 percentage points except for the OECD, which overestimated the debt-to-GDP ratio by 0.642 percentage points (Table 1). On the *t+0* forecast horizon, the results were mixed as there were signs of under- and overestimations by -0.025 to 0.425 percentage points. The year-ahead forecast errors are driven by the large outliers where the debt-to-GDP ratio level is remarkably underestimated in the crisis years 2009 and 2020. The under- or overestimations are, however, not statistically significant for the MoF or the other forecasters in the full sample. The MoF's forecasts do not contain more bias than the other forecasters' projections.

When the outliers are omitted (reduced sample), the tendency to overestimate the debt-to-GDP ratio emerges more clearly than in the full sample for both forecast horizons as all the mean errors are positive. The means and the medians give mostly similar results, except for the negative medians in Etla's *t+1* and the European Commission's *t+0* forecasts. However, the overestimations are not statistically significant. In the reduced sample as well, the OECD stands out from the other forecasters with more overestimation than the others, but no statistically significant bias is detected (Table 1). The reason is the use of HAC standard errors, which widens the confidence intervals remarkably. This reflects the problems with statistical testing of short time series.

No signs of systematic errors are found in the error distributions, either, since they are centred close to zero with the peak on the negative side and many overestimations up to 9 percentage points (Appendix 3). The large negative outliers are from 2009 and 2020. Also, the means and the medians give similar signs in 19 out of 24 cases, and the scale of the medians is roughly the same as in the mean errors (Tables 1 and 2). These findings indicate robustness of the results.

The official Finnish debt-to-GDP ratio forecasts have rarely been assessed in the literature. Using SCP forecasts for 26 EU members, including Finland, Cronin and McQuinn (2023) found evidence of underestimation in the EU in the forecast horizons from *t+1* to t+4 during 2012–2019. In current-year forecasts (*t+0*), they found signs of overestimation. My results are partly in line with the EU's general picture. In the *t+0* horizon, I get signs of overestimations as in Cronin and McQuinn (2023) for both the reduced and the full sample. In the *t+1* forecasts, I get evidence of underestimation in the full sample in the same manner as in Cronin and McQuinn (2023). However, omitting the outliers gives opposite results to those of Cronin and McQuinn. The reason is that their sample also covers the years 2016–2018, when Finland faced rapid economic growth, during which the MoF overestimated the debt-to-GDP ratio (Figures 2, 9). However, my longer sample captures several business cycles and reveals the typical forecasting behaviour of the MoF and the reference forecasters.

The effect of business cycles and surprise recessions on debt-to-GDP forecast errors have been studied before with wider datasets and more sophisticated methods. Estefania–Flores et al. (2023) found, using

[22] The Bank of Finland produces the fiscal forecasts together with the European System of Central Banks (ESCB), which have together agreed on the principal of prudency of forecasts.

[23] The countries were Belgium, Germany, Greece, Ireland, Spain, France, Italy, Luxembourg, Netherlands, Austria, Portugal, Finland, Denmark, Sweden, and the United Kingdom.

the IMF and EIU forecasts of 174 countries, that the forecast errors of debt-to-GDP ratios are associated with surprise recessions in the forecast horizon in advanced economies. Also, Hadzi–Vaskov et al. (2023) found debt-to-GDP ratio forecast errors to be associated with errors in GDP growth forecasts. They showed, based on data of six Latin American countries, that the forecast errors for public debt-to-GDP ratios were negatively associated with surprises in GDP growth. This means that optimistic GDP forecasts are associated with underestimated debt-to-GDP ratios. I obtained similar indications using basic scatter plots (Appendix 1), but more detailed analysis of this is left for further research.

General government expenditure-to-GDP ratio

In the full sample, the MoF and the other forecasters underestimated the general government expenditure-to-GDP ratio at the *t+1* forecast horizon but overestimated it in *t+0* forecasts (Table 1). The negative mean error of the expenditure forecasts for the year ahead is partly driven by the large underestimations in the years 2009 and 2020. This can be judged from the fact that in the reduced sample, there are less signs of underestimations. The medians, which are not as sensitive to outliers as mean errors, are closer to zero than the mean errors (Table 1). The underestimations are not statistically significant for the MoF or the other forecasters in any of the cases or samples, which means that the expenditure-to-GDP forecasts are not biased. The MoF's forecasts do not contain more bias than the other forecasters' projections.

The distributions of the general government expenditure-to-GDP forecast errors support my findings (Appendix 3). The error distributions by the Ministry of Finance and other forecasters are centred close to zero for both the *t+1* and *t+0* forecasts. However, the small number of observations starts to interfere with the interpretation of the histograms (sample from the year 2000 onwards only).

There are few research papers on possible bias in the general government expenditure forecasts by the Finnish MoF. The IMF (2015) concluded that the Finnish MoF's expenditure forecasts in the SCPs have been underestimating the outturn in the *t+0*–t+2 horizon during 2000–2013. They also find that the MoF's expenditure forecasts are relatively biased compared with other European countries. In line with the IMF, the results show signs of underestimations in the *t+1* horizon although my sample is eight years longer. But on the *t+0* horizon, I get opposite results to those of the IMF. When the years 2000–2021 are considered and the data of autumn forecasts are used, the MoF's *t+0*-forecasts overestimate the expenditure-to-GDP ratio but not statistically significantly.

Giuriato et al. (2016) found that the expenditure forecasts presented by 13 national governments in the EU, including Finland, were on average overestimated by 0.37 and 0.62 percentage points on the *t+1* and *t+0* horizon, respectively, during 1999–2013. In line with the results of Giuriato et al., I find signs of overestimation in the *t+0* horizon. But on the *t+1* horizon, my results are opposite to those of Giuriato. The reason could be the slowdown in GDP growth which Finland experienced at the beginning of the 21st century. This also resulted in a higher expenditure-to-GDP ratio than expected.

There is statistically significant bias in the general government balance-to-GDP forecasts of the MoF but not in the expenditure-to-GDP forecasts. Since the fiscal balance is calculated as a net of government revenue and expenditure, and the expenditure forecasts are unbiased, the underestimation of the fiscal balance is driven by the underestimation of revenue. Cronin and McInerney (2023) obtained similar results using an EU-wide dataset of forecasts produced by national governments and the European Commission during 2013–2019. This shows that the forecasting behaviour of the MoF is similar in this sense to the EU-wide behaviour.

A slightly pessimistic approach to forecasting general government revenue and fiscal balance is more desirable than over-optimistic forecasts from the perspective of the precautionary principle that the Budgetary Frameworks Directive (2011/85/EU) has required from the MoF since 2013. Similarly, it is better to overestimate expenditure and debt rather than to underestimate them. For this reason, a slight systematic error in forecasts skewed in this direction is acceptable as it may lead to more responsible fiscal policy planning and budgeting.

4.2 Accuracy

In the full sample, the mean absolute error (MAE) of the Ministry of Finance's forecasts for the year ahead are for the balance-to-GDP ratio 1.463 percentage points (other forecasters 1.434–1.483, naïve forecast

2.571), for the debt-to-GDP ratio 3.201 percentage points (other forecasters 3.363–4.183, naïve forecast 5.704), and for the expenditure-to-GDP ratio 1.967 percentage points (other forecasters 1.724–1.795, naïve forecast 2.362) (Table 3, Figure 4). Almost all the mean absolute errors are statistically significant, except for the cases where HAC standard errors had to be used due to autocorrelation or heteroskedasticity[24]. Omitting the crisis years as outliers and using the reduced sample decreases the mean absolute error of the MoF by 0.331–0.981 percentage points (and of the other forecasters by 0.261–0.762 percentage points) (Table 3). This finding is intuitive since in the *t+1* forecasts, the largest errors are made in the crisis years (Figures 1–3). After the outliers have been omitted, most of the errors are still large compared to the previous literature and statistically significant. For the current period forecasts, leaving out the crisis years does not improve the accuracy as much since the differences between the errors of the crisis years and non-crisis-years are smaller (Figures 1–3). In some cases, omitting the outliers even makes the *t+0* forecasts less accurate on average.[25] In the current-year forecast horizon, the MoF's forecast accuracy improves by 0.016–0.090 percentage points (other forecasters from -0.040 to 0.317).

In both samples, the MAEs of the *t+0* forecasts are around half (40%–50%) of the size of the corresponding *t+1* forecasts, which means that the forecasts improve over time. The finding is intuitive, as it is easier to produce forecasts on the current year: statistical data on the year have already been collected by the time the autumn forecasts are produced[26]. The finding is in line with the previous literature. It is often found that the shorter the forecast horizon, the more accurate the forecasts get. This applies for fiscal forecasts (see e. g. Frankel, 2011; Estefania-Flores, 2023) but also for other macroeconomic variables (see e. g. analysis of GDP forecasts in Fortin et al., 2020).

In the previous forecast evaluation literature, there are some estimates available of the Finnish fiscal forecast accuracy for different time periods. Fioramanti et al. (2016) evaluated the European Commission's forecasts and calculated the mean absolute forecast errors (MAE) for the balance-to-GDP ratio for several EU countries. For Finland, the MAE of the year-ahead (current-year) forecasts for the period 1995–2014 was 1.40 (0.93) percentage points. The estimates are close to the values I obtained for *t+1* forecasts but higher for *t+0* forecasts. This could mean that the European Commission has improved the *t+0* forecasting for Finland over time since my sample ends later. For the period 1998–2014, the corresponding numbers for the euro area are 0.78 (0.50), and for the period 1995–2014, the corresponding numbers for Sweden are 1.31 (1.02). This means that the Finnish mean absolute errors are higher than the average of the euro area countries and higher compared with Sweden for the *t+1* forecasts, but the difference is only around one decimal. This could mean that it is more difficult for the European Commission to forecast the Finnish fiscal balance than that of the rest of the euro zone countries.

An et al. (2018) calculated forecast errors from the IMF's and Consensus Economics forecasts of the fiscal balance-to-GDP ratio for 29 countries, including both advanced and emerging market economies. The sample covered the years 2001–2014. For advanced economies, such as Finland, the mean absolute error of the year-ahead forecasts prepared by the IMF (Consensus) was around 1.450 (1.263), and for the current-year forecasts around 0.825 (0.637). According to my results, the mean absolute errors of the forecasts for the fiscal balance of Finland are close to the findings of An et al. (2018), even though my sample is longer, and the forecasters are not the same. Somewhat older data were analyzed in Jalles et al. (2015), where the forecasts prepared by Consensus Economics for the years 1993–2009 were analyzed. For fiscal balance, the value of the MAE for advanced economies for the year ahead is 1.33 and for the current year 0.68, which is close to my results. These findings show that the accuracy of the Finnish fiscal forecasts is close to the levels of other advanced economies.

[24] In the *t+1* forecasts, there was autocorrelation in the full sample in all the naïve forecasts and the debt forecast of the OECD. In the reduced sample, there was autocorrelation in the debt forecast of Etla, the European Commission the OECD, and the naïve forecast, and in the expenditure forecast of the MoF and the European Commission, and the naïve forecast. After the HAC standard errors were used, the confidence intervals increased to cover several percentage points. This reflects the uncertainty associated with the statistical testing using short samples and leads to the conclusion that some mean absolute errors do not differ from zero statistically significantly (Appendix 4, Tables 4–9).

[25] This means that in some cases (the debt forecasts of Etla and the European Commission and the expenditure forecasts of the Bank of Finland), the forecasters have been able to predict the current year during the crisis quite well compared with years of more normal economic growth.

[26] The quarterly national accounts data of the second quarter have typically been available from the Statistics Finland at the beginning of September and those of the third quarter at the beginning of December. The general government expenditure and revenue data by quarter are available a few weeks after the GDP figures.

Next, I compare the accuracy of the Ministry of Finance's forecasts with that of the other forecasts (Table 3, Figure 4). The differences between the MoF and the other forecasters in the *t+1* forecast horizon are for the balance-to-GDP ratio only 0.004–0.029 percentage points (naïve 1.108), for the debt-to-GDP 0.162–0.982 percentage points (naïve 2.503), and for the expenditure-to-GDP ratio around 0.172–0.243 percentage points (naïve 0.395). At the *t+0* forecast horizon, the differences between the forecasters increase by some decimals.

The OECD's forecast for the debt-to-GDP ratio is clearly more inaccurate at both forecast horizons. This is also seen in the error distributions of the OECD, which are flatter than those of the other forecasters (Figures 13–14), and in the forecast errors over time (Figure 2). One possible reason for this could be that it is more difficult for the OECD to produce fiscal forecasts for a small open economy like Finland because the OECD is not a local institution. The result is in line with the finding by Leal et al. (2008), who summarized the previous literature of fiscal forecasts and found that outside forecasters tend to be less accurate than inside forecasters. Also, Batchelor (2001) had similar findings in his study using macroeconomic data on real GDP, consumer expenditure, business investment, industrial production, consumer price index, and unemployment rate. However, in some studies, the international organizations perform better. An et al. (2018) compared the IMF's fiscal forecasts with the private sector Consensus forecasts and found that the IMF's forecasts were not significantly less accurate than the Consensus forecasts.

The naïve forecasts for the debt-to-GDP ratio and fiscal balance-to-GDP ratio are clearly more inaccurate than those of the MoF's or the other forecasters, but the forecasters have difficulties in outperforming the naïve forecast in the expenditure-to-GDP forecast especially when the crisis years are omitted[27] (Table 3 and Figure 4). Theil's U2 statistics show similar findings: it is below unity in most cases except for the expenditure-to-GDP forecasts for the *t+0* horizon (Tables 4–9). The reason could be that the expenditure-to-GDP ratio does not fluctuate much over time, and hence, the naïve no-change forecast is more difficult to beat. The findings show that the key forecasters produce most added value when they put effort in preparing the forecasts for the debt-to-GDP and balance-to-GDP ratios.

To judge whether the differences between forecasters are statistically significant, I use confidence intervals (Tables 4–9). I find that the Ministry of Finance's forecasts are as accurate as those produced by the other forecasters, apart from the debt-to-GDP *t+0* forecast, where the MoF is more accurate than the naïve forecast. The point estimate of the Ministry of Finance's forecast error falls within the 95% confidence intervals of the other forecasters in all cases, except for the *t+0* debt forecast, in which the Ministry's forecast is more accurate than the naïve forecast.

The results obtained using other indicators show a rather similar picture, which indicates the robustness of the results (Appendix 4). The results obtained using the RMSE do not deviate too much from the results based on the MAE. The MSE differs from the MAE in that due to squared value, it highlights large forecast errors made by the OECD and the naïve forecast. Otherwise, the MSE results are mostly in line with previous findings. The mean absolute percentage error (MAPE) reveals that the errors of fiscal balance forecasts are highly proportional to the realized values[28]: the MAPE of fiscal balance ranges from 22.454 to 64.539 (naïve forecasts 59.761-111.433), whereas the MAPE values of the debt and expenditures forecasts range from 1.283 to 8.767 and 1.283 to 3.662, respectively.

[27] It is a quite typical finding in the literature that the naïve forecasts perform poorly during crisis years, as the naïve forecast cannot predict the turning points of time series. See a recent comparison of IMF forecasts and the naïve forecasts in Eicher et al. (2019), who showed that the IMF has outperformed the naïve forecasts in times of crises.
[28] When calculating the MAPE of the fiscal balance-to-GDP ratio, the denominator is often close to zero, which results in high values of the total measure.

Table 3: Mean absolute errors (MAE) of the fiscal forecasts.

Full sample	balance-to-GDP	debt-to-GDP	expenditure-to-GDP
t+1	***t+1***	***t+1***	***t+1***
Ministry of Finance	1.463**	3.201**	1.967**
Etla Economic Research	1.467**	3.455**	-
Bank of Finland	1.434**	3.363**	1.724**
European Commission	1.452**	3.507**	1.795**
OECD	1.483**	4.183	-
Naïve forecast	2.571	5.704	2.362
t+0	***t+0***	***t+0***	***t+0***
Ministry of Finance	0.656**	1.369	1.055**
Etla Economic Research	0.856	1.657**	-
Bank of Finland	0.576**	1.281	1.064**
European Commission	0.633**	1.462**	0.795**
OECD	0.806**	2.587	-
Naïve forecast	1.636**	3.092**	1.445**
Reduced sample	balance-to-GDP	debt-to-GDP	expenditure-to-GDP
t+1	***t+1***	***t+1***	***t+1***
Ministry of Finance	1.132**	2.220**	1.484
Etla Economic Research	1.114**	2.751	-
Bank of Finland	1.173**	2.733**	1.374**
European Commission	1.116**	2.745	1.347
OECD	1.161**	3.611	-
Naïve forecast	2.231**	5.262	1.932
t+0	***t+0***	***t+0***	***t+0***
Ministry of Finance	0.613	1.353**	0.965**
Etla Economic Research	0.783	1.672**	-
Bank of Finland	0.539	1.262**	1.085**
European Commission	0.588	1.467**	0.675**
OECD	0.766	2.270**	-
Naïve forecast	1.279	2.496	1.040**

*Notes: All values are in percentage points. Reduced sample refers to a sample where the years 2009 and 2020 have been omitted as outliers. Balance-to-GDP refers to the net lending / net borrowing of general government calculated as a net of total revenues and total expenditures. Debt-to-GDP refers to the general government debt concept used in the Excessive Deficit Procedure (EDP) reporting. Expenditure-to-GDP refers to the consolidated total expenditure concept used in European System of Accounts (ESA) accounting standards. ** denotes significant at 5%.*

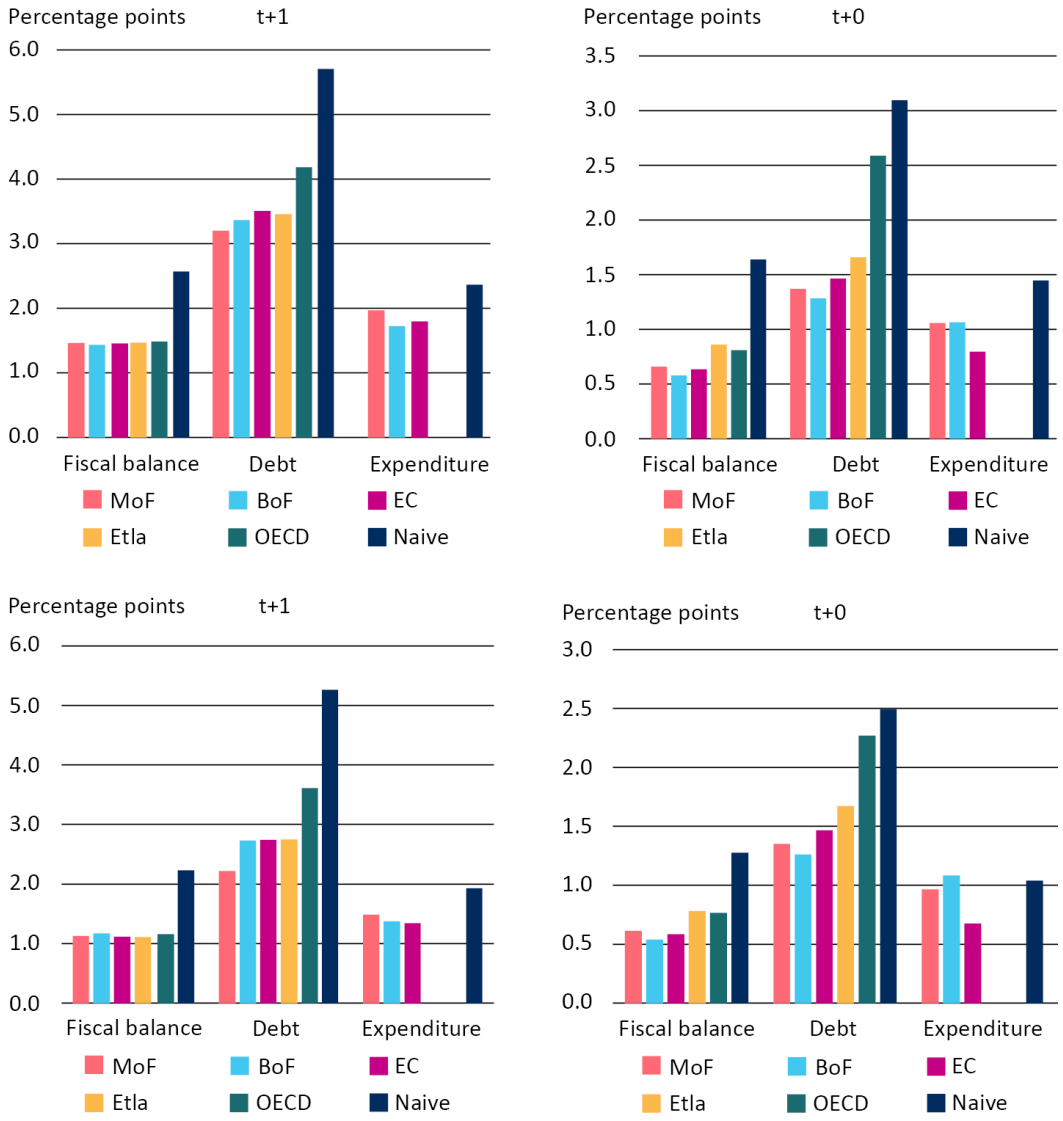

Figure 4: Mean absolute errors (MAE). Upper row: full sample. Lower row: reduced sample. MoF = Ministry of Finance, BoF = Bank of Finland, Etla = Etla Economic Research, OECD = Organisation for Economic Co-operation and Development, Naïve = Naïve forecast.

5 Conclusions

This paper provides systematic evidence of the quality of the Ministry of Finance's official fiscal forecasts used in the planning of the state budget of Finland. The dataset covered the forecasts of the GDP ratios of general government balance, debt, and expenditure by the Ministry of Finance and other key forecasters during the years 1997–2021. For each variable, I evaluated the year-ahead (*t+1*) and the current-year (*t+0*) forecasts published in autumn. The analyses were based on traditional, univariate measures, which are common in the forecast evaluation literature.

I found that the periods of under- and overestimations vary over time and are associated with the errors made in the macroeconomic business cycle forecast. This finding is in line with the recent literature. On average, the Ministry of Finance underestimates the general government balance-to-GDP ratio in both forecast horizons. The results do not change even if I leave out the outliers (years 2009 and 2020).

However, the underestimation is statistically significant only in the year-ahead (*t+1*) forecasts after omitting the outliers and in the current-year (*t+0*) forecasts in the full sample. The other key forecasters also underestimate the balance, and the MoF's forecasts do not contain more bias than the others. The previous literature has also documented underestimation of the Finnish fiscal balance-to-GDP ratio, even though there are differences in the time periods, and some studies use the spring forecast data.

There is no statistically significant bias in the debt-to-GDP ratio or expenditure-to-GDP ratio forecasts by the MoF or the other key forecasters. Even though the underestimations or overestimations are not statistically significant, I found that, especially in the one-year-ahead forecasts, the MoF has typically overestimated the debt-to-GDP ratio but at times of economic crisis, underestimated it. Regarding the expenditure-to-GDP, the MoF underestimates the expenditures on the one-year-ahead forecast horizon but overestimated it contemporaneously. The MoF's debt forecasts are not more biased than the forecasts of other key forecasters.

A slightly pessimistic approach to forecasting the general government balance is more desirable than over-optimistic forecasts from the perspective of the precautionary principle that the Budgetary Frameworks Directive (2011/85/EU) has required from the MoF since 2013. Similarly, it is better to overestimate expenditure and debt rather than to underestimate them. For this reason, a slight systematic error in the forecasts skewed in this direction is acceptable as it may lead to more responsible fiscal policy planning and budgeting. However, despite the observed prudent forecasting behaviour, the Finnish public finances have run large deficits and accumulated a large amount of debt after the financial crisis.

The forecasters made large forecast errors particularly in the years of economic crises. Furthermore, in the period of 1997–2021, the mean absolute errors of the Ministry of Finance and the other forecasters were large on average and statistically significant in most cases. The results indicate large inaccuracies in the forecasts and even omitting the outliers does not change this result remarkably. The Ministry of Finance is as inaccurate as the other forecasters, except for the naïve debt forecast and the debt forecasts by the OECD, which are more inaccurate. Compared with the previous literature, my results are close to the estimates obtained before, even if the time periods differ. The reason could be that either the forecasting methods have not developed remarkably, or they have but at the same time forecasting has become more difficult. However, my analysis also provided new information since the forecasts of expenditure-to-GDP ratio and debt-to-GDP ratio are rarely analyzed in the literature.

The forecasters were compared with the naïve forecast because it is one of the simplest forecasting models. However, the naïve forecast is inaccurate, especially at the turning points of the time series. The naïve forecasts for the debt-to-GDP ratio and fiscal balance-to-GDP ratio are clearly more inaccurate than those of the MoF's or the other forecasters, but the forecasters have difficulties in outperforming the naïve forecast in the expenditure-to-GDP forecasts especially when the crisis years are omitted. However, in the full sample, all the forecasters produce more accurate forecasts than the naïve forecast. Forecasters therefore generate added value when preparing predictions compared with a situation in which only the naïve forecast is used.

The reported findings show that, on average, the Ministry of Finance's forecasts are as reliable as the forecasts produced by the other key forecasters from the perspective of unbiasedness and accuracy. The Ministry of Finance has produced its fiscal forecasts in accordance with the provisions of the Budgetary Frameworks Directive. Still, there is room for improvement of the forecasts.

Appendix 1: Scatter plots between the forecast errors of real GDP growth and the fiscal variables

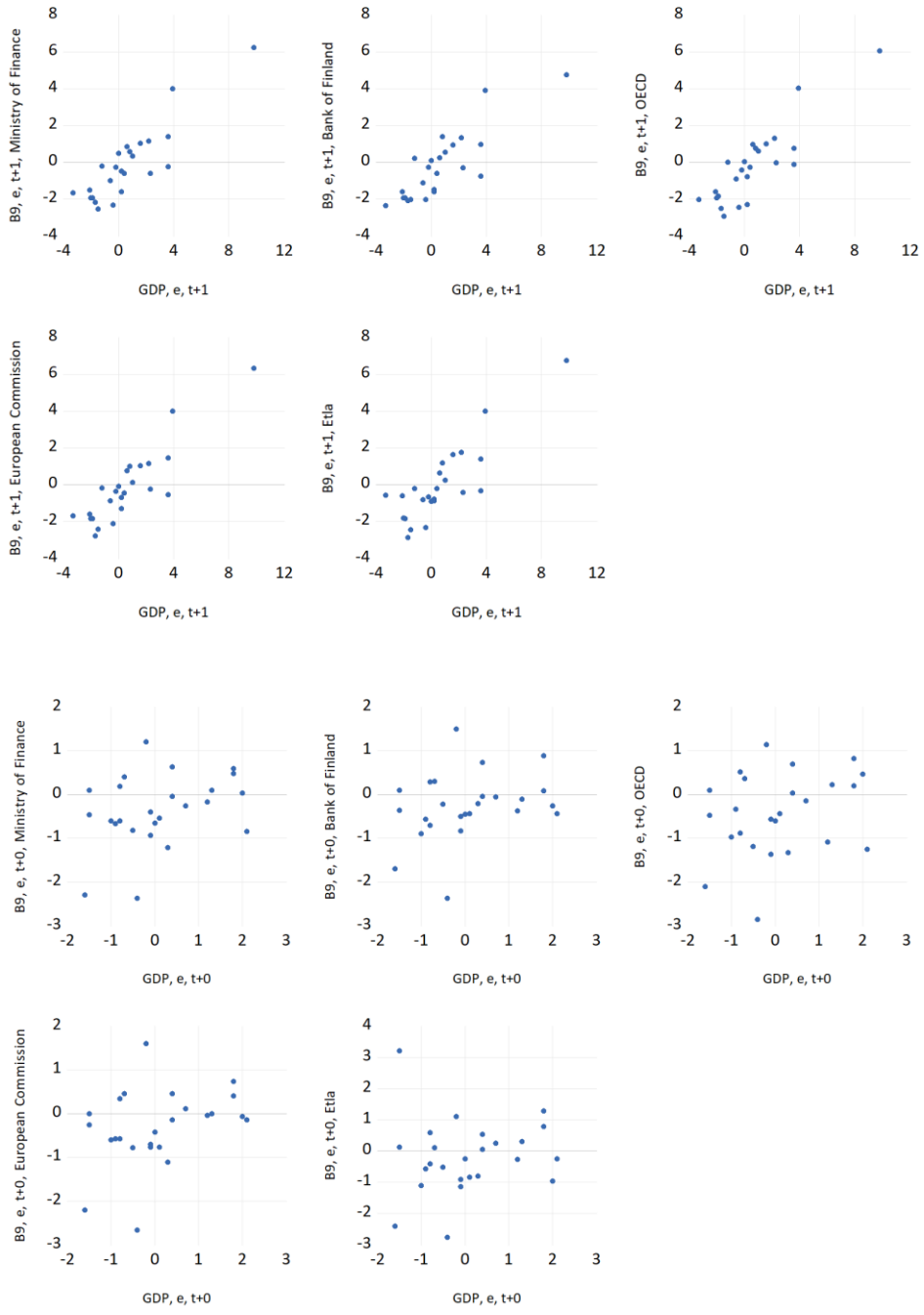

Figure 5: Scatter plots of the general government fiscal balance forecast error (vertical axis) and the GDP growth forecast error (horizontal axis).

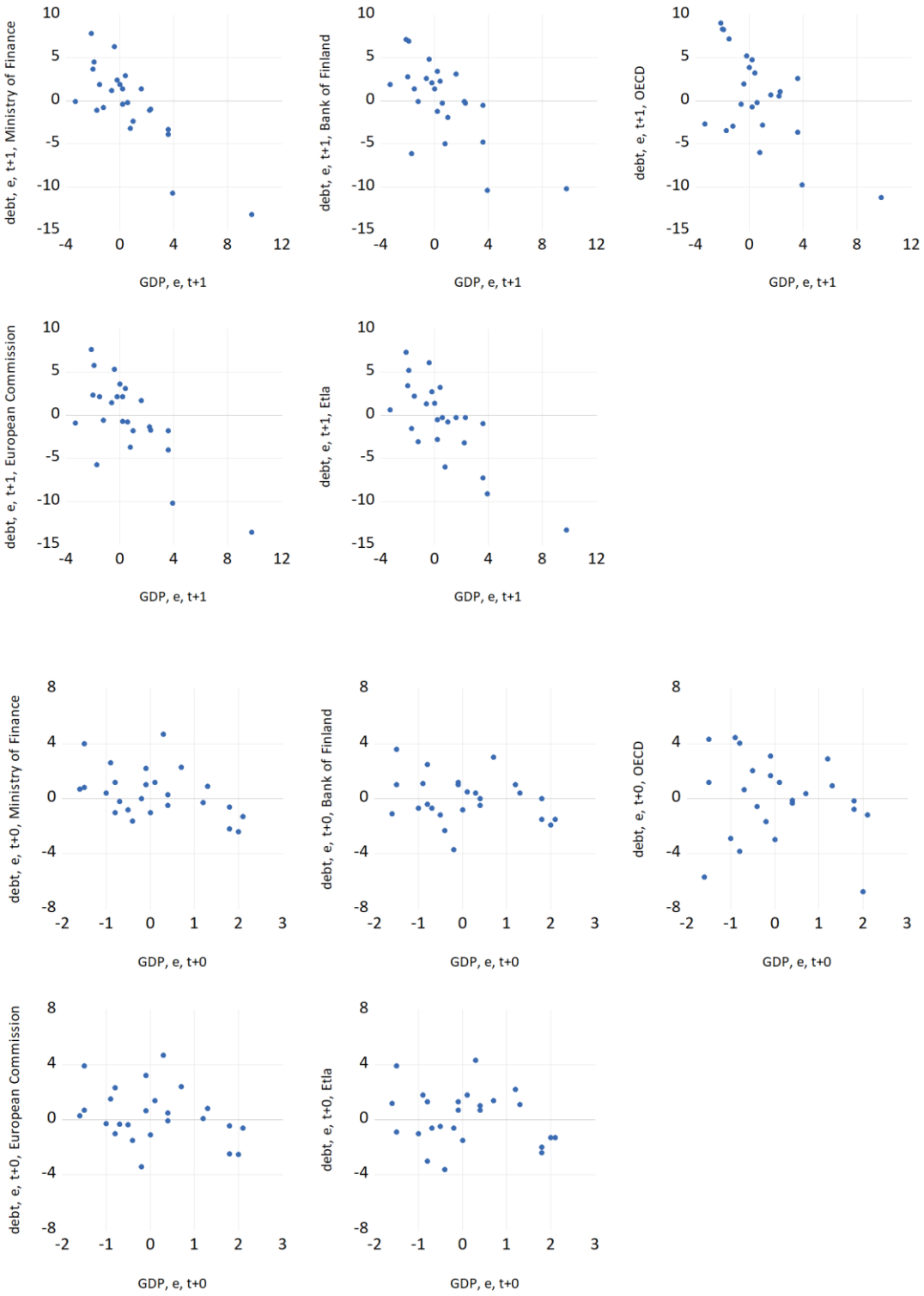

Figure 6: Scatter plots of the general government debt forecast error (vertical axis) and the GDP growth forecast error (horizontal axis).

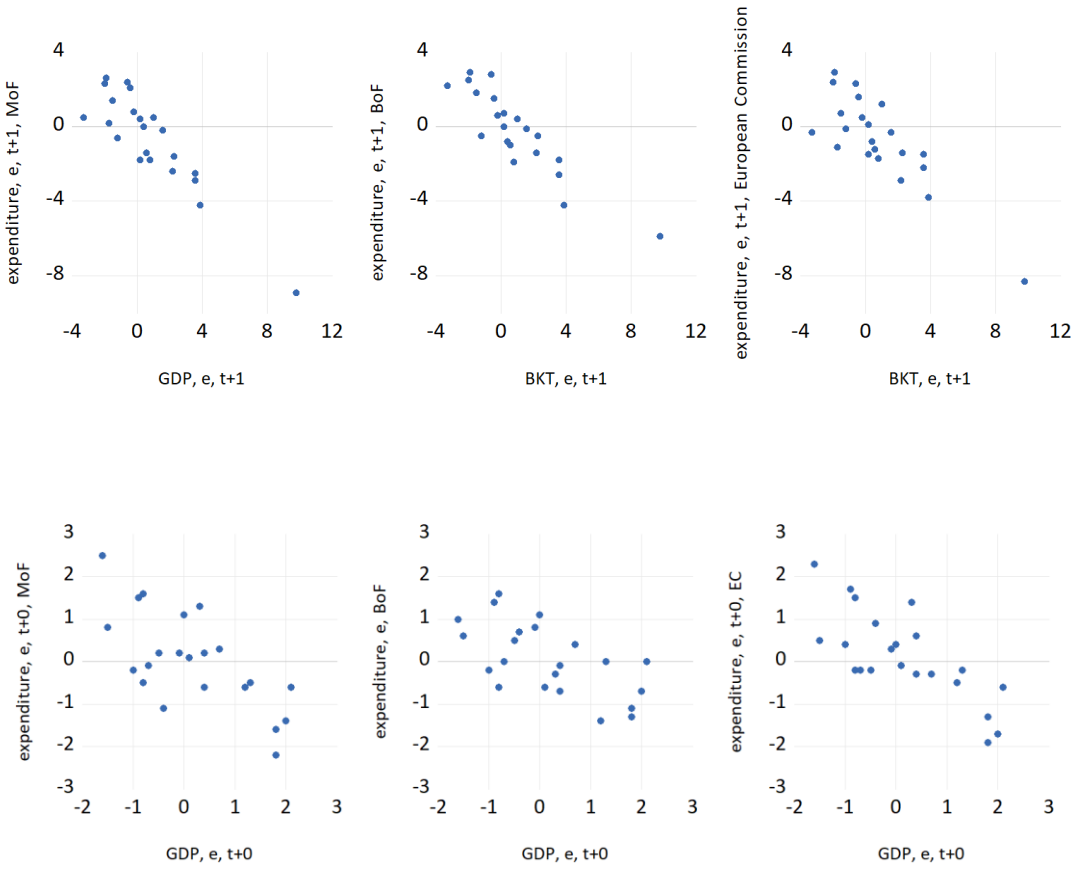

Figure 7: Scatter plots of the general government expenditure forecast error (vertical axis) and the GDP growth forecast error (horizontal axis).

Appendix 2: Forecasts and outturns

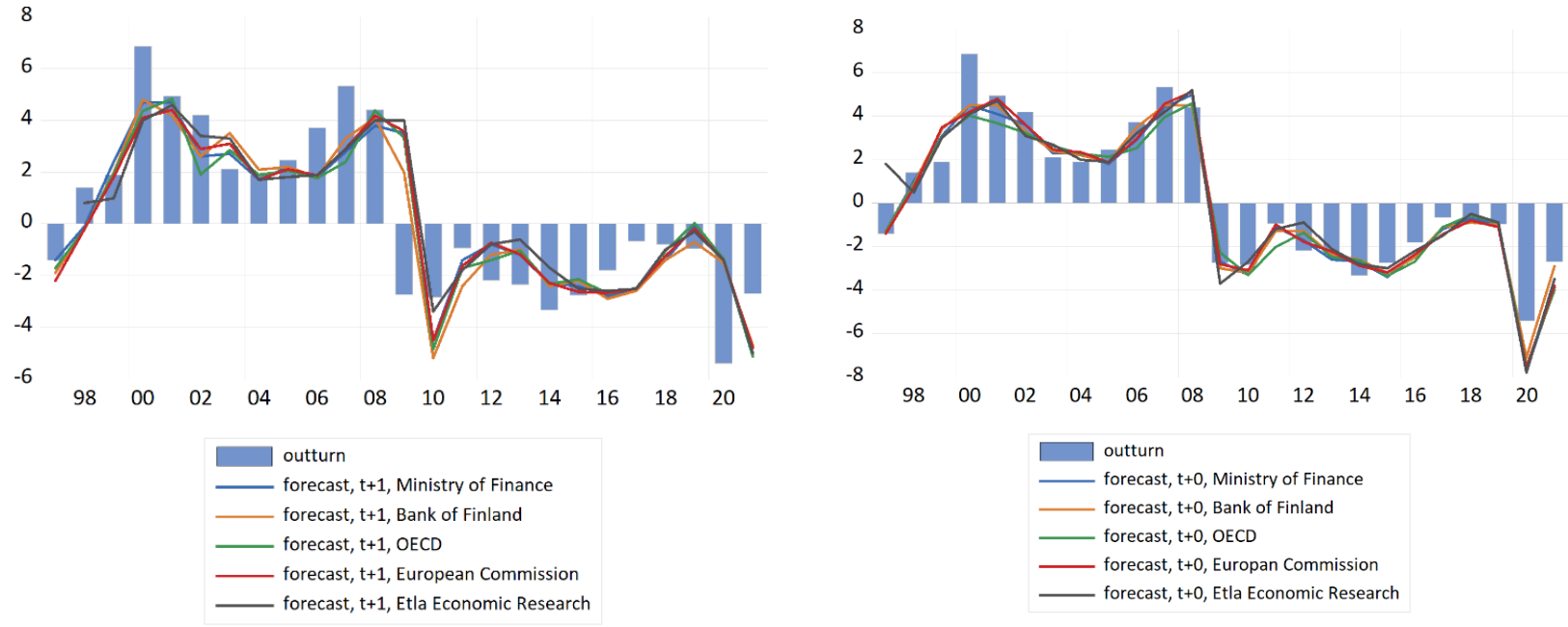

Figure 8: General government fiscal balance, % in relation to GDP. The outturns and forecasts for the year ahead (left panel) and for the current year (right panel).

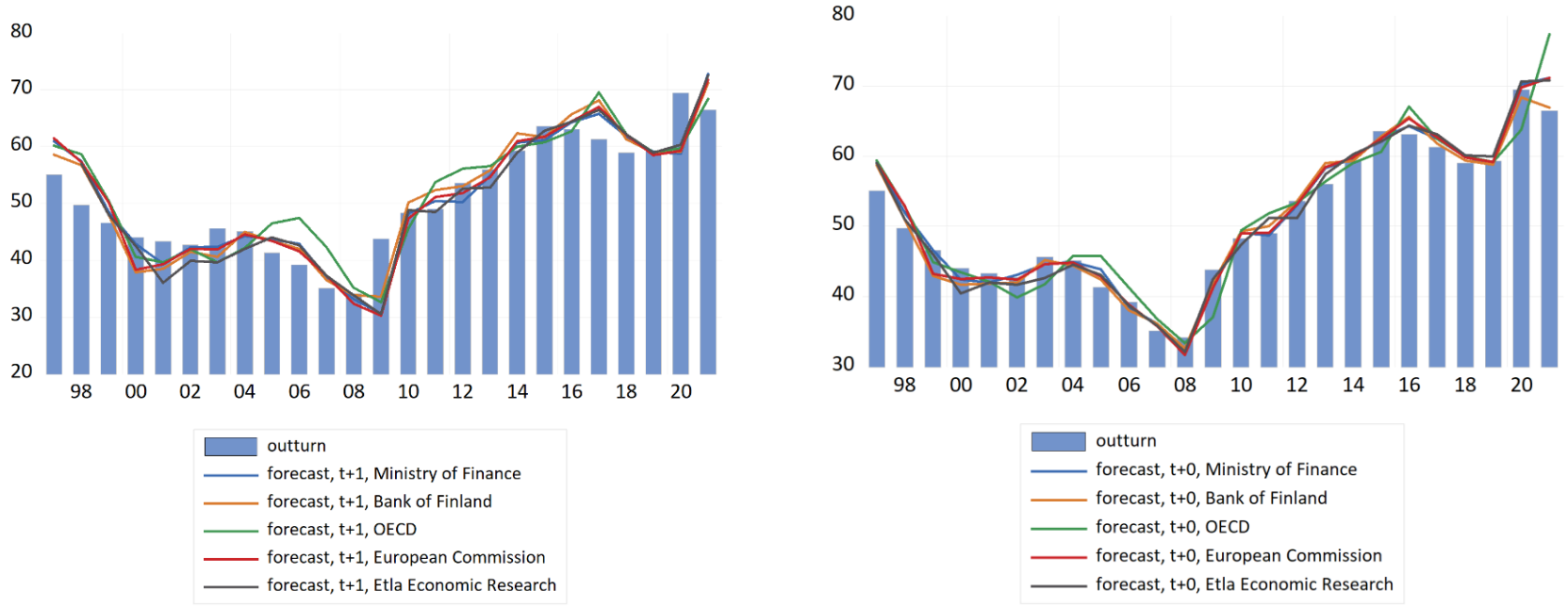

Figure 9: General government debt, % in relation to GDP. The outturns and forecasts for the year ahead (left panel) and for the current year (right panel).

Figure 10: General government expenditure, % in relation to GDP. The outturns and forecasts for the year ahead (left panel) and for the current year (right panel).

Appendix 3. Forecast error distributions

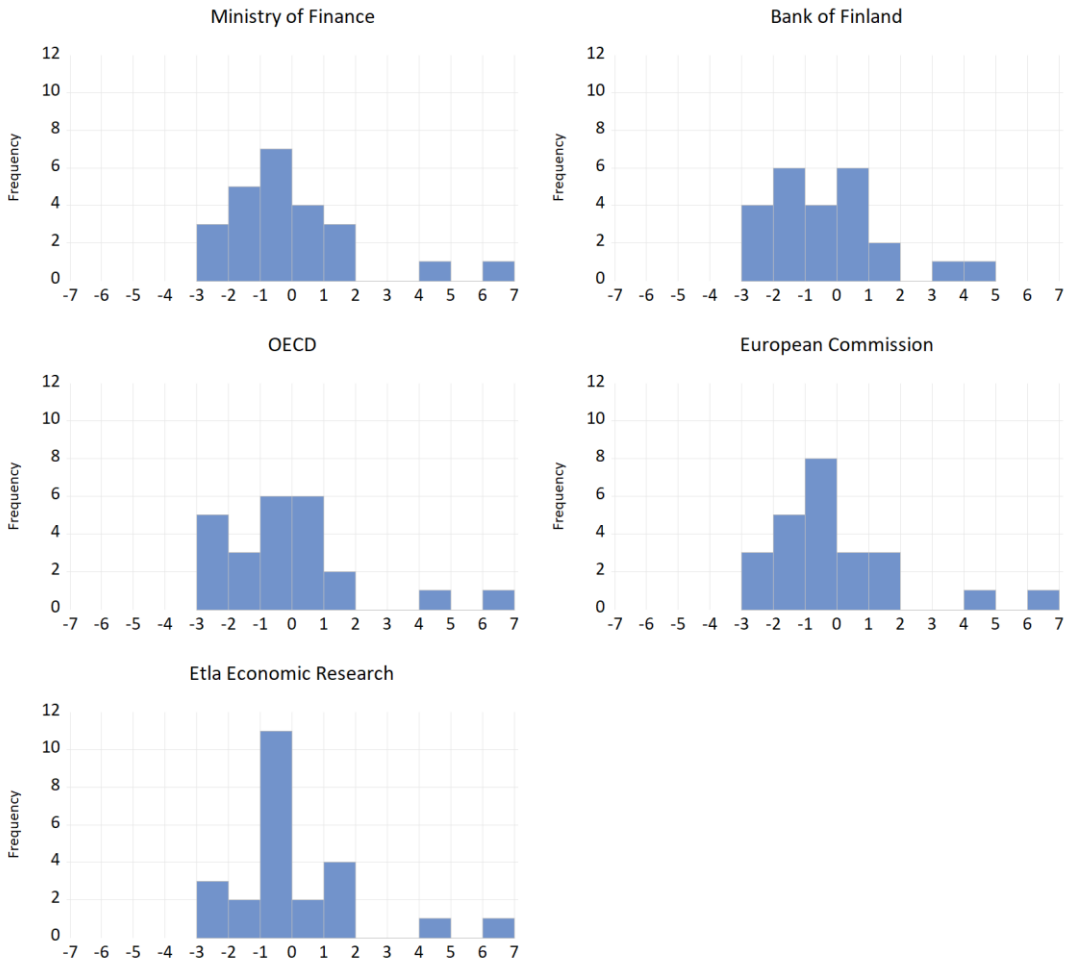

Figure 11: Distributions of forecast errors. General government fiscal balance, % in relation to GDP. Horizontal axes show the percentage point values. Forecasts for the year ahead (*t+1*).

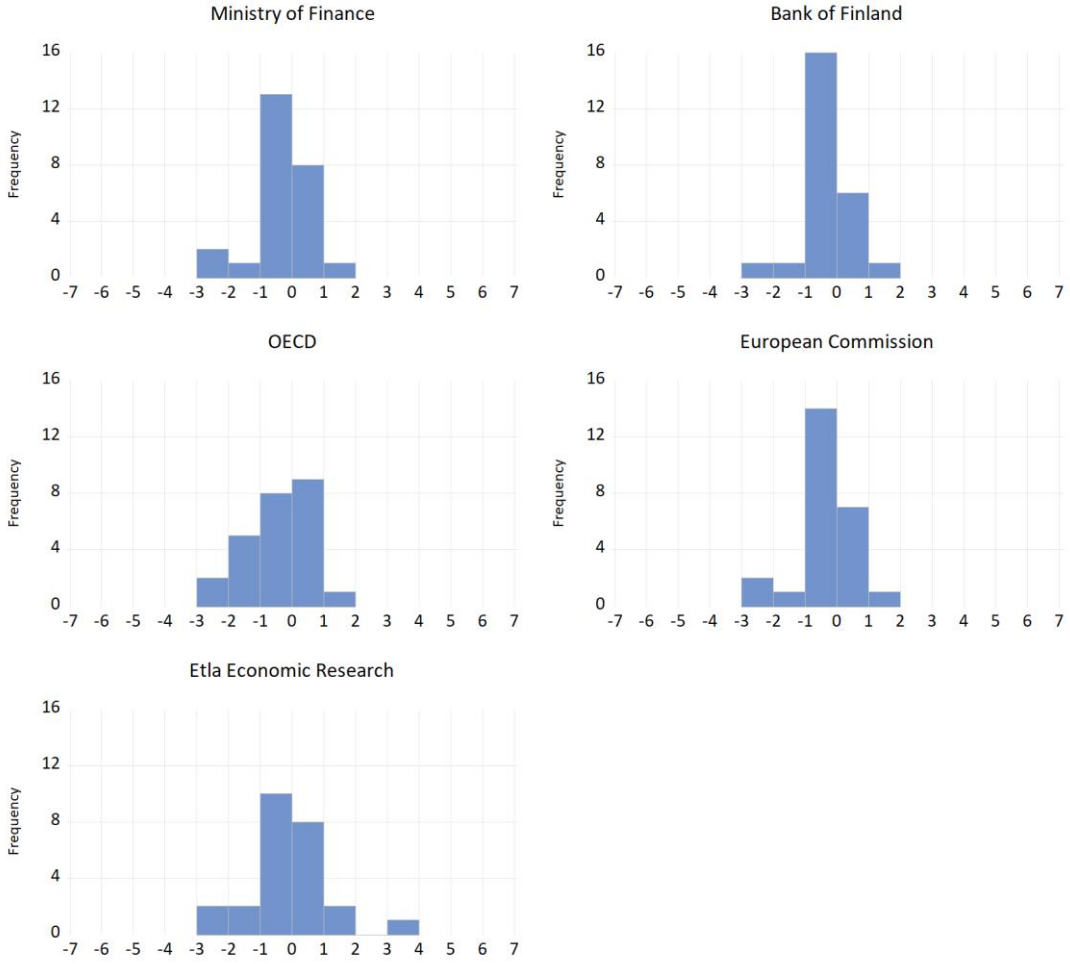

Figure 12: Distributions of forecast errors. General government fiscal balance, % in relation to GDP. Horizontal axes show the percentage point values. Forecasts for the current year (*t+0*).

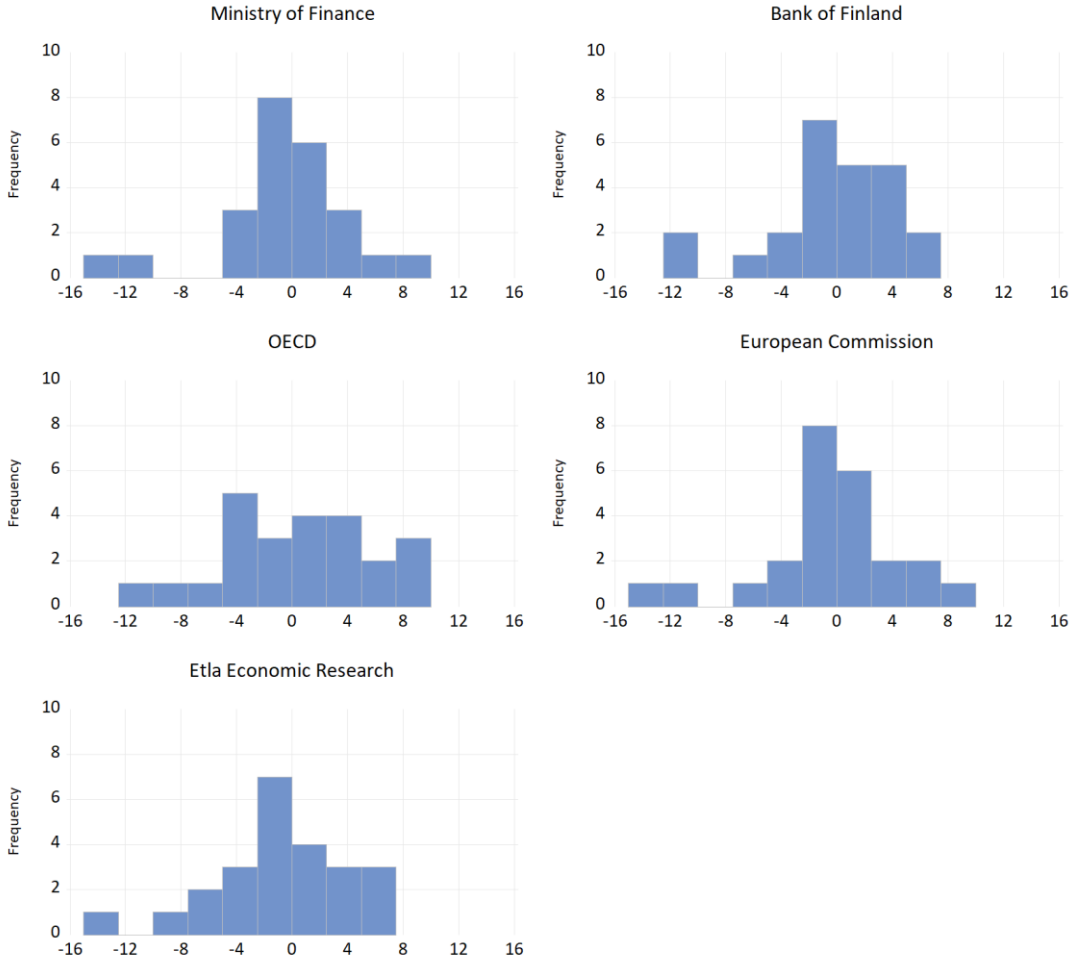

Figure 13: Distributions of forecast errors. General government debt, % in relation to GDP. Horizontal axes show the percentage point values. Forecasts for the year ahead (*t+1*).

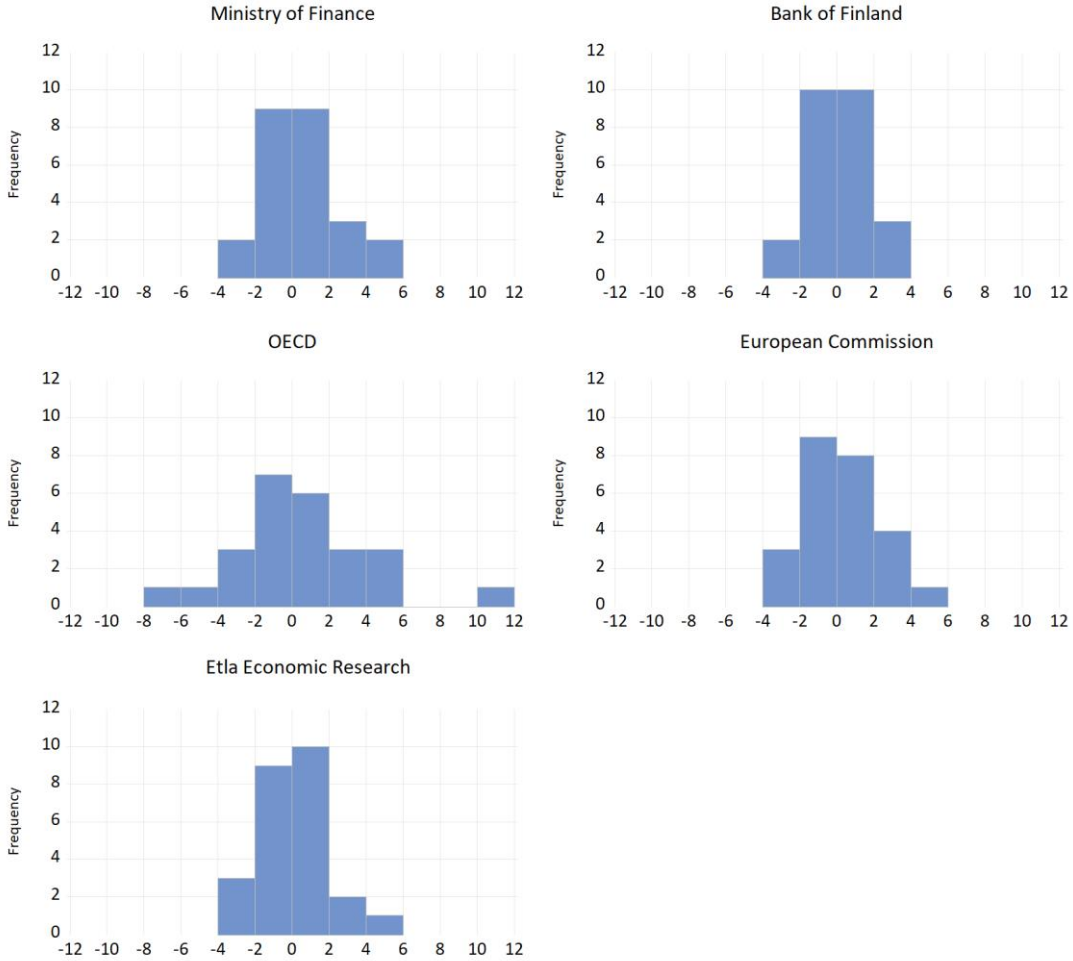

Figure 14: Distributions of forecast errors. General government debt, % in relation to GDP. Forecasts for the current year (*t+0*).

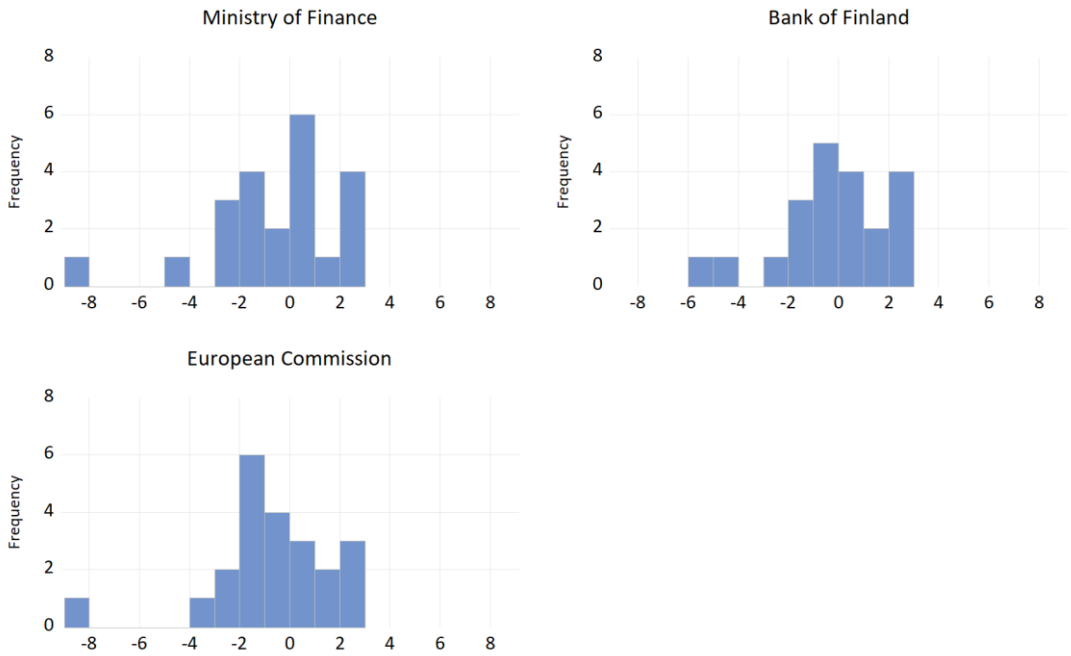

Figure 15: Distributions of forecast errors. General government expenditure, % in relation to GDP. Forecasts for the year ahead (*t+1*).

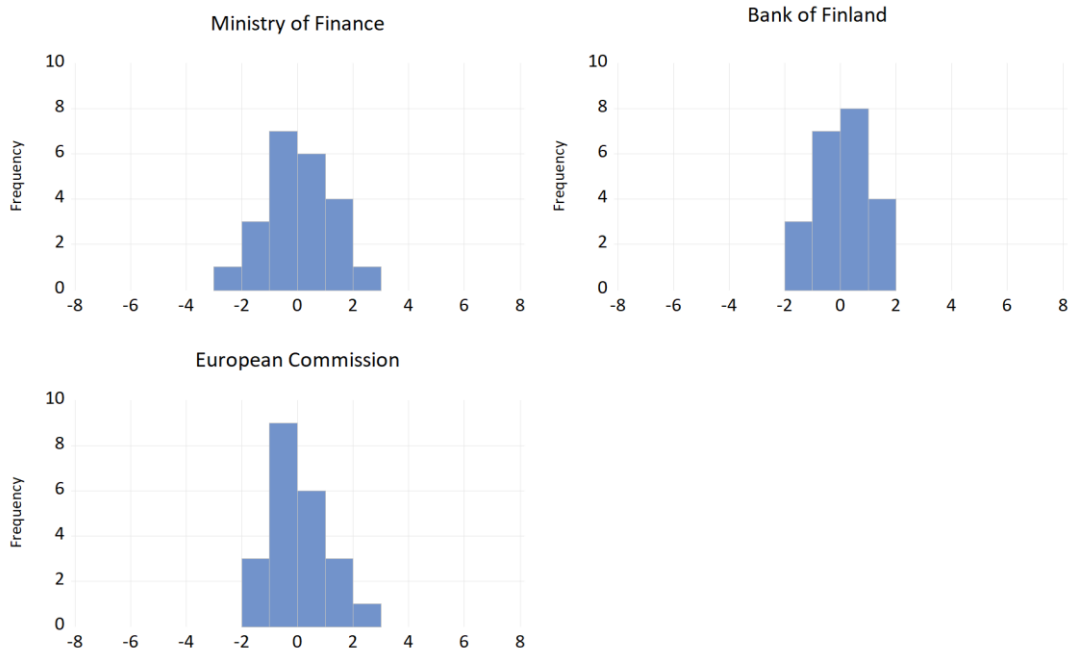

Figure 16: Distributions of forecast errors. General government expenditure, % in relation to GDP. Forecasts for the current year (*t+0*).

Appendix 4. Tables 4–9.

Table 4: Key figures on errors in general government fiscal balance forecasts, *t+1*.

Full sample:	**Naïve**	**MoF**	**Etla**	**BoF**	**EC**	**OECD**
Mean error (ME)	0.146	-0.121	0.000	-0.234	-0.130	-0.184
confidence interval (lower limit; upper limit)	-6.931; 7.222, H	-0.954; 0.711	-0.865; 0.865	-0.990; 0.523	-0.980; 0.710	-1.039; 0.671
Median	-0.500	-0.366	-0.487	-0.450	-0.413	-0.164
Mean squared error (MSE)	10.753	3.903	4.197	3.262	3.974	4.132
Root mean squared error (RMSE)	3.279	1.976	2.049	1.806	1.993	2.033
Mean absolute error (MAE)	2.571	1.463**	1.467**	1.434**	1.452**	1.483**
confidence interval (lower limit; upper limit)	-4.506; 9.647, H	0.630; 2.295	0.602; 2.332	0.677; 2.190	0.612; 2.292	0.628; 2.338
Mean absolute percentage error (MAPE)	111.433	63.521	61.618	64.539	62.295	62.729
Theil U2	1.000	0.746	0.778	0.707	0.747	0.772
e < 0	12	15	16	14	16	14
e >= 0	13	9	8	10	8	10
Observations	24	24	24	24	24	24
Reduced sample without outliers (2009, 2020):	**Naïve**	**MoF**	**Etla**	**BoF**	**EC**	**OECD**
Mean error (ME)	-0.414	-0.596**	-0.487	-0.646**	-0.609**	-0.658**
confidence interval (lower limit; upper limit)	-1.643; 0.815	-1.126; -0.066	-1.043; 0.070	-1.179; -0.113	-1.142; -0.077	-1.234; -0.082
Median	-0.670	-0.536	-0.583	-0.673	-0.494	-0.338
Mean squared error (MSE)	7.860	1.783	1.811	1.863	1.813	2.119
Root mean squared error (RMSE)	2.804	1.335	1.346	1.365	1.347	1.456
Mean absolute error (MAE)	2.231**	1.132**	1.114**	1.173**	1.116**	1.161**
confidence interval	1.002; 3.461	0.602; 1.662	0.558; 1.67	0.640; 1.706	0.584; 1.649	0.585; 1.736
Mean absolute percentage error (MAPE)	104.223	55.491	52.573	59.211	54.015	54.895
Theil U2	1.000	0.619	0.685	0.567	0.620	0.671
e < 0	13	15	16	14	16	15
e >= 0	9	7	6	8	6	7
Observations	22	22	22	22	22	22

Notes: Forecasts for the period 1998–2021. Forecasts produced during the years 1997 and 2020. The confidence intervals calculated based on HAC standard errors are marked with H. MoF = Ministry of Finance; Etla = ETLA Economic Research; BoF = Bank of Finland; EC = European Commission; OECD = Organization for Economic Co-operation and Development. A different unit of measurement is used for the mean squared error (MSE) because forecast errors are squared when the MSE is calculated.

Table 5: Key figures on errors in general government debt forecasts, *t+1*.

Full sample:	**Naïve**	**MoF**	**Etla**	**BoF**	**EC**	**OECD**
Mean error (ME)	-0.954	-0.251	-0.672	-0.047	-0.402	0.534
confidence interval (lower limit; upper limit)	-14.379; 12.471, H	-2.168; 1.667	-2.648; 1.305	-1.930; 1.836	-2.380; 1.575	-10.742; 11.811, H
Median	-0.489	-0.150	-0.300	0.650	-0.642	0.642
Mean squared error (MSE)	41.215	20.682	22.360	19.887	22.090	27.611
Root mean squared error (RMSE)	6.420	4.548	4.729	4.459	4.700	5.255
Mean absolute error (MAE)	5.704	3.201**	3.455**	3.363**	3.507**	4.183
confidence interval (lower limit; upper limit)	-7.721; 19.129, H	1.283; 5.118	1.479; 5.432	1.48; 5.246	1.53; 5.484	-7.093; 15.46, H
Mean absolute percentage error (MAPE)	11.536	6.311	6.986	6.618	7.014	8.767
Theil U2	1.000	0.701	0.740	0.652	0.727	0.792
e < 0	12	13	14	12	13	11
e >= 0	12	11	10	12	11	13
Observations	24	24	24	24	24	24
Reduced sample without outliers (2009, 2020):	**Naïve**	**MoF**	**Etla**	**BoF**	**EC**	**OECD**
Mean error (ME)	0.167	0.551	0.285	0.885	0.642	1.536
confidence interval (lower limit; upper limit)	-12.575; 12.908, H	-0.678; 1.781	-6.915; 7.485, H	-0.603; 2.373	-6.451; 7.734, H	-7.564; 10.635, H
Median	1.661	0.550	-0.300	1.400	0.420	0.894
Mean squared error (MSE)	35.780	7.998	12.588	12.049	10.997	20.086
Root mean squared error (RMSE)	5.982	2.828	3.548	3.471	3.316	4.482
Mean absolute error (MAE)	5.262	2.220**	2.751	2.733**	2.745	3.611
confidence interval (lower limit; upper limit)	-7.480; 18.004, H	0.990; 3.450	-4.449; 9.951, H	1.245; 4.221	-4.347; 9.838, H	-5.489; 12.711, H
Mean absolute percentage error (MAPE)	11.003	4.594	5.646	5.481	5.576	7.763
Theil U2	1.000	0.394	0.514	0.489	0.458	0.673
e < 0	10	11	12	10	11	9
e >= 0	12	11	10	12	11	13
Observations	22	22	22	22	22	22

Notes: Forecasts for the period 1998–2021. Prepared between 1997 and 2020. The confidence intervals calculated based on HAC standard errors are marked with H. MoF = Ministry of Finance; Etla = ETLA Economic Research; BoF = Bank of Finland; EC = European Commission; OECD = Organisation for Economic Co-operation and Development. A different unit of measurement is used for the mean squared error (MSE) because forecast errors are squared when the MSE is calculated.

Table 6: Key figures on errors in expenditure forecasts, *t+1*.

Full sample:	**Naïve**	**MoF**	**BoF**	**EC**
Mean error (ME)	-0.686	-0.729	-0.238	-0.681
confidence interval (lower limit; upper limit)	-7.325; 5.954, H	-1.986; 0.529	-1.303; 0.827	-1.846; 0.484
Median	-0.200	-0.200	-0.100	-0.300
Mean squared error (MSE)	9.621	7.340	4.938	6.305
Root mean squared error (RMSE)	3.102	2.709	2.222	2.511
Mean absolute error (MAE)	2.362	1.967**	1.724**	1.795**
confidence interval (lower limit; upper limit)	-4.278; 9.001, H	0.709; 3.224	0.659; 2.789	0.630; 2.960
Mean absolute percentage error (MAPE)	4.364	3.662	3.200	3.316
Theil U2	1.000	0.880	0.690	0.797
e < 0	11	11	11	13
e >= 0	10	10	10	8
Observations	21	21	21	21
Reduced sample without outliers (2009, 2020):	**Naïve**	**MoF**	**BoF**	**EC**
Mean error (ME)	-0.079	-0.116	0.268	-0.116
confidence interval (lower limit; upper limit)	-6.498; 6.34, H	-3.801; 3.569, H	-0.537; 1.074	-3.506; 3.274, H
Median	0.100	0.000	0.000	-0.300
Mean squared error (MSE)	5.722	3.016	2.697	2.583
Root mean squared error (RMSE)	2.392	1.737	1.642	1.607
Mean absolute error (MAE)	1.932	1.484	1.374**	1.347
confidence interval (lower limit; upper limit)	-4.487; 8.350, H	-2.201; 5.169, H	0.568; 2.179	-2.042; 4.737, H
Mean absolute percentage error (MAPE)	3.620	2.825	2.596	2.536
Theil U2	1.000	0.715	0.643	0.627
e < 0	9	9	9	11
e >= 0	10	10	10	8
Observations	19	19	19	19

Notes: Forecasts for the period 2001–2021. Prepared between 2000 and 2020. The confidence intervals calculated based on HAC standard errors are marked with H. MoF = Ministry of Finance; Etla = ETLA Economic Research; BoF = Bank of Finland; EC = European Commission; OECD = Organization for Economic Co-operation and Development. A different unit of measurement is used for the mean squared error (MSE) because forecast errors are squared when the MSE is calculated.

Table 7: Key figures on errors in general government fiscal balance forecasts, *t+0*.

Full sample:	**Naïve**	**MoF**	**Etla**	**BoF**	**EC**	**OECD**
Mean error (ME)	-0.020	-0.360**	-0.192	-0.264	-0.308	-0.440**
confidence interval (lower limit; upper limit)	-0.984; 0.944	-0.693; -0.027	-2.67; 2.285, H	-0.575; 0.047	-0.661; 0.044	-0.823; -0.056
Median	0.096	-0.400	-0.250	-0.262	-0.148	-0.441
Mean squared error (MSE)	5.455	0.781	1.385	0.638	0.825	1.058
Root mean squared error (RMSE)	2.336	0.884	1.177	0.799	0.908	1.028
Mean absolute error (MAE)	1.636**	0.656**	0.856	0.576**	0.633**	0.806**
confidence interval (lower limit; upper limit)	0.672; 2.600	0.323; 0.989	-1.621; 3.334, H	0.265; 0.887	0.28; 0.985	0.422; 1.19
Mean absolute percentage error (MAPE)	68.758	23.525	38.175	22.454	24.176	31.704
Theil U2	1.000	0.480	0.585	0.438	0.495	0.503
e < 0	12	15	14	17	16	15
e >= 0	13	10	11	8	9	10
Observations	25	25	25	25	25	25
Reduced sample without outliers (2009, 2020):	**Naïve**	**MoF**	**Etla**	**BoF**	**EC**	**OECD**
Mean error (ME)	-0.522	-0.292	-0.061	-0.200	-0.235	-0.405
confidence interval (lower limit; upper limit)	-3.891; 2.847, H	-1.818; 1.235, H	-2.882; 2.760, H	-1.764; 1.363, H	-1.906; 1.436, H	-2.277; 1.468, H
Median	-0.202	-0.400	-0.241	-0.222	-0.148	-0.441
Mean squared error (MSE)	2.896	0.619	1.211	0.564	0.686	0.949
Root mean squared error (RMSE)	1.702	0.787	1.100	0.751	0.828	0.974
Mean absolute error (MAE)	1.279	0.613	0.783	0.539	0.588	0.766
confidence interval	-2.09; 4.648, H	-0.913; 2.14, H	-2.038; 3.604, H	-1.024; 2.103, H	-1.084; 2.259, H	-1.106; 2.639, H
Mean absolute percentage error (MAPE)	59.761	23.719	37.952	22.555	24.346	32.081
Theil U2	1.000	0.464	0.670	0.525	0.529	0.570
e < 0	12	14	12	15	14	14
e >= 0	11	9	11	8	9	9
Observations	23	23	23	23	23	23

Notes: Forecasts for the period 1998–2021. Forecasts produced during the years 1997 and 2020. The confidence intervals calculated based on HAC standard errors are marked with H. MoF = Ministry of Finance; Etla = ETLA Economic Research; BoF = Bank of Finland; EC = European Commission; OECD = Organisation for Economic Co-operation and Development. A different unit of measurement is used for the mean squared error (MSE) because forecast errors are squared when theMSE is calculated.

Table 8: Key figures on errors in general government debt forecasts, *t+0*.

Full sample:	**Naïve**	**MoF**	**Etla**	**BoF**	**EC**	**OECD**
Mean error (ME)	-0.340	0.415	0.159	-0.025	0.330	0.425
confidence interval	-1.971; 1.291	-2.803; 3.633, H	-0.644; 0.962	-3.524; 3.474, H	-0.460; 1.119	-5.472; 6.322, H
Median	0.600	0.300	0.700	0.000	0.062	0.372
Mean squared error (MSE)	15.723	3.174	3.811	2.652	3.769	12.691
Root mean squared error (RMSE)	3.965	1.782	1.952	1.629	1.941	3.563
Mean absolute error (MAE)	3.092**	1.369	1.657**	1.281	1.463**	2.587
confidence interval	1.461; 4.723	-1.849; 4.587, H	0.854; 2.460	-1.900; 4.462, H	0.673; 2.253	-3.31; 8.484, H
Mean absolute percentage error (MAPE)	6.150	2.758	3.301	2.649	2.918	5.040
Theil U2	1.000	0.401	0.441	0.380	0.433	0.801
e < 0	10	11	12	12	12	12
e >= 0	15	14	13	13	13	13
Observations	25	25	25	25	25	25
Reduced sample without outliers (2009, 2020):	**Naïve**	**MoF**	**Etla**	**BoF**	**EC**	**OECD**
Mean error (ME)	0.496	0.525	0.220	0.103	0.455	1.004
confidence interval	-5.658; 6.649, H	-0.212; 1.262	-0.629; 1.069	-3.435; 3.642, H	-6.618; 7.528, H	-3.832; 5.84, H
Median	0.661	0.150	0.100	0.000	-0.019	0.512
Mean squared error (MSE)	8.476	3.178	3.900	2.673	3.814	10.399
Root mean squared error (RMSE)	2.911	1.783	1.975	1.635	1.953	3.225
Mean absolute error (MAE)	2.496	1.353**	1.672**	1.262**	1.467**	2.270**
confidence interval	-3.658; 8.649, H	0.616; 2.090	0.823; 2.520	0.556; 1.967	0.646; 2.288	0.945; 3.595
Mean absolute percentage error (MAPE)	5.084	2.715	3.315	2.621	2.901	4.451
Theil U2	1.000	0.565	0.569	0.548	0.614	0.960
e < 0	8	10	11	10	11	10
e >= 0	15	13	12	13	12	13
Observations	23	23	23	23	23	23

Notes: Forecasts for the period 1998–2021. Prepared between 1997 and 2020. The confidence intervals calculated based on HAC standard errors are marked with H. MoF = Ministry of Finance; Etla = ETLA Economic Research; BoF = Bank of Finland; EC = European Commission; OECD = Organisation for Economic Co-operation and Development. A different unit of measurement is used for the mean squared error (MSE) because forecast errors are squared when the MSE is calculated.

Table 9: Key figures on errors in expenditure forecasts, *t+0*.

Full sample:	**Naïve**	**MoF**	**BoF**	**EC**
Mean error (ME)	-0.264	0.200	0.473	0.114
confidence interval	-1.246; 0.719	-0.545; 0.945	-0.463; 1.408	-0.38; 0.607
Median	0.250	0.000	0.000	-0.150
Mean squared error (MSE)	4.226	2.428	3.990	1.062
Root mean squared error (RMSE)	2.056	1.558	1.997	1.031
Mean absolute error (MAE)	1.445**	1.055**	1.064**	0.795**
confidence interval	0.463; 2.428	0.310; 1.799	0.128; 1.999	0.302; 1.289
Mean absolute percentage error (MAPE)	2.681	1.956	1.971	1.487
Theil U2	1.000	0.732	0.831	0.468
e < 0	10	11	10	12
e >= 0	12	11	12	10
Observations	22	22	22	22
Reduced sample without outliers (2009, 2020):	**Naïve**	**MoF**	**BoF**	**EC**
Mean error (ME)	0.260	0.165	0.505	0.095
confidence interval	-0.296; 0.816	-0.555; 0.885	-0.466; 1.476	-0.323; 0.513
Median	0.550	0.000	0.000	-0.150
Mean squared error (MSE)	1.399	2.261	4.315	0.759
Root mean squared error (RMSE)	1.183	1.503	2.077	0.871
Mean absolute error (MAE)	1.040**	0.965**	1.085**	0.675**
confidence interval	0.484; 1.596	0.245; 1.685	0.114; 2.056	0.257; 1.093
Mean absolute percentage error (MAPE)	1.975	1.809	2.018	1.283
Theil U2	1.000	1.279	1.400	0.645
e < 0	8	10	7	11
e >= 0	12	10	13	9
Observations	20	20	20	20

Notes: Forecasts for the period 2001–2021. Prepared between 2000 and 2020. The confidence intervals calculated based on HAC standard errors are marked with H. MoF = Ministry of Finance; Etla = ETLA Economic Research; BoF = Bank of Finland; EC = European Commission; OECD = Organisation for Economic Co-operation and Development. A different unit of measurement is used for the mean squared error (MSE) because forecast errors are squared when the MSE is calculated.

References

An, Z., Jalles, J. T., Loungani, P., & Sousa, R. M. (2018). Do IMF Fiscal Forecasts Add Value?. *Journal of Forecasting*, *37*(6), 650–665.

Auerbach, A. J. (1995). Tax projections and the budget: lessons from the 1980's. *American Economic Review, 85*, 165–9.

Auerbach, A. J. (1999). On the performance and use of government revenue forecasts. *National Tax Journal*, *52*, 765–82.

Bank of Finland (2021). Bank of Finland's institutional repository. Retrieved on 25 August 2021 from https://publications.bof.fi/

Batchelor, R. (2001). How Useful Are the Forecasts of Intergovernmental Agencies? The IMF and OECD versus the Consensus. *Applied Economics*, *33*(2), 225–235.

Beaudry, P. & Willems, T. (2022). On the Macroeconomic Consequences of Over-Optimism. *American Economic Journal: Macroeconomics, 14*(1), 38–59.

Beetsma, R., Busse, M. Germinetti, L., Giuliodori, M. & Larch, M. (2023). Is the road to hell paved with good intentions? An empirical analysis of budgetary follow-up in the EU. *Journal of International Money and Finance, 135*, 1–16. https://doi.org/10.1016/j.jimonfin.2023.102854

Beetsma, R., Debrun, X., Fang, X., Kim, Y., Lledó, V., Mbaye, S. & Zhang, X. (2019). Independent fiscal councils: Recent trends and performance. *European Journal of Political Economy*, *57*, 53-69.

Boot, T. & Pick, A. (2020). Does modeling a structural break improve forecast accuracy?. *Journal of Econometrics*, *215*(1), 35–59. https://doi.org/10.1016/j.jeconom.2019.07.007

Council Directive 2011/85/EU of 8 November 2011 on requirements for budgetary frameworks of the Member States.

Cronin, D. & McInerney, N. (2023). Official fiscal forecasts in EU member states under the European Semester and Fiscal Compact – An empirical assessment. *European Journal of Political Economy, 76*, art. no. 102227. https://doi.org/10.1016/j.ejpoleco.2022.102227

Cronin, D. & McQuinn, K. (2021). Fiscal policy and growth forecasts in the EU: are official forecasters still misestimating fiscal multipliers?. *Review of World Economics, 157*(3), 453–462. https://doi.org/10.1007/s10290-021-00428-z

Cronin, D. & McQuinn, K. (2023). Government debt forecast errors and the net expenditure rule in EU countries: Undue optimism at a cost. *Journal of Policy Modeling, 45*(6), 1113–1131. https://doi.org/10.1007/s10290-021-00428-z

Croushore, D. & van Norden, S. (2018). Fiscal forecasts at the FOMC: Evidence from the Greenbooks. *Review of Economics and Statistics, 100*(5), 933–945. https://doi.org/10.1162/rest_a_00664

Duncan, R. & Martinez-Garcia, E. (2019). New perspectives on forecasting inflation in emerging economies: An empirical assessment. *International Journal of Forecasting, 35*(3), 1008–1031.

Eicher, T. S., Kuenzel, D. J., Papageorgiou, C., & Christofides, C. (2019). Forecasts in Times of Crises. *International Journal of Forecasting, 35*(3), 1143–1159.

Estefania-Flores, J., Furceri, D., Kothari, S. & Ostry, J.D. (2023). Worse than you think: Public debt forecast errors in advanced and developing economies. *Journal of Forecasting, 42*(3), 685–714.

European Commission (2000). Commission Regulation (EC) No 1500/2000, of 10 July 2000, implementing Council Regulation (EC) No 2223/96 with respect to general government expenditure and revenue. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32000R1500:EN:PDF

European Commission, Eurostat (2014). European system of accounts : ESA 2010, Publications Office. https://data.europa.eu/doi/10.2785/16644

European Commission (2022a). Economic forecasts. Retrieved on 28 April 2022 from https://economy-finance.ec.europa.eu/economic-forecast-and-surveys/economic-forecasts_en

European Commission (2022b). Index of Economic Forecasts. Retrieved on 28 April 2022 from https://ec.europa.eu/economy_finance/publications/european_economy/forecasts/index_en.htm

ETLA Economic Research (Etla) (2021). Etla's database. Retrieved on 10 May 2021 from http://etladb.etla.fi/

Fioramanti, M., González Cabanillas, L., Roelstraete, B. & Ferrandis Vallterra, S. A. (2016). European Commission's Forecasts Accuracy Revisited: Statistical Properties and Possible Causes of Forecast Errors. Discussion paper 027. https://doi.org/10.2765/283693

Fortin, I., Koch, S. P., & Weyerstrass, K. (2020). Evaluation of economic forecasts for Austria. *Empirical Economics, 58*(1), 107–137. https://doi.org/10.1007/s00181-019-01814-1

Frankel, J. (2011). Over-optimism in forecasts by official budget agencies and its implications. *Oxford Review of Economic Policy*, *27*(4), 536–562. https://doi.org/10.1093/oxrep/grr025

Frankel, J., & Schreger, J. (2013). Over-optimistic official forecasts and fiscal rules in the eurozone. *Review of World Economics, 149*(2), 247–272.

Gatti, R., Lederman, D., Islam, A.M., Nguyen, H., Lotfi, R. & Emam Mousa, M. (2024). Data transparency and GDP growth forecast errors. *Journal of International Money and Finance*, *140,* art. no. 102991. https://doi.org/10.1016/j.jimonfin.2023.102991

Gilbert, N, D. & de Jong, J. F. M (2017). Do European fiscal rules induce a bias in fiscal forecasts? Evidence from the Stability and Growth Pact. *Public Choice*, *170*, 1–32. https://doi.org/10.1007/s11127-016-0372-1

Giovannelli, A. & Pericoli, F.M. (2020). Are GDP forecasts optimal? Evidence on European countries. *International Journal of Forecasting*, *36*(3), 963–973.

Giuriato, L., Cepparulo, A., & Barberi, M. (2016). Fiscal forecasts and political systems: a legislative budgeting perspective. *Public Choice, 168*(1/2), 1–22. https://doi-org.pc124152.oulu.fi:9443/10.1007/s11127-016-0345-4

Hadzi-Vaskov, M., Ricci, L.A., Werner, A.M. & Zamarripa, R. (2023). Authorities' Fiscal Forecasts in Latin America: Are they Optimistic?. *Economia LACEA Journal, 22*(1), 135–152. https://doi.org/10.31389/eco.416

IMF (2015). *Finland: Fiscal Transparency Evaluation*. Country Report No. 2015/060.

Jalles, J., T., Karibzhanov, I. & Loungani, P. (2015). Cross-country evidence on the quality of private sector fiscal forecasts. *Journal of Macroeconomics, 45*, 186–201. https://doi.org/10.1016/j.jmacro.2015.04.009

Jiao, X., Song, Y., Kong, Y. & Tang, X. (2022). Volatility forecasting for crude oil base on text information and deep learning PSO-LSTM model. *Journal of Forecasting, 41*(5), 933–944.

Jonung, L., & Larch, M. (2006). Improving fiscal policy in the EU: the case for independent forecasts. *Economic Policy*, *21*(47), 491–534.

Keereman, F. (1999). The track record of the Commission Forecasts. *European Economy, Economic Papers*, no. 137.

Kellokumpu, J., Pesola, S., Savolainen, L., Grünbaum, J. (2023). Automatic Fiscal Stabilizers in Finland 1993–2021. Working paper. Fiscal Policy Monitoring Function. National Audit Office of Finland.

Krol, R. (2014). Forecast bias of government agencies. *Cato Journal,* 34(1), 99–112.

Lahtinen, M., Mäki-Fränti, P., Määttä, K. & Volk, R. (2009). Valtion talousarvioiden verotuloennusteiden osuvuus. Eduskunnan tarkastusvaliokunnan julkaisu, 1/2009.

Leal, T., Perez, J. J., Tujula, M. & Vidal, J. P. (2008). Fiscal Forecasting: Lessons from the Literature and Challenges. *Fiscal Studies, 29*(3), 347–386.

Merola, R. & Perez, J. J. (2013). Fiscal forecast errors: Governments versus independent agencies? *European Journal of Political Economy, 32*, 285–299. https://doi.org/10.1016/j.ejpoleco.2013.09.002

Ministry of Finance (2021). Economic Surveys. Retrieved on 13 May 2021 from https://vm.fi/en/economic-forecasts

Morikawa, M. (2022). Uncertainty in long-term macroeconomic forecasts: Ex post evaluation of forecasts by economics researchers. *Quarterly Review of Economics and Finance, 85*, 8–15.

Musso, A. & S. Phillips (2002), Comparing Projections and Outcomes of IMF-Supported Programs, IMF Staff Papers, Vol. 49, No. 1.

National Audit Office (2016). Reliability of macroeconomic forecasts. Fiscal policy audit report. Fiscal policy audit report of the National Audit Office, 11/2016.

National Audit Office (2018). Fiscal policy evaluation on the reliability of the Ministry of Finance macroeconomic forecasts. The short-term forecasts of GDP growth, unemployment rate and inflation for the years 1976-2016 under review. Fiscal policy monitoring report. 7/2018.

OECD (2022). OECD.Stat. Retrieved on 13 December 2022 from https://stats.oecd.org/

Reifschneider, D. & Tulip, P. (2019). Gauging the uncertainty of the economic outlook using historical forecasting errors: The Federal Reserve's approach. *International Journal of Forecasting*, *35*(4), 1564 –1582.

Rybacki, J. (2021). Fiscal deficit forecasts in Europe: evidence for a double standard?. *Baltic Journal of Economics*, *21*(1), 26–42. https://doi.org/10.1080/1406099X.2021.1888439

Strauch, R., Hallerberg, M., & von Hagen (2004). Budgetary Forecasts in Europe—The Track Record of Stability and and Convergence Programmes, ECB WP No. 30.

Theil (1966). Applied economic forecasting. North-Holland, Amsterdam.

von Hagen, J. (2010). Sticking to fiscal plans: the role of institutions. *Public Choice*, *144*(3/4), 487–503. http://www.jstor.org/stable/40835625

NATIONAL AUDIT OFFICE OF FINLAND (NAOF)
PORKKALANKATU 1, PO BOX 1119, FI-00101 HELSINKI

tel. +358 9 4321 | www.vtv.fi | @VTV_fi