Evaluating the Macroeconomic Forecasting Performance of the Ministry for Finance and Employment*

Working Paper 01/2024

Kurt Davison†
Gilmour Camilleri‡
Kylie Spiteri§

* The views expressed in the paper are solely those of the authors and do not necessarily reflect those of the MFAC. Any errors and omissions are the authors' own. Working Papers describe research in progress by the authors and are published to elicit comments and to further debate. The authors would like to thank Dr Moira Catania, Dr Stephanie Vella and Dr Stephanie Fabri for their comments and suggestions.
† Kurt Davison is a Senior Economist at the Malta Fiscal Advisory Council. E-mail: kurt.davison@mfac.org.mt.
‡ Gilmour Camilleri is the Chief Economist of the Malta Fiscal Advisory Council. He is also a Visiting Lecturer at the University of Malta. E-mail: gilmour.camilleri.2@mfac.org.mt.
§ Kylie Spiteri is an Economist at the Malta Fiscal Advisory Council. E-mail: kylie.spiteri@mfac.org.mt.

Abstract

This working paper evaluates the macroeconomic forecasting performance of the MFE for forecasts produced between 2004 and 2022 in the Update of the Stability/Convergence Programmes (spring vintage) and in the Draft Budgetary Plans (autumn vintage). The authors' assessment of this evaluation rests on three pillars: forecasting accuracy, unbiasedness, and forecasting efficiency. Inter-institutional forecasting performance evaluation is also undertaken by benchmarking the results obtained for MFE with those carried out for the COM and CBM forecasts.

The authors find that the MFE underestimates real and nominal GDP growth and expenditure components in the spring and autumn forecasting vintages. This is also the case when evaluating the forecasting accuracy of the COM and the CBM forecasts. When considering forecasting biasedness, the authors find evidence of downward biases in real and nominal GDP forecasts of the MFE, especially when excluding crises years. Indeed, some bias was evident in all expenditure components save for gross fixed capital formation. Evidence of forecast biasedness was also present for the COM and the CBM.

The authors did not find evidence of strong inefficiencies in MFE macroeconomic forecasts, indicating that the information available when the forecasts were produced was adequately factored in. Nonetheless, when assessing the relationship of past data outturns with forecast errors, the authors identified elements of forecasting inefficiencies in private and public consumption expenditure in year *t.*

An important finding from this study is the effect of statistical revisions on the results presented in this paper. While the authors did not estimate the causal effect of statistical data revisions on forecasting error performance, it is noted that statistical data tends to be revised upwards. Indeed, the authors find an upward systematic bias in most of the components analysed, which leads to a downward bias in forecast errors.

1 Introduction

"The macroeconomic and budgetary forecasts for fiscal planning shall be subject to regular, unbiased, and comprehensive evaluation based on objective criteria, including ex-post evaluation. The result of that evaluation shall be made public and taken into account appropriately in future macroeconomic and budgetary forecasts. If the evaluation detects a significant bias affecting macroeconomic forecasts over a period of at least four consecutive years, the Member State concerned shall take the necessary action and make it public."

Council Directive 2011/85/EU – Article 4(6)

Macroeconomic forecasts are key in framing government policies, particularly the budget process. Given their pivotal role, regular assessment and evaluation of forecast performance are vital to improving forecast quality and accuracy. Reliable economic forecasts build economic certainty and confidence and allow economic agents to make more efficient decisions. On the contrary, inaccurate forecasts beyond certain margins, whether overpredicted or underpredicted, have negative consequences by misleading decisions.

This analysis evaluates the forecasting performance of the macroeconomic projections presented by the Ministry for Finance and Employment (MFE) for spring (2004 – 2022) and autumn (2013 – 2022), focusing on forecasts for real and nominal GDP growth and its expenditure components in real terms. This analysis also compares the results with those of the European Commission (COM) and the Central Bank of Malta (CBM). The Economic Policy Department within the MFE is responsible for producing macroeconomic projections for the Government of Malta through its Short-term Quarterly Forecasting Econometric Model for Malta (STEMM).[1] Macroeconomic forecasts made by the MFE serve as important inputs in several key documents and policymaking decisions of the Government of Malta, including:

a. the annual Update of the Stability Programme (USP), presented by the Government of Malta to the European Commission every April, highlighting its macroeconomic and fiscal projections for years t up to t+3 following European Union Council regulations[2],

b. the annual Medium-Term Fiscal Strategy for Malta covering years t up to t+3 by the requirements of Article 15(8) of the Fiscal Responsibility Act,

[1] STEMM is an expenditure-driven model developed in collaboration with Cambridge Econometrics in 2002. For a detailed technical report on this macroeconomic model, please refer here.

[2] European Union Council Regulations – Council Regulation (EC) 1466/97 on the strengthening of the surveillance of budgetary positions and the surveillance and coordination of economic policies, as amended by Council regulation (EC) 1055/2005 of 27 June 2005 and Regulation (EU) No 1175/2011 – the preventive arm of the Stability and Growth Pact.

c. the Annual and Half-Yearly reports published by the MFE, and

d. the Draft Budgetary Plan (DBP) presented by the Government of Malta to the European Commission every year, including macroeconomic and budgetary projections covering years t and t+1.

In all of the above, the Malta Fiscal Advisory Council (MFAC) is mandated by law to:

a. endorse, as it considers appropriate, the macroeconomic and fiscal forecasts prepared by the Ministry for Finance and provide an assessment of the official forecasts;

b. analyse and assess whether the Government's Medium Term Fiscal Policy Statement and Medium-Term Fiscal Policy Strategy are compliant with the provisions of the Act, issue an opinion and any appropriate recommendations;

c. in relation to each National Medium Term Fiscal Plan, Stability Programme, Annual Draft Budget and Annual Budget, provide an assessment of whether the fiscal stance for the year or years concerned is, in the opinion of the Fiscal Council, conducive to prudent economic and budgetary management, and in conformity with the provisions of this Act, including by reference to the provisions of the Stability and Growth Pact;

d. assess the Government's budgetary performance against the fiscal targets and policies specified in the fiscal strategy and its compliance with the provisions of this Act;

e. analyse and issue an opinion and any recommendations pursuant to the Government's publication of the half-yearly and the annual report on the execution of the budget.

Fiscal Responsibility Act (Cap. 534 of the Laws of Malta) – Article 13(3), *ad verbatim*

2 Data and Methodology

The analysis undertaken in this paper builds upon and updates a similar exercise by Camilleri and Vella (2015), who test for forecast accuracy and unbiasedness and present the uncertainty surrounding the macroeconomic projections using fan charts for the period 2004-2013.[3] The main findings from this study indicate that the forecast performance by the Ministry compares favourably to other small open economies and even across other independent institutions' forecasts. They also do not find any systematic bias in the spring forecasts for nominal and real GDP but do find bias in the separate expenditure components.

In this paper, the authors update the sample period to include macroeconomic projections up to 2022 and extend the analysis to include both the forecasts published in the USP and the DBP. The study also adopts a broader rigorous assessment to evaluate forecast performance. In fact, forecast performance is assessed on four pillars: accuracy, unbiasedness, efficiency, and benchmarking.

The objectives of this research paper are twofold. First, it is being carried out in the context of Council Directive 2011/85 of the European Union on the requirements for budgetary frameworks of the Member States on the evaluation of forecast biasedness. Second, the MFAC believes that assessments of the forecast performance of the projections produced by the MFE are crucial to identifying areas of improvement and issuing recommendations and advice in that regard.

2.1 Data Sources

This section delineates the data sources utilised to assess the forecasting performance of the MFE. Time series for nominal GDP, real GDP and its components were collected for the spring forecast rounds (Update of Stability / Convergence Programme) covering the 2004-2022 period and for the autumn forecast round (Draft Budgetary Plan) covering the 2013 - 2022 period (See Table 2.1).[4] Data was also obtained from the Quarterly Reviews of the CBM and the spring and autumn forecasts of the COM to enable a benchmarking exercise.

[3] The paper titled 'Interpolating forecast errors for assessing uncertainty in macroeconomic forecasts: an analysis for Malta' may be accessed here.

[4] The first publication of Malta's Draft Budgetary Plan was on 15 October 2013. This following regulation (EU) No 473/2013 of the European Parliament and of the Council of 21 May 2013 on common provisions for monitoring and assessing draft budgetary plans and ensuring the correction of excessive deficit of the Member States in the euro area.

From 2004 to 2009, the Stability/Convergence Programmes used to be published in November, while since 2011, it began to be published in April.[5,6] For comparability purposes, data published by the COM for 2004-2009 were collected from the autumn forecast round issue (published in November) while the rest of the sample was collected from the spring issue (published in May).[7] In 2010, due to changes in timing following the introduction of the European Semester, the Stability Programme was not published by the MFE. However, forecast data for this year were still available internally. This study used this data to have a complete time series.

The projections of the CBM are available from 2008 onwards. To maintain data comparability with the MFE forecasts, data for 2008 and 2009 were collected from the fourth Quarterly Review issue (published in December), while the rest of the data was collected from the second Quarterly review issue (published in May). In addition, the forecasts from the DBP are compared to the COM's autumn forecasts (published in November) and the fourth Quarterly Review by the CBM (published in December).

For both the COM and the CBM, the nominal GDP forecast is derived using the published GDP deflator. Since the CBM did not publish nominal GDP nor GDP deflator forecasts before 2014, the comparison is only carried out on a limited data set from 2015 to 2022. In addition, while forecast data for the MFE in the Stability/Convergence Programmes are available up to year t+3, the COM and the CBM only report forecasts for years t and t+1. In the autumn vintage, forecast data for all institutions is available for years t and t+1. Additional analysis was also undertaken, excluding years of crises, namely the global financial crises (2009 & 2010) and the COVID-19 pandemic (2020 & 2021), to evaluate the evidence of any systematic errors made in 'normal' times.

[5] EU member states that share the euro currency issue are mandated to issue annual updates of the 'Stability Programme', whereas those that do not share the euro currency issue a different publication known as the 'Convergence Programme'. Subsequently, before the adoption of the Euro, Malta published Convergence Programmes, and once the Euro was adopted, Stability Programmes began to be published.

[6] The change in publication date reflects developments at the EU level in relation to the European Semester.

[7] For ease of reference, the forecasts published in the Stability/Convergence Programme will be referred to as the spring forecast vintage.

Table 2.1: Variables and Sample Period[8]

Variable	Definition	Forecast Round	Sample Period
Nominal GDP growth	The growth in the estimate of the value of goods and services produced in the economy year-on-year measured in market prices.	USP	2004-2022
		DBP	2013-2022
Real GDP growth	The growth in the estimate of the value of goods and services produced in the economy year-on-year measured in real terms.	USP	2004-2022
		DBP	2013-2022
Real Private Consumption Growth	The growth in the goods or services used without further transformation in production by households and non-profit institutions year-on-year measured in real terms.	USP	2004-2022
		DBP	2013-2022
Real Public Consumption Growth	The growth in the goods or services used without further transformation in production by the government year-on-year measured in real terms.	USP	2004-2022
		DBP	2013-2022
Real Gross Fixed Capital Formation	The growth in the total value of producers' acquisitions, less disposal of fixed assets during the accounting period, plus certain additions to the value of non-produced assets measured in real terms.	USP	2004-2022
		DBP	2013-2022
Real Exports	The growth in the products of local origin sold to other countries measured in real terms.[9]	USP	2004-2022
		DBP	2013-2022
Real Imports	The growth in the products of foreign origin brought into the country measured in real terms.[10]	USP	2004-2022
		DBP	2013-2022

Source: NSO, MFE

2.2 Limitations of the Data Sources

At the outset, the authors note that, as with all statistical exercises, some inherent limitations are worth mentioning. Forecast error evaluations generally use at least 20 years of data. Our sample size, especially for the autumn forecast period, is notably smaller, with only ten data points. Furthermore, data limitations exist along with the different cut-off points across institutions, which constrain cross-institution benchmarking. Therefore, the benchmarking results should be interpreted with caution.

[8] Definitions were adapted from the metadata section of NSO which can be accessed here.

[9] The products exported consist of sales, barter, gifts or grants of good and services from residents to non-residents.

[10] The products imported include goods, intended for consumption which enter the territory of a particular country from another foreign country and are placed under the Customs procedure for free circulation within the European Union or for inward processing.

It is important to note that statistical errors in national accounts data can also affect forecast errors. Forecast errors are influenced by the vintage of input data used in the forecasting model and the vintage of data used as a benchmark to estimate the forecast errors. While the most up-to-date statistical data gives a more accurate and reliable estimate of forecast errors, it may underestimate the efficiency of a forecasting model by diluting statistical errors with the pure forecast errors of a given economic model. Using the latest national accounts release as a benchmark on which to compute forecast errors, coupled with the relatively small sample size, adds a further degree of uncertainty to the analysis in that any further revisions to national accounts data will have an impact on the accuracy of forecast projections both ex-ante and ex-post.[11] The ex-ante effect is mainly through the trajectories of the forecast projections, while ex-post forecast errors may emerge since the base on which the forecast projections were estimated in the first place would have changed.

2.3 Methodology

The assessment of the MFE's predicting performance is based on the following four pillars:

- Closeness of forecasts to the actual outcome (accuracy).
- Whether the forecasts produced by MFE have been more or less accurate than those produced by other institutions, namely CBM and COM (benchmarking exercise).
- Whether forecasts have been consistently optimistic or conservative (unbiasedness).
- Whether forecasts have reflected information available at the time they were produced (efficiency).

The authors acknowledge that these measures of performance are interrelated: for instance, if forecast accuracy is high, there is less scope for forecast biasedness. Similarly, forecast accuracy improves as available information is used more efficiently. Nonetheless, each indicator offers different perspectives on forecast performance worth evaluating.

2.3.1 Forecast accuracy and benchmarking

Forecast accuracy is measured by calculating the mean error, the mean absolute error, the root mean squared error, the mean relative absolute error and Theil's U Statistic. These are explained below:

- The mean error (ME) is the average of forecast errors. More formally,

$$\text{ME} = \frac{1}{T}\sum_{i=1}^{T} \hat{y}_t - y_t = \frac{1}{T}\sum_{i=1}^{T} e_t \qquad (1)$$

[11] The actual data used as a benchmark to compare with the forecast vintages is NSO News Release 095/2023 available here.

where the forecast of variable y for period t is denoted by $\hat{y}_t$ and the actual value by y_t, and e_t is a forecast error defined as forecast minus outturn.

The ME must be interpreted cautiously because a small result does not necessarily indicate good forecast accuracy, as negative forecast errors offset positive ones. Moreover, it is not meant for comparing and evaluating a method's forecast accuracy across different data sets.

- The mean absolute error (MAE) is the average of the absolute error, which is the deviation of forecasts from actual points, disregarding the sign of the error. Formally,

$$\text{MAE} = \frac{1}{T}\sum_{i=1}^{T}|\hat{y}_t - y_t| = \frac{1}{T}\sum_{i=1}^{T}|e_t| \qquad (2)$$

again, the forecast of variable y for period t is denoted by $\hat{y}_t$ and the actual value by y_t, and e_t is a forecast error, defined as forecast minus outturn.

- The mean relative absolute error (MRAE) is an alternative to the MAE as a scale-dependent measure. The MRAE implies taking an average of the absolute value of the relative share of errors, i.e., the forecast error based on the forecasts published by the institutions of interest as a share of the forecast error obtained from the benchmark method. Usually, the benchmark method is the random walk without drift model, where f_t^* is equal to the last observation. The MRAE is calculated by using this formula:

$$\text{MRAE} = \frac{1}{T}\sum_{i=1}^{T}\left|\frac{\hat{y}_t - y_t}{f_t^* - y_t}\right| \qquad (3)$$

A concern for this measure is that if the forecasting error obtained from the benchmark method is zero, using the random walk without drift model as a benchmark method would no longer be possible.

- The root mean squared error (RMSE) is a common forecast accuracy measure calculated as the standard deviation of the forecast errors. This measure disproportionately penalises forecast accuracy according to the magnitude of the forecast errors, i.e., it accounts for the fact that large forecast errors are considered more problematic than small ones by using a quadratic loss function. More formally,

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{i=1}^{T}(\hat{y}_t - y_t)^2} = \sqrt{\frac{1}{T}\sum_{i=1}^{T}(e_t)^2} \qquad (4)$$

The sensitivity of the RMSE to data outliers is the most prevalent justification for employing this scale-dependent statistic. As a result, this forecast accuracy indicator is considered superior to the other measures outlined in this chapter.

- Theil's U statistic (U) is a relative accuracy measure that compares the predictions with a naïve forecast.[12] This forecast accuracy measure can be interpreted as the ratio of the RMSE to the standard deviation of the forecast errors from the naïve model. Analogous to the RMSE, this measure also attributes more weight to large errors by squaring the deviations. If Theil's U statistic exceeds 1, it implies that the forecast from the model is no more accurate than a naïve forecast. Theil's U statistic is calculated by using the following formula:

$$U = \sqrt{\frac{\frac{1}{T}\sum_{i=1}^{T}(\hat{y}_t - y_t)^2}{\frac{1}{T}\sum_{i=1}^{T}(y_t - y_{t-1})^2}} \qquad (5)$$

A simple comparison of the different forecast accuracy measures was also conducted across institutions (MFE, COM and CBM).

2.3.2 Unbiasedness

In the history of the EU's Stability and Growth Pact, some governments have justified fiscal expansions or postponed fiscal adjustments by being more optimistic when forecasting medium-term growth (Larch et al., 2021; Frankel, 2011). In this context, the authors assessed the real and nominal GDP growth forecasts along with the GDP components in real terms produced by the MFE for any potential upward or downward bias. To carry out this evaluation, the Least Squares methodology was employed whereby the forecast errors were regressed on a constant, with a null hypothesis that the constant was zero. In case of biasedness, the constant would be significant and take a non-zero value. Formally, the following regression was estimated:

$$e_t = \beta_0 + \varepsilon_t \qquad (6)$$

where e_t is the forecast error of variable y for period t and where ε_t is a zero-mean error term. Under the null hypothesis of unbiasedness, $\beta_0 = 0$. If $\beta_0 < (>) 0$, this means that the forecast has been systematically too low (high). In line with the methodology employed by the Bank of England (2015) when evaluating for the presence of bias in macroeconomic forecasts, the authors estimated the regression using OLS with Heteroscedasticity and Autocorrelation (HAC) standard errors to account for any potential autocorrelation and heteroscedasticity issues.[13]

2.3.3 Efficiency

The efficiency of forecasts is assessed by regressing forecast errors on information that was known when the forecasts were made. Ideally, forecast errors will be uncorrelated with any such information. Otherwise, those errors could have been reduced by incorporating that information when the forecasts were made. To perform

[12] Naïve Forecast is a forecasting technique in which the forecast for the current period (y_t) is set to the actual value of the previous period (y_{t-1}).

[13] The results are based on a HAC adjustment using Andrew's Automatic bandwidth method.

tests of strong efficiency, the authors estimated the following equation using OLS with HAC standard errors:[14]

$$e_t^{t-h} = \beta_0 + \beta_1 z_{t-h} + u_t$$

where e_t^{t-h} represents forecast errors from time t to time t-h, z_{t-h} is a variable that was known to the forecaster at time t-h, and u_t is a zero-mean error term. Under the null hypothesis of strong efficiency, the authors test whether $\beta_1 = 0$, if the hypothesis is accepted (i.e., $\beta_1 = 0$), we can deduce that the forecasts are strongly efficient.

For each of the variables in our sample, strong efficiency tests were performed using the following for z: the previous forecast error (for the same variable) known to the forecasters at t-h and past data outturns (based on the most recent actual data). Each z variable was tested separately. The first test evaluates whether forecast errors could have been avoided by not repeating past errors, while the second tests for how efficient the inclusion of past data was to produce forecasts for each variable.

[14] Forecasts are deemed to be 'strongly' efficient if forecast errors are uncorrelated with information known at the time the forecasts were made.

3 Forecast errors growth projections

This section describes the forecast errors observed for nominal GDP, real GDP growth and its expenditure components for the projections undertaken by the MFE. Forecast errors in this study are defined as the forecast at time t less the actual data for time t.[15] More formally,

$$e_{t,t} = \hat{y}_t - y_t \text{ for the current year; and}$$
$$e_{t+1,t} = \hat{y}_{t+1} - y_{t+1} \text{ for the following year,}$$

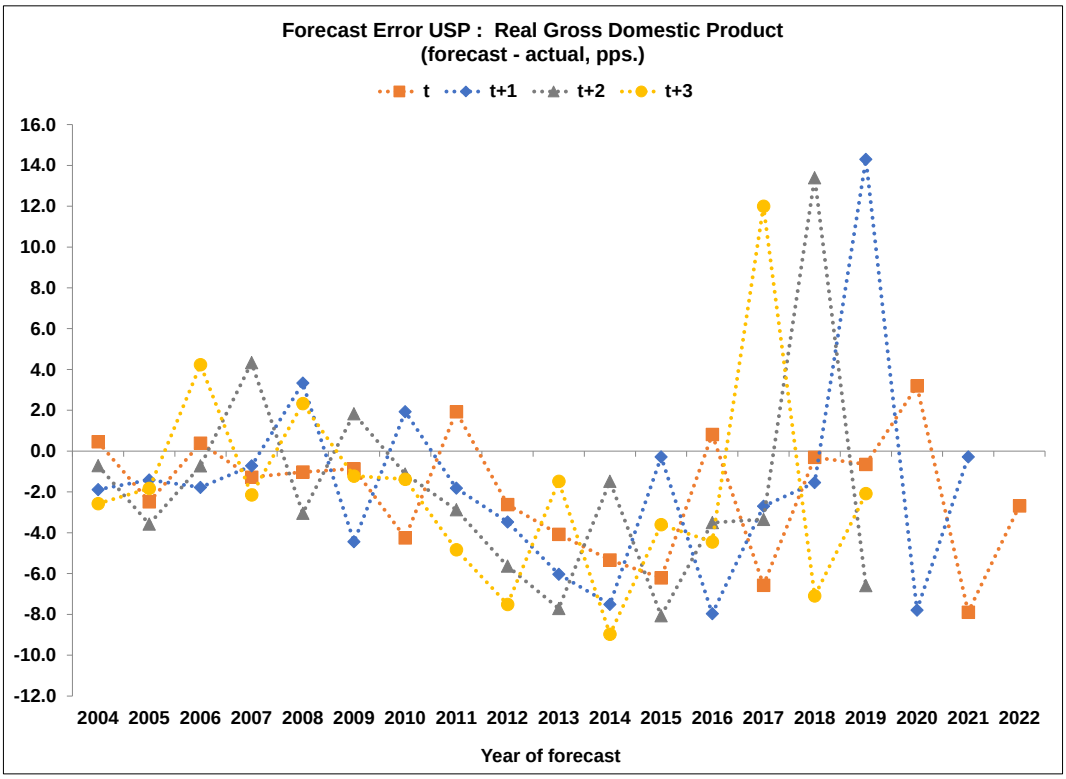
where $\hat{y}_t$and $\hat{y}_{t+1}$ are the projections made at time t for t and t+1 respectively, y_t is the actual data of variable y for year t, and y_{t+1} is actual data for variable y for year t+1. Therefore, a positive forecast error for real GDP growth implies an overestimation of the growth rate, while a negative value means an underestimation of the growth rate.

Forecast errors for real GDP growth in the USP range between -9.0 and +14.3 percentage points (pp) (See Chart 3.1). The overestimation of growth is largely attributed to unexpected economic shocks, since when removing the 2008/2009 financial crisis years and the COVID-19 crisis years, the forecast errors for real GDP range between -9.0 and 2.3 pp (See Charts A1 in Appendix A). The CBM and the COM's spring forecast errors have also been affected by the COVID-19 crisis, as the overestimation substantially decreases to a maximum of 2.0 pp when these years are excluded from the sample (See Charts A9 and A17 in Appendix A).

Overall, the data shows that there is more tendency to underestimate growth. Indeed, all components show a degree of underestimation of the rate of growth. This could be due to structural changes in the economy not incorporated in the model and statistical revisions. The authors also note that the tendency to underestimate real GDP growth has increased post-2010, with forecast errors excluding the crises years for the period 2004 to 2010 ranging from -3.6 to +2.3 pp while the forecast errors from 2011 to 2022 ranged from -9.0 to +1.9 pp (See Charts A1 in Appendix A).

[15] The actual data refers to the latest release available on the NSO website at the time of compilation of this report being NSO News Release 095/2023.

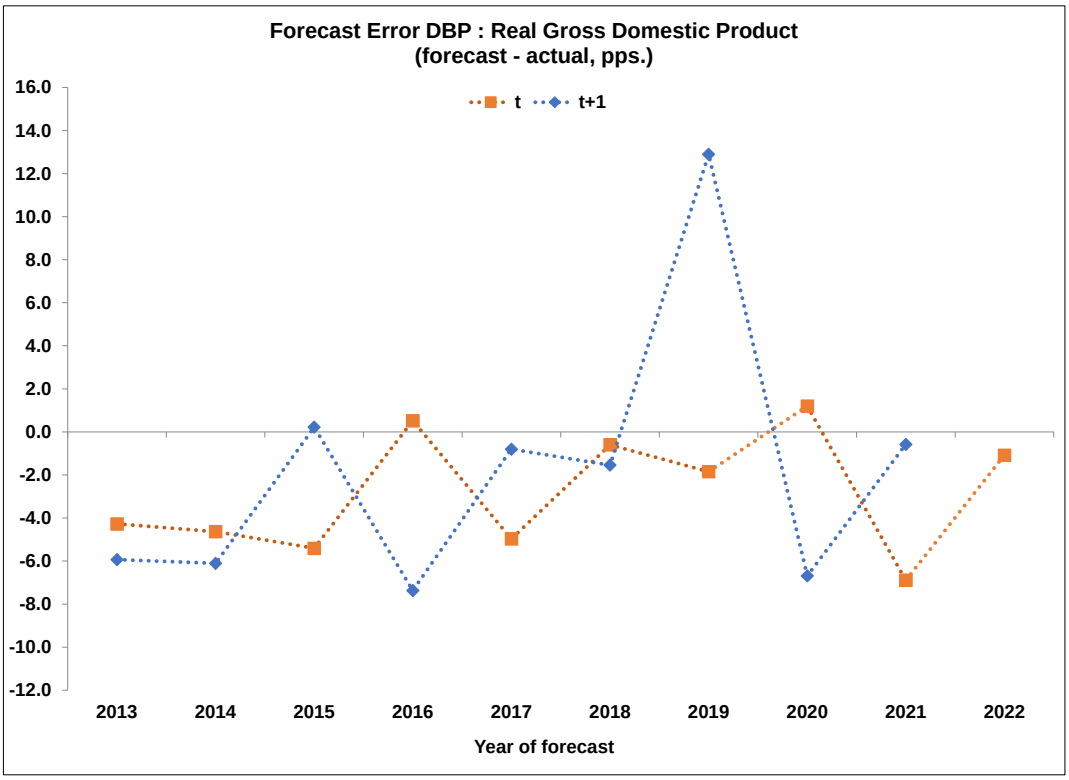
Chart 3.1: Forecast Error: Real GDP (forecast – actual, pps.) (USP)[16]



Forecast errors for real GDP growth for the entire sample in the DBP range from -7.4 to 12.9 pp, which are similar to the range of the CBM autumn forecast errors, which varies between -6.8 and +12.7 pp (See Chart 3.2 and Chart A25 in Appendix A). On the other hand, the COM has a wider range of forecast errors, on the pessimistic side: from -8.7 to +12.8 pp. The overestimation for the autumn forecast round drops to an average of 0.7 pp when the crises periods are removed, confirming that the overestimation is mainly caused by unanticipated shocks (See Charts A33 in Appendix A).

[16] Forecast errors for all components in the USP are categorised by a forecasting vintage since 2004 for current (t), one-year (t+1), two-year (t+2) and three-year (t+3) ahead forecasts.
Forecast errors for all components in the DBP are categorised by forecasting vintage since 2013 for current (t) and one-year (t+1) ahead forecasts.
The horizontal axis represents the year in which the forecast was undertaken.
Example: Forecast Error for 2020 in 2020 (t) = Forecast for 2020 made in 2020 – Actual Data for 2020 (orange line), Forecast Error for 2020 in 2019 (t+1) = Forecast for 2020 made in 2019 – Actual Data for 2020 (blue line), Forecast Error for 2020 in 2018 (t+2) = Forecast for 2020 made in 2018 – Actual Data for 2020 (grey line), Forecast Error for 2020 in 2017 (t+3) = Forecast for 2020 made in 2017 – Actual Data for 2020 (yellow line).

Chart 3.2: Forecast Error: Real GDP (forecast – actual, pps.) (DBP)



A similar pattern is evident for nominal GDP growth, although forecast errors are marginally higher than for real GDP (See Chart 3.3 and Chart 3.4). Indeed, in the Update of the Stability Programme, forecast errors for nominal GDP growth range from -10.0 to +15.4 pp, while in the DBP, these range between -9.4 and +13.6 pp. With the exclusion of the crises years, the overestimation falls to a maximum of 3.4 pp in the Stability programme and 0.5 pp in the DBP (See Charts A2 in Appendix A). The same pattern is evident in the spring and autumn forecasts of the CBM and COM (See Charts A10, A18, A26 and A34 in Appendix A). The figures show that MFE also tends to underestimate nominal GDP growth. This tendency has been more pronounced in the more recent period between 2011 and 2022 in the case of the Stability Programme. Indeed, the forecast errors for nominal GDP growth excluding crises years from 2004 to 2010 are in the region of -3.3 to +3.4 pp, while from 2011 to 2022, the forecast errors range between -10.0 and +3.1 pp (See Chart A2 in Appendix A).

Chart 3.3: Forecast Error: Nominal GDP (forecast – actual, pps.) (USP)

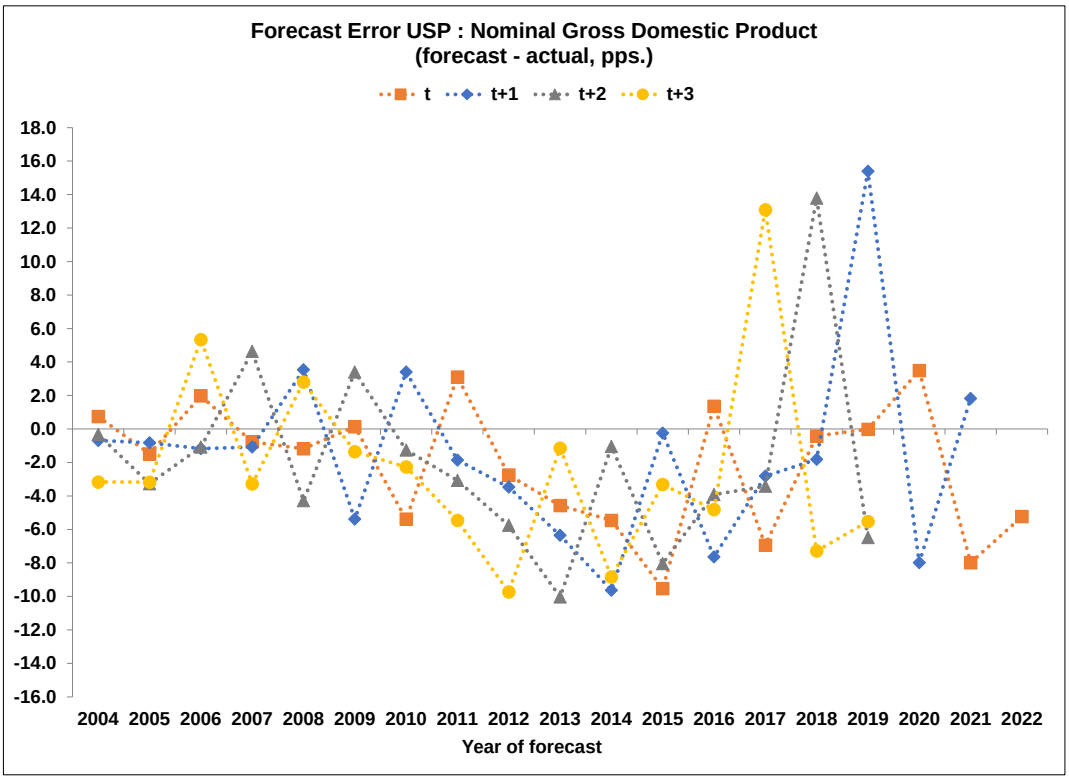
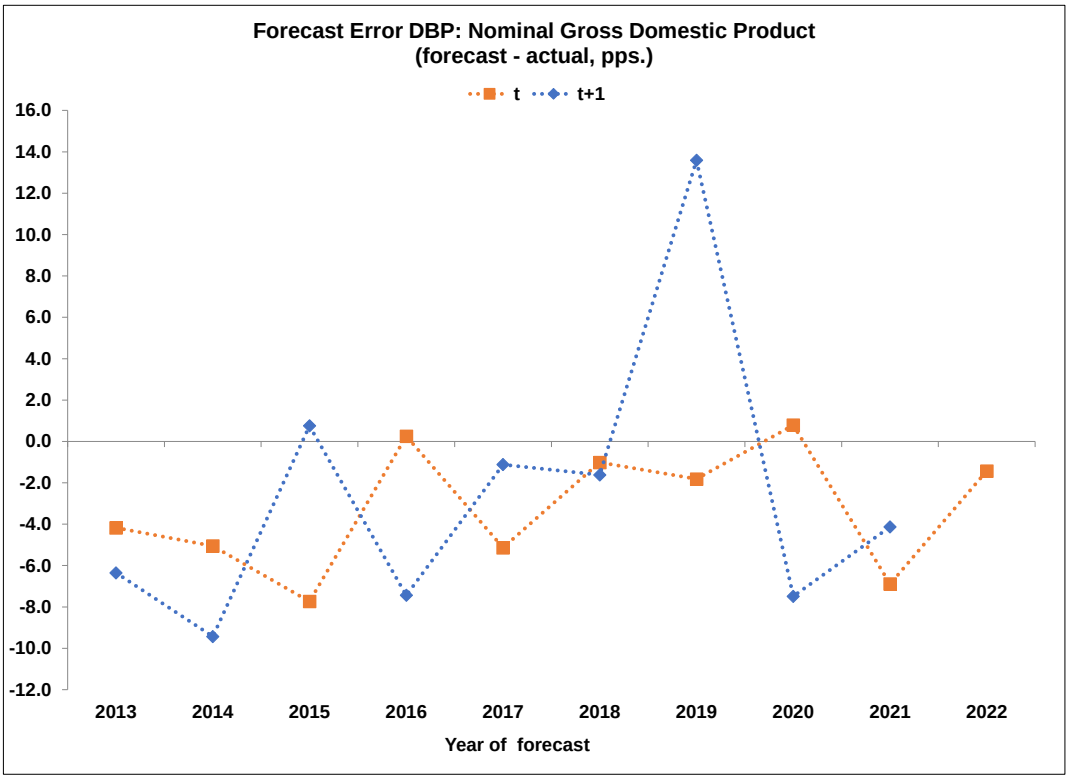


Chart 3.4: Forecast Error: Nominal GDP (forecast – actual, pps.) (DBP)



Private consumption is one of the most important contributors to real GDP growth. Furthermore, by constituting an important tax base, it has direct implications on the

fiscal position, as highlighted by Camilleri and Vella (2015), who state that its forecast accuracy can have a substantial impact on the accuracy of indirect tax revenue projections as it is a very tax-rich expenditure component. The forecast errors for this component in the USP vary from -5.9 to +14.5 pp; in the DBP, the forecast errors range from -7.7 to +14.6 pp. Charts 3.5 and 3.6 show that the overestimation in this case is due to the unexpected shock caused by the COVID-19 crisis, as the removal of crises periods decreases the overestimation to a maximum of 3.9 pp in the USP and 1.2 pp for the DBP. Indeed, the most evident overestimation was for 2020, capturing the start of the COVID-19 period, whereby at time t, MFE was expecting real private consumption growth to be -0.8% in the USP and -6.0 % in the DBP. Actual results for 2020 show that private consumption contracted by 10.6%, resulting in a forecast error of 9.8 pp in the case of the USP and 4.6 pp for the DBP.

The forecast error is the largest for the one-year ahead forecast in 2019 (approximately 14.5 pp), as MFE was forecasting private consumption growth to average 4.0% in the USP and the DBP, given the economy's robustness at the time. Forecasts made in 2017 (t+3) and 2018 (t+2) for 2020 show similar forecast errors in the USP but to a slightly lesser extent. The spring forecast errors by the COM are comparable to those of the USP, while those of the CBM are slightly lower. The CBM and COM autumn forecast errors are akin to those undertaken by MFE in the DBP however, forecast errors for 2022 indicate an overestimation of the rate of growth rather than an underestimation as found for the DBP (See Charts A11, A19, A27 and A35 in Appendix A).

Chart 3.5: Forecast Error: Private Consumption (forecast – actual, pps.) (USP)

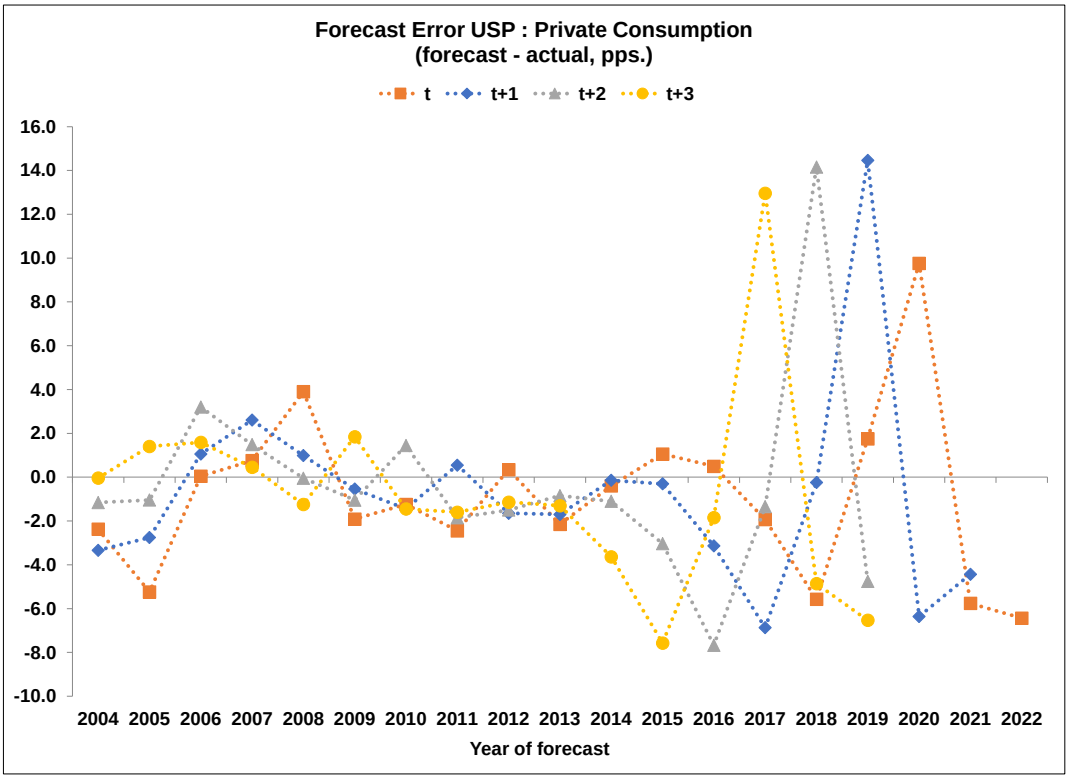
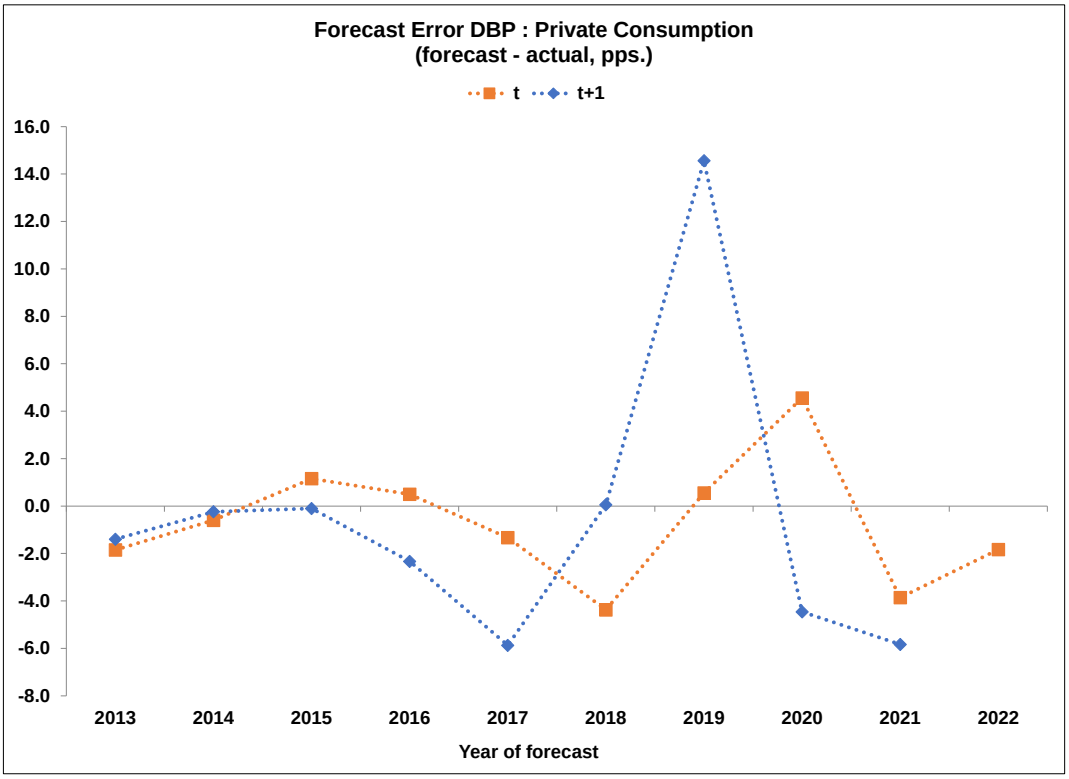
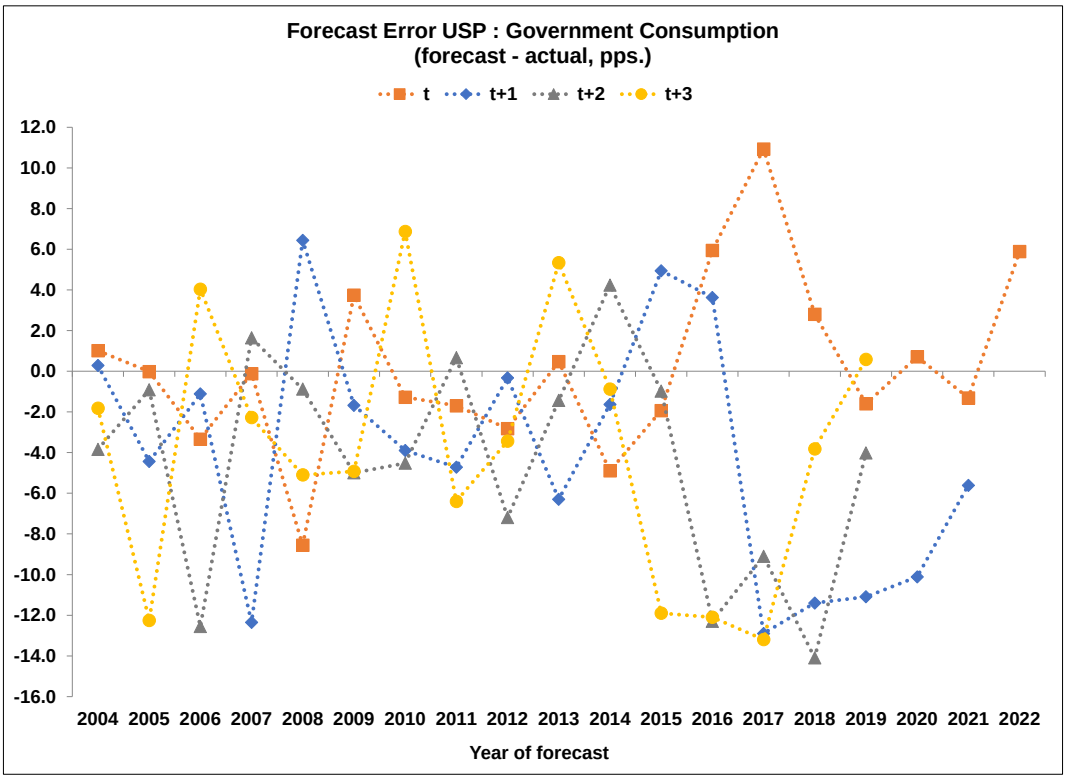


Chart 3.6 Forecast Error: Private Consumption (forecast – actual, pps.) (DBP)



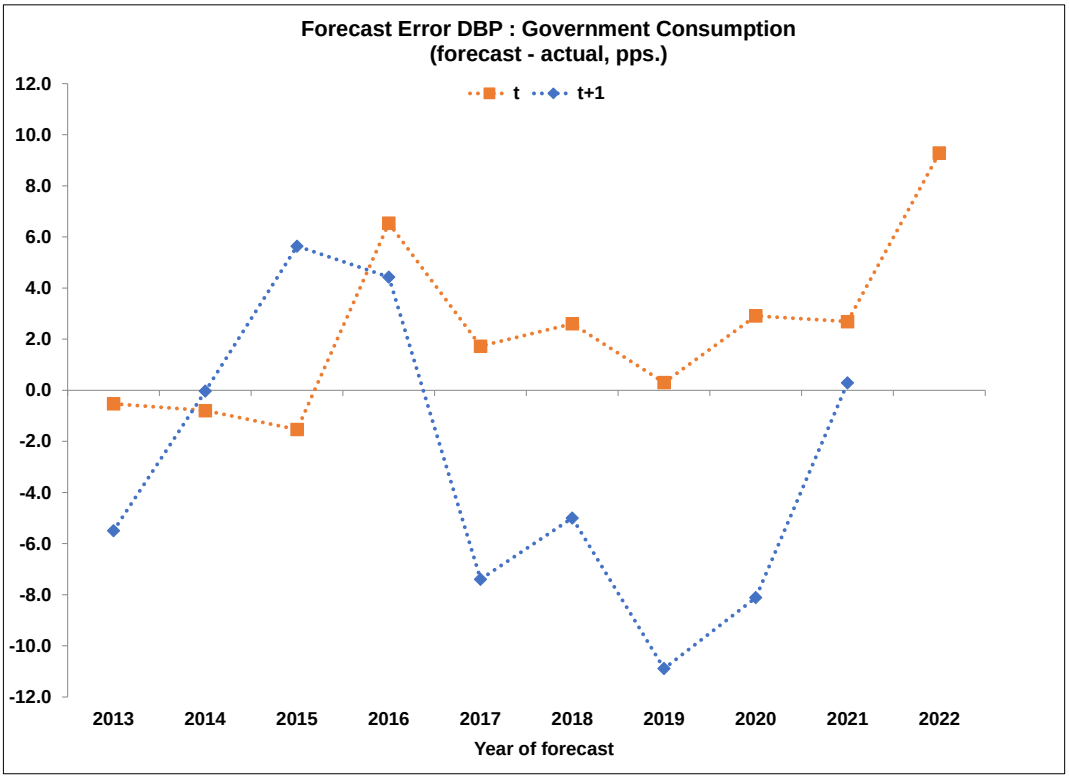
The government expenditure component is an exogenous variable in the MFE's model. It is determined by the fiscal forecasts for compensation of employees, intermediate consumption, consumption of fixed capital, social transfers in kind and market output for own final use and payments for non-market output. Charts 3.7 and 3.8 demonstrate how the MFE, in both the USP and the DBP, tend to overestimate government consumption at time t, implying higher public consumption than is realised. Consequently, the overestimation range is higher than that for private consumption in both the USP (10.9 pp) and the DBP (9.3 pp).

Chart 3.7: Forecast Error: Government Consumption (forecast – actual, pps.) (USP)



The extent of underestimation of public consumption is similar to that of private consumption, being the most significant in the forecasts prepared for 2008, 2018 and 2020. The need for government intervention to address the financial crisis was unforeseen in 2007. Indeed, in 2007, public consumption stood at 1.3%, while in 2008, it grew by 11.6 pp to 12.9%, causing substantial forecast errors. After that, public consumption expenditure decreased again, to increase again by 10.9 pp to 12.4% in 2018, resulting in a significant forecast error in 2017 for t+1 of 12.9 pp. Such increase in expenditure in 2018 can be attributed to factors such as the Valletta Capital of Culture and the EU presidency. The COVID-19 pandemic further increased public consumption expenditure in 2020 as the government offered support measures to keep the economy afloat. Forecasts from 2017 onward for 2020 exhibit high negative forecast errors because COVID-19 was unknown when the forecasts were being undertaken. It can be noted from both the DBP and the USP that government consumption for year t is generally overestimated by MFE, while the one-year ahead forecast is usually underestimated. The DBP's forecasts undertaken for 2022 exhibit positive forecast errors, in contrast to the CBM and COM's autumn estimates, which show significant underestimation in the one-period ahead forecast in 2021 and the current year forecast in 2022 (See Charts A28 and A36 in Appendix A).

Chart 3.8: Forecast Error: Government Consumption (forecast – actual, pps.) (DBP)



Charts 3.9 and 3.10 demonstrate that due to the component's volatility, gross fixed capital formation forecast errors are substantially larger than those for other components. Forecast errors for this component in the USP range from -59.3 to +20.9 pp, while those in the DBP range from -57.5 to +16.0 pp, showing significant underestimation over time. A single extraordinary investment in a small, open economy such as Malta's can significantly affect the actual growth rate in a given year. Therefore, it is difficult to anticipate the extent of growth beforehand unless the investment is known in advance. This explains the large magnitude of underestimation in the forecasts. Both charts show that the forecasts for 2015 were the most underestimated in terms of growth. The actual figures reveal an investment growth of 62.3%, while the projection at time t was 23.6%, resulting in a forecast error of 38.7 pp.

Chart 3.9: Forecast Error: GFCF (forecast – actual, pps.) (USP)

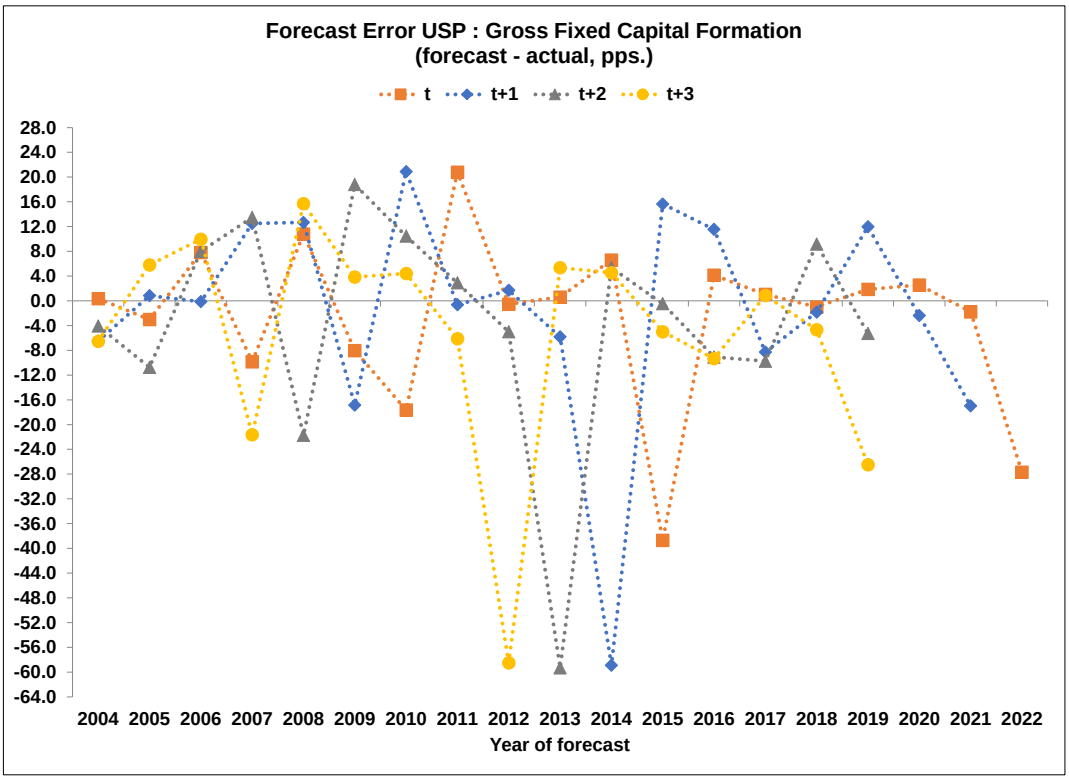
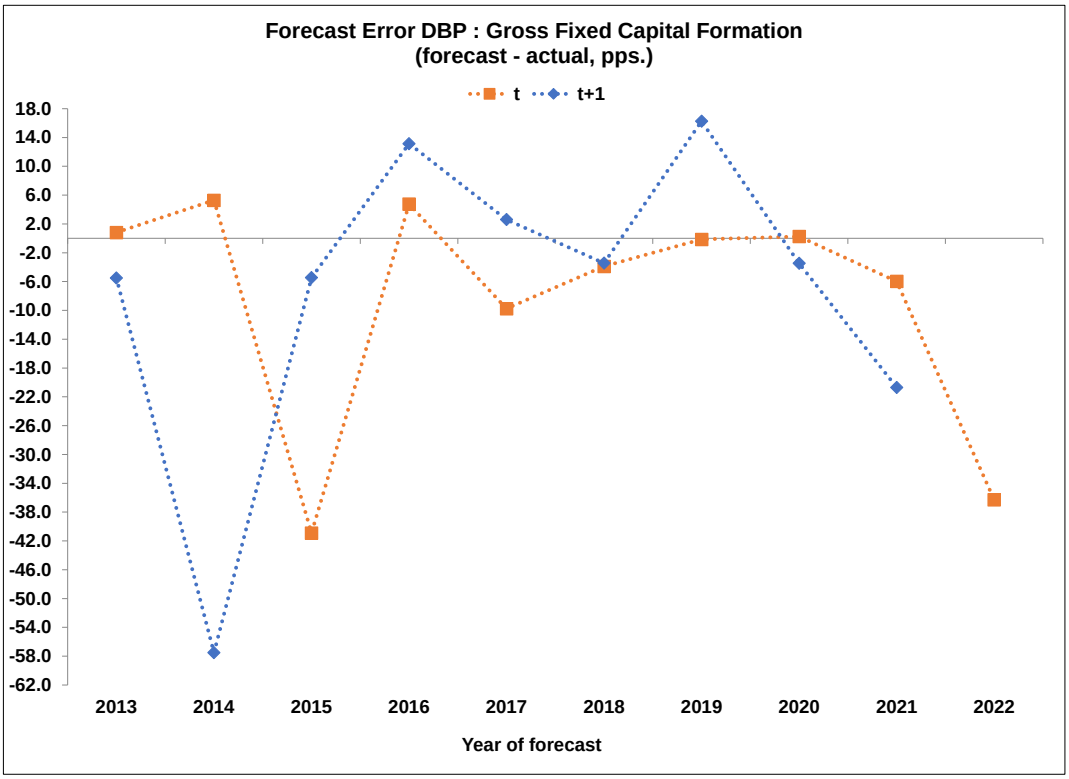


Chart 3.10: Forecast Error: GFCF (forecast – actual, pps.) (DBP)



Forecasts for 2015 made in previous years (2012, 2013, 2014) show an even greater forecast inaccuracy due to the lack of information about prospective investments. Most

of this increase was brought on by high investment from the aviation industry, which was not captured in the economic model by MFE. This relatively high investment, which took place in 2015, had a significant base effect on subsequent years. Another significant investment in the aviation industry was made in 2022, leading to an underestimation of 27.7 pp in the USP and 36.3 pp in the DBP. The autumn forecasts by the CBM and the COM show less variation at time t post-2017 compared to the DBP (See Charts A29 and A37 in Appendix A). It is also important to highlight that the influence on the accuracy of GDP growth forecasts is far less significant than these estimates suggest given the significant import content of such investments.

Malta, being an open economy, is substantially reliant on its exports. Thus, this component contributes significantly to GDP growth. While the forecast errors for the DBP range from -15.4 to 3.3 pp, the USP range for exports runs from -24.8 to 6.8 pp. Charts 3.11 to 3.14 show that the export and import components are underestimated to a greater extent than the domestic components of GDP. This stem from the fact that exports of goods and services are susceptible to international trade developments, commodity price movements and exchange rate movements. Additionally, the volatility in oil prices and its effects on fuel bunkering activity and offshore oil transhipment presents another hurdle to the forecast of exports of goods and services. The CBM and COM forecast errors align with those of the MFE (See Charts A14, A22, A30 and A38 in Appendix A).

The import component variation ranges from -22.2 to +4.3 pp for the USP, while for the DBP, the range is from -17.8 to +2.7 pp, implying a similar pattern to that of exports of goods and services. This is primarily due to the high import content of Maltese exports. Thereby, industrial goods and oil imports are determined mainly by exports of related goods in STEMM with an elasticity close to unitary. Imports are also driven by domestic demand, primarily investment, which is linked with imports of capital goods, and private consumption associated with imports of consumer goods. Charts 3.13 and 3.14 show that the forecasts of imports have been quite conservative throughout the years, stemming from the underestimation of investment and exports. Both the CBM and the COM forecast errors reveal comparable underestimation of this component, with the sole difference being that the autumn forecast errors of the CBM and the COM for 2022 are positive, while those of the DBP are negative (See Charts A31 and A39 in Appendix A).

Chart 3.11: Forecast Error: Exports (forecast – actual, pps.) (USP)

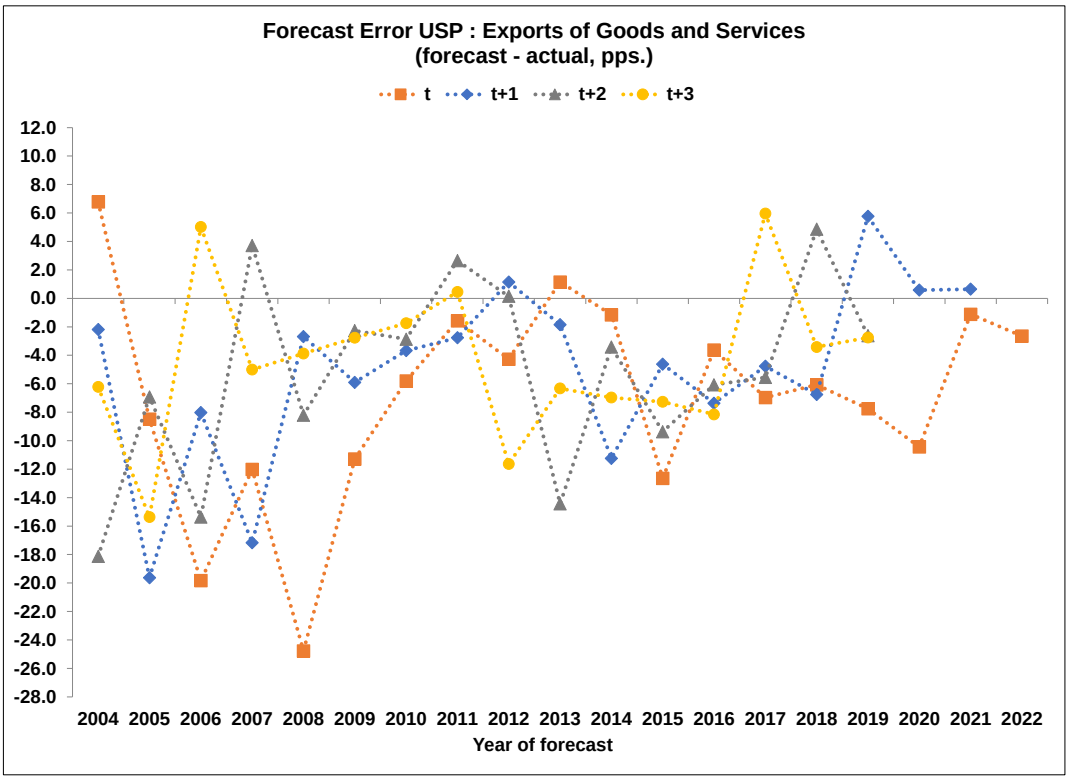


Chart 3.12: Forecast Error: Exports (forecast – actual, pps.) (DBP)

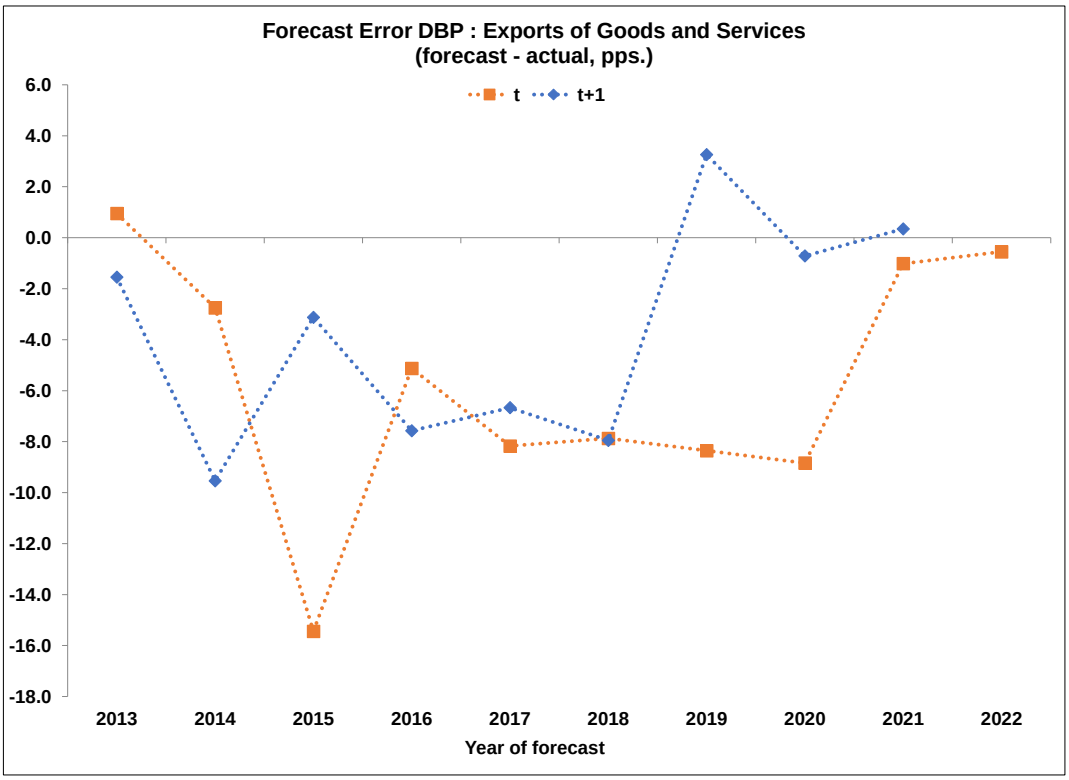
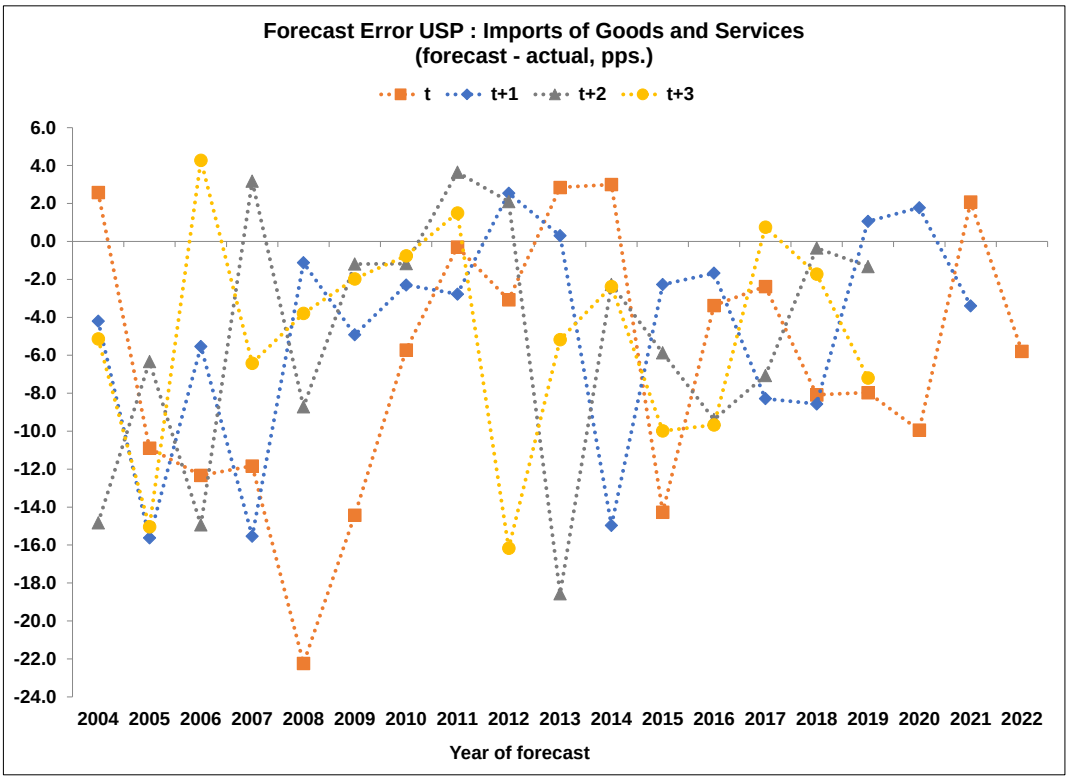
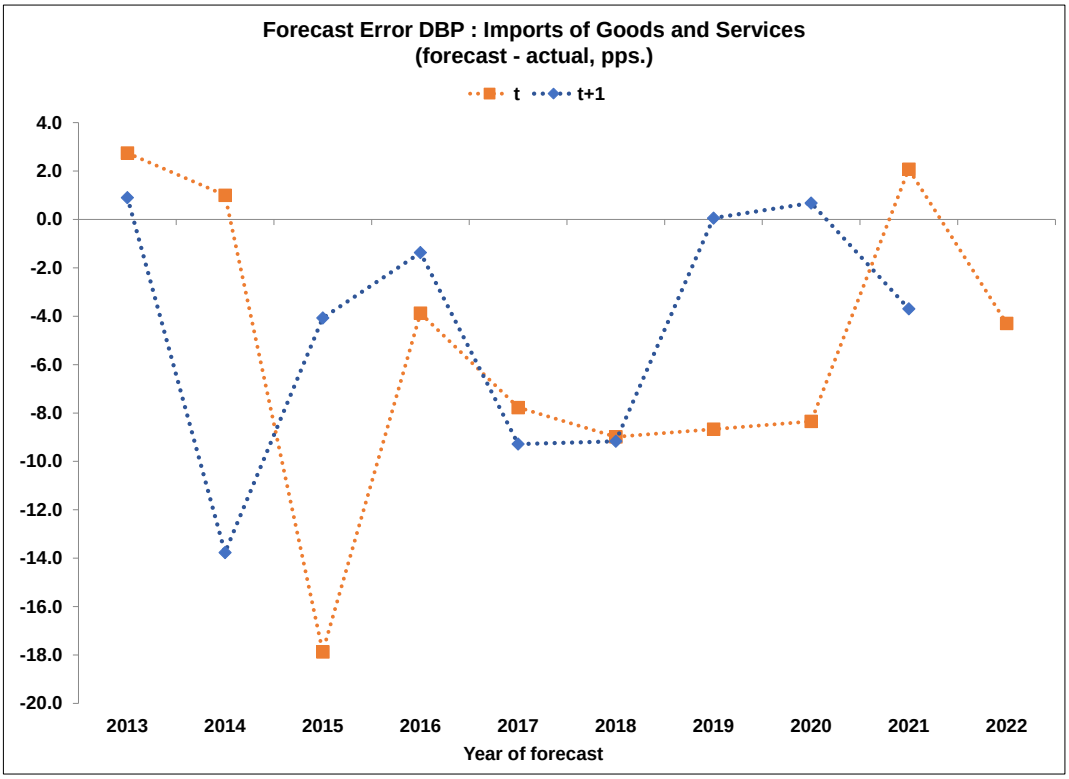


Chart 3.13: Forecast Error: Imports (forecast – actual, pps.) (USP)



Charts 3.14: Forecast Error: Imports (forecast – actual, pps.) (DBP)



The impact on real GDP is the net exports (imports less exports). Charts 3.15 and 3.16 do not clearly illustrate any trend regarding the prevalence of forecast errors for net

exports, as this component is dependent on the composition of the growth of exports and imports for each year. The range of forecast errors for net exports varies much less than that for exports and imports individually, as errors offset each other. Indeed, the range of forecast errors for the USP is -7.5 to +5.2 pp, whereas the DBP ranges from -6.2 to +4.2 pp. It is pertinent to note that the forecasts of autumn for the CBM and the COM exhibit positive forecast errors for 2022, contrary to what is revealed by the DBP (See Charts A32 and A40 in Appendix A).

The forecast error growth estimates have demonstrated an overall tendency to underestimate real and nominal GDP, as all components show a degree of underestimation with a more pronounced inclination post-2010. The data also indicates that overestimation tends to happen due to unforeseen events such as the COVID-19 crisis, especially in the private consumption component. On the other hand, government consumption tends to be overestimated at time t but underestimated in the forecast for the outer years. Gross fixed capital formation emerges as a highly volatile component with substantial forecast errors attributed to the unpredictability of significant investments in a small, open economy like Malta. The lack of information about prospective investments, especially in sectors like the aviation industry, contributes to sizable inaccuracies in forecasting GFCF, consequently affecting the forecasts for imports. Although both export and import components exhibit significant underestimation, net exports are more balanced as errors generally offset each other.

Chart 3.15: Forecast Error: Net Exports (forecast – actual, pps.) (USP)

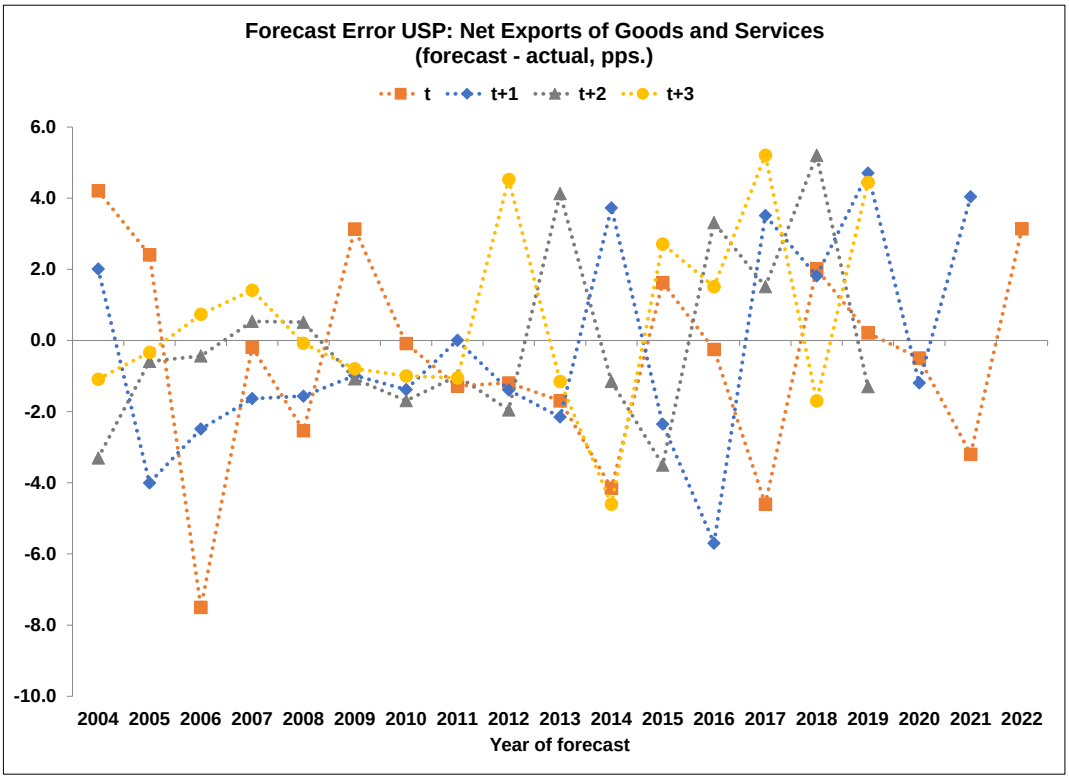
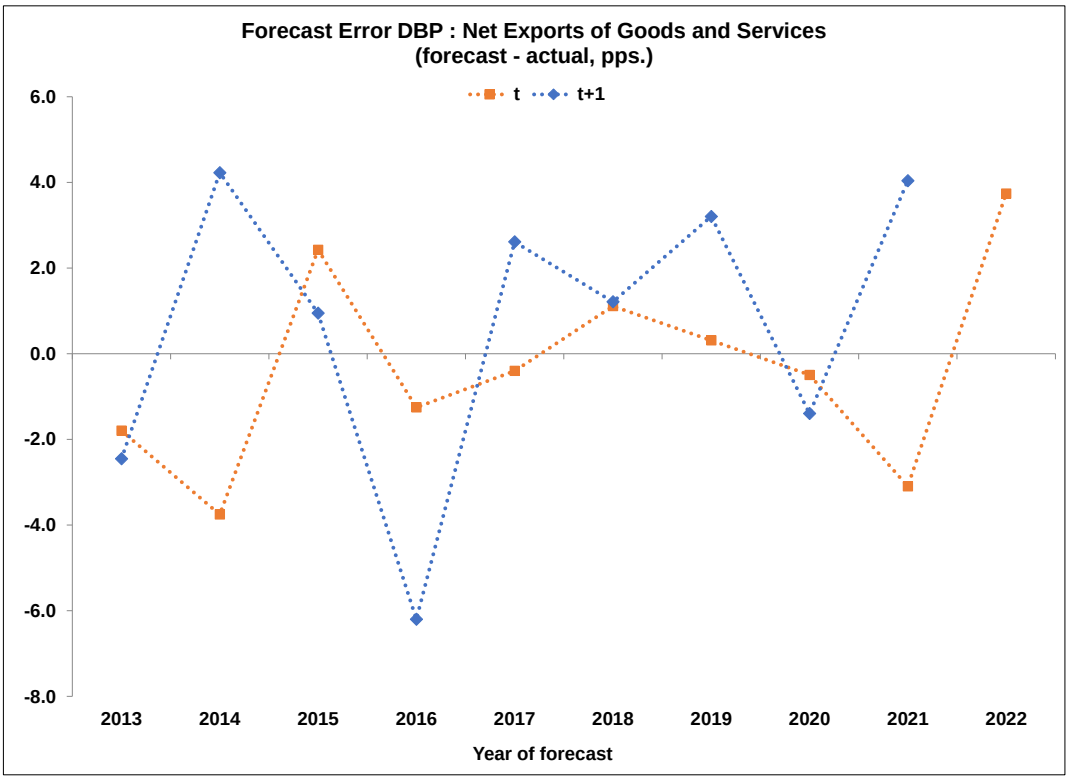


Chart 3.16: Forecast Error: Net Exports (forecast – actual, pps.) (DBP)



4 Empirical results and analysis

In this section, the main empirical findings are set out and explained. While going through this section, it is important to keep certain limitations in mind. In particular, the sample size available is relatively small, especially for the autumn forecast round. Furthermore, comparisons to other institutions should be made cautiously as these may have different sample sizes.

This section outlines the accuracy, unbiasedness, and efficiency of MFE forecasts for nominal GDP, real GDP, and its expenditure components (including net exports). The accuracy and unbiasedness tests are compared with those derived from tests on the COM and the CBM forecasts. Finally, the authors assess the statistical revisions since an important caveat is that the cause of the forecast errors, inaccuracy and biases reported in this analysis are not strictly all explained by errors made by the forecaster at the time of the forecasting exercise but are also partly attributable to statistical revisions in the data.

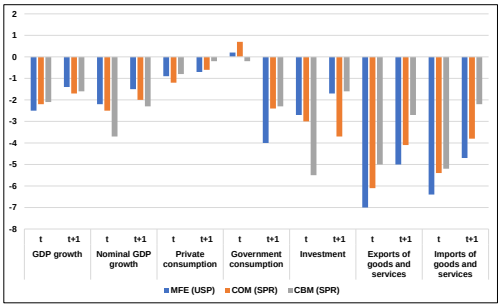
4.1 Accuracy of MFE forecasts and benchmarking

As described in Section 2.3, forecast accuracy is measured over a given period by measuring several statistics. The mean error is not the best measure available to gauge forecast accuracy. However, it is the only measure which provides a sense of the direction of the error. As the methodology indicates, the RMSE is considered superior to the other summary statistics in measuring forecast accuracy and is given more weight in the analysis. Notwithstanding this, the Mean Error and Theil's U statistic are also discussed (results obtained for the MAE and the MRAE are shown in Appendix B).

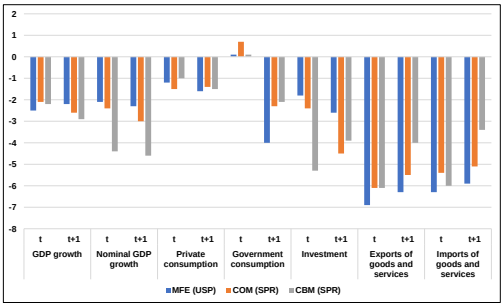
The mean error was calculated for MFE in both the spring and the autumn forecast rounds, including all the samples and excluding the crises periods. This is also benchmarked to the forecasting performance of the COM and the CBM for the current and one-year ahead forecasts. The results obtained for nominal GDP, real GDP and the components of real GDP can be reviewed in Chart 4.1. In general, the forecast error appears to be, on average, negative for nominal GDP, real GDP, and the majority of its components. Thus, the outturn is generally higher than initially projected for that year. This indicates an element of underestimation which is present for all institutions. In the case of the MFE, some element of overestimation is observed for the forecast of government consumption for the same year the forecast is being published. This is also observed for the COM in both spring and autumn and the CBM in autumn, when including the full sample.

Chart 4.1: Mean error

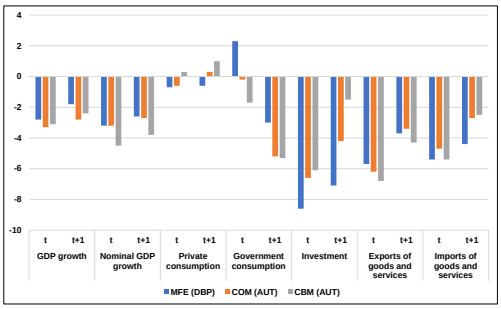
Spring forecast round – full sample



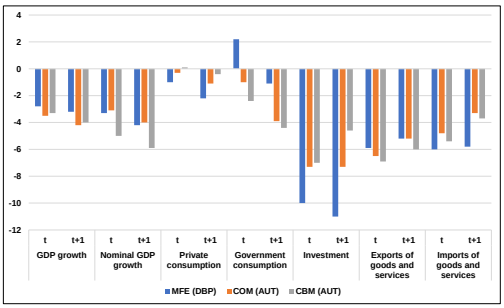
Spring forecast round – excl. crises



Autumn forecast round – full sample



Autumn forecast round – excl. crises



Focusing on the RMSE statistic, it is noted that the RMSE is generally lower in the spring forecast round when compared to the autumn vintage. Also, the forecast error in the one-year ahead forecast tends to be slightly higher across all institutions (See Chart 4.2).[17] This is expected as assumptions considered in the forecast are more susceptible to revisions, and information is more limited the further away the forecast is from the year in which it is made.

When comparing the RMSE of the MFE forecast with that of the other institutions, similar patterns and magnitudes of inaccuracy are observed. In general, the forecast for investment is the least accurate which is indicated by a higher RMSE. At the same time, exports and imports of goods and services also have a relatively high RMSE.[18] At time t, private consumption appears to be the most accurate component. The patterns indicated by MFE are also reflected in those of other institutions.

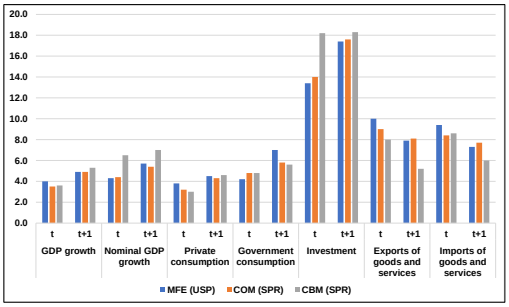
When excluding the crises periods, most components indicate a more accurate forecast over time. This suggests that forecast accuracy tends to improve when eliminating the crises years (both the financial and COVID-19 crises) from the sample.

[17] A low RMSE implies a more accurate forecast, while a high RMSE signifies a less accurate forecast.

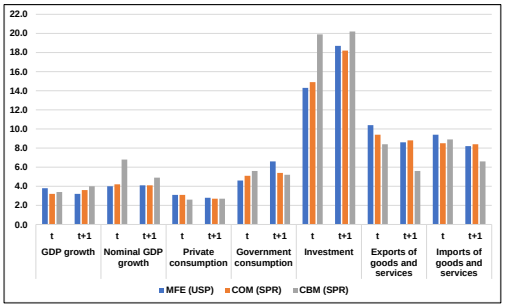
[18] As indicated in section 4.4 investment, exports of goods and services and imports of goods and services also have the highest amount of revisions, which should be kept in mind when interpreting these results.

Chart 4.2: Root Mean Squared Error

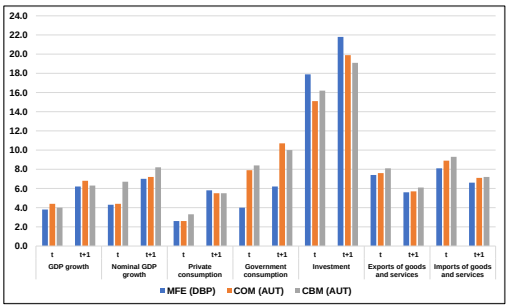
Spring forecast round – full sample



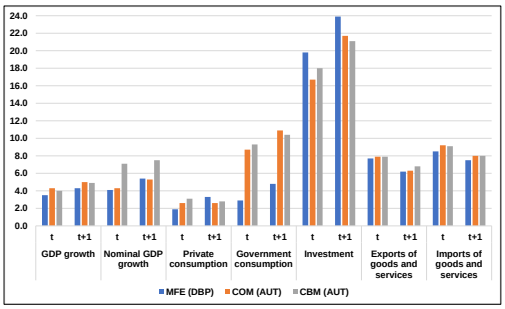
Spring forecast round – excl. crises



Autumn forecast round – full sample



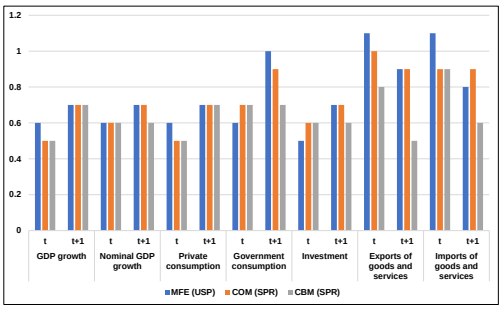
Autumn forecast round – excl. crises



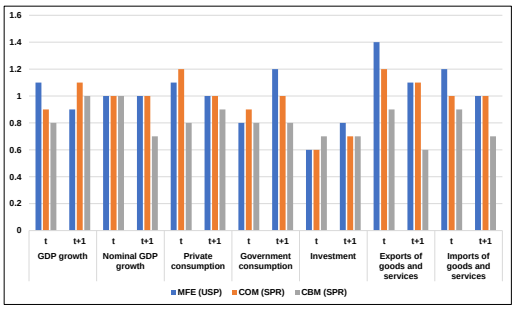
Another accuracy measure which provides useful insights is Theil's U statistic. From this statistic, it is observed that the estimations carried out by the various institutions across both forecast rounds are, in general, more accurate than a simple naïve forecast, when analysing the full sample, as Theil's U statistic is much closer to one (See Chart 4.3). In the sample which excludes the crises years, in some instances the Theil's U statistic also exceeds one, meaning that it is less accurate than a simple forecasting model. It is also noted that the forecasts of the one-year ahead forecasts are generally less accurate than the current year forecast when compared to a naïve forecast.

Chart 4.3: Theil's U-statistic

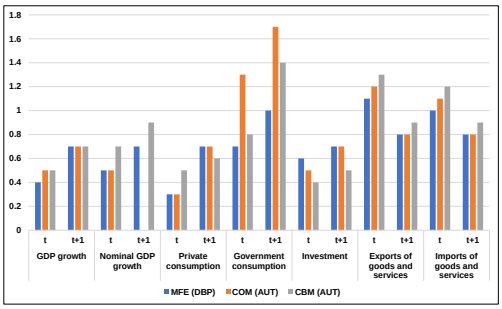
Spring forecast round – full sample



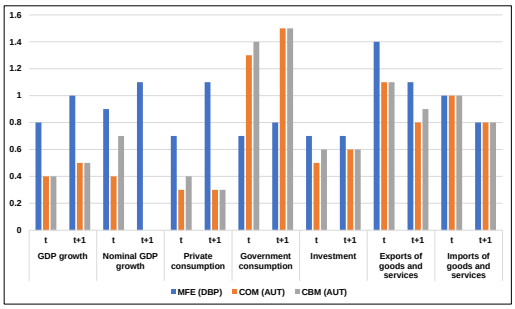
Spring forecast round – excl. crises



Autumn forecast round – full sample



Autumn forecast round – excl. crises



An interesting result is that in both the spring and the autumn forecast round, when using the full sample, government consumption at time t tends to be more accurate when compared to a naïve forecast and to the other institutions. With the removal of the crisis periods, the forecast for government consumption produced by the COM and CBM turns out less accurate when compared to a naïve forecast in both the current year and the one-year ahead forecasts. In contrast, the MFE forecast remains more accurate than the naïve forecast.[19]

4.2 Unbiasedness of MFE forecasts and benchmarking

Results for forecast unbiasedness are presented in Tables 4.1 to 4.7. These tables present regression results for the MFE, COM and CBM for both the spring and the autumn forecast rounds across different sample periods and include results with crises and without the crises years.

The results for both real and nominal GDP show significant instances of bias based on a confidence interval at the 95% and 99% level (See Table 4.1 and 4.2). Indeed, the implication is that forecasts appear to have been systematically too low overall.

When analysing the unbiasedness in the spring forecast round for the MFE considering the full sample without excluding the crises periods, the authors find evidence of downward biases for year t in real and nominal GDP forecasts, whereas the forecasts for year t+1 to year t+3 did not feature any bias. Results for the Draft Budgetary Plan forecast period show a similar result: over the full sample, both real and nominal GDP were underestimated in year t, with no bias detected in year t+1. Similar results were observed for the other institutions under review, with the only exception being the CBM forecast for nominal GDP at time t as the spring forecast round did not feature any bias, while for real GDP, the CBM autumn forecast round indicates no bias in the current year forecast but does indicate a bias in the one-year ahead forecast.

[19] The fact that the CBM and COM are in general producing less accurate forecasts for government consumption may be the result of less information being available to these institutions as compared to the government.

Table 4.1 Nominal GDP unbiasedness test

		Full sample				Excl. crises			
	Sample	t	t+1	t+2	t+3	t	t+1	t+2	t+3
MFE (USP)	2004 2022	-2.16**	-1.49	-1.89	-2.39	-2.08**	-2.31**	-3.15**	-3.83***
		(0.02)	(0.28)	(0.20)	(0.11)	(0.04)	(0.03)	(0.01)	(0.00)
COM (Spring)	2004 2022	-2.55**	-2.03	Na	Na	-2.44**	-2.96***	Na	Na
		(0.01)	(0.11)			(0.02)	(0.00)		
CBM (Spring)	2015 2022	-2.18	-3.71	Na	Na	-4.37	-4.52**	Na	Na
		(0.39)	(0.11)			(0.12)	(0.01)		
MFE (DBP)	2013 2022	-3.21**	-2.57	Na	Na	-3.25**	-4.17**	Na	Na
		(0.01)	(0.30)			(0.01)	(0.03)		
COM (Autumn)	2013 2022	-3.16**	-2.68	Na	Na	-3.11**	-4.00**	Na	Na
		(0.01)	(0.29)			(0.03)	(0.03)		
CBM (Autumn)	2013 2022	-3.10**	-2.36	Na	Na	-3.30**	-4.01**	Na	Na
		(0.01)	(0.29)			(0.01)	(0.01)		

Note: p-value in parenthesis; a p-value of less than 0.01 indicates the presence of bias at the 99% confidence interval (***), a p-value of less than 0.05 indicates the presence of bias at the 95% confidence interval (**), while a p-value of less than 0.1 indicates the presence of bias at the 90% confidence interval (*).

Notably, when excluding the crises years from the sample, the downward bias in the forecasts also tends to persist in the outer years. Indeed, in the spring forecast period, nominal GDP is biased downwards from t to t+3. On the other hand, in the case of real GDP, the year t to year t+2 forecasts appear to have been biased downwards, while no bias was found in year t+3.

Table 4.2 Real GDP unbiasedness test

		Full sample				Excl. crises			
	Sample	t	t+1	t+2	t+3	t	t+1	t+2	t+3
MFE (USP)	2004 2022	-2.53***	-1.42	-1.86	-1.91	-2.55***	-2.21***	-3.15***	-3.13***
		(0.00)	(0.23)	(0.18)	(0.14)	(0.00)	(0.00)	(0.00)	(0.00)
COM (Spring)	2004 2022	-2.21***	-1.73	Na	Na	-2.15***	-2.65***	Na	Na
		(0.00)	(0.14)			(0.00)	(0.00)		
CBM (Spring)	2008 2022	-2.07**	-1.60	Na	Na	-2.24**	-2.87**	Na	Na
		(0.02)	(0.27)			(0.02)	(0.01)		
MFE (DBP)	2013 2022	-2.80**	-1.77	Na	Na	-2.79**	-3.16**	Na	Na
		(0.01)	(0.42)			(0.01)	(0.04)		
COM (Autumn)	2013 2022	-3.34**	-2.80	Na	Na	-3.50**	-4.19**	Na	Na
		(0.01)	(0.24)			(0.01)	(0.01)		
CBM (Autumn)	2013 2022	-3.56	-4.44**	Na	Na	-4.94**	-5.87**	Na	Na
		(0.16)	(0.03)			(0.04)	(0.02)		

Note: p-value in parenthesis; a p-value of less than 0.01 indicates the presence of bias at the 99% confidence interval (***), a p-value of less than 0.05 indicates the presence of bias at the 95% confidence interval (**), while a p-value of less than 0.1 indicates the presence of bias at the 90% confidence interval (*).

The DBP forecast period shows comparable results, whereby year t and t+1 forecasts were biased downwards. The increased biasedness noted in the forecasts by the three institutions when excluding the crises years is, to a certain extent, understandable. Indeed, a reason why with the inclusion of all the data bias seems to be less could be that since, in general the MFE tend to underestimate their forecasts for nominal and real GDP, the inclusion of years where the actual data would have been lower due to a crisis year could mean that the MFE forecast would have been closer to actual data. Thus, the element of biasedness would decrease. In general, the increased biasedness indicated by the MFE forecasts when excluding the crises years seems to be replicated by the other institutions.

Turning to the separate components of real GDP, no bias for the MFE forecasts for private consumption was observed when considering the full sample. However, when excluding the crises periods similar to what we identified in real and nominal GDP, the intercept widens, becoming more negative, and forecasts at t+1 to t+3 appear to be biased downwards in the spring forecast round. This is also replicated across all the other institutions. On the other hand, the Autumn forecast round shows less bias, with only forecasts at t+1 produced by the MFE indicating a bias at the 90% confidence level.

Table 4.3 Private consumption unbiasedness test

		Full sample				Excl. crises			
	Sample	t	t+1	t+2	t+3	t	t+1	t+2	t+3
MFE (USP)	2004 2022	-0.91	-0.73	-0.31	-0.80	-1.20	-1.55**	-1.33*	-1.92**
		(0.31)	(0.51)	(0.79)	(0.50)	(0.13)	(0.03)	(0.10)	(0.04)
COM (Spring)	2004 2022	-1.23*	-0.54	Na	Na	-1.45*	-1.38**	Na	Na
		(0.09)	(0.61)			(0.06)	(0.05)		
CBM (Spring)	2008 2022	-0.82	-0.17	Na	Na	-1.04	-1.75**	Na	Na
		(0.30)	(0.90)			(0.20)	(0.03)		
MFE (DBP)	2013 2022	-0.70	-0.61	Na	Na	-0.96	-2.23*	Na	Na
		(0.42)	(0.77)			(0.18)	(0.06)		
COM (Autumn)	2013 2022	-0.57	0.31	Na	Na	-0.33	-1.09	Na	Na
		(0.53)	(0.88)			(0.75)	(0.31)		
CBM (Autumn)	2013 2022	0.29	1.04	Na	Na	0.14	-0.43	Na	Na
		(0.80)	(0.60)			(0.91)	(0.72)		

Note: p-value in parenthesis; a p-value of less than 0.01 indicates the presence of bias at the 99% confidence interval (***), a p-value of less than 0.05 indicates the presence of bias at the 95% confidence interval (**), while a p-value of less than 0.1 indicates the presence of bias at the 90% confidence interval (*).

When assessing the government consumption forecasts for unbiasedness, the authors find that the forecasts by the MFE appear to be more biased when compared to other institutions. The downward bias indicated in the outer years suggests that government consumption forecasted in t+1, t+2 and t+3 generally tends to be higher than the actual consumption. When removing the crises periods in our analysis, the biasedness is similar to that indicated with the full sample and the intercept remains similar. The reasoning behind this is that contrary to the other components, this component tends to increase at times of crises as the government spends more to maintain economic activity in the country; the concept of counter-cyclical fiscal policy.

Table 4.4 Government consumption unbiasedness test

		Full Sample				Excl. crises			
	Sample	t	t+1	t+2	t+3	t	t+1	t+2	t+3
MFE (USP)	2004 2022	0.21	-4.02**	-4.39**	-3.83**	0.14	-3.99**	-4.41***	-3.83*
		(0.84)	(0.01)	(0.01)	(0.03)	(0.91)	(0.02)	(0.01)	(0.06)
COM (Spring)	2004 2022	0.68	-2.38*	Na	Na	0.69	-2.31	Na	Na
		(0.55)	(0.08)			(0.62)	(0.11)		
CBM (Spring)	2008 2022	-0.16	-2.35	Na	Na	0.05	-2.06	Na	Na
		(0.90)	(0.12)			(0.98)	(0.23)		
MFE (DBP)	2013 2022	2.32*	-2.96	Na	Na	2.20	-1.09	Na	Na
		(0.06)	(0.16)			(0.15)	(0.59)		
COM (Autumn)	2013 2022	-0.24	-5.23	Na	Na	-0.96	-3.94	Na	Na
		(0.93)	(0.15)			(0.78)	(0.38)		
CBM (Autumn)	2013 2022	-1.70	-5.28	Na	Na	-2.43	-4.43	Na	Na
		(0.55)	(0.12)			(0.50)	(0.29)		

Note: p-value in parenthesis; a p-value of less than 0.01 indicates the presence of bias at the 99% confidence interval (***), a p-value of less than 0.05 indicates the presence of bias at the 95% confidence interval (**), while a p-value of less than 0.1 indicates the presence of bias at the 90% confidence interval (*).

Turning to gross fixed capital formation, no degree of biasedness is detected. However, this does not necessarily indicate that forecast errors are not present within the forecasts. Indeed, this component has several extremely positive forecast errors and extremely negative forecast errors with a high variance. Thus, in this case, these extreme values seem to be cancelling each other and as a result, this test does not indicate any bias.

Table 4.5 Gross fixed capital formation unbiasedness test

		Full Sample				Excl. crises			
	Sample	t	t+1	t+2	t+3	t	t+1	t+2	t+3
MFE (USP)	2004 2022	-2.71	-1.69	-3.57	-5.48	-1.77	-2.56	-4.40	-6.02
		(0.39)	(0.69)	(0.45)	(0.24)	(0.65)	(0.63)	(0.45)	(0.31)
COM (Spring)	2004 2022	-2.98	-3.72	Na	Na	-2.37	-4.54	Na	Na
		(0.37)	(0.39)			(0.56)	(0.37)		
CBM (Spring)	2008 2022	-5.54	-1.61	Na	Na	-5.25	-3.87	Na	Na
		(0.25)	(0.76)			(0.41)	(0.57)		
MFE (DBP)	2013 2022	-8.58	-7.11	Na	Na	-10.01	-10.97	Na	Na
		(0.13)	(0.36)			(0.17)	(0.25)		
COM (Autumn)	2013 2022	-6.61	-4.21	Na	Na	-7.26	-7.29	Na	Na
		(0.18)	(0.56)			(0.24)	(0.42)		
CBM (Autumn)	2013 2022	-6.05	-1.51	Na	Na	-6.96	-4.57	Na	Na
		(0.26)	(0.83)			(0.30)	(0.61)		

Note: p-value in parenthesis; a p-value of less than 0.01 indicates the presence of bias at the 99% confidence interval (***), a p-value of less than 0.05 indicates the presence of bias at the 95% confidence interval (**), while a p-value of less than 0.1 indicates the presence of bias at the 90% confidence interval (*).

On the other hand, exports and imports of goods and services have the most significant downward bias among the components, mainly at the 99% confidence level, with a high negative intercept value (See Table 4.6 and 4.7). Both the COM and the CBM have a similar level of bias but mainly at the 95% confidence level.

Table 4.6 Exports of goods and services unbiasedness test

		Full sample				Excl. crises			
	Sample	t	t+1	t+2	t+3	t	t+1	t+2	t+3
MFE (USP)	2004 2022	-6.98***	-5.03***	-5.24**	-4.39**	-6.93***	-6.31***	-6.81***	-6.07***
		(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
COM (Spring)	2004 2022	-6.10***	-4.10**	Na	Na	-6.11**	-5.54**	Na	Na
		(0.00)	(0.03)			(0.01)	(0.01)		
CBM (Spring)	2008 2022	-4.99**	-2.75**	Na	Na	-6.15**	-3.97**	Na	Na
		(0.01)	(0.04)			(0.01)	(0.01)		
MFE (DBP)	2013 2022	-5.72**	-3.73**	Na	Na	-5.93**	-5.17**	Na	Na
		(0.01)	(0.03)			(0.02)	(0.01)		
COM (Autumn)	2013 2022	-6.18***	-3.44*	Na	Na	-6.51**	-5.19**	Na	Na
		(0.00)	(0.07)			(0.01)	(0.01)		
CBM (Autumn)	2013 2022	-6.82***	-4.29**	Na	Na	-6.93***	-6.01***	Na	Na
		(0.00)	(0.02)			(0.00)	(0.00)		

Note: p-value in parenthesis; a p-value of less than 0.01 indicates the presence of bias at the 99% confidence interval (***), a p-value of less than 0.05 indicates the presence of bias at the 95% confidence interval (**), while a p-value of less than 0.1 indicates the presence of bias at the 90% confidence interval (*).

Table 4.7 Imports of goods and services unbiasedness test

		Full sample				Excl. crises			
	Sample	t	t+1	t+2	t+3	t	t+1	t+2	t+3
MFE (USP)	2004 2022	-6.43***	-4.75***	-5.19**	-4.93***	-6.28***	-5.89***	-6.33***	-6.33***
		(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
COM (Spring)	2004 2022	-5.44***	-3.78**	Na	Na	5.39**	-5.06**	Na	Na
		(0.00)	(0.03)			(0.01)	(0.02)		
CBM (Spring)	2008 2022	-5.21**	-2.24	Na	Na	-6.05**	-3.47*	Na	Na
		(0.01)	(0.17)			(0.02)	(0.10)		
MFE (DBP)	2013 2022	-5.41**	-4.42**	Na	Na	-5.99**	-5.80**	Na	Na
		(0.02)	(0.04)			(0.03)	(0.02)		
COM (Autumn)	2013 2022	-4.71*	-2.71	Na	Na	-4.86	-3.31	Na	Na
		(0.10)	(0.27)			(0.15)	(0.31)		
CBM (Autumn)	2013 2022	-5.40*	-2.51	Na	Na	-5.43*	-3.70	Na	Na
		(0.06)	(0.32)			(0.09)	(0.25)		

Note: p-value in parenthesis; a p-value of less than 0.01 indicates the presence of bias at the 99% confidence interval (***), a p-value of less than 0.05 indicates the presence of bias at the 95% confidence interval (**), while a p-value of less than 0.1 indicates the presence of bias at the 90% confidence interval (*).

4.3 Efficiency of MFE forecasts

As described in Section 2.3.3, this study conducts two tests of strong MFE forecast efficiency with the sample size excluding the crises periods. Forecasts are deemed 'strongly' efficient if forecast errors are uncorrelated with information known at the time the forecasts were made. The tests carried out are on the relationship between forecast errors and past forecast errors and past data outturns. The results for these tests are shown in tables 4.8 to 4.11. The tables show the estimated β_1 coefficients and associated p-values as defined in the methodology section.

For the one-year, two-year and three-year ahead forecasts we could not reject the hypothesis that the forecast errors were unrelated to past forecast errors i.e., that β1 is equal to 0 in the case of the MFE spring forecasts across all components (See Table 4.8). In the case of the current year forecasts, only real GDP was found to be positively related to past forecast errors, at the 90% confidence level. Chart 4.4 indicates that, for real GDP, the positive relationship between forecast errors for the current year and

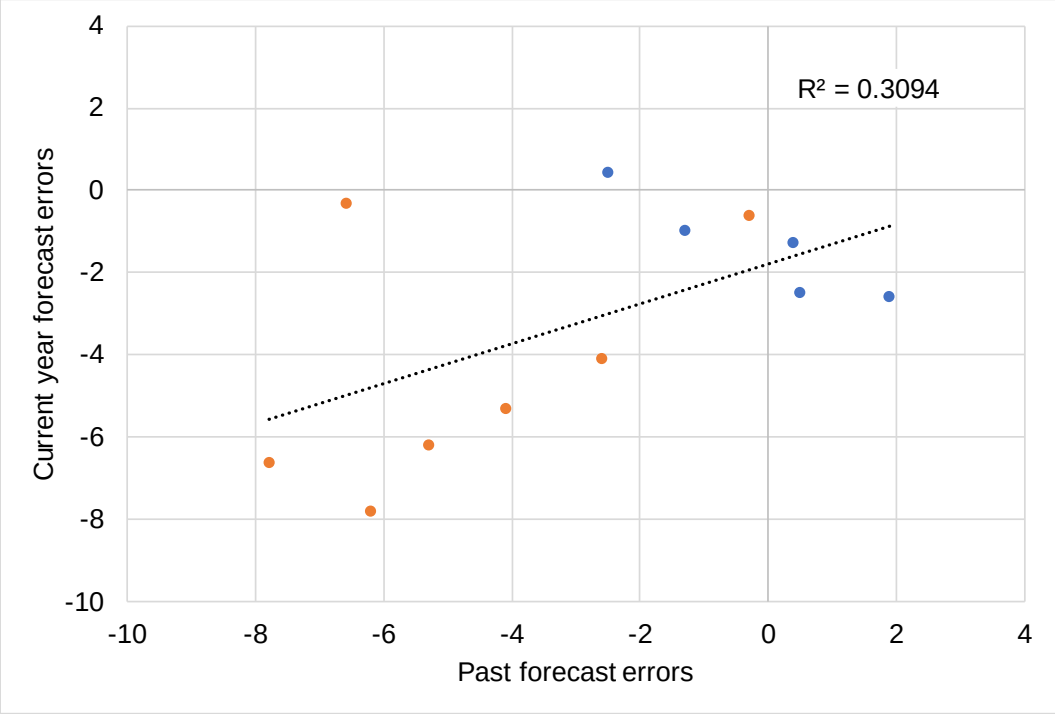
past forecast errors were mainly driven by post-2013 data. A positive relationship indicates that when the previous year forecast error was high the current year forecast error tended to be high as well. In contrast, a negative relationship indicated that when the previous forecast error was high, the current year forecast error was low.

Table 4.8 Strong efficiency of spring MFE forecasts (excl. crises periods)

	β_1 (p-value) from regressions of forecast errors on: Previous forecast errors			
	t	t+1	t+2	t+3
Nominal GDP	-0.03	0.31	0.05	-0.03
	(0.92)	(0.25)	(0.87)	(0.91)
Real GDP	0.49*	0.19	0.09	0.05
	(0.06)	(0.52)	(0.77)	(0.88)
Private consumption	0.09	0.23	0.22	0.35
	(0.83)	(0.54)	(0.61)	(0.38)
Government consumption	0.44	0.20	0.08	0.09
	(0.22)	(0.64)	(0.85)	(0.85)
Investment	-0.30	-0.14	-0.03	-0.04
	(0.30)	(0.67)	(0.94)	(0.91)
Exports of goods and services	0.39	0.08	0.07	0.25
	(0.24)	(0.83)	(0.81)	(0.64)
Imports of goods and services	0.38	-0.02	0.02	-0.06
	(0.29)	(0.97)	(0.96)	(0.89)

Note: p-value in parenthesis; a p-value of less than 0.01 indicates the presence of bias at the 99% confidence interval (***), a p-value of less than 0.05 indicates the presence of bias at the 95% confidence interval (**), while a p-value of less than 0.1 indicates the presence of bias at the 90% confidence interval (*).

Chart 4.4 MFE current-year real GDP growth forecast errors and past forecast errors



Note: orange dots indicate post 2013 relationship between current-year forecast errors with past forecast errors, while blue dots indicate the same relationship pre-2013.

A similar exercise was carried out using autumn data. Contrary to spring, for the current year forecasts, the hypothesis that the forecast errors were unrelated to past forecast

errors across all components could not be rejected. In the one-year ahead forecasts only exports of goods and services was found to be negatively related to past forecast errors at the 90% confidence level (See Table 4.9). In this case, this indicates that when the forecast error for the previous year was low, the current year forecast error tended to be high.

Table 4.9 Strong efficiency of Autumn MFE forecasts (excl. crises periods)

	β_1 (p-value) from regressions of forecast errors on: Previous forecast errors	
	t	t+1
Nominal GDP	-0.32	-0.41
	(0.55)	(0.49)
Real GDP	-0.35	-0.46
	(0.49)	(0.43)
Private consumption	-0.04	-0.14
	(0.93)	(0.84)
Government consumption	-0.08	0.20
	(0.87)	(0.73)
Investment	-0.49	0.02
	(0.32)	(0.98)
Exports of goods and services	-0.09	-0.61*
	(0.82)	(0.08)
Imports of goods and services	-0.13	-0.43
	(0.77)	(0.36)

Note: p-value in parenthesis; a p-value of less than 0.01 indicates the presence of bias at the 99% confidence interval (***), a p-value of less than 0.05 indicates the presence of bias at the 95% confidence interval (**), while a p-value of less than 0.1 indicates the presence of bias at the 90% confidence interval (*).

The same test for 'strong' efficiency was also conducted on past data outturns. For the one-year, two-year and three-year ahead forecasts we could not reject the hypothesis that the forecast errors were unrelated to past data outturns in the case of the MFE spring forecasts across all components (See Table 3.10). In the current-year forecasts we only find evidence of inefficiencies in the forecast for private consumption which forecast errors are found to be positively related to past data outturns at the 90% confidence interval. Similarly, for autumn, t+1 forecast errors were found to be unrelated with past data outturns, however the null hypothesis of 'strong' efficiency was rejected at the 95% confidence interval in the case of current year government consumption forecasts (+ve intercept). These indicate that when actual data for private consumption and government consumption was high, the forecast error for the following year tended to be high.

Table 4.10 Strong efficiency of spring MFE forecasts (excl. crises periods)

	β_1 (p-value) from regressions of forecast errors on: Previous data outturn			
	t	t+1	t+2	t+3
Nominal GDP	-0.05	0.02	0.08	-0.33
	(0.78)	(0.94)	(0.79)	(0.33)
Real GDP	-0.22	0.05	-0.07	0.31
	(0.19)	(0.83)	(0.82)	(0.53)
Private consumption	0.35*	-0.13	-0.30	-0.27
	(0.10)	(0.79)	(0.61)	(0.61)
Government consumption	0.32	0.20	-0.17	0.65
	(0.28)	(0.54)	(0.69)	(0.21)
Investment	0.02	-0.04	0.03	-0.53
	(0.92)	(0.89)	(0.93)	(0.28)
Exports of goods and services	-0.30	0.38	0.27	0.26
	(0.32)	(0.18)	(0.29)	(0.19)
Imports of goods and services	-0.24	-0.05	0.20	0.17
	(0.47)	(0.87)	(0.54)	(0.58)

Note: p-value in parenthesis; a p-value of less than 0.01 indicates the presence of bias at the 99% confidence interval (***), a p-value of less than 0.05 indicates the presence of bias at the 95% confidence interval (**), while a p-value of less than 0.1 indicates the presence of bias at the 90% confidence interval (*).

Table 4.11 Strong efficiency of autumn MFE forecasts (excl. crises periods)

	β_1 (p-value) from regressions of forecast errors on: Previous data outturn	
	t	t+1
Nominal GDP	-0.06	0.29
	(0.71)	(0.50)
Real GDP	-0.06	0.34
	(0.72)	(0.48)
Private consumption	0.08	-0.76
	(0.60)	(0.49)
Government consumption	0.50**	0.27
	(0.03)	(0.54)
Investment	0.04	0.23
	(0.90)	(0.60)
Exports of goods and services	-0.18	0.01
	(0.67)	(0.98)
Imports of goods and services	0.04	-0.23
	(0.93)	(0.50)

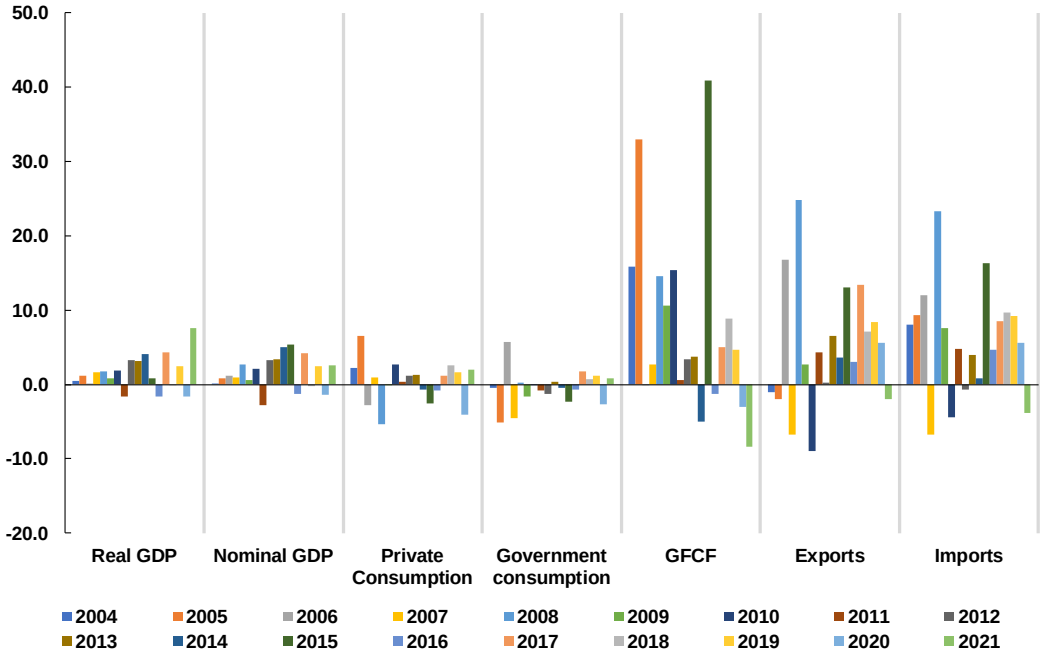
Note: p-value in parenthesis; a p-value of less than 0.01 indicates the presence of bias at the 99% confidence interval (***), a p-value of less than 0.05 indicates the presence of bias at the 95% confidence interval (**), while a p-value of less than 0.1 indicates the presence of bias at the 90% confidence interval (*).

4.4 Statistical revisions

It is important to note that statistical revisions can considerably influence forecast accuracy, unbiasedness, and efficiency. Indeed, revisions in the data will be included as part of the forecast error identified in this paper since forecast errors are mainly made up of errors attributed to statistical revisions and forecaster errors. Indeed, both real and nominal GDP and the components of GDP have undergone several revisions across the years, as can be visually identified from Chart 4.5 for the spring forecast round and Chart 4.6 for the autumn forecast round. The charts show the real growth in year t less the actual growth rate for that same year as reported in the latest national accounts news release. It is noted that most revisions have tended to be upward in

both the spring and the autumn forecast rounds; that is, the latest data has generally been revised upwards.[20]

Chart 4.5 Statistical revisions from the NSO release available in Spring to NR095/2023 – (pp)[21]



In the spring forecast round, the authors note that most of the revisions carried out were on gross fixed capital formation, exports of goods and services and imports of goods and services. The highest revision in investment was carried out for 2015, whereby growth was revised by +40.9 pp. In terms of revisions, exports of goods and services and imports of goods and services tend to mirror each other, with the highest revision for both carried out in 2008 and revised upwards by +24.9 pp and +23.3 pp, respectively. It is also interesting to note that, in general, government consumption tends to have more balanced revisions, averaging -0.1 pp across the period analysed. Overall, real and nominal GDP are revised upwards, averaging +1.5 pp and +1.6 pp, respectively.

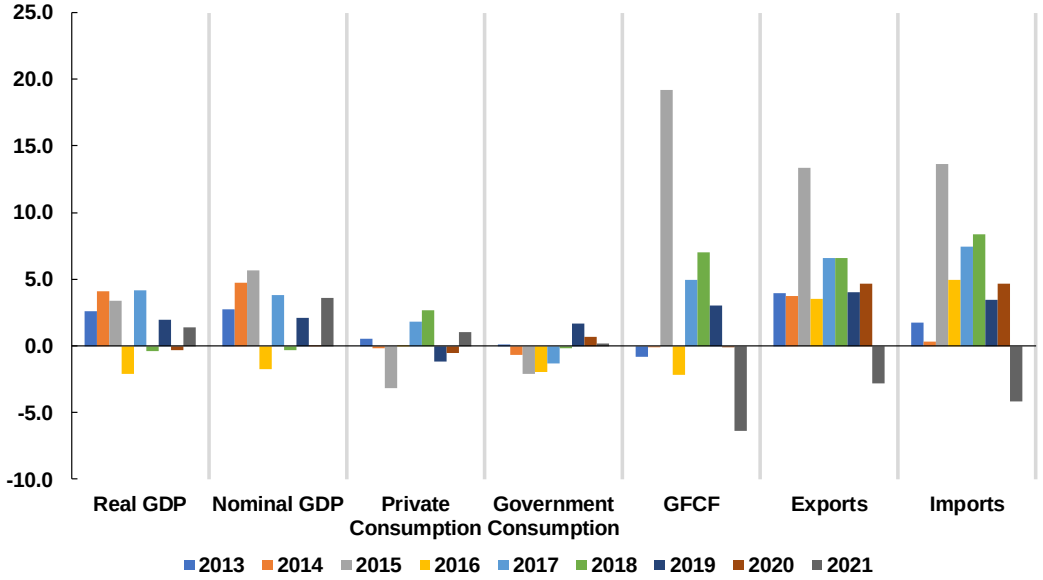
In the autumn forecast round revisions are also noted to be generally upwards, however the scale of revisions tended to be lower in magnitude. Nonetheless, the highest revisions were also carried out on gross fixed capital formation, exports of goods and services and imports of goods and services. The highest revision in these components was carried out for 2015 data, increasing the actual figures for 2015 by 19.2 pp, 13.3 pp and 13.7 pp, respectively. Similar to the data for spring, government consumption was quite balanced in its revisions, averaging -0.4 pp. For autumn data,

[20] A positive value means that the data published in the latest NSO release is higher than the data contained in the release available in Spring/Autumn.

[21] The NSO releases that were available in Spring/ Autumn refer to the last NSO release that was published prior to the publication of the USP/ DBP and used in the preparation of the forecasts.

real GDP and nominal GDP were also on average revised upwards more than the spring forecasts for the period under analysis, averaging 1.6 pp and 2.3 pp, respectively.

Chart 4.6 Statistical revisions from the NSO release available in Autumn to NR095/2023 – (pp)



These upward revisions in GDP statistics are supported by evidence of systematic bias at the 99% confidence interval (see Table 4.12).[22] This means that in most cases the latest data release shows a higher growth than what was available when the forecasting exercise was carried out. In the spring forecast round statistical revisions, the authors find that both real GDP and nominal GDP, investment and imports of goods and services are biased upwards at the 99% confidence interval, while exports of goods and services are biased upwards at the 95% confidence interval. It is interesting to note that no systematic bias is identified for private consumption and government consumption. On the other hand, systematic bias is also found present in the statistical revisions in the autumn forecast rounds, albeit less strong, which may be the result of lower sample size. Indeed, only exports of goods and services is found to be biased at the 99% confidence interval, while imports of goods and services and nominal GDP is biased at the 95% confidence interval. Real GDP is only found to be biased at the 90% confidence interval and investment, private consumption and government consumption are not found to be systematically biased.

[22] This was calculated by applying the unbiasedness test on the statistical revisions.

Table 4.12 Unbiasedness test of actual data

	Real GDP	Nominal GDP	Private Consumption	Public Consumption	Investment	Exports	Imports
Spring	1.53***	1.56***	1.01	-0.12	9.32***	4.19**	5.81***
	(0.01)	(0.01)	(0.26)	(0.86)	(0.01)	(0.05)	(0.00)
Autumn	1.64*	2.29**	0.11	-0.39	2.74	4.84***	4.51**
	(0.06)	(0.03)	(0.85)	(0.37)	(0.29)	(0.01)	(0.03)

Note: p-value in parenthesis; a p-value of less than 0.01 indicates the presence of bias at the 99% confidence interval (***), a p-value of less than 0.05 indicates the presence of bias at the 95% confidence interval (**), while a p-value of less than 0.1 indicates the presence of bias at the 90% confidence interval (*).

Since statistical revisions are sizable and tend to be inclined on the upside this may be a contributing factor in the bias, accuracy and efficiency results previously identified. It is also interesting to note that the current year forecast errors of real GDP and nominal GDP and all components for year t are found to be inversely related to the statistical revision of real and nominal GDP for year t-1. This relatively high inverse correlation indicates that statistical revisions are indeed affecting the accuracy of the forecasts. The extent to which forecast errors have been affected by statistical revisions will be explored further in forthcoming publications.

5 Conclusions and recommendations

This study aims to assess the forecasting performance of the MFE's macroeconomic projections, with a particular focus on the forecasts for real and nominal GDP growth and its expenditure components in real terms. The MFE's forecasting ability was evaluated by comparing the results to those of the COM and the CBM in terms of accuracy, unbiasedness, and efficiency.

The empirical results presented in this paper for accuracy show that MFE forecast errors for real GDP, nominal GDP, and most of its components tend to be underestimated in both the spring and the autumn forecast rounds. These results for the MFE are also evidenced by the other institutions.

This study also finds robust evidence of downward bias in MFE forecasts for exports and imports of goods and services. Real GDP and nominal GDP figures are also found to be biased downwards when removing the crises years. Gross fixed capital formation is the only component with no bias detected which may be the result of high variance present in the forecast errors for this variable. In general, unbiasedness tests carried out on the other institutions' forecasts delineated comparable results.

Efficiency tests did not indicate major 'strong' inefficiencies in MFE forecasts, thus suggesting that past data is sufficiently used in the production of forecasts. For the spring forecast, we only find real GDP at time t to be statistically significant at the 90% confidence interval when examining the relationship between past forecast errors and forecast errors at time t, t+1, t+2, and t+3. However, the same test performed on the autumn forecasts indicates that there is a degree of inefficiency in the forecast for exports of goods and services at time t+1 at the 90% confidence interval. According to the test of "strong" efficiency performed on previous outturns, only private consumption at time t in the spring forecast round and government consumption at time t in the autumn forecast round were identified to have an element of inefficiency in their projections at the 90% and 95% confidence intervals.

Nonetheless, statistical revisions which were carried out over time by the NSO, need to be taken into consideration. Although for the purposes of this paper we did not delve into deriving the actual forecaster error by eliminating the statistical revision error, we find an upward systematic bias in most of the components analysed. These upward revisions tend to push forecast errors to be more biased downwards and is thus important to factor this in when analysing the empirical results presented in this paper.

This analysis of the MFE's forecasting performance provides valuable insights for the MFAC. Indeed, regular assessments of the forecast performance of the projections produced by the MFE are key to identifying areas of improvement both in macroeconomic and fiscal projections and issuing recommendations and advice in that regard. To this end, such analysis needs to be carried out regularly and steps need to

be taken to rectify any inaccuracy, unbiasedness or inefficiencies identified during such exercise.

Bibliography

Bank of England. (2015). *Evaluating forecast performance*. Independent Evaluation Office. https://www.bankofengland.co.uk/-/media/boe/files/independent-evaluation-office/2015/evaluating-forecast-performance-november-2015.pdf

Camilleri, G., & Vella, K. J. (2015). *Interpolating forecast errors for assessing uncertainty in macroeconomic forecasts: An analysis for Malta*. (Working Paper 01/2015). Economic Policy Department. Ministry for Finance and Employment. https://economicpolicy.gov.mt/wp-content/uploads/2022/11/WP01-2015-Working_Paper_Full.pdf

Council Directive 2011/85/EU – Article 4(6) (2011). *On requirements for budgetary framework of the Member States.* Official Journal of the European Union.

Council Regulation (EU) No 1466/97 (1997). On the strengthening of the surveillance of budgetary positions and the surveillance and coordination of economic policies. Official Journal of the European Union.

Council Regulation (EU) No 473/2013 (2013). *On common provisions for monitoring and assessing draft budgetary plans and ensuring the correction of excessive deficit of the Member States in the euro area.* Official Journal of the European Union.

Fiscal Responsibility Act (2014), Malta. Government Press. Retrieved February 19, 2015.

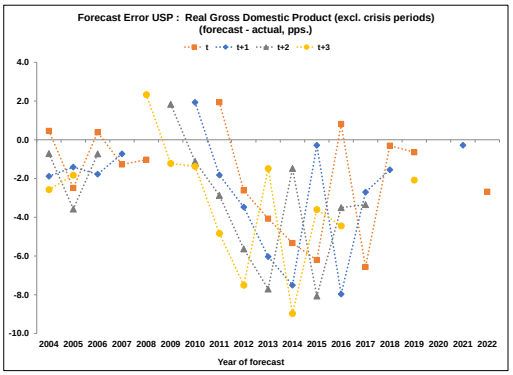
Frankel, J. (2011). *Over-optimism in Forecasts by Official Budget Agencies and its implications*. Oxford Review of Economic Policy 27(4): 536 – 562.

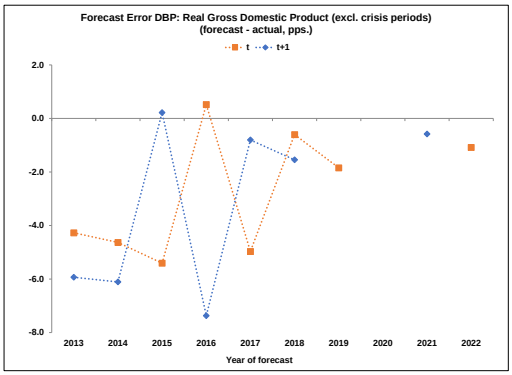
Larch, M., Orseau, E., and Van der Wielen, W. (2021). *Do EU fiscal rules support or hinder countercyclical fiscal policy?* Journal of International Money and Finance 112.

Appendices

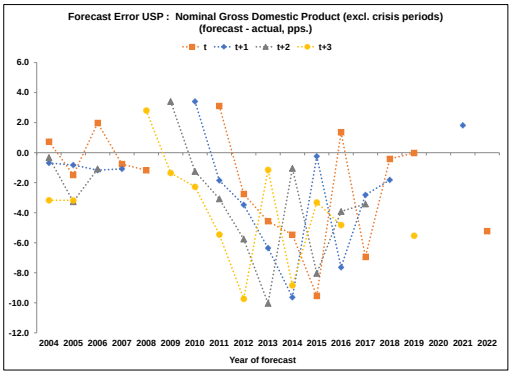
Appendix A

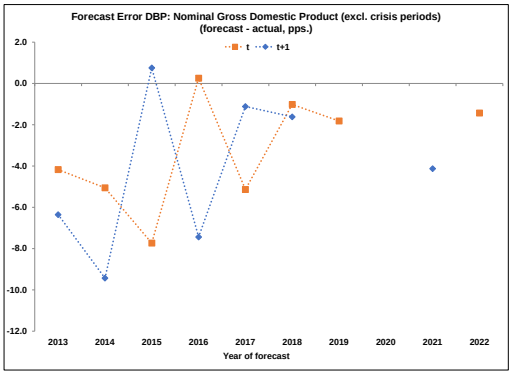
Charts A1: Forecasts Errors for Real GDP by MFE excluding crises



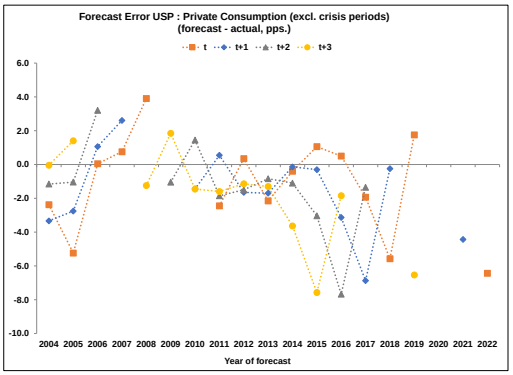


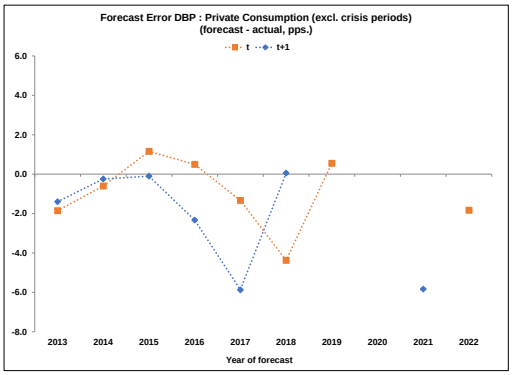
Charts A2: Forecasts Errors for Nominal GDP by MFE excluding crises



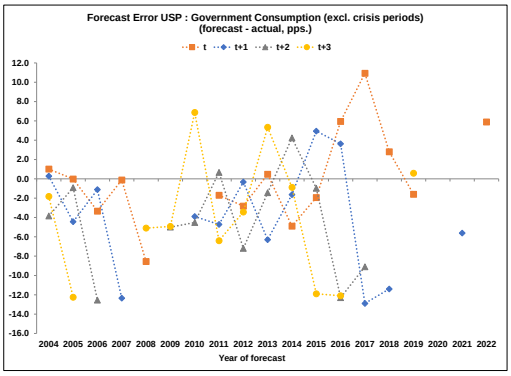


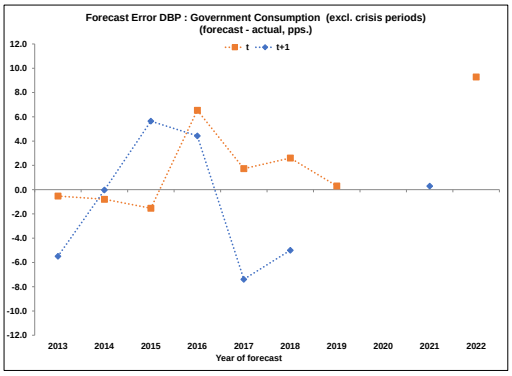
Charts A3: Forecasts Errors for Private Consumption by MFE excluding crises



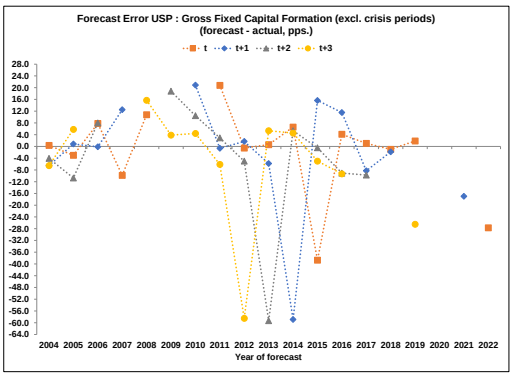


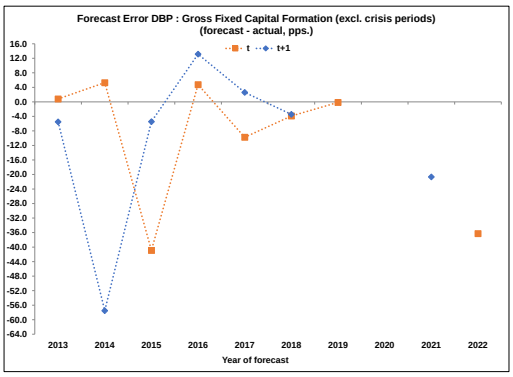
Charts A4: Forecasts Errors for Government Consumption by MFE excluding crises



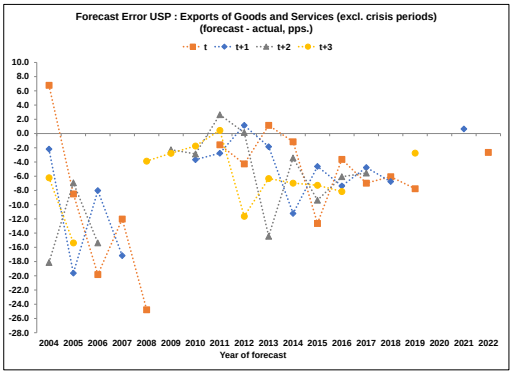


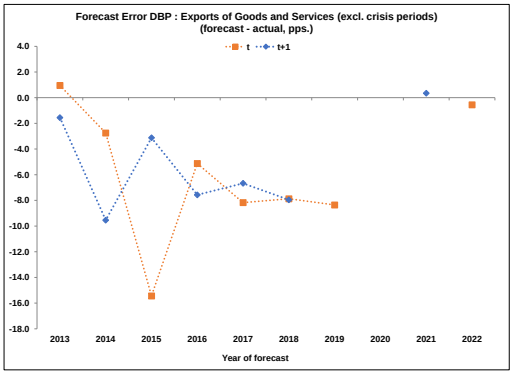
Charts A5: Forecasts Errors for Gross Fixed Capital Formation by MFE excluding crises



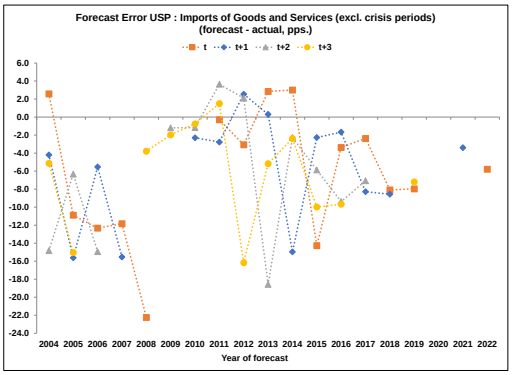


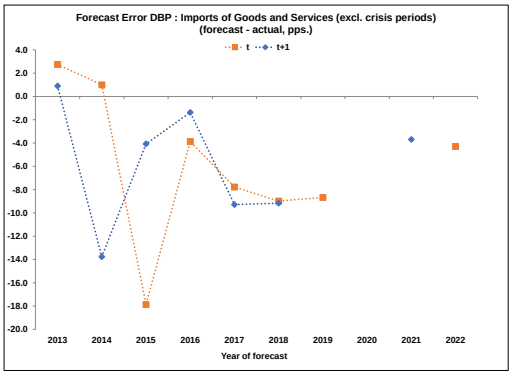
Charts A6: Forecasts Errors for Exports by MFE excluding crises



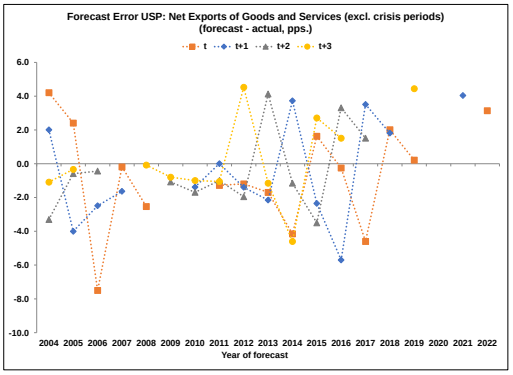


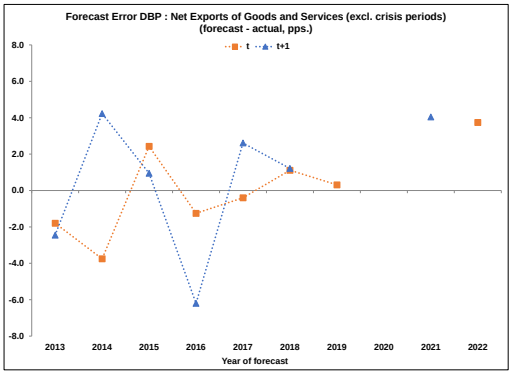
Charts A7: Forecasts Errors for Imports by MFE excluding crises



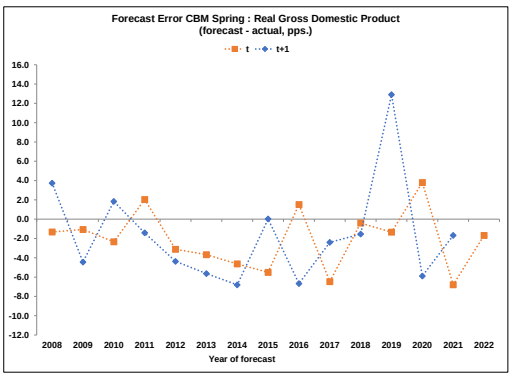


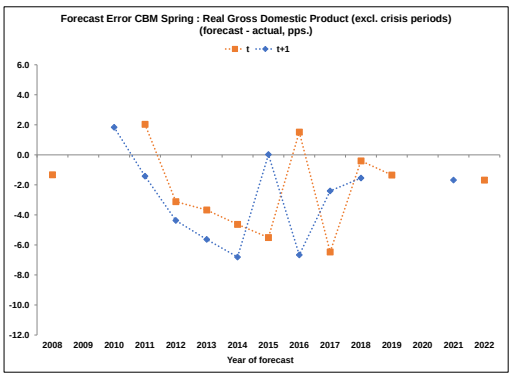
Charts A8: Forecasts Errors for Net Exports by MFE excluding crises



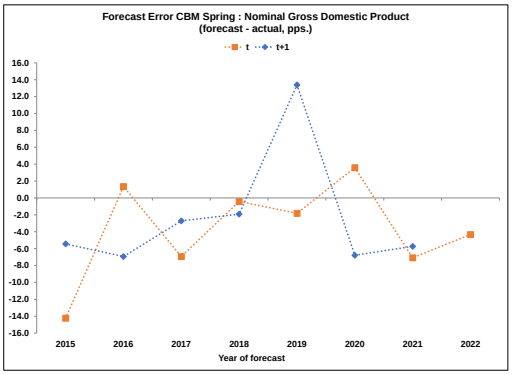


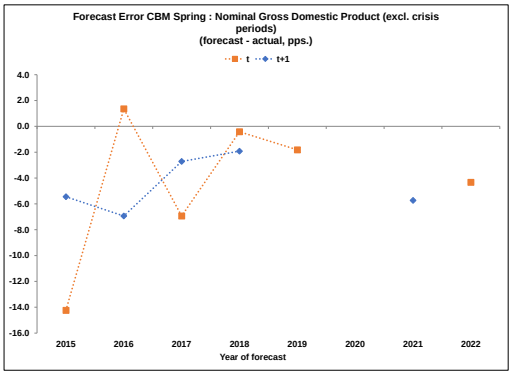
Charts A9: Forecasts Errors for Real GDP by CBM for spring including and excluding crises



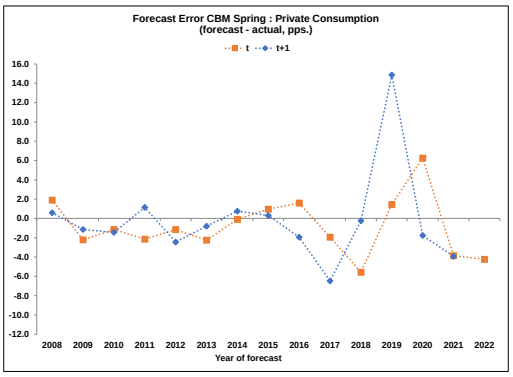


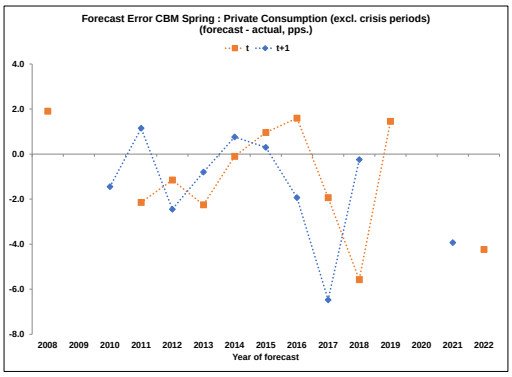
Charts A10: Forecasts Errors for Nominal GDP by CBM for spring including and excluding crises



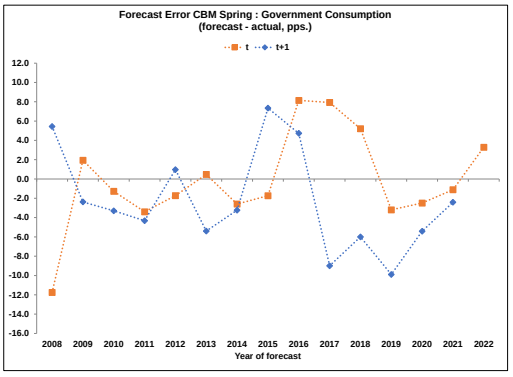


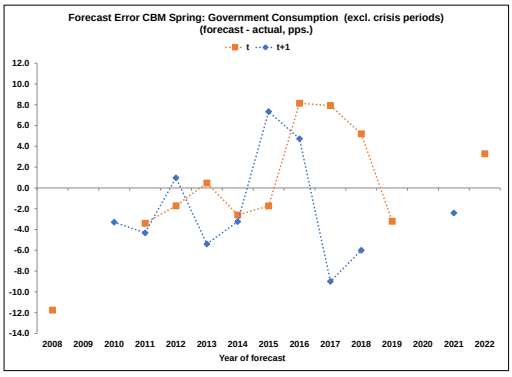
Charts A11: Forecasts Errors for Private Consumption by CBM for spring including and excluding crises



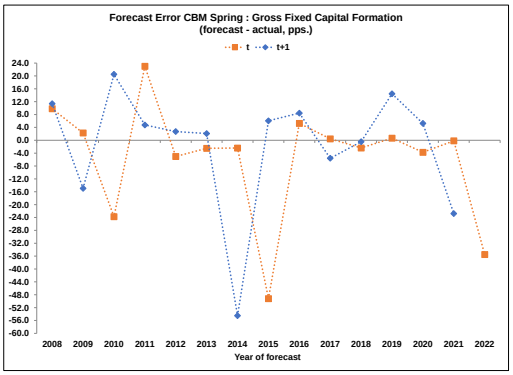


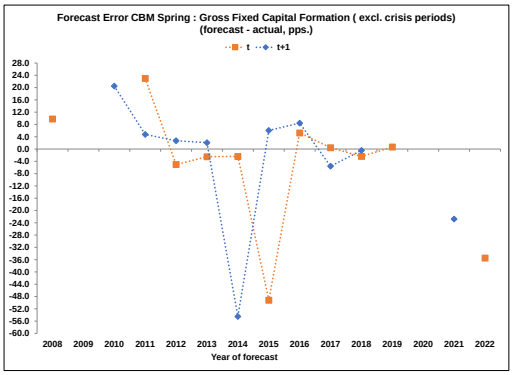
Charts A12: Forecasts Errors for Government Consumption by CBM for spring including and excluding crises



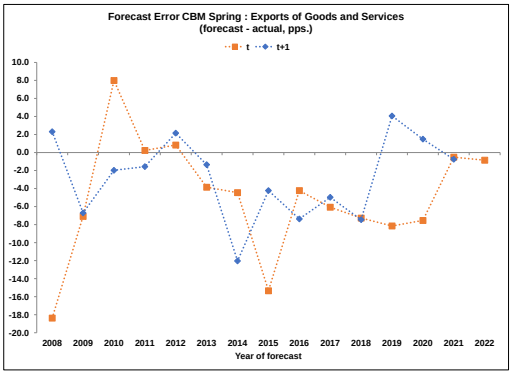


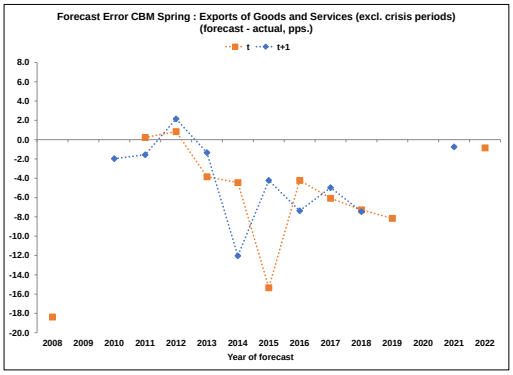
Charts A13: Forecasts Errors for Gross Fixed Capital Formation by CBM for spring including and excluding crises



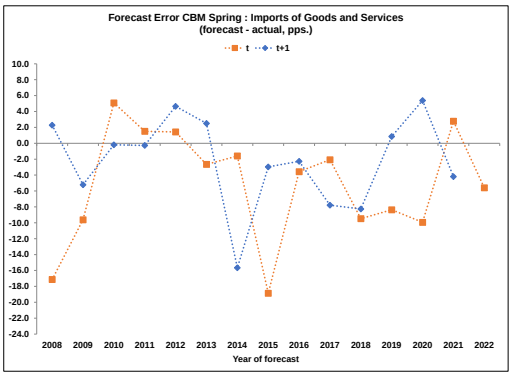


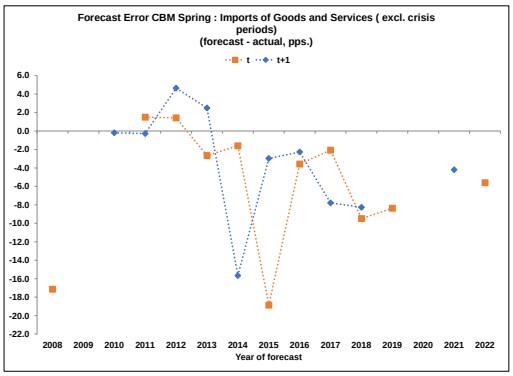
Charts A14: Forecasts Errors for Exports by CBM for spring including and excluding crises



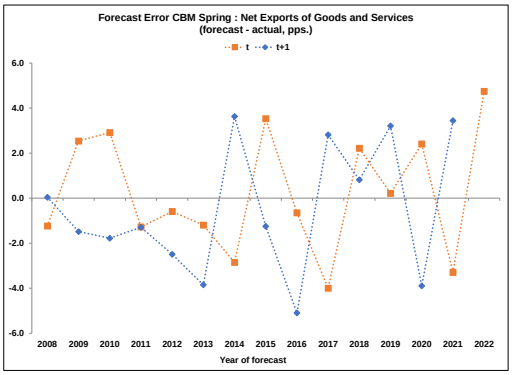


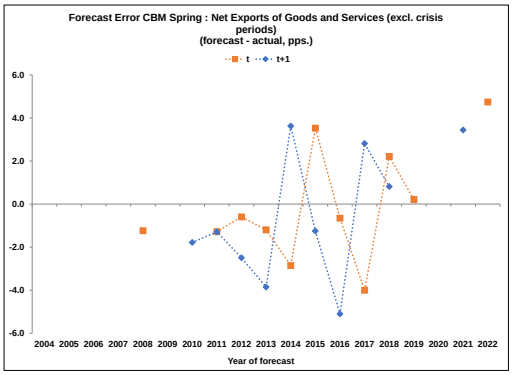
Charts A15: Forecasts Errors for Imports by CBM for spring including and excluding crises



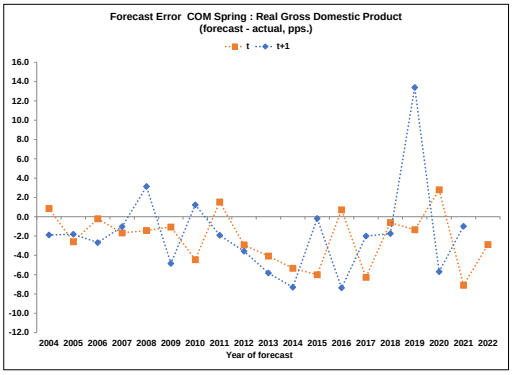


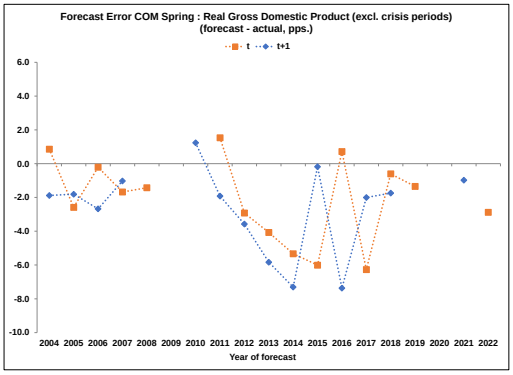
Charts A16: Forecasts Errors for Net Exports by CBM for spring including and excluding crises



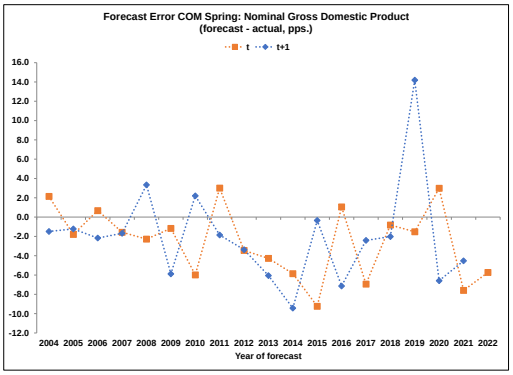


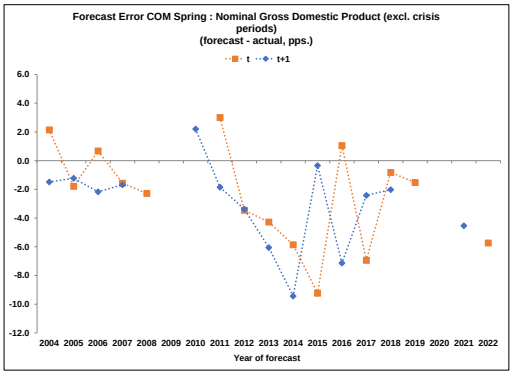
Charts A17: Forecasts Errors for Real GDP by COM for spring including and excluding crises



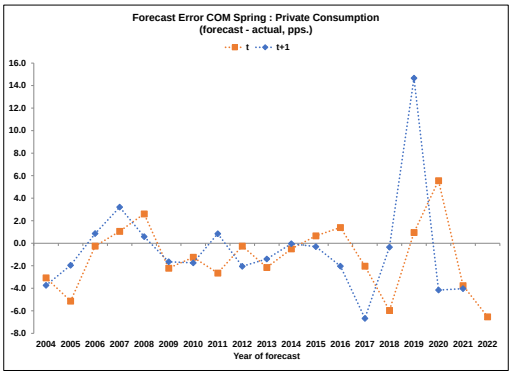


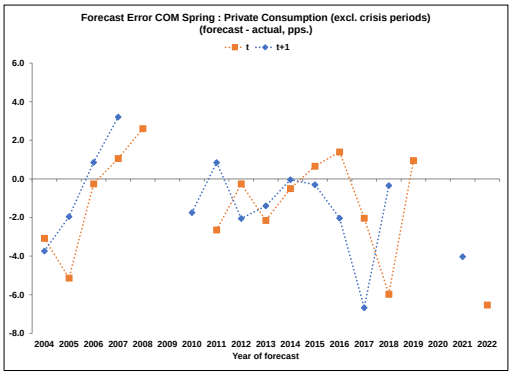
Charts A18: Forecasts Errors for Nominal GDP by COM for spring including and excluding crises



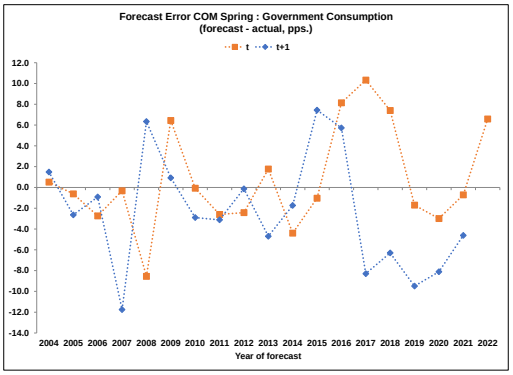


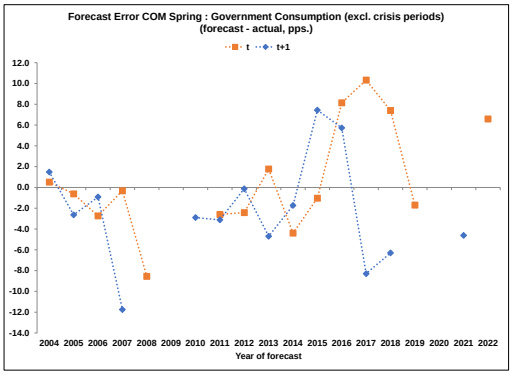
Charts A19: Forecasts Errors for Private Consumption by COM for spring including and excluding crises



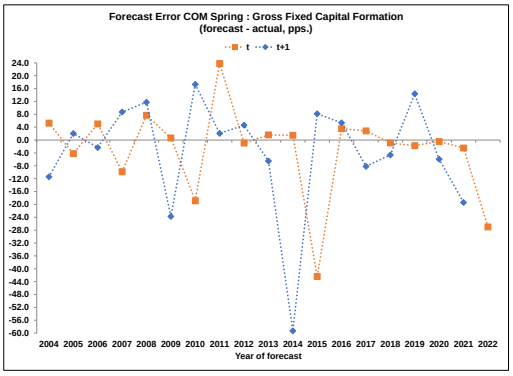


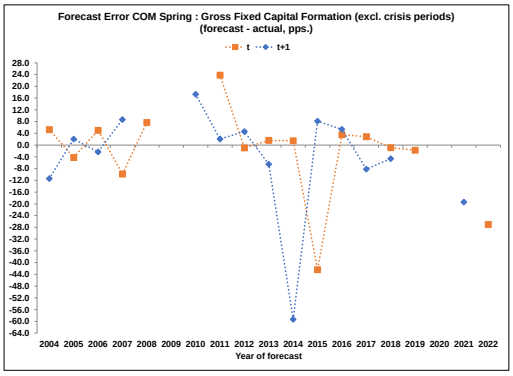
Charts A20: Forecasts Errors for Government Consumption by COM for spring including and excluding crises



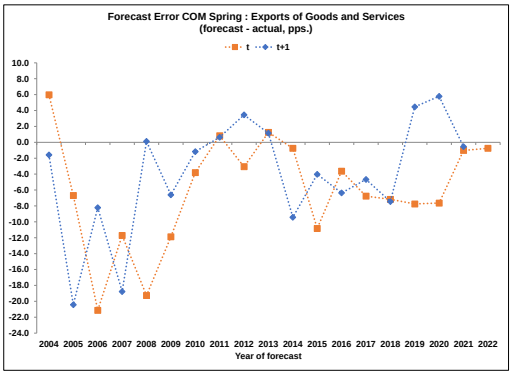


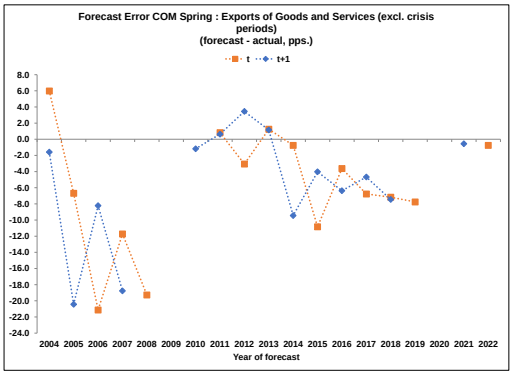
Charts A21: Forecasts Errors for Gross Fixed Capital Formation by COM for spring including and excluding crises



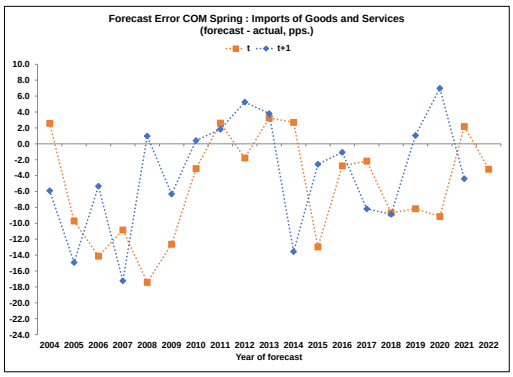


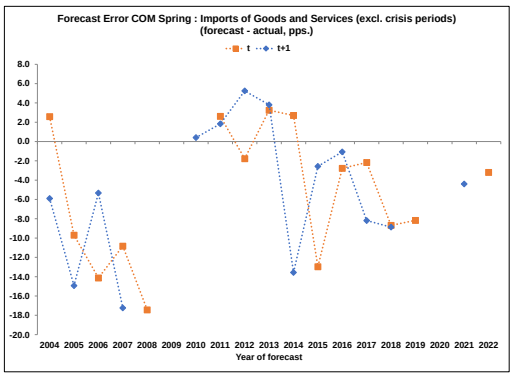
Charts A22: Forecasts Errors for Exports by COM for spring including and excluding crises



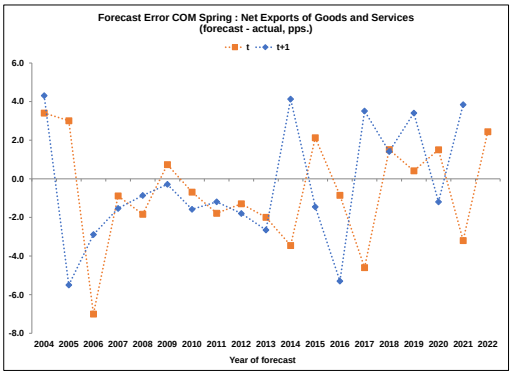


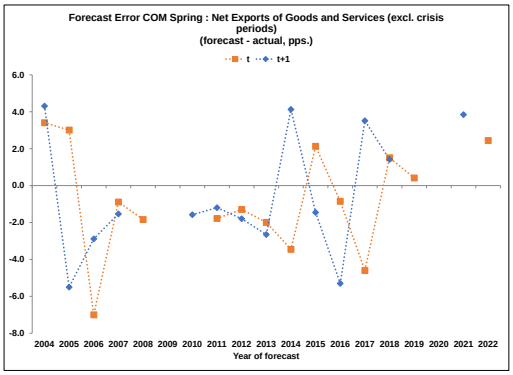
Charts A23: Forecasts Errors for Imports by COM for spring including and excluding crises



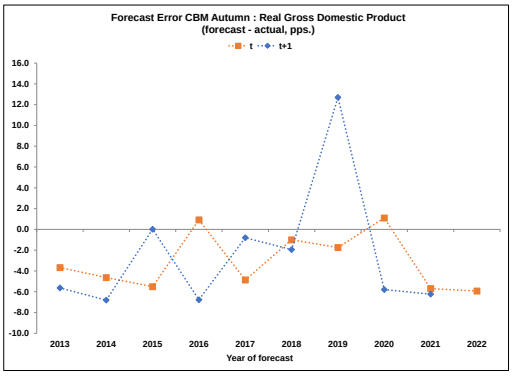


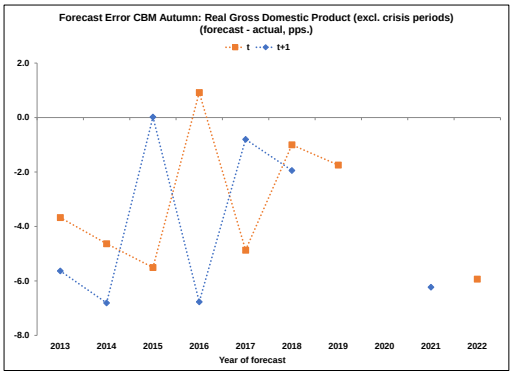
Charts A24: Forecasts Errors for Net Exports by COM for spring including and excluding crises



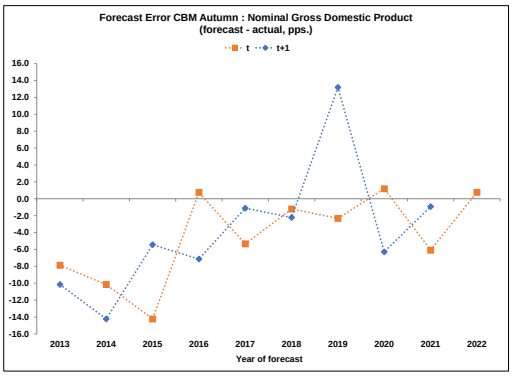


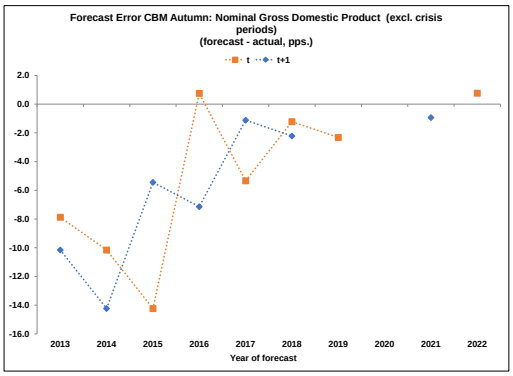
Charts A25: Forecasts Errors for Real GDP by CBM for autumn including and excluding crises



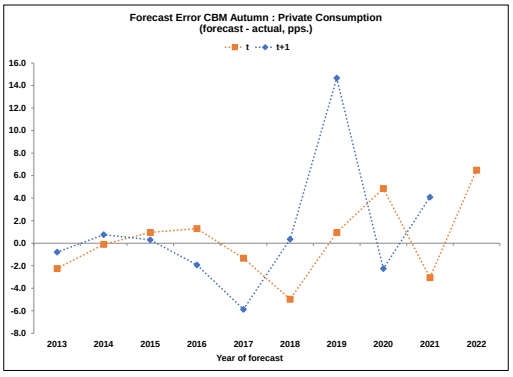


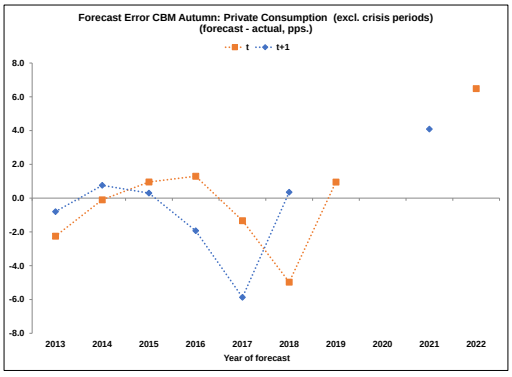
Charts A26: Forecasts Errors for Nominal GDP by CBM for autumn including and excluding crises



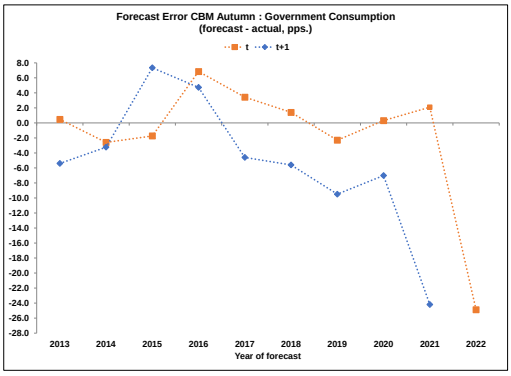


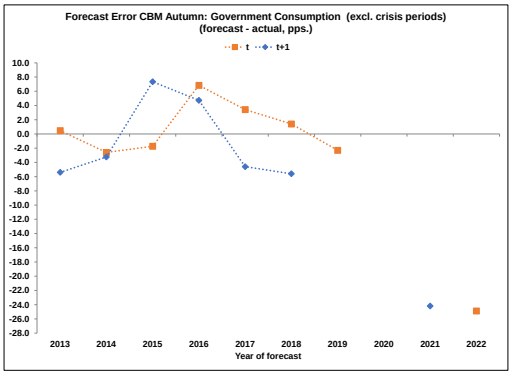
Charts A27: Forecasts Errors for Private Consumption by CBM for autumn including and excluding crises



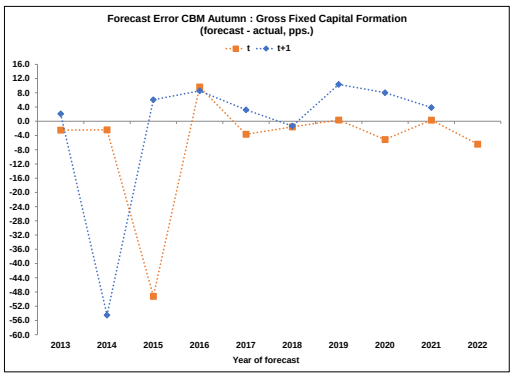


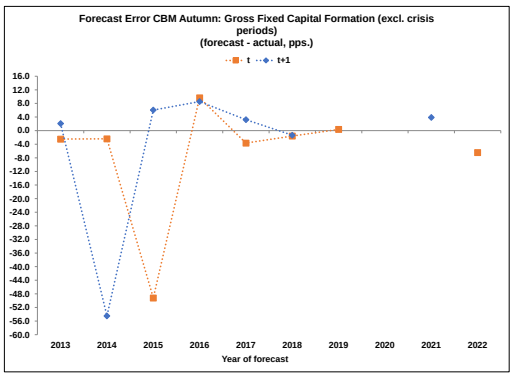
Charts A28: Forecasts Errors for Government Consumption by CBM for autumn including and excluding crises



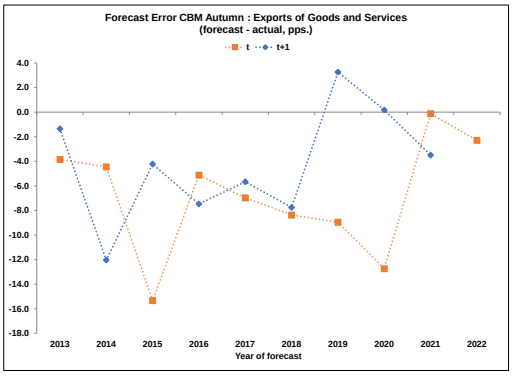


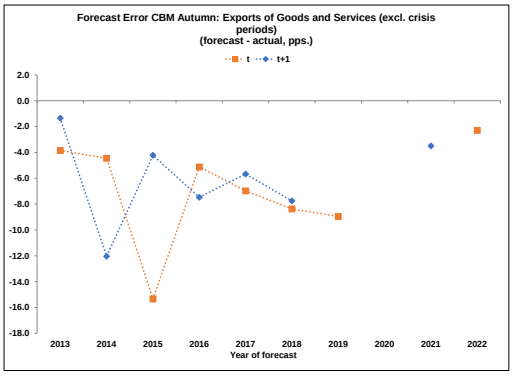
Charts A29: Forecasts Errors for Gross Fixed Capital Formation by CBM for autumn including and excluding crises



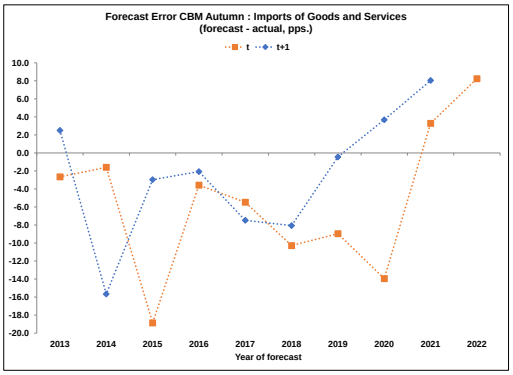


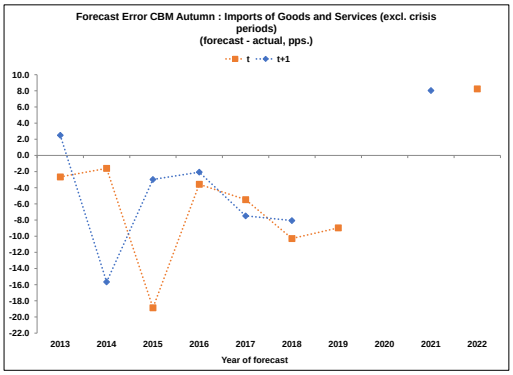
Charts A30: Forecasts Errors for Exports by CBM for autumn including and excluding crises



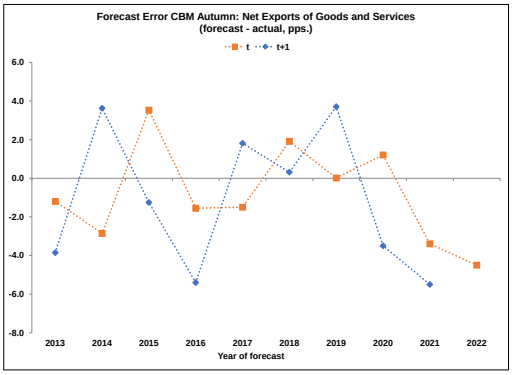


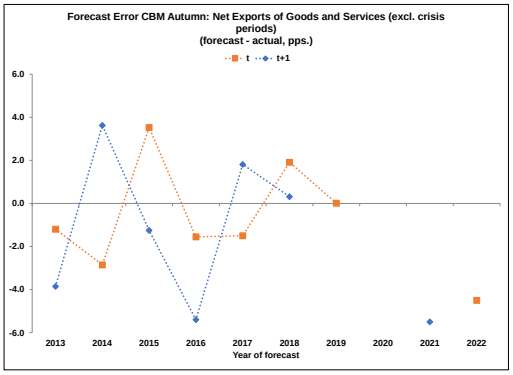
Charts A31: Forecasts Errors for Imports by CBM for autumn including and excluding crises



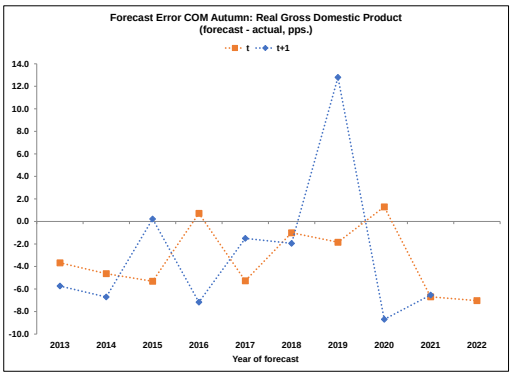


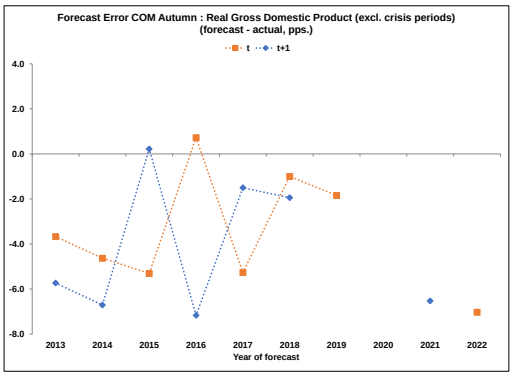
Charts A32: Forecasts Errors for Net Exports by CBM for autumn including and excluding crises



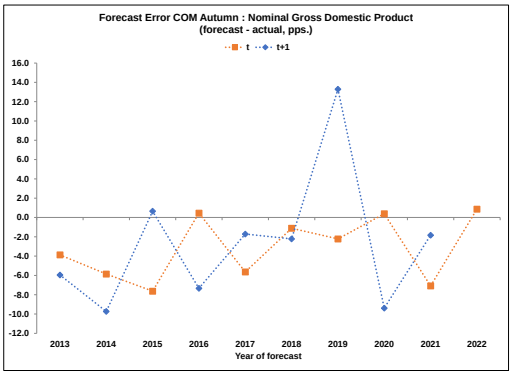


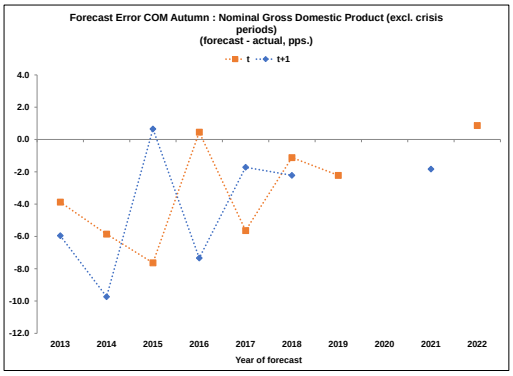
Charts A33: Forecasts Errors for Real GDP by COM for autumn including and excluding crises



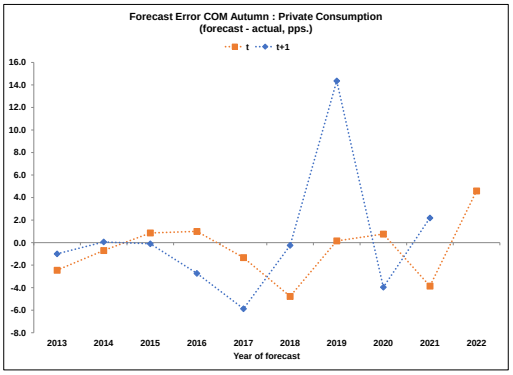


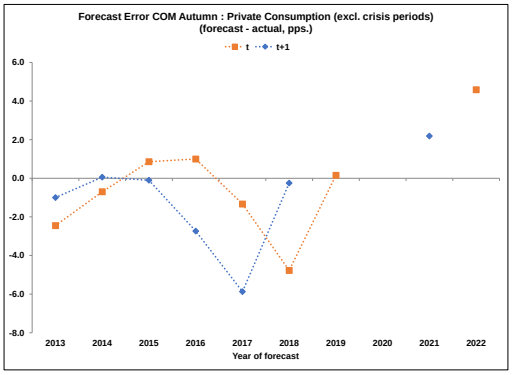
Charts A34: Forecasts Errors for Nominal GDP by COM for autumn including and excluding crises



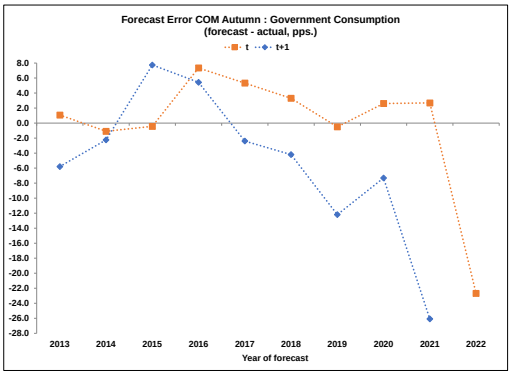


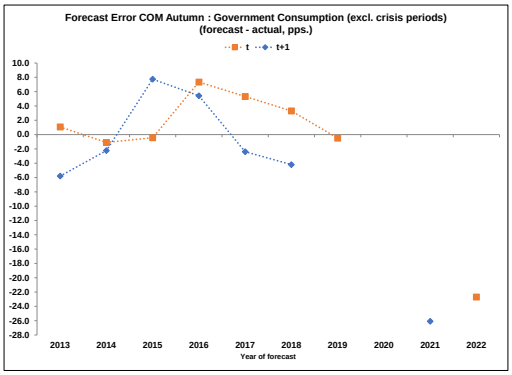
Charts A35: Forecasts Errors for Private Consumption by COM for autumn including and excluding crises



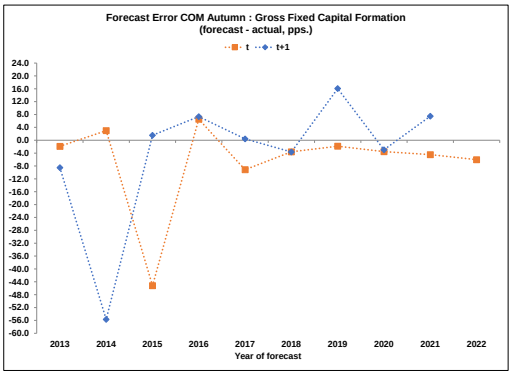


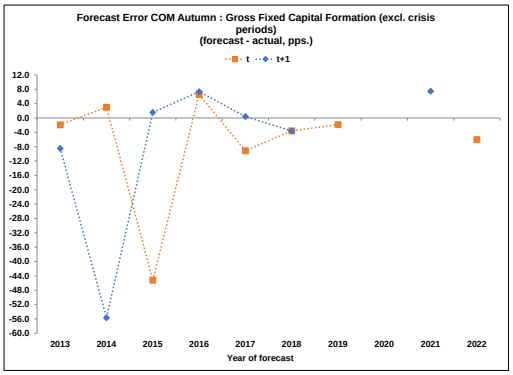
Charts A36: Forecasts Errors for Government Consumption by COM for autumn including and excluding crises



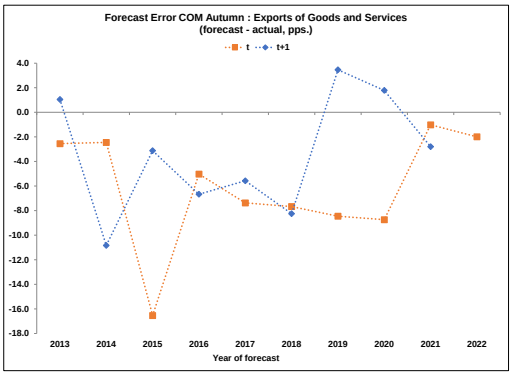


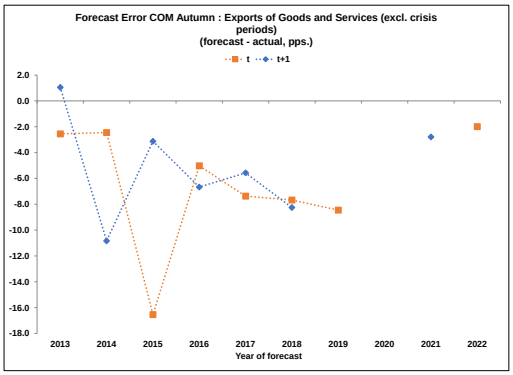
Charts A37: Forecasts Errors for Gross Fixed Capital Formation by COM for autumn including and excluding crises



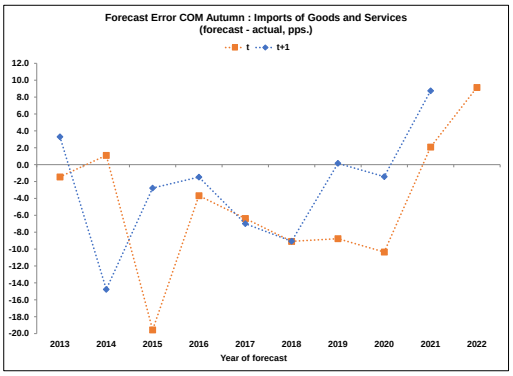


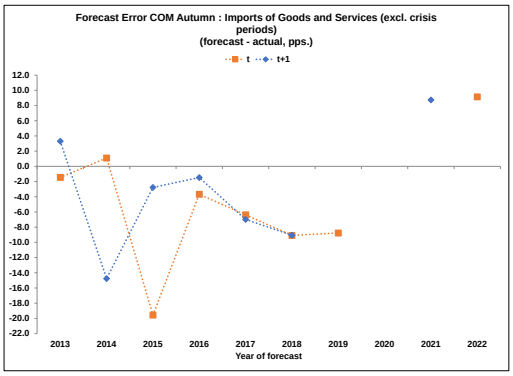
Charts A38: Forecasts Errors for Exports by COM for autumn including and excluding crises



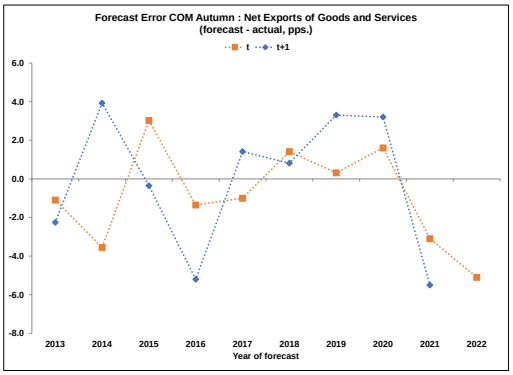


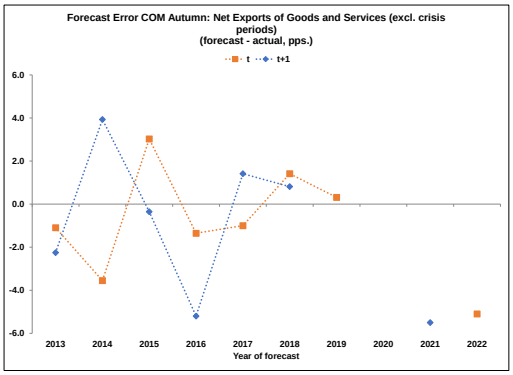
Charts A39: Forecasts Errors for Imports by COM for autumn including and excluding crises





Charts A40: Forecasts Errors for Net Exports by COM for autumn including and excluding crises





Appendix B

Accuracy measures – all sample

Mean Error

	GDP growth		Nominal GDP growth		Private consumption		Government consumption		Investment		Exports of goods and services		Imports of goods and services	
	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1
MFE (USP)	-2.5	-1.4	-2.2	-1.5	-0.9	-0.7	0.2	-4.0	-2.7	-1.7	-7.0	-5.0	-6.4	-4.7
COM (SPR)	-2.2	-1.7	-2.5	-2.0	-1.2	-0.6	0.7	-2.4	-3.0	-3.7	-6.1	-4.1	-5.4	-3.8
CBM (SPR)	-2.1	-1.6	-3.7	-2.3	-0.8	-0.2	-0.2	-2.3	-5.5	-1.6	-5.0	-2.7	-5.2	-2.2
MFE (DBP)	-2.8	-1.8	-3.2	-2.6	-0.7	-0.6	2.3	-3.0	-8.6	-7.1	-5.7	-3.7	-5.4	-4.4
COM (AUT)	-3.3	-2.8	-3.2	-2.7	-0.6	0.3	-0.2	-5.2	-6.6	-4.2	-6.2	-3.4	-4.7	-2.7
CBM (AUT)	-3.1	-2.4	-4.5	-3.8	0.3	1.0	-1.7	-5.3	-6.1	-1.5	-6.8	-4.3	-5.4	-2.5

Mean Absolute Error

	GDP growth		Nominal GDP growth		Private consumption		Government consumption		Investment		Exports of goods and services		Imports of goods and services	
	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1
MFE (USP)	3.2	3.8	3.3	4.2	2.8	2.9	3.1	5.7	8.7	11.4	7.8	5.9	9.4	7.3
COM (SPR)	2.8	3.7	3.6	4.2	2.5	2.8	3.6	4.8	8.5	12.0	6.9	5.8	6.8	6.0
CBM (SPR)	2.4	3.3	2.1	2.4	1.9	2.1	3.0	3.9	8.7	9.7	4.9	3.2	5.2	3.5
MFE (DBP)	3.1	4.7	3.4	5.8	2.1	3.9	2.9	5.3	10.8	14.2	5.9	4.5	6.6	4.8
COM (AUT)	3.7	5.7	3.5	5.8	2.0	3.4	4.7	8.2	8.5	11.5	6.2	4.8	7.2	5.4
CBM (AUT)	3.5	5.2	5.0	6.7	2.6	3.4	4.6	8.0	8.1	10.9	6.8	5.0	7.7	5.7

Root Mean Squared Error

	GDP growth		Nominal GDP growth		Private consumption		Government consumption		Investment		Exports of goods and services		Imports of goods and services	
	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1
MFE (USP)	4.0	4.9	4.3	5.7	3.8	4.5	4.2	7.0	13.4	17.4	10.0	7.9	9.4	7.3
COM (SPR)	3.5	4.9	4.4	5.4	3.2	4.3	4.8	5.8	14.0	17.6	9.0	8.1	8.4	7.7
CBM (SPR)	3.6	5.3	6.5	7.0	3.0	4.6	4.8	5.6	18.2	18.3	8.0	5.2	8.6	6.0
MFE (DBP)	3.8	6.2	4.3	7.0	2.6	5.8	4.0	6.2	17.9	21.8	7.4	5.6	8.1	6.6
COM (AUT)	4.4	6.8	4.4	7.2	2.6	5.5	7.9	10.7	15.1	19.9	7.6	5.7	8.9	7.1
CBM (AUT)	4.0	6.3	6.7	8.2	3.3	5.5	8.4	10.0	16.2	19.1	8.1	6.1	9.3	7.2

Theil's U Statistic

	GDP growth		Nominal GDP growth		Private consumption		Government consumption		Investment		Exports of goods and services		Imports of goods and services	
	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1
MFE (USP)	0.6	0.7	0.6	0.7	0.6	0.7	0.6	1.0	0.5	0.7	1.1	0.9	1.1	0.8
COM (SPR)	0.5	0.7	0.6	0.7	0.5	0.7	0.7	0.9	0.6	0.7	1.0	0.9	0.9	0.9
CBM (SPR)	0.5	0.7	0.6	0.6	0.5	0.7	0.7	0.7	0.6	0.6	0.8	0.5	0.9	0.6
MFE (DBP)	0.4	0.7	0.5	0.7	0.3	0.7	0.7	1.0	0.6	0.7	1.1	0.8	1.0	0.8
COM (AUT)	0.5	0.7	0.5	0.0	0.3	0.7	1.3	1.7	0.5	0.7	1.2	0.8	1.1	0.8
CBM (AUT)	0.5	0.7	0.7	0.9	0.5	0.6	0.8	1.4	0.4	0.5	1.3	0.9	1.2	0.9

Mean Relative Absolute Error

	GDP growth		Nominal GDP growth		Private consumption		Government consumption		Investment		Exports of goods and services		Imports of goods and services	
	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1
MFE (USP)	0.6	0.7	0.5	0.7	0.8	0.8	0.6	1.0	0.5	0.6	1.1	0.8	1.1	0.7
COM (SPR)	0.6	0.7	0.7	0.8	0.7	0.8	0.7	0.9	0.4	0.7	1.0	0.8	1.0	0.8
CBM (SPR)	0.5	0.6	0.4	0.4	0.5	0.6	0.6	0.7	0.5	0.5	0.7	0.4	0.7	0.5
MFE (DBP)	0.4	0.7	0.5	0.8	0.4	0.7	0.5	0.9	0.5	0.7	1.1	0.9	1.0	0.8
COM (AUT)	0.5	0.8	0.5	0.8	0.4	0.6	0.8	1.5	0.4	0.5	1.2	0.9	1.1	0.9
CBM (AUT)	0.5	0.7	0.7	0.9	0.5	0.6	0.8	1.4	0.4	0.5	1.3	0.9	1.2	0.9

Accuracy measures – excl. crises

Mean Error excl. 2009, 2010 and 2020,2021

	GDP growth		Nominal GDP growth		Private consumption		Government consumption		Investment		Exports of goods and services		Imports of goods and services	
	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1
MFE (USP)	-2.5	-2.2	-2.1	-2.3	-1.2	-1.6	0.1	-4.0	-1.8	-2.6	-6.9	-6.3	-6.3	-5.9
COM (SPR)	-2.1	-2.6	-2.4	-3.0	-1.5	-1.4	0.7	-2.3	-2.4	-4.5	-6.1	-5.5	-5.4	-5.1
CBM (SPR)	-2.2	-2.9	-4.4	-4.6	-1.0	-1.5	0.1	-2.1	-5.3	-3.9	-6.1	-4.0	-6.0	-3.4
MFE (DBP)	-2.8	-3.2	-3.3	-4.2	-1.0	-2.2	2.2	-1.1	-10.0	-11.0	-5.9	-5.2	-6.0	-5.8
COM (AUT)	-3.5	-4.2	-3.1	-4.0	-0.3	-1.1	-1.0	-3.9	-7.3	-7.3	-6.5	-5.2	-4.8	-3.3
CBM (AUT)	-3.3	-4.0	-5.0	-5.9	0.1	-0.4	-2.4	-4.4	-7.0	-4.6	-6.9	-6.0	-5.4	-3.7

Mean Squared Error excl. 2009, 2010 and 2020, 2021

	GDP growth		Nominal GDP growth		Private consumption		Government consumption		Investment		Exports of goods and services		Imports of goods and services	
	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1
MFE (USP)	2.9	2.5	3.0	3.1	2.3	2.2	3.5	4.9	9.0	10.4	8.0	6.6	7.4	6.3
COM (SPR)	2.6	2.8	3.4	3.3	2.3	2.1	3.9	4.4	9.2	11.4	7.2	6.3	6.9	6.7
CBM (SPR)	2.1	2.3	1.9	1.6	1.6	1.4	3.3	3.3	9.1	9.1	4.6	3.1	4.8	3.5
MFE (DBP)	2.9	3.2	3.3	4.4	1.5	2.3	2.9	4.0	12.7	15.5	6.2	5.3	6.9	6.0
COM (AUT)	2.9	3.3	2.8	3.3	1.6	1.4	4.2	6.0	7.7	9.4	5.2	4.3	5.9	5.2
CBM (AUT)	2.8	3.1	4.3	4.6	1.8	1.6	4.4	6.1	7.6	8.8	5.5	4.7	6.0	5.2

Root Mean Squared Error excl. 2009, 2010 and 2020,2021

	GDP growth		Nominal GDP growth		Private consumption		Government consumption		Investment		Exports of goods and services		Imports of goods and services	
	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1
MFE (USP)	3.8	3.2	4.0	4.1	3.1	2.8	4.6	6.6	14.3	18.7	10.4	8.6	9.4	8.2
COM (SPR)	3.2	3.6	4.2	4.1	3.1	2.7	5.1	5.4	14.9	18.2	9.4	8.8	8.5	8.4
CBM (SPR)	3.4	4.0	6.8	4.9	2.6	2.7	5.6	5.2	19.9	20.2	8.4	5.6	8.9	6.6
MFE (DBP)	3.5	4.3	4.1	5.4	1.9	3.3	2.9	4.8	19.8	23.9	7.7	6.2	8.5	7.5
COM (AUT)	4.3	5.0	4.3	5.3	2.6	2.6	8.7	10.9	16.7	21.7	7.9	6.3	9.2	8.0
CBM (AUT)	4.0	4.9	7.1	7.5	3.1	2.8	9.3	10.4	18.0	21.1	7.9	6.8	9.1	8.0

Theil's U Statistic excl. 2009, 2010 and 2020,2021

	GDP growth		Nominal GDP growth		Private consumption		Government consumption		Investment		Exports of goods and services		Imports of goods and services	
	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1
MFE (USP)	1.1	0.9	1.0	1.0	1.1	1.0	0.8	1.2	0.6	0.8	1.4	1.1	1.2	1.0
COM (SPR)	0.9	1.1	1.0	1.0	1.2	1.0	0.9	1.0	0.6	0.7	1.2	1.1	1.0	1.0
CBM (SPR)	0.8	1.0	1.0	0.7	0.8	0.9	0.8	0.8	0.7	0.7	0.9	0.6	0.9	0.7
MFE (DBP)	0.8	1.0	0.9	1.1	0.7	1.1	0.7	0.8	0.7	0.7	1.4	1.1	1.0	0.8
COM (AUT)	0.4	0.5	0.4	0.0	0.3	0.3	1.3	1.5	0.5	0.6	1.1	0.8	1.0	0.8
CBM (AUT)	0.4	0.5	0.7	0.0	0.4	0.3	1.4	1.5	0.6	0.6	1.1	0.9	1.0	0.8

Mean Relative Absolute Error excl. 2009, 2010 and 2020,2021

	GDP growth		Nominal GDP growth		Private consumption		Government consumption		Investment		Exports of goods and services		Imports of goods and services	
	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1	t	t+1
MFE (USP)	0.8	0.8	0.9	0.9	1.1	1.0	0.7	1.1	0.5	0.7	1.4	1.1	1.1	0.9
COM (SPR)	0.7	0.9	0.9	.0	1.0	1.0	0.7	0.9	0.5	0.7	1.1	1.0	0.9	1.0
CBM (SPR)	0.6	0.8	0.5	0.5	0.6	0.7	0.6	0.7	0.5	0.5	0.7	0.5	0.6	0.5
MFE (DBP)	0.7	0.8	0.8	1.1	0.7	1.0	0.5	0.7	0.6	0.7	1.5	1.3	1.1	0.9
COM (AUT)	0.4	0.5	0.4	0.4	0.3	0.2	0.7	1.1	0.4	0.4	1.0	0.8	0.9	0.8
CBM (AUT)	0.4	0.4	0.6	0.6	0.3	0.3	0.8	1.1	0.3	0.4	1.0	0.9	0.9	0.8