Fiscal policy monitoring report

Fiscal policy evaluation on the reliability of the Ministry of Finance macroeconomic forecasts

The short-term forecasts of GDP growth, unemployment rate and inflation for the years 1976-2016 under review

7/2018

Fiscal policy evaluation on the reliability of the Ministry of Finance macroeconomic forecasts

The short-term forecasts of GDP growth, unemployment rate and inflation for the years 1976-2016 under review

Translation of fiscal policy monitoring report 7/2018

HELSINKI 2018

Fiscal policy monitoring and evaluation report of the National Audit Office

Registry no. 127/56/2018

The National Audit Office monitors and evaluates fiscal policy in its role as an independent national fiscal policy evaluation body under the Stability Pact (Fiscal Compact) and within the meaning of Europe-an Union law. Provisions on the evaluation task are laid down in the Act on the National Audit Office of Finland (676/2000) and the Act on the implementation and application of the provisions governed by the Treaty on Stability, Coordination and Governance in the Economic and Monetary Union and on requirements concerning multiannual budg-etary frameworks (869/2012, 'Fiscal Policy Act'). Evaluation comprises the assessment of the setting and implementation of the fiscal policy rules steering the fiscal policy. It also covers monitoring the prepara-tion and implementation of the General Government Fiscal Plan, assessing the reliability of the macroeconomic forecasts and budget-ary projections used as a basis for the fiscal policy and evaluating compliance with the Stability and Growth Pact. By evaluating fiscal policy, the National Audit Office promotes transparent and easy-to-understand regulations and stable and sustainable general govern-ment finances.

This report assesses the reliability of short-term forecasts by the Ministry of Finance on GDP growth, unemployment rate and inflation in 1976–2016. The key legal basis for this duty consists of the European Union Budgetary Frameworks Directive (2011/85/EU) and the Government Decree on the General Government Fiscal Plan (120/2014, as amended by decree 601/2017), which is based on the Budgetary Frameworks Directive.

Helsinki 5 April 2018

Marko Männikkö
Deputy Auditor General

Arto Kokkinen
Senior Economist

Summary of the results

Summary of the results

As part of its fiscal policy evaluation, the National Audit Office (NAOF) assesses the reliability of forecasts by the Ministry of Finance (FM). This report focuses on forecasts for the current year and one-year ahead fiscal year on the growth of the Finnish gross domestic product (GDP), the unemployment rate and inflation in 1976–2016. The autumn economic forecasts by the FM were selected for the review, because these are the forecasts that are used when planning the state budget for the year ahead. The ministry's economic forecasts were compared with forecasts produced by the Research Institute of the Finnish Economy (Etla), the Organisation for Economic Co-operation and Development (OECD), the Labour Institute for Economic Research (PT) and Pellervo Economic Research (PTT). *No characteristics (bias) that would require corrective action based on the Government Decree on the General Government Fiscal Plan were detected in the GDP, unemployment rate and inflation forecasts by the Ministry of Finance.*

Reliability of the forecasts was studied using the following questions:

1. How accurately have the FM's forecasts predicted future economic development when compared to forecasts by other key forecasters (accuracy)?
2. Have errors in the FM's forecasts been temporally independent and non-persistent?
3. Have the FM's forecasts been unbiased?
4. Have the FM's forecasts encompassed the naïve forecast?
5. Have the FM's forecasts covered all the pertinent information at the time of forecasting?

Based on the results, reliability of the FM's forecasts on GDP growth, unemployment rate and inflation for the year ahead (*t+1*) and the current year (*t+0*) was, together with the forecasts by Etla, among the best of the compared forecasts.

The accuracy was studied by comparing the mean absolute error in the FM's forecast with those of the other forecasters. The FM's mean was, as a general rule, within the 95% confidence intervals of the two best forecasters. Only Etla was statistically slightly better than the Ministry of Finance in the current year's unemployment rate forecasts. In accuracy of the current year's GDP growth and inflation as well as in all of the year ahead variables, the FM was statistically among the most accurate forecasters.

In the case of forecasts for the budget year (*t+1*), the accuracy of the FM in 1976–2016 was statistically as good as that of the other forecasters. The FM, PTT and Etla were the most accurate in forecasting the GDP and unemployment rate during the sample period. In inflation forecasts, the FM was – together with the OECD – the least accurate, but the differences were minor. Statistically speaking, the FM has been as good as the best Finnish forecasters in the reference group in the forecasting accuracy of the current year's (*t+0*) GDP growth and inflation. In the sample between 1976 and 2016, Etla was the most accurate in forecasting the unemployment rate, followed by PTT. Also in the case of GDP growth, Etla was, together with PTT, the most accurate, while the FM was the third in both categories. PTT was the most accurate in the current year's inflation forecasts, closely followed by the FM and Etla.

Questions 2–5 were answered by studying the reliability of the forecasts by the FM and the reference group with statistical tests corresponding to the tests used by the European Commission in assessing its own forecasts. The FM's budget forecasts on the year ahead (*t+1*) passed, together with the forecasts by Etla, the most reliability tests: both institutions passed 13 out of 15 tests. The FM's forecasts proved to be statistically unbiased in the case of all the three variables, and there was no temporal persistence in the forecast errors. Forecasts by the Ministry of Finance and Etla most often passed the test on encompassing the naïve forecast and the test on coverage of all pertinent information at the time of the forecast.

Summary of the results

The results on the current year's forecasts (*t+0*) were almost the same: the FM and Etla passed 14 out of 15 reliability tests, and PTT and PT were not far behind, passing 13 out of 15 tests. The FM's forecasts were unbiased and there was no temporal persistence of forecast errors. Comprehensiveness of the information included in the forecasts on the current year's GDP, unemployment rate and inflation by the FM and Etla proved to be good. Results of the statistical tests have been compiled into a summary table, which is available in the Summary and conclusions section of this report.

Reliability of forecasts and independence of the preparation of forecasts used in state budget planning should be continuously monitored in Finland as well. According to the European Commission, the arrangement used by Finland in the preparation of official forecasts is exceptional in the EU: the forecasts used as the basis for the fiscal policy are prepared by the Ministry of Finance, and they are not confirmed by an independent third party. Elsewhere in the euro area, macroeconomic forecasts are either produced by independent institutions or the forecasts prepared by a country's finance ministry are approved by an independent body. The arrangement in Finland is, however, in accordance with the EU regulations.

Contents

1 Introduction

The National Audit Office (NAOF) monitors and evaluates fiscal policy in its role as an independent national fiscal policy evaluation body under the Stability Pact (Fiscal Compact) and within the meaning of European Union law. Provisions on the evaluation task are laid down in the Act on the National Audit Office of Finland (676/2000)[1] and the Act on the implementation of the Treaty on Stability, Coordination and Governance in the Economic and Monetary Union, the implementation of Treaty provisions of a legislative nature as well as requirements concerning multi-annual budgetary frameworks (the "Fiscal Policy Act"; 869/2012)[2]. Evaluation comprises the assessment of the setting and implementation of the fiscal policy rules steering the fiscal policy. It also covers monitoring the preparation and implementation of the General Government Fiscal Plan, assessing the reliability of the macroeconomic forecasts and budgetary projections used as a basis for the fiscal policy and evaluating compliance with the Stability and Growth Pact.

This report assesses the reliability of the economic forecasts by the Ministry of Finance. The key legal basis for this duty consists of the European Union Budgetary Frameworks Directive (2011/85/EU)[3] and the Government Decree on the General Government Fiscal Plan (120/2014, as amended by decree 601/2017)[4], which is based on the Budgetary Frameworks Directive.

According to the EU Budgetary Frameworks Directive, "Member States shall ensure that fiscal planning is based on realistic macroeconomic and budgetary forecasts using the most up-to-date information. Budgetary planning shall be based on the most likely macro fiscal scenario or on a more prudent scenario. The macroeconomic and budgetary forecasts shall be compared with the most updated forecasts of the Commission and, if appropriate, those of other independent bodies." (2011/85/EU, Chapter III, Article 4, paragraph 1) The Directive also states "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." (2011/85/EU, Chapter III, Article 4, paragraph 6)

Paragraphs 8, 9 and 15 in the justification part of the Budgetary Frameworks Directive state the following: "Biased and unrealistic macroeconomic and budgetary forecasts can considerably hamper the effectiveness of fiscal planning and consequently impair commitment to budgetary discipline, while transparency and discussion of forecasting methodologies can significantly increase the quality of macroeconomic and budgetary forecasts for fiscal planning. (...) A crucial element in ensuring the use of realistic forecasts for the conduct of budgetary policy is transparency, which should entail the public availability not only of the official macroeconomic and budgetary forecast prepared for fiscal planning, but also of the methodologies, assumptions and relevant parameters on which such forecasts are based. (...) The quality of official macroeconomic and budgetary forecasts is critically enhanced by regular, unbiased and comprehensive evaluation based on objective criteria. Thorough evaluation includes scrutiny of the economic assumptions, comparison with forecasts prepared by other institutions, and evaluation of past forecast performance."

Furthermore, according to section 5a of the Government Decree on the General Government Fiscal Plan (120/2014, as amended by decree 601/2017), "The Ministry of Finance shall take into consideration the conclusions of the National Audit Office on the macroeconomic and fiscal forecasts produced by the Ministry when preparing forecasts. If, according to the conclusions, the macroeconomic forecasts have included a bias that has had a major impact on the forecasts during four consecutive years, the Ministry of Finance shall publish the actions taken to correct the bias or issue a public opinion, insofar as it does not concur with the conclusions of the National Audit Office."

The Economics Department of the Ministry of Finance (FM) publishes its forecasts on the development of the Finnish GDP, components of supply and demand, unemployment rate, inflation and general government finances four times per year in its economic surveys. Furthermore, future development of the general government finances is assessed in more detail in the General Government Fiscal Plan that is published in the spring. The economic

surveys published in the spring and autumn are more extensive and detailed than the surveys published in the summer and winter, and they also include an assessment of the medium-term outlook of the national economy and the general government finances (next 4–5 years). In addition, the FM publishes a report on the accuracy of its forecasts in the spring.

This report studies the reliability of the short-term forecasts produced by the Economics Department of the Ministry of Finance. The evaluation focuses on the autumn forecasts for the year ahead (*t+1*), which are used as the basis of the state budget planning, and the autumn forecasts for the current year (*t+0*).

The data used in the evaluation consists of forecasts on Finland's GDP growth, unemployment rate and inflation for the years 1976–2016 by the Economics Department of the Ministry of Finance, the Research Institute of the Finnish Economy (Etla), the Organisation for Economic Co-operation and Development (OECD), the Labour Institute for Economic Research (PT) and Pellervo Economic Research (PTT). To calculate forecast errors, the forecasts have been compared with the GDP growth in the national accounts, labour force survey unemployment rate and consumer price index for the same reference period of time recorded by Statistics Finland.

Structure of the report is as follows: Section 1.1 in the introduction includes a brief review of previous studies of the accuracy and reliability of Finnish macroeconomic forecasts, while Section 1.2 focuses on the statistical data used in this report. Chapter 2 reviews the basic concepts and statistical methods used in the assessment. Chapter 3 includes a versatile assessment of the reliability of the FM's autumn forecasts in 1976–2016: first, the FM's forecasts are assessed in relation to forecasts by other forecasting institutions by comparing the mean errors, the mean absolute errors and the statistical confidence intervals in the mean and mean absolute errors of the different forecasters. Next, the unbiasedness of the FM's forecasts and the coverage of the information therein are tested using five statistical methods. The coverage of the information is investigated by studying first whether the forecast has encompassed the naïve forecast information and second whether the forecast has covered all pertinent information at the time of producing the forecast. To enable a comparison of the FM's results with the other forecasters, reliability of the forecasts by the other forecasters (Etla, OECD, PT and PTT) have been subjected to the same tests (the detailed results are given in Appendices 3–6). Chapter 4 includes a summary of the report and the key results.

1.1 About previous research[5]

The reliability of macroeconomic forecasts on Finland have previously been studied mostly from the viewpoint of accuracy and unbiasedness. The most commonly used statistical methods include mean error, mean absolute error and the root of mean squared error.

Forecasts by the Ministry of Finance

The earliest previous studies based on statistical inference are from the 1980s. Konttinen (1986)[6] studied short-term forecasts and their accuracy in his doctoral thesis. The research results state that forecast deviations are fairly high at times, but contemporaneous forecasts by different forecasters are fairly similar. According to a study by Mutikainen and Suvanto (1986)[7], the FM has been fairly successful as a whole in forecasting the different items of the goods and services account. However, Viren (1994)[8] states that unsuccessful forecasts have inevitably led to errors in fiscal policy. Viren emphasises the fact that the overheating of the economy in the late 1980s was not forecasted and the economic downturn in the early 1990s was also a surprise for the forecasters. The NAOF (1999)[9] has also paid attention to the accuracy of state revenue estimates. The NAOF based its assessment of forecast errors on the above-mentioned literature (Mutikainen and Suvanto, 1986; Viren, 1994).

The turn in the economic cycle in 2001 reawakened the discussion about the accuracy of economic forecasts. Pehkonen (2002)[10] states that the accuracy of the forecasts has only been rather good at best, and that the same accuracy could have been easily reached with simple foresight methods.

Later in the 2000s, the long-term upswing in the economy and the downswing following it gave rise to a new discussion about the accuracy of economic forecasts, tax revenue forecasts in particular.

A study commissioned by Parliament assessed the accuracy of state budgets. The results of the study by Lahtinen et al. (2009)[11] stated that no significant systematic bias was observed in the review of the long-term economic forecasts. It was stated, however, that forecast errors are dependent on the economic situation, because the GDP growth is often underestimated during an upswing and overestimated during a downswing. Lanne (2009)[12] assessed the symmetry of the "forecaster's loss function" instead of unbiasedness. According to the results, the FM has been successful with its forecasts when assessed based on the symmetric loss function. Symmetric loss function refers to the forecaster emphasising both under- and overestimates using the same weight. On the basis of an asymmetric loss function, a forecaster may, for example, attempt to avoid overestimating the economic development and consequently make more prudent forecasts.

The International Monetary Fund (IMF)[13] assessed the transparency of Finland's general government finances in 2015. The assessment dealt with budget economy forecasts, including short- and long-term macroeconomic forecasts. The IMF has assessed that the medium-term forecasts on the development of the Finnish GDP are relatively accurate and the Finnish official forecasts are among the ten most accurate in the EU. According to the IMF, the GDP forecasts on the year ahead by the FM have been more accurate than forecasts by other institutions, which, in other words, means that the FM's forecasts have not been inaccurate when compared to other forecasters.

The assessment by IMF also dealt with the accuracy of revenue, expenditure and net lending forecasts. The analysis was based on mean errors proportioned to the GDP. The assessment stated that forecasts on the balance of the general government finances have been unbiased on average, but both revenue and expenditure forecasts were underestimated between 2000 and 2013. Underestimation has occurred in all the three sectors: central government, local government and social security funds. The assessment compared the accuracy of general government finances forecasts with forecasts by other forecasting institutions, and the FM was deemed the most accurate of all the institutions. The comparison was based on observations made during a period of two years, however, which means that the conclusions cannot be considered highly reliable.

Before preparing this report, the NAOF has studied the reliability of macroeconomic forecasts by the Ministry of Finance in an unpublished audit memorandum, Tilastoanalyysi makroennusteiden osuvuudesta (NAOF 2016a)[14] and in a fiscal policy audit report, Reliability of macroeconomic forecasts (NAOF 2016b)[15]. These studied forecasts on items of the goods and services account and forecasts on general government revenue, expenditure and net lending in 1997–2014. Deviation of the forecast error from zero was studied using the mean error (ME) and accuracy of the forecast was studied based on the mean absolute error (MAE; the inspection memorandum also used the root of the mean squared error). Forecast errors by the Ministry of Finance were compared with the forecast errors by the other principal forecasters by using statistical confidence intervals as well. Furthermore, the unbiasedness of the forecasts was studied in a preliminary statistical *t* test (considering, though, the small number of observations in the statistical testing). In general, the short-term forecasts on GDP and components of the goods and services account during the mentioned period of time were found statistically unbiased when studied with the fairly limited data set. Repeated underestimation of the current year's import forecasts (*t+0*) was detected, but the same trend was also detected in the forecasts of the other assessed forecasters. The underestimation was probably caused by the fact that the statistical data of foreign trade, particularly the data on the exports and imports of services, has been significantly revised in the statistics by Statistics Finland. Underestimation was also detected in the forecasts on general government revenue and expenditure. Even though the underestimation of the expenditure was only slightly higher than the underestimation of the revenue, the

errors in the net lending forecasts that are based on the difference between these two figures did not, however, statistically deviate from zero on average. In 1997–2014, the accuracy of the FM's forecasts was statistically as good as that of the other forecasters.

Forecasts by other forecasting institutions

The accuracy of forecasts by forecasters other than the FM, such as the Bank of Finland, has also been assessed. Newby and Orjasniemi (2011)[16] assessed the accuracy of forecasts by the Bank of Finland using the mean error, the mean absolute error and the mean squared error. They were unable to assess the unbiasedness of the forecasts due to the short time series, however. The authors state that the forecasting of turning points in economic development is particularly difficult. They add that it is impossible to forecast surprises.

Forecasts by the European Commission have been assessed on four different studies. The first assessment was published in 1999. In this publication, the Commission assesses the accuracy of its own forecasts using three indicators (ME, MAE and RMSE). (European Commission, 1999)[17] According to the results, the forecasts by the European Commission have been fairly accurate. The results were updated in 2007 (European Commission, 2007)[18] and in 2012 (European Commission, 2012)[19]. The results are compared with forecasts by the IMF and the OECD, for example, stating that the forecasts are in line with each other. No bias was detected in the current year's forecasts, but problems were detected in the case of some GDP components in the year ahead forecasts. In 2016, the European Commission[20] published its latest update on the reliability of its own forecasts (European Commission, 2016). The conclusions on the accuracy of the forecasts were similar to those of the above-mentioned assessments. In addition to analysing and comparing the accuracy of its forecasts, the Commission studies different aspects of the reliability of its forecasts using seven statistical tests (European Commission 2016, pp. 36–38).

The methods used by the Commission are used as the starting point for the statistical analysis of the reliability of the forecasts in this assessment report, and five of the tests used by the Commission are used to assess the Finnish macroeconomic forecasts. The Commission has tested temporal independence of forecast errors, unbiasedness, whether the forecast errors have included equal number of positive and negative signs, and information efficiency of forecasts. The last-mentioned includes testing whether the forecast includes all the information included in the naïve forecast and whether the forecast covers all the current, pertinent information at the time of the forecast and is therefore efficient. In addition to the above-mentioned three indicators (ME, MAE and RMSE), the Commission has used the forecast loss function approach and the Diebold-Mariano test when assessing the accuracy of forecasts. The 2016 update also studied whether adding forecast observations from 2012–2014 to the data as of 2000 improved or deteriorated the average results. One approach was comparing forecasts made before and after the 2008–2009 financial crisis. Adding the latest years either did not influence the average reliability or slightly improved the reliability.

1.2 Research data

To study the reliability of forecasts, one needs to gather, first, data on forecasts on the variables to be studied and second, statistics on actual values of these variables. When such data has been gathered, one soon notices that forecasts on the current year and forecasts on the year ahead are published several times per year. Furthermore, national accounts statistics on macroeconomic variables are revised and published several times again after the first publication date. Therefore, for the comparison of the forecast and the actual value, the forecast vintage and the publication vintage of the statistics must be selected.

The Economics Department of the Ministry of Finance publishes four economic surveys per year that include macroeconomic forecasts. In the summer and right before Christmas, the FM publishes a concise forecast on the development of the current year and the two years following the current year. The economic surveys published in the spring and autumn

are more extensive and include detailed analyses of the state of the economy and the economic outlook. In addition to short-term forecasts, the spring and autumn economic surveys include an assessment of the medium-term outlook (the next 4–5 years).

On the other hand, Statistics Finland publishes several versions of the national accounts per one statistical year. The most specific statistics on the GDP and the detailed components of the goods and services account can only be generated after basic information about the economic actors has been collected and compiled, which takes time. The supply and demand statistics in the national accounts, which most accurately describe the development during a specific year, are based on product group specific supply and use tables, which are published in Finland 18–24 months after the end of the statistical year.

Economic decision-makers require information about the economic development more quickly, which is why advance information on the national accounts that is based on preliminary and more concise data is published. In addition to the annual data, Statistics Finland publishes quarterly national accounts two months (previously 65 days) after the end of each quarter. The first advance information on the development of the GDP during the previous year, or the "preliminary annual GDP estimate", is published at the turn of February and March. It is based on the quarterly accounts. Several source statistics that are based on the previous year's final closed book keeping accounts by enterprises and other institutions are published during the spring. Using this annual data, Statistics Finland publishes more specific information containing annual accounts in July. Even though the data is not as detailed as the data for the supply and use tables, the more specific information containing annual accounts with GDP and the components published in July is revised to a lesser extent than the preliminary annual estimate. Thus, the information available in July better describes the development of the final figures.

This report focuses on assessing the reliability of the FM's economic growth, unemployment rate and inflation forecasts that are published in the autumn. The autumn forecast or the "budget forecast" was selected as the forecast time as in the NAOF 2016 reliability of the forecasts memorandum. The autumn forecast was selected because the autumn forecast for the year ahead (*t+1*) GDP at current prices forms the basis for the assessment of the tax revenue available for general government expenditure and is thus used as the basis for the planning of the next year's budget. The unemployment rate development assessments are closely connected on the other hand to assessments on the development of an important input to economic production, the number of employed (and working hours), and on the other hand, unemployment expenditure is one of the major cyclical expenditure items in the state budget. Forecasts on the development of prices (the consumer price index and the entire GDP price index) are used when converting (inflating) the development of the GDP volume into GDP at current prices, as the GDP volume change is always originally forecasted at fixed prices of the previous year.

The currently available statistics on macroeconomic variables are used as the basis when preparing the forecasts. When the year ahead (*t+1*) is forecasted in the autumn, the annual (*t-1*) national account statistics that were published in July are available for the first time. Furthermore, when forecasting both the current year (*t+0*) and the year ahead (*t+1*) in the autumn, information from one quarter of the current year (*t+0*), or in the case of an election year, information from two quarters of the current year is also available. Therefore, the autumn forecast includes more recent statistical data than the forecasts made in the winter, spring and summer. Using the autumn forecast when assessing the reliability of forecasts is also justified for this reason, as it is fairer for the forecasters.

Statistics Finland submitted in electronic format the published information on the annual national accounts from 1997 to 2016. Therefore, for each year the annual preliminary estimate published at the turn of February and March, the annual national accounts that had been revised for the first time in July, all later revised and more detailed information including versions of annual national accounts published at a later time and finally, the version of national accounts which is compliant to the most recent EU national accounting standard (ESA2010) was available. Similarly to the NAOF audit memorandum from 2016, for the GDP growth the national account publications selected for this forecast evaluation are the first time revised annual national accounts publications in July. Thus, the current year forecast published by the FM in the autumn (*t+0*) was compared with the annual accounts

version published in July of the next year (*t+1*). Similarly, FM's autumn forecast on the year ahead (*t+1*) was compared with the annual GDP statistics published in July of the year following that (t+2). The revised and more detailed annual source data including GDP published in July is considered to offer a sufficiently accurate picture of the development during the previous year. It is also clearly closer to the information available at the time the forecast was made than the information available when compiling the most accurate GDP in national accounts in the supply and use tables. Therefore, the revisions in the GDP statistics influence the GDP forecast error calculated in this report less than when one would have used the final, supply and use table based GDP statistics.

Statistics Finland was not able to offer digital versions of all the versions of the accounts for all the early years between 1976 and 1996. Instead, final information for each of these years, compliant with each accounting standards and each base years (skt80, skt85, skt90, skt95 = ESA95) was provided. In addition, growth figures for 1976–1996 according to the most recent EU accounting standard (ESA2010), which was introduced in 2014, were available. The GDP growth figures with each base year used at that time were used to create a continuous series of growth rates, and the forecasts were compared with this series. The forecast errors of all the institutions that were calculated in this manner were finally compared with the forecast errors based on the ESA2010 growth rates of the said years. Perhaps somewhat surprisingly, the mean errors calculated based on each accounting version at time of these years experienced only minimal changes, and whenever the mean errors changed, particularly in the case of the (*t+1*) forecasts, they somewhat decreased in the case of all the forecasters when using the latest ESA2010 data for calculating the forecast errors. Thus, the GDP volume changes based on the latest European System of Integrated Economic Accounts (ESA2010) were finally selected for 1976–1996 against which the forecasts were compared. These figures are publicly available to anyone on the website of Statistics Finland and, as stated above, their use did not seem to systematically impair the forecasters' position.

The actual unemployment rate is based on figures published by Statistics Finland in a sample-based labour force survey and the inflation figures are based on changes of Statistics Finland's sample-based consumer price index. Labour force survey concepts and definitions in Finland were harmonised in 1997 and 1998 to better correspond to the guidelines and recommendations of the EU and the International Labour Organization (ILO). The unemployment rate figures in Finland for 1981–1997 were thus afterwards revised according to the change in concepts, changes in the collection methods of unemployment statistics and other changes. In this assessment report, the unemployment rate observations for 1981–1997 were gathered from the originally published statistical yearbooks and labour force statistics publications (in the case of 1997, the figure is the mean for the four quarters). Therefore, these figures are comparable with the unemployment rate forecasts made at each year in accordance with the unemployment concept used at that time.

The forecast data was collected from an economic forecast database maintained by Etla. Originally, forecasts produced by all the available economic forecasters were taken from the database. After a preliminary analysis, it was decided that in addition to the autumn forecasts of the FM, forecasts for the year ahead (*t+1*) and the current year (*t+0*) from all forecasters for whom a sufficient number of annual forecasts was available in the Etla database to run the statistical tests correctly would be used. Therefore, the reliability of the forecasts by FM in this assessment are compared with the forecasts of Etla, the OECD, PT and PTT. The idea was to gather long time series data to make sure that the number of observations will be sufficient for statistical testing. The years 1976–2016 were selected, first, to serve that purpose and second, because the Etla database includes forecast observations for the years (*t+1*) and (*t+0*) for most of the selected forecasters.

According to the storage data of the forecasts selected from the Etla database, the forecasts by FM for each year (1976–2015, 41 forecasts in total) are from September. The forecasts by Etla for 1976–1987 (12 in total) are from November and the forecasts by Etla for 1988–2016 (29 in total) are from September. The forecasts by PTT (1981–2015) are from September (2005–2006, 2008–2016, 11 in total) or October (1981–2004, 2007, 25 in total). Based on the time of the year of the forecasts, the forecasts by Etla and PTT have the most equal or/and perhaps a slightly better position in relation to FM. The early forecasts by PT

(1975–1991, 17 in total) are from March, April or June and the rest of PT's forecasts are from October (1992–2001, 10 in total) or September (2002–2015, 14 in total). Of the forecasts by the OECD, 34 (1981–1998, 18 forecasts and 2000–2015, 16 forecasts) are from a point earlier than that of the Ministry of Finance's forecasts, from May or June, and seven of the forecasts have been made at a later point in time (1975–1980 in December and 1999 in October). On the basis of the forecast time, PT has an inferior position when compared to the FM: 17 of its early forecasts are from the spring months or June, which means that less statistical data for the current year was available when the forecasts were being prepared. [21] The PT forecasts starting from 1992 are from September or October, however. The internal forecast times of the OECD for the Finnish economy are the worst when compared to the FM: 34 of the forecasts are from May or June, while all of FM's forecasts are from September. This difference can be reflected as larger forecast errors and otherwise poorer test results for the OECD and PT.

2 Concepts and methods

2.1 Basic concepts[22]

The quality of forecasts is most commonly assessed based on forecast errors. The difference between the actual figure y_t and the forecasted value $\hat{y}_t$, e_t, is called the forecast error, which in the year *t* can be expressed as:

$$e_t = \hat{y}_t - y_t$$

A positive forecast error is generated when the forecasted figure is higher than the actual figure. In such a case, the development of the economy has been overestimated, while a negative forecast error means that the forecaster has underestimated the development of the economy.

The quality of forecasts cannot be very well assessed on the basis of a forecast error during one year. Therefore, forecasts should always be assessed in the longer term.

Mean error – unbiasedness

An unbiased forecast is correct on average. This means that forecasts both exceed and remain below the actual value as often. In a longer-term assessment, the mean forecast error (mean error, ME) is generated during the years *t* and *T*:

$$ME = \frac{1}{T}\sum_{t=1}^{T} e_t ,$$

where e_t is the forecast error in the year t.

The unbiasedness of the forecast becomes questionable if the mean error deviates from zero in a longer period. Deviation of the mean error from zero can be tested using statistical tests, such as the *t* test (and its modified version which takes into account the temporal dependence of forecast errors) or with a regression-based method (where the temporal dependence of forecast errors must also be taken into account when necessary, see Section 2.2).

Mean absolute error – a comparison of the accuracy of forecasts

The mean error cannot be used to compare the accuracy of different forecasts. One reason for this is that positive and negative forecast errors cancel each other. A better starting point for a comparison is taking into account the actual deviation between the forecast and the actual figure in the indicator.

The most common approach to such an assessment is using the following two indicators. The first is the mean absolute error (MAE):

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

The mean absolute error indicates how far from the actual value the forecasted figures were on average. The smaller the figure, the more successful the forecast.

The second indicator that is often used in comparisons is the mean squared error (MSE). The forecast error for each year is squared and an average is calculated from the squared error, which means that the direction of the forecast error is insignificant also in this approach. Due to the squaring, the indicator penalises individual large forecast errors more strongly than the mean absolute error. Often, a square root of the indicator is finally taken to make the unit of measure comparable with the other indicators. The most commonly used indicator is the root of mean squared error (RMSE):

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T} e_t^2}$$

2.2 Statistical tests in testing reliability of forecasts

The European Commission tested the reliability of the forecasts it provides for the Member States in 2016 with seven statistical tests (European Commission, 2016, pp. 36–38)[23]. This assessment report uses the following five tests in the testing of the reliability of Finland's GDP growth, unemployment rate and inflation forecasts, of which the tests 1, 3, 4 and 5 correspond to tests used by the European Commission:

1 Testing temporal independence and non-persistence of forecast errors

Forecast errors should be random and observed errors should be temporally independent of each other. In other words, they should not include any systematic variation along time. If an error is repeated in a forecast or if an error is always corrected with a similar error of the opposite sign, the forecast error is persistent and (perfectly) autocorrelated (for more information on autocorrelation, please see, e.g. Chatfield, 2003[24] and Hamilton, 1994[25]). In such a case, the forecast error can be modelled to forecast future forecast errors. Temporal independence and forecast errors not being persistent and permanent can be studied by testing that the forecast error observations with different lags are not autocorrelated. This can be tested with the Ljung-Box test, for example, which is also used by the European Commission (Ljung and Box, 1978)[26]:

$$Q = n(n+2)\sum_{k=1}^{h} \frac{\hat{\rho}_k^2}{n-k}$$

where n is the sample size, $\hat{\rho}_k$ is the sample autocorrelation with lag k and h indicates how many lags the autocorrelation test covers. The test studies simultaneously whether all the autocorrelation coefficients, from lag one to the specified lag, are statistically significantly different from zero. This report uses three lags, partly because FM's short-term forecast horizon is three years (the years *t+0*, *t+1* and *t+2*) and partly because the forecast of an earlier year (such as *t+0*) can be considered to influence the consequent (*t+1*, *t+2*) forecasts. Furthermore, the European Commission also used three lags in the corresponding test. The null hypothesis is that the observations are first, identically distributed and second, independently distributed (in other words, all the observed autocorrelations are zero). The Ljung-Box test follows the χ^2 distribution with the degrees of freedom h.

2 Testing unbiasedness with the *t* test

The unbiasedness of forecasts can be studied with the *t* test to infer whether the mean error does not deviate statistically significantly from zero. The common form of the *t* test when testing the deviation of the parameter $\hat{\beta}$ estimated based on the sample data from the H_0 hypothesis value β_{H0}:

$$t = \frac{\hat{\beta} - \beta_{H0}}{s.e.(\hat{\beta})},$$

where $s.e.(\hat{\beta})$ refers to the standard error of the parameter estimate $\hat{\beta}$.

Next, let $\overline{x}_{e_t}$ stand for the mean of the forecast error. When the deviation from zero of the mean of temporally independent forecast errors is tested based on a sample, the *t* test can be written in the form

$$t = \frac{\overline{x}_{e_t} - 0}{s.e.(\overline{x}_{e_t})} = \frac{\overline{x}_{e_t}}{\frac{s_{e_t}^2}{n}},$$

where $s_{e_t}^2$ is the variance of the forecast errors calculated based on the sample and *n* is the number of observations in the sample. One should also note that the forecast errors are assumed to be temporally independent. If the forecast errors are not temporally independent, the following modified *t* test variable should be used, according to time series analyst Dr. Chris Chatfield (see Kokkinen and Wouters, 2016, p. 101)[27]:

$$t = \frac{\overline{x}_{e_t} - 0}{s.e.(\overline{x}_{e_t})\frac{(1+\alpha)}{(1-\alpha)}} = \frac{\overline{x}_{e_t}}{\frac{s_{e_t}^2(1+\alpha)}{n(1-\alpha)}},$$

where α is the autocorrelation coefficient of the forecast errors between lags *t* and *t-1*. The above-mentioned modified *t* test statistic must be compared with the *t* distribution with the degrees of freedom $n\frac{(1-\alpha^2)}{(1+\alpha^2)}$. Whenever the above-mentioned modified standard error and *t* test have been used in the results of this assessment report, there is a note "*modified*" after the standard error.

3 Testing unbiasedness with the regression method

The unbiasedness of a forecast can also be studied from another perspective: a forecast is unbiased when the forecast's expected value corresponds to the expected value of the actual variable (that is observed afterwards):

$E(y_{t+0,t}) = E(y_t)$ in the case of a year (*t+0*) forecast
$E(y_{t+1,t}) = E(y_{t+1})$ in the case of a year (*t+1*) forecast

The European Commission (2016) tested the unbiasedness of its forecast for the same year (*t+0*) and the year ahead (*t+1*) by regressing the forecast error e_t with the constant term α:

$$e_{t+0,t} = \alpha + \varepsilon_{t+0,t}$$
$$e_{t+1,t} = \alpha + \varepsilon_{t+1,t},$$

where $e_{t+0,t}$ refers to an error in the same year's forecast (*t+0*) in the forecast made in the year *t* and similarly $e_{t+1,t}$ refers to an error in the forecast for the year (*t+1*) in a forecast made in the year *t*. ε_t is assumed to be an independently and identically distributed residual term. In the case of an unbiased forecast, $\alpha = 0$.

One must note that if the forecast errors are autocorrelated (and/or possibly heteroscedastic), the residual after regressing with a constant term is also autocorrelated (and/or heteroscedastic). In such a case, standard errors that take into account heteroscedasticity and autocorrelation of the residual, must be used in the testing of the statistical significance of the regression parameters (such as α here, see Newey-West, 1987)[28]. Whenever the regression analysis results in this assessment report have used the heteroscedasticity and autocorrelation consistent (HAC) standard errors by Newey-West and the *t* test , this is marked with "*HAC*" after the standard error.

4 Forecast encompassing test

Whether the forecast for the current year (*t+0*) encompasses the naïve forecast can be tested using the following model proposed by Fair and Schiller (Fair and Schiller, 1990)[29]:

$$y_t = \alpha + \beta y_{t-1} + \delta y_{t+0,t} + \varepsilon_t,$$

where y_t is the observed actual value at the time t, y_{t-1} is the previous year's actual value (which is used as the naïve forecast), $y_{t+0,t}$ is the current year's (*t+0*) forecast at year *t* and ε_t is independently, identically distributed residual term. If the forecast encompasses the naïve forecast and also includes significant additional information, $\beta = 0$ and $\delta > 0$.

A similar test to test whether the forecast for the year (*t+1*) encompasses the naïve forecast can be presented in the format

$$y_{t+1} = \alpha + \beta y_{t-1} + \gamma y_{t+0,t} + \delta y_{t+1,t} + \varepsilon_{t+1},$$

where y_{t+1} is the observed actual value at the time $t+1$, y_{t-1} is the previous year's actual value, $y_{t+0,t}$ is the forecast for the current year (*t+0*) made at year *t*, $y_{t+1,t}$ is the forecast for the year ahead (*t+1*) made at year *t* and ε_{t+1} is an independently and identically distributed residual term. If the forecast for the year (*t+1*) encompasses the naïve forecast and also includes significant additional information, $\beta=\gamma=0$ and $\delta > 0$.

5 Testing forecast information efficiency

Above, it was studied whether the forecast encompasses naïve forecast and whether it in addition includes other significant information. A stronger requirement for a forecast is whether the forecast contains all available information at the time of the forecast. To test this, the following regressions were performed on the (*t+0*) and (*t+1*) forecasts (European Commission, 2016, p. 38):

$$y_t = \alpha + \beta y_{t+0,t} + \varepsilon_{1t}$$

$$y_{t+1} = \alpha + \beta y_{t+1,t} + \varepsilon_{1t+1},$$

where $y_t (/ y_{t+1})$ is the actual value of y at the time $t\ (/t+1)$ and $y_{t+0,t}\ (/ y_{t+1,t})$ is the forecast for the current year (/the year ahead) at the time t and $\varepsilon_t\ (/\varepsilon_{t+1})$ is the term used for independently, identically distributed residuals. If the forecast is efficient, $\alpha = 0$ and $\beta = 1$, which indicates that the mean realised outcome equals to the mean forecast.

3 Were the economic growth, unemployment rate and inflation forecasts by the FM for the years 1976–2016 reliable?

In this chapter, the reliability of the autumn forecasts on GDP growth, unemployment rate and inflation by the Ministry of Finance is studied by comparing the accuracy of the forecasts on the Finnish economy by four other major forecasters. Furthermore, temporal independence and non-persistence of the FM's forecast errors, unbiasedness of the forecasts and whether the forecasts encompassed the naïve forecast are studied by means of statistical tests. In addition, it is tested whether it can be proven that the forecasts did not include any available pertinent information when the target variable was being forecasted. If a forecast included all the pertinent information, it was information efficient. Section 3.1 reports these studies for the year ahead (*t+1*) forecasts and Section 3.2 for the current year (*t+0*) forecasts.

3.1 The reliability of the FM's year ahead forecasts (*t+1*)

The forecasts for the year ahead in the autumn are especially important, because they are used when planning the state budget, i.e. when assessing the general government revenue and expenditure. The reliability of the forecasts is studied first by describing the average bias of forecast errors, i.e. the deviation from zero either to the negative or to the positive. Secondly, reliability is studied by investigating the accuracy of the different forecasters. This is done by comparing the mean absolute errors. This variable indicates how far from zero the difference between the forecast and the actual value is on average. Next, the reliability of the forecasts by the Ministry of Finance is studied with five statistical tests.

The year ahead forecasts (*t+1*) by the FM were unbiased in the case of all the variables, and the forecast errors were temporally non-persistent. Statistically speaking, the accuracy of the Ministry's forecasts did not deviate from the other forecasters. As a general rule, the forecasts by the FM encompassed the naïve forecast (with the unemployment rate forecast being the only exception) and the pertinent information at the time of the forecast (with the inflation forecast being the only exception). In light of the results below, one cannot claim the FM to be any more unreliable than the reference forecasters in the case of the year ahead GDP growth, unemployment rate and inflation forecasts.

Comparison of the accuracy of the year ahead forecasts

Figures 11–13 in Appendix 1 indicate the forecasted GDP growth rates, unemployment rates and consumer price inflation as percentages, and the actual values. The figures below (Figs 1–3) indicate the forecast errors of the different forecasters (as percentages).

Both the figures in Appendix 1 and Figures 1–3 show that the forecast errors in the year ahead forecasts have been fairly high. The forecasts and forecast errors of the different forecasters were rather similar, even though there were some differences between the forecasters. The forecast errors of Finnish forecasters are often close to each other. It seems that PTT overestimated the GDP growth the least in 2011–2013. The OECD's forecast error seems to be higher than the others most often. One should keep in mind, though, that the forecasts by the OECD have been prepared earlier inside the calendar year than the forecasts by the other institutions.

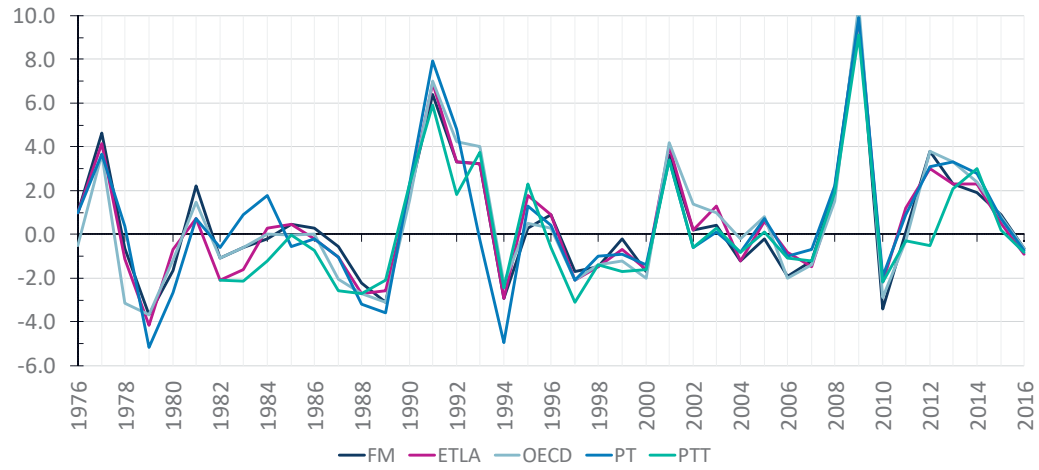

Figure 1: Forecast errors in GDP growth (*t+1*)

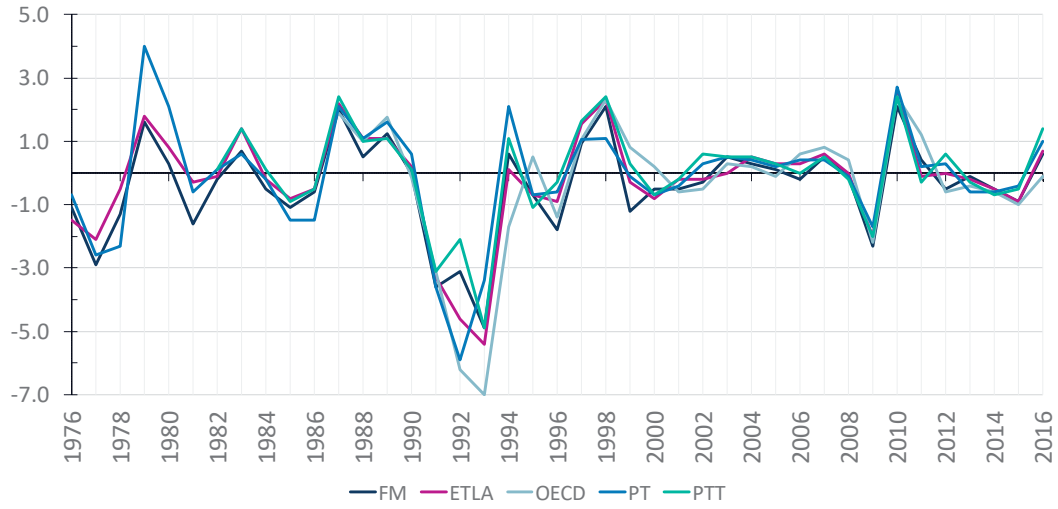

Figure 2: Forecast errors in unemployment rate (*t+1*)

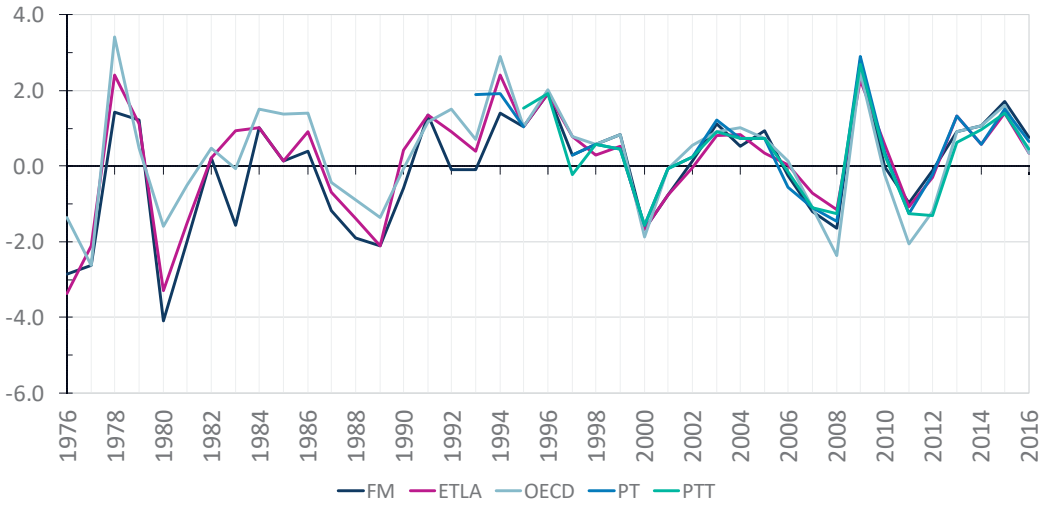

Figure 3: Forecast errors in consumer inflation (*t+1*)

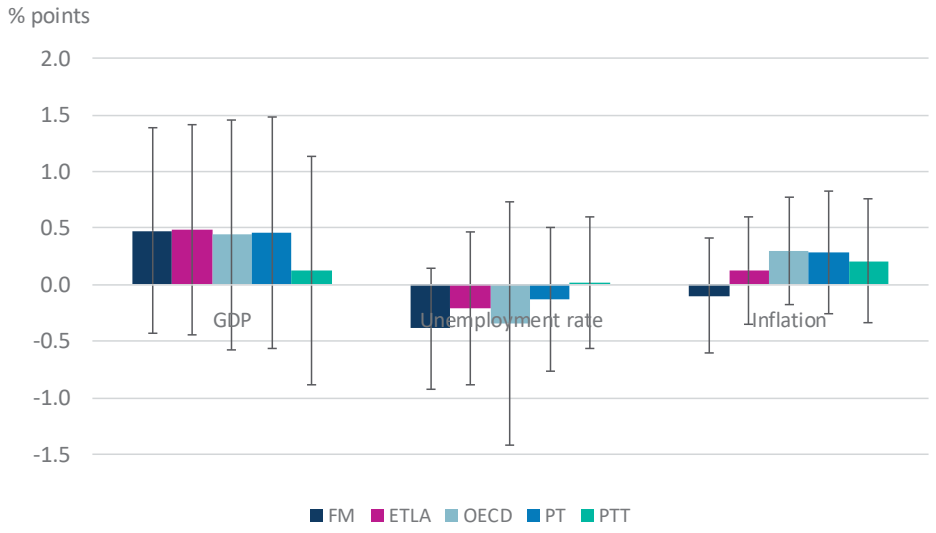

Figure 4: Mean error (percentage points) for the year ahead forecast (*t+1*) and the 95% confidence interval (t distribution with HAC standard errors)

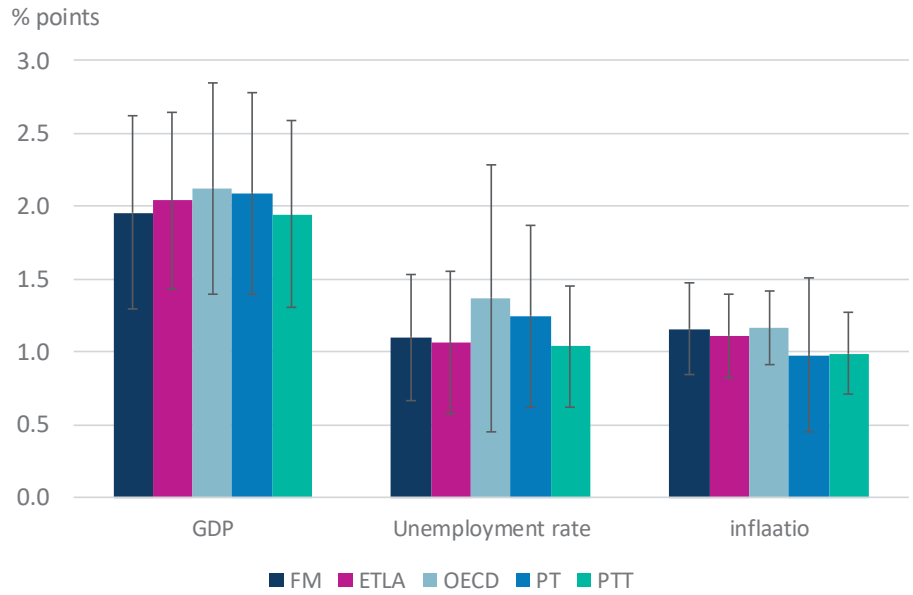

Figure 5: Mean absolute error (percentage points) for the year ahead forecast (*t+1*) and the 95% confidence interval (t distribution with HAC standard errors)

Fig. 4 indicates the mean errors for the year ahead (*t+1*) forecasts, i.e. the average bias for each variable in the case of each forecaster. The figure shows that the mean error of the Ministry of Finance for GDP growth is roughly at the same level as the mean error of the other forecasters. PTT's forecast error was the least biased on average. The forecast errors of the FM (and Etla) on inflation are the closest to zero. The FM's unemployment rate forecast errors deviate on average the most from zero, but the differences between the forecasters are very small. One should note that the mean forecast errors of the FM were within the 95% confidence interval of the other forecasters[30]: Thus, the average deviations of FM's forecast errors from zero (plus and minus) were not statistically significantly higher than in the forecast errors of the reference forecasters. Meanwhile, the entire group's year ahead forecast errors with these variables does not seem statistically significantly to deviate from zero in the *t* test (with *HAC* standard errors). The unbiasedness of the forecasts is investigated in more detail below.

Fig. 5 illustrates the forecasters' average accuracy with the mean absolute error. This variable indicates how far from zero the forecast error of each forecaster was between 1976 and 2016 on average. As the mean absolute error (MAE) is a measure of the average distance from zero, this figure can be better used when comparing the accuracy of the forecasters (another alternative would be using the root of the mean squared error of the forecasts). The figure shows the magnitude of the forecasts' mean deviations: the mean absolute error of the year ahead GDP growth forecasts has been, depending on the forecaster, a little under or a little over two per centage points; the unemployment rate MAE has been between one and 1.4 per cent; and the inflation MAE has been between one and 1.2 per cent.

Based on the MAE of the sample from 1976–2016, the Ministry of Finance has been the second most accurate in forecasting GDP growth in the year ahead forecasts. The FM is very close to the best forecaster, PTT, and Etla is a very close third right after the FM. The situation in the case of the accuracy of the year ahead unemployment forecasts is very similar, except that the second most accurate forecaster after PTT is Etla and FM is the third right after Etla. In the case of year ahead forecasts on inflation, the FM and the OECD were the least accurate and PT and PTT the most accurate. All in all, the forecasters have been very close to each other in terms of the average forecast accuracy: in the sample, the mean absolute errors in the GDP and inflation forecasts of all the forecasters were within a couple of decimals of percentage points. The difference in the case of the unemployment rate forecasts was around three decimals (around four decimals for the OECD).

Fig. 5 also indicates the 95% confidence interval for the MAE of each forecaster, which offers a wider perspective for the assessment of the forecast accuracy than only the mean values calculated from the sample. The MAE of the FM in all of the variables falls within the 95% confidence interval of the other forecasters. On the basis of this, *one cannot claim that the accuracy of the FM's year ahead GDP, unemployment rate or inflation forecasts statistically significantly deviates from the forecasts of the other forecasters*.

Studying unbiasedness and reliability of the year ahead forecasts by the Ministry of Finance with statistical tests

As Section 2.2 indicates, the unbiasedness and reliability of forecasts can be studied by means of statistical tests. An estimator is unbiased if an estimate calculated from a sample hits on average the population parameter value. In such a case, the difference between the estimate and the population parameter, i.e. the bias, does not statistically significantly deviate from zero.

The unbiasedness of a forecast suggests that the forecast is, on average, sufficiently close to the actual value published at a later time. In such a case, the difference between the forecast and the actual value, i.e. the forecast error, does not statistically significantly deviate from zero. As stated in Chapter 2, unbiasedness can be studied with, for example, a *t* test or by regressing the forecast error with a constant term.

Table 1: Testing temporal independence of the forecast errors in the Ministry of Finance forecasts for the year (*t+1*). A Ljung-Box autocorrelation test on GDP growth (*t+1*) forecast errors in 1976–2016.

	GDP growth *t+1*	*Unemployment rate* *t+1*	Inflation *t+1*
H0: autocorrelation coefficients with lags 1–3 = 0			
Q ~ X^2(3),	3.17	2.48	3.81
p value	*p* = 0.37	*p* = 0.48	*p* = 0.28
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Number of lags	3	3	3
Degrees of freedom	3	3	3
Observations	41	41	41

Before performing the unbiasedness test and other tests, the temporal independence of forecast errors must be studied. If any temporal dependence is detected, it must be taken into account in the other tests. The summarised results in Table 1 indicate that the test hypothesis (the H0 hypothesis) of the FM's (*t+1*) year forecast errors being temporally independent cannot be rejected in the case of any of the studied variables.

Table 2: Testing unbiasedness of the year ahead forecast (*t+1*) with the t test, 1976–2016

	GDP growth *t+1*	Unemployment rate *t+1*	Inflation *t+1*
H0: mean forecast error = 0			
Mean value	0.48	-0.39	-0.10
(standard error)	(0.42)	(0.23)	(0.23)
t value	1.12	-0.168	-0.43
Pr(>|t|)	0.27	0.101	0.67
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	41

Table 3: Testing unbiasedness of (*t+1*) forecast errors of the Ministry of Finance in 1976–2016 with the regression model

$e_{t+1,t} = \alpha + \varepsilon_{t+1,t}$

	GDP growth *t+1*	Unemployment rate *t+1*	Inflation *t+1*
H0: alpha = 0			
Alpha	0.48	-0.39	*-0.11*
(standard error)	(0.42)	(0.23)	*(0.23)*
t value	*1.12*	*-1.67*	*-0.48*
Pr(>|t|)	*0.27*	*0.104*	*0.64*
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	41

The results in Tables 2 and 3 indicate that the test hypothesis on the unbiasedness of the (*t+1*) year forecasts by the FM cannot be rejected on the basis of the *t* test and the regression test for the variables studied from 1976–2016.

The results in Table 4 indicate that the test hypothesis "The forecast encompasses the naïve forecast" cannot be rejected in the case of the FM's GDP growth and inflation (*t+1*) forecasts, but the hypothesis must be rejected in the case of the unemployment rate forecasts. Table 5 indicates that the test hypothesis "The forecast covers all the current pertinent information" cannot be rejected in the case of the FM's GDP growth and unemployment rate (*t+1*) forecasts, but the hypothesis must be rejected in the case of the inflation forecasts.

All of the forecast errors in the FM's year ahead forecasts proved to be temporally non-persistent and the forecasts were proven unbiased. The GDP growth and inflation forecasts encompassed the naïve forecast, and the GDP growth and unemployment rate forecasts covered all the pertinent information at the time of the forecast. To achieve an overall idea of how the above-mentioned results are proportioned to the results from a larger group of forecasters, Tables 17–44 in Appendices 3–6 include the results of the same tests on the *(t+1)* forecasts of other studied economic forecasters. Furthermore, there is a summary of the results for the FM and the other forecasters in Chapter 4.

Table 4: Test on encompassing the naïve forecast 1976–2016, (*t+1*) forecast

$$y_{t+1} = \alpha + \beta y_{t-1} + \gamma y_{t+0,t} + \delta y_{t+1,t} + \varepsilon_{t+1}$$

	GDP growth *t+1*	Unemployment rate *t+1*	Inflation *t+1*
1) Beta **(standard error)** *H0: beta = 0*	0.02 (1.30)	-0.23 (0.29)	*-0.18* *(0.14)*
t value	0.13	-0.80	*-1.32*
Pr(>|t|)	0.90	0.43	*0.20*
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
2) Gamma **(standard error)** H0: gamma = 0	-0.55 (0.30)	0.38 (0.78)	0.15 (0.26)
t value	-1.88	0.49	0.55
Pr(>|t|)	0.068	0.63	0.58
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
3) Delta **(standard error)** H0: delta = 0, delta > 0	2.45*** (0.60)	0.79 (0.58)	1.13*** (0.31)
t value	4.01	1.37	3.62
Pr(>|t|)	0.000***	0.18	0.0009***
Statistical inference	Delta ≠ 0, delta > 0	**Delta = 0,** ***H0 must be rejected***	Delta ≠ 0, delta > 0
Observations	39	39	39
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

Table 5: Testing information efficiency of GDP growth forecast errors (*t+1*) with the regression model, 1976–2016

$$y_{t+1} = \alpha + \beta y_{t+1,t} + \varepsilon_{1t+1}$$

	GDP growth *t+1*	Unemployment rate *t+1*	Inflation *t+1*
Alpha **(standard error)** *H0: alpha = 0*	-1.14 (1.05)	0.82 (0.57)	-0.56 (0.31 *HAC*)
t value	*-1.09*	*1.43*	*-1.81*
Pr(>|t|)	*0.28*	*0.16*	*0.08*
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Beta – 1 **(standard error)** *H0: (beta-1) = 0*	0.25 (0.36)	-0.05 (0.064)	0.17* (0.08 *HAC*)
t value	0.70	-0.82	2.13
Pr(>|t|)	0.49	0.42	0.04*
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	***H0 must be rejected***
Observations	41	41	41

HAC heteroscedasticity and autocorrelation consistent standard error

3.2 The reliability of the FM's current year forecasts (*t+0*)

As above, the reliability of the current year forecasts is first studied by means of the mean error of a sample of the forecasts in 1976–2016. The unbiasedness of the different forecasters is described with mean error diagrams and accuracy is compared with mean absolute errors. Next, the independence, unbiasedness and reliability of the Ministry of Finance's current year forecasts is studied with statistical tests. The accuracy of the current year forecasts was somewhat better than the accuracy of the year ahead forecasts, which were discussed above. The results were fairly similar, however.

All in all, the FM cannot be deemed to be statistically any worse than the best forecasters in the reference group in the case of the reliability of the current year forecasts either: the forecasts were unbiased in the case of all the studied variables and the forecast errors were temporally non-persistent. Based on statistical testing, the accuracy of the Ministry's GDP and inflation forecasts did not statistically deviate from the other forecasters. Etla was statistically somewhat more accurate in its unemployment rate forecasts. As a general rule, the forecasts by the FM encompassed the naïve forecast (with the GDP forecast being the only exception). According to the statistical test, the Ministry's forecasts covered all the pertinent information at the time of forecasting in the case of all the variables.

In light of all the results, one cannot claim the FM to be statistically any more unreliable than the reference forecasters in the case of the current year GDP growth, unemployment rate and inflation forecasts. However, the accuracy of Etla's unemployment rate forecasts was statistically somewhat better than the accuracy of FM's forecasts.

Comparison of the accuracy of the current year forecasts

As in the case of the (*t+1*) forecasts, the accuracy of the current year forecasts (*t+0*) is studied by comparing the forecasts of the FM and the four other forecasters. Figures 14–16 in Appendix 2 present the current year forecasts on the growth rate, unemployment rate and consumer price inflation (as percentages) and the actual values. Figures 6–8 below present the forecast errors of the different forecasters' (*t+0*) GDP growth, unemployment rate and inflation forecasts (as percentages).

The figures show that the forecasts and forecast errors were yet again roughly similar, even though there were some differences between the forecasters. One should note that the OECD's inflation forecast error rate (although being unbiased) was higher than that of the other forecasters, probably because most of the OECD's forecasts were generated earlier in the calendar year than the forecasts of the other institutions, in May or June (almost half of PT's forecasts are also from the spring or June).

Fig. 9 presents the mean errors of the forecasters' current year forecasts, i.e. the average bias. The FM's GDP growth mean forecast error in 1976–2016 was the least biased and hardly deviated from zero. Etla and PT were not far behind. In the case of the unemployment rate forecasts, the FM and the OECD were the second least biased, while the forecasts by PT were the least biased. The mean error of the FM's inflation forecasts was very close to the least biased forecasters, which were PTT, Etla and PT.

Fig. 10 presents the mean absolute errors for comparison of the accuracy of the sample from 1976–2016. The FM was the third most accurate in forecasting the current year unemployment rate and GDP growth; only Etla and PTT were more accurate than the FM. In the case of inflation, the FM and Etla were the second most accurate and PTT was the most accurate. Accuracy of the GDP and inflation forecasts by the FM does not statistically significantly deviate from the other most successful forecasters, PTT and Etla. Statistically speaking, Etla has been slightly more accurate in forecasting the unemployment rate, but the MAE of the FM is only slightly outside the Etla confidence interval. The OECD's accuracy in forecasting the GDP and inflation and PT's accuracy in forecasting inflation seem to have been statistically poorer, however, as already mentioned, they have produced a good amount of their forecasts earlier in the calendar year than the others.

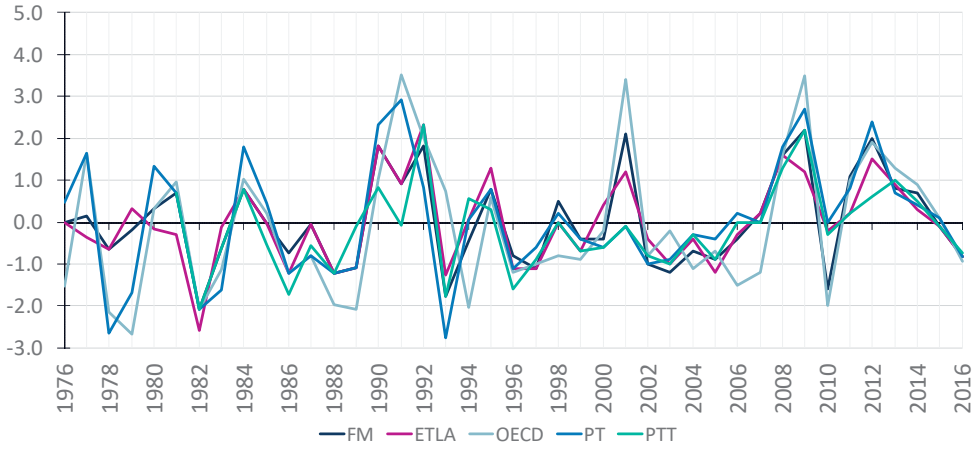

Figure 6: Forecast errors in GDP Growth (*t+0*)

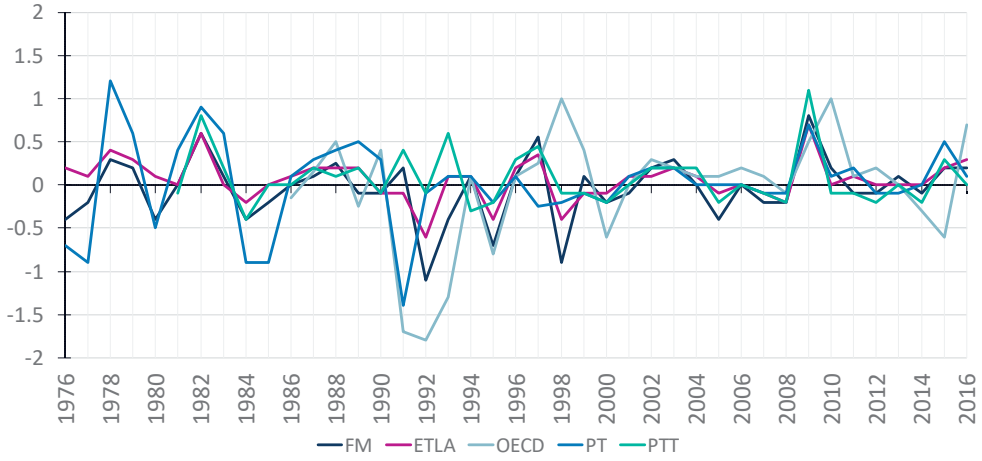

Figure 7: Forecast errors in unemployment rate (*t+0*)

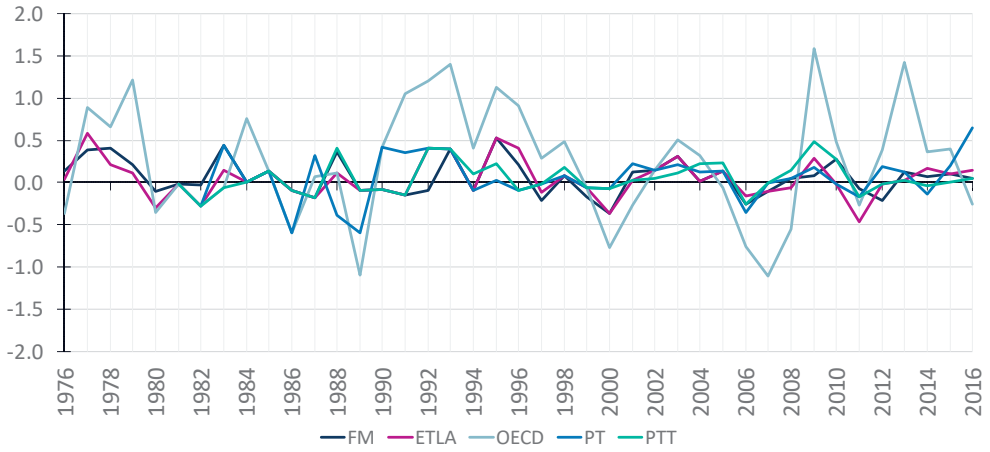

Figure 8: Forecast errors in consumer inflation (*t+0*)

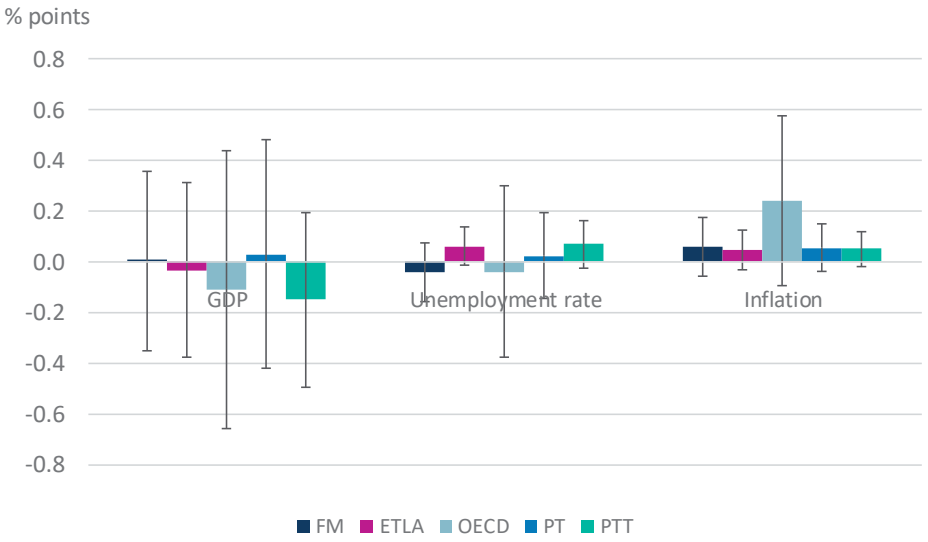

Figure 9: Mean error (percentage points) for the current year forecast (*t+0*) and the 95% confidence interval (*t* distribution with *HAC* standard errors)

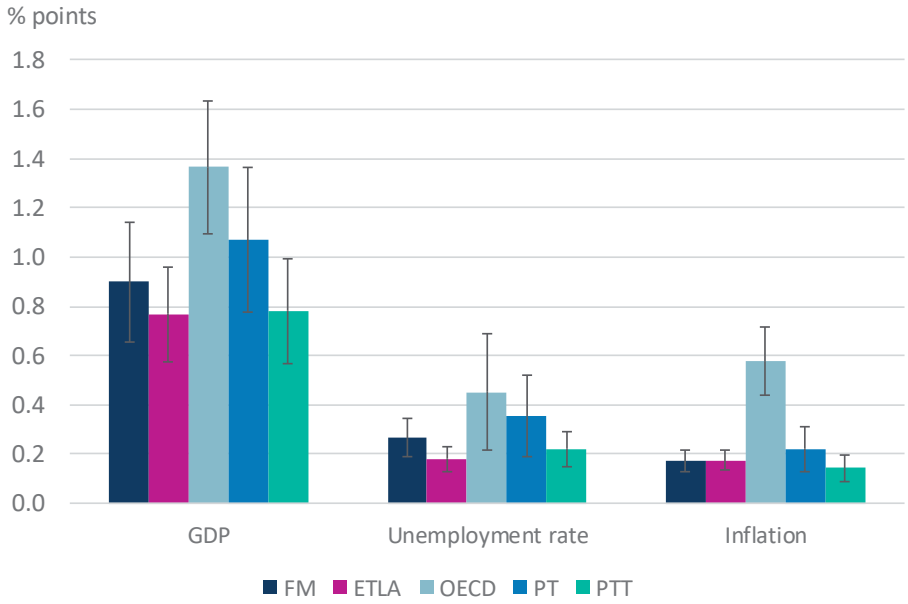

Figure 10: Mean absolute error (percentage points) for the current year forecast (*t+0*) and the 95% confidence interval (*t* distribution with *HAC* standard errors)

Studying unbiasedness and reliability of the current year forecasts by the Ministry of Finance with statistical tests

A forecast is considered unbiased if the forecast error mean does not statistically significantly deviate from zero. This is tested with the *t* test and the regression test.

As above, the temporal independence of forecast errors is studied first, because any autocorrelation must be taken into account in the unbiasedness tests. The results in Table 6 indicate that the hypothesis on temporal independence of the FM's (*t+0*) forecasts cannot be rejected in the case of GDP growth, unemployment rate or inflation.

Table 6: Testing temporal independence of the forecast errors in the Ministry of Finance forecast for the year (*t+0*). A Ljung-Box autocorrelation test on GDP growth (*t+0*) forecast errors in 1976–2016

	GDP growth *t+0*	Unemployment rate *t+0*	Inflation *t+0*
H0: autocorrelation coefficients with lags 1–3 = 0			
Q ~ X²(3),	2.20	5.92	1.39
p value	*p* = 0.53	*p* = 0.12	*p* = 0.71
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Number of lags	3	3	3
Degrees of freedom	3	3	3
Observations	41	41	41

Table 7: Testing unbiasedness of the current year forecast (*t+0*) with the t test, 1976–2016

	GDP growth *t+0*	Unemployment rate *t+0*	Inflation *t+0*
H0: mean forecast error = 0			
Mean value	0.005	-0.04	0.06
(standard error)	(0.17)	(0.057)	(0.03)
t value	0.03	-0.73	1.84
Pr(>|t|)	0.98	0.47	0.07
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	41

Table 8: Testing unbiasedness of (*t+0*) forecast errors of the Ministry of Finance in 1976–2016 with the regression model

$e_{t+0,t} = \alpha + \varepsilon_{t+0,t}$			
	GDP growth *t+0*	Unemployment rate *t+0*	Inflation *t+0*
Alpha	0.005	-0.039	*0.06*
(standard error)	(0.17)	(0.058)	*(0.033)*
H0: alpha = 0			
t value	*0.028*	*-0.68*	*1.84*
Pr(>|t|)	*0.98*	*0.50*	*0.07*
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	41
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

Table 9: Test on encompassing the naïve forecast 1976–2016, (*t+0*) forecast

$$y_t = \alpha + p y_{t-1} + o y_{t+0,t} + \varepsilon_t$$

	GDP growth *t+0*	Unemployment rate *t+0*	Inflation *t+0*
1) Beta **(standard error)** *H0: beta = 0*	-0.14* (0.07)	0.0349 (0.035)	*-0.024* *(0.019 **HAC**)*
t value	-2.20	0.99	*-1.29*
Pr(>|t|)	0.0341*	0.33	*0.21*
Statistical inference	***H0 must be rejected***	*H0 cannot be rejected*	*H0 cannot be rejected*
2) Delta **(standard error)** H0: delta = 0, delta > 0	1.17*** (0.07)	0.98 (0.04)	1.01*** (0.017 ***HAC***)
t value	15.73	26.85	60.28
Pr(>|t|)	0.000***	0.000***	0.000***
Statistical inference	Delta ≠ 0, delta > 0	Delta ≠ 0, delta > 0	Delta ≠ 0, delta > 0
Observations	40	40	40
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

HAC heteroscedasticity and autocorrelation consistent standard error

Table 10: Testing information efficiency of current year forecast errors (*t+0*) with the regression model, 1976–2016

$$y_t = \alpha + \beta y_{t+0,t} + \varepsilon_{1t}$$

	GDP growth *t+0*	Unemployment rate *t+0*	Inflation *t+0*
Alpha **(standard error)** *H0: alpha = 0*	-0.19 (0.22)	-0.03 (0.14)	-0.013 (0.049)
t value	*-0.86*	*-0.19*	*-0.28*
Pr(>|t|)	*0.40*	*0.85*	*0.78*
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Beta – 1 **(standard error)** *H0: (beta-1) = 0*	0.08 (0.06)	0.01 (0.016)	-0.01 (0.009)
t value	1.30	0.52	-1.06
Pr(>|t|)	0.20	0.61	0.29
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	41

The results of Tables 7 and 8 above indicate that the test hypothesis of the forecasts being unbiased cannot be rejected in the case of the FM's current year GDP, unemployment rate and inflation forecasts. Tables 9 and 10 indicate that the test hypotheses "The forecast encompasses the naïve forecast" and "The forecast covers all the pertinent information at

the time of forecasting" cannot be rejected in the case of the unemployment rate and inflation. The strictest hypothesis on coverage of all pertinent information cannot be rejected in the case of the GDP growth, even though there is some indication suggesting that the test hypothesis "The forecast encompasses the naïve forecast" should be rejected at the five per cent level of significance ($p < 0.05$). To proportion the typicality of the results, the test results for the other forecasters are given in Tables 17–44 in Appendices 3–6. Furthermore, Chapter 4 includes a summary of the test results of the FM and the other forecasters.

The current year forecasts of all the forecasters proved to be unbiased and the forecast errors of the FM, Etla and PT were temporally non-persistent. The forecasts by the FM and Etla passed the information efficiency tests in the case of all the variables. The FM's current year GDP forecasts did not pass the information coverage test that was originally expected to be the weaker of the two tests, but they did pass the stricter test. The results for PTT's unemployment rate and Etla's inflation forecasts were similar.

4 Summary and conclusions

This report assessed the reliability of forecasts by the Ministry of Finance on GDP growth, unemployment rate and inflation in 1976–2016. Reliability was studied, first, by comparing the average bias and accuracy of the FM's forecasts to those by four other forecasting institutions. Next, it was studied whether the forecast errors were temporally non-persistent and whether the forecasts were statistically significantly unbiased. Finally, the information covered by the forecasts and the information efficiency of the forecasts were studied by testing whether the forecasts encompassed the naïve forecast and whether they included all pertinent information at the time of forecasting. *No such characteristics (bias) that would require corrective action based on the Government Decree on the General Government Fiscal Plan were detected in the GDP, unemployment rate and inflation forecasts by the Ministry of Finance.*

The average forecast error or bias was explored as the average deviation from zero either to the negative or to the positive direction. This average bias of the FM forecasts was compared with the results for the other forecasters. Next, the accuracy of the FM's forecasts was compared with the accuracy of the reference forecasters using the mean absolute error, i.e. by calculating how far from zero (regardless of the direction), the forecasts were from the actual values on average.

When analysing both the average bias and the accuracy, a 95% statistical confidence interval was calculated for each forecaster. The FM's average bias and accuracy in the case of year ahead *t+1*) forecasts were within the 95% confidence interval of all the other four forecasters. Therefore, the average bias and accuracy of the year ahead GDP, unemployment rate and inflation forecasts of the Ministry of Finance cannot be deemed to have statistically significantly deviated from the reference forecasters.

In the case of all the variables, the current year (*t+0*) forecast mean error or bias of the FM was within the confidence intervals of the best forecasters, i.e. the Finnish forecasters. (The OECD's (*t+0*) inflation forecast was more biased than the other forecasters, to the positive direction). In terms of average accuracy, the FM's current year GDP and inflation forecasts were statistically speaking among the three best forecasters, together with Etla and PTT. The FM's accuracy was statistically as good as the second best forecaster, PTT. On the basis of the results, one cannot claim that the accuracy of the current year GDP and inflation forecasts by the Ministry of Finance statistically deviated from the two other top forecasters. In current year unemployment rate forecast accuracy Etla was statistically speaking slightly better than the FM.

In addition to the above-mentioned examinations, temporal independence and unbiasedness of forecast errors in the case of all the three variables, whether the forecasts encompassed the naïve forecast and whether the forecasts included all the current pertinent information were studied by means of statistical tests.

The FM's year ahead (*t+1*) forecasts passed 13 out of the 15 reliability tests together with Etla. None of the above-mentioned statistical tests could be rejected in the case of the FM's GDP growth forecasts. The FM's unemployment rate forecasts passed all the tests except for the naïve forecast encompassing test, and the inflation forecasts failed only the strictest test, i.e. the information efficiency test. Taking into account all of the tests, one cannot claim that the GDP, unemployment rate and inflation forecasts of any of the other forecasters were more reliable than the forecasts by the FM. At the same time, Etla proved to be as reliable as the FM.

Together with the forecasts by Etla, the FM's current year (*t+0*) GDP, unemployment rate and inflation forecasts passed 14 out of the 15 reliability tests, while the forecasts by PTT passed 13 tests. Forecasts by PT passed 12 and by OECD 11 tests, having both produced a good amount of their forecasts earlier in the calendar year than others. In the case of the FM's current year forecasts, only passing the naïve forecast encompassing test in GDP was not fully confirmed, while the strictest test on information efficiency was passed. Based on all the test results for the current year forecasts, one cannot claim that the FM was not overall as reliable as the other forecasters in the forecasting of the GDP, unemployment rate and inflation.

The assessment is concluded by presenting summary tables including the results of all the statistical tests to which the FM and the four reference forecasters were subjected.

Table 11: Statistical testing of the reliability of GDP growth forecasts (*t+1*) in 1976–2016

GDP growth (*t+1*) forecast	1. Temporal independence of forecast errors	2. Unbiasedness of forecast – *t* test	3. Unbiasedness of forecast – *regression* test	4. Does the forecast encompass the naïve forecast?	5. Forecast information efficiency
Test hypothesis	Forecast errors are independent with lags 1–3	Forecast is unbiased	Forecast is unbiased	Forecast encompasses the naïve forecast	Forecast covers all pertinent information
FM	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected
ETLA	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected
OECD	Cannot be rejected	Cannot be rejected	Cannot be rejected	*Must be rejected*	Cannot be rejected
PT	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected
PTT	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected

Table 12: Statistical testing of the reliability of unemployment rate forecasts (*t+1*) in 1976–2016

Unemployment rate (*t+1*) forecast	1. Temporal independence of forecast errors	2. Unbiasedness of forecast – *t* test	3. Unbiasedness of forecast – *regression* test	4. Does the forecast encompass the naïve forecast?	5. Forecast information efficiency
Test hypothesis	Forecast errors are independent with lags 1–3	Forecast is unbiased	Forecast is unbiased	Forecast encompasses the naïve forecast	Forecast covers all pertinent information
FM	Cannot be rejected	Cannot be rejected	Cannot be rejected	*Must be rejected*	Cannot be rejected
ETLA	*Cannot be rejected* (at the rejection limit)	Cannot be rejected	Cannot be rejected	*Must be rejected*	Cannot be rejected
OECD	*Must be rejected*	Cannot be rejected	Cannot be rejected	*Must be rejected*	*Must be rejected (residual non-random)*
PT	*Must be rejected*	Cannot be rejected	Cannot be rejected	*Must be rejected*	*Must be rejected (residual non-random)*
PTT	Cannot be rejected	Cannot be rejected	Cannot be rejected	*Must be rejected*	Cannot be rejected

Table 13: Statistical testing of the reliability of inflation forecasts (*t+1*) in 1976–2016

Inflation (*t+1*) forecast	1. Temporal independence of forecast errors	2. Unbiasedness of forecast – *t* test	3. Unbiasedness of forecast – *regression* test	4. Does the forecast encompass the naïve forecast?	5. Forecast information efficiency
Test hypothesis	Forecast errors are independent with lags 1–3	Forecast is unbiased	Forecast is unbiased	Forecast encompasses the naïve forecast	Forecast covers all pertinent information
FM	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected	*Must be rejected*
ETLA	Cannot be rejected	Cannot be rejected	Cannot be rejected	*Must be rejected*	Cannot be rejected
OECD	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected
PT	Cannot be rejected	Cannot be rejected	Cannot be rejected	(no test, n = 23)	(no test, n = 23)
PTT	Cannot be rejected	Cannot be rejected	Cannot be rejected	(no test, n = 22)	(no test, n = 22)

Table 14: Statistical testing of the reliability of GDP growth forecasts (*t+0*) in 1976–2016

GDP growth (*t+0*) forecast	1. Temporal independence of forecast errors	2. Unbiasedness of forecast – *t* test	3. Unbiasedness of forecast – *regression* test	4. Does the forecast encompass the naïve forecast?	5. Forecast information efficiency
Test hypothesis	Forecast errors are independent with lags 1–3	Forecast is unbiased	Forecast is unbiased	Forecast encompasses the naïve forecast	Forecast covers all pertinent information
FM	Cannot be rejected	Cannot be rejected	Cannot be rejected	*Must be rejected*	Cannot be rejected
ETLA	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected
OECD	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected	*Must be rejected*
PT	*Must be rejected*	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected
PTT	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected

Table 15: Statistical testing of the reliability of unemployment rate forecasts (*t+0*) in 1976–2016

Unemployment rate (*t+0*) forecast	1. Temporal independence of forecast errors	2. Unbiasedness of forecast – *t* test	3. Unbiasedness of forecast – *regression* test	4. Does the forecast encompass the naïve forecast?	5. Forecast information efficiency
Test hypothesis	Forecast errors are independent with lags 1–3	Forecast is unbiased	Forecast is unbiased	Forecast encompasses the naïve forecast	Forecast covers all pertinent information
FM	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected
ETLA	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected
OECD	*Must be rejected*	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected
PT	*Must be rejected*	Cannot be rejected	Cannot be rejected	Cannot be rejected	*Must be rejected (residual non-random)*
PTT	Cannot be rejected	Cannot be rejected	Cannot be rejected	*Must be rejected*	Cannot be rejected

Table 16: Statistical testing of the reliability of inflation forecasts (*t+0*) in 1976–2016

Inflation (*t+0*) forecast	1. Temporal independence of forecast errors	2. Unbiasedness of forecast – *t* test	3. Unbiasedness of forecast – *regression* test	4. Does the forecast encompass the naïve forecast?	5. Forecast information efficiency
Test hypothesis	Forecast errors are independent with lags 1–3	Forecast is unbiased	Forecast is unbiased	Forecast encompasses the naïve forecast	Forecast covers all pertinent information
FM	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected
ETLA	Cannot be rejected	Cannot be rejected	Cannot be rejected	*Must be rejected* (p = 0.045)	Cannot be rejected
OECD	Cannot be rejected	*Must be rejected*	*Must be rejected*	Cannot be rejected	Cannot be rejected
PT	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected
PTT	Cannot be rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected	*Must be rejected*

Appendix 1: Year ahead (*t+1*) forecasts and actual figures

Appendix 1: Year ahead (*t+1*) forecasts and actual figures

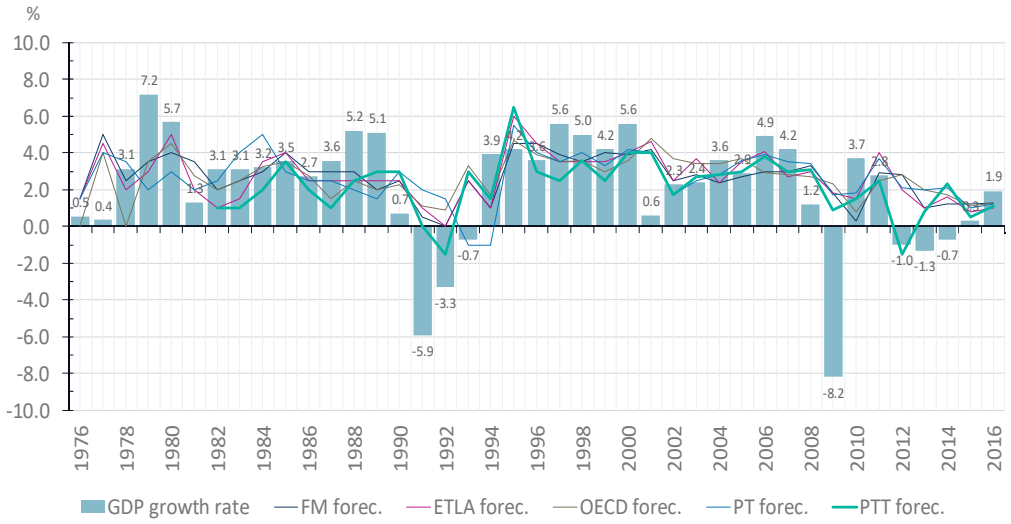

Figure 11: Actual GDP growth and year ahead (*t+1*) forecasts in 1976–2016

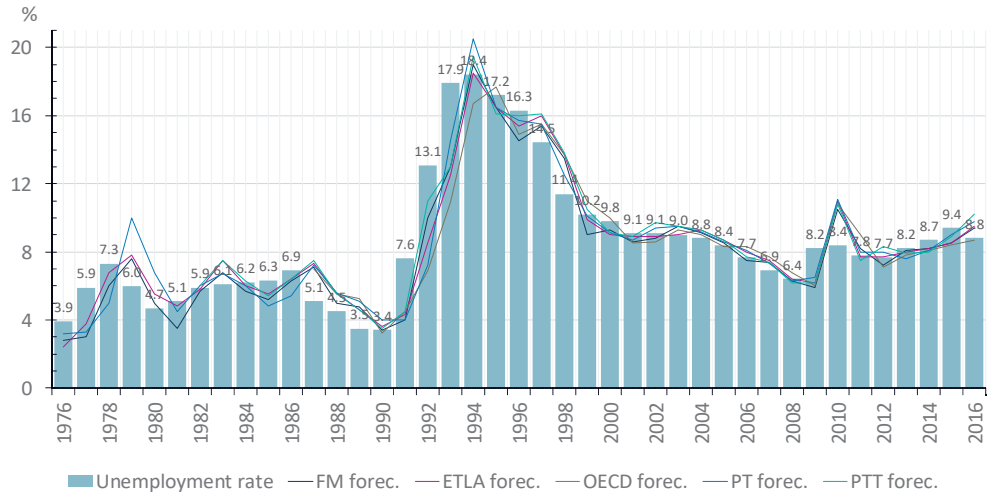

Figure 12: Actual unemployment rate and year ahead (*t+1*) forecasts in 1976–2016

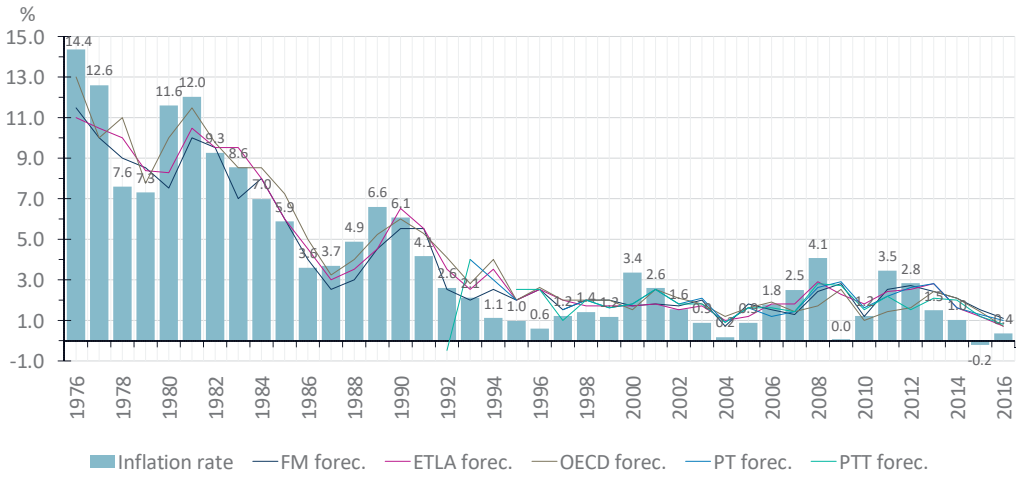

Figure 13: Actual inflation and year ahead (*t+1*) forecasts in 1976–2016

Appendix 2: Current year (*t+0*) forecasts and actual figures

Appendix 2: Current year (*t+0*) forecasts and actual figures

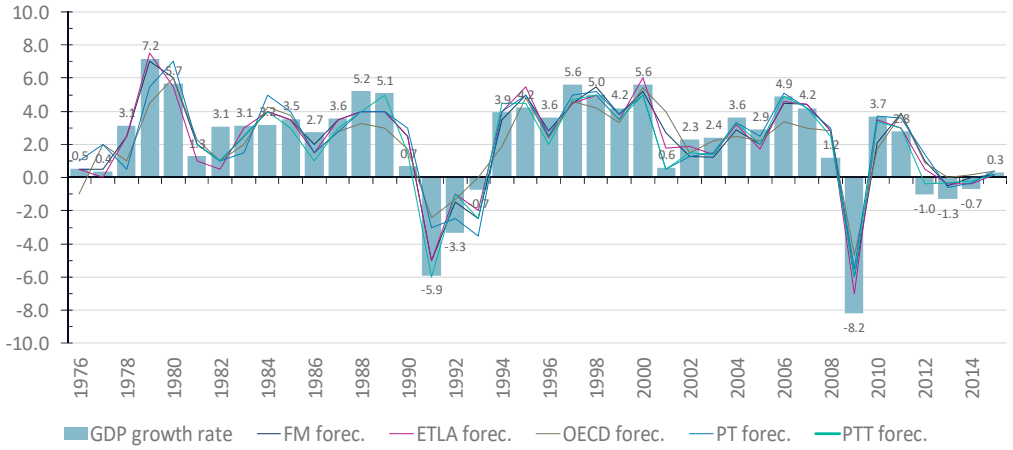

Figure 14: Actual GDP growth and current year (*t+0*) forecasts in 1976–2016

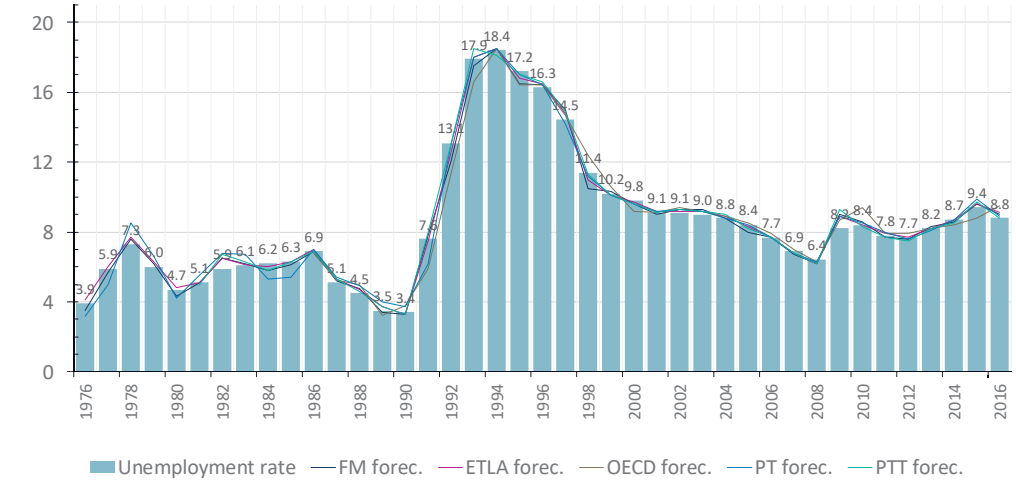

Figure 15: Actual unemployment rate and current year (*t+0*) forecasts in 1976–2016

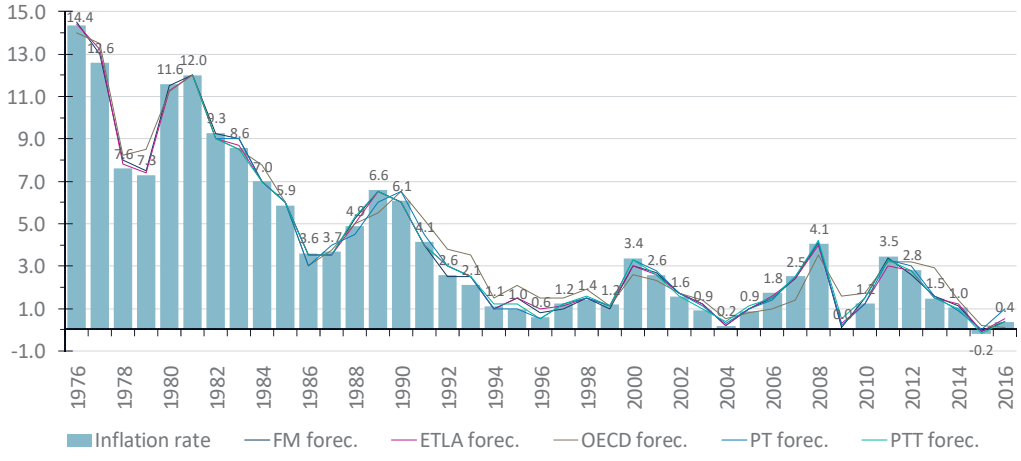

Figure 16: Actual inflation and year ahead (*t+1*) forecasts in 1976–2016

Appendix 3: Statistical tests, forecasts by Etla

Appendix 3: Statistical tests, forecasts by Etla

Table 17: Testing temporal independence of Etla's (*t+1*) and (*t+0*) forecast errors in 1976–2016

Year ahead (t+1)	*GDP growth (t+1)*	*Unemployment rate (t+1)*	*Inflation (t+1)*
H0: autocorrelation coefficients with lags 1–3 = 0			
Q ~ $X^2(3)$	1.62	7.80	*6.87*
	p = 0.66	*p = 0.05*	*p = 0.08*
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected – close to the limit value*	*H0 cannot be rejected*
Observations	41	41	41
Current year (t+0)	*GDP growth (t+0)*	*Unemployment rate (t+0)*	*Inflation (t+0)*
H0: autocorrelation coefficients with lags 1–3 = 0			
Q ~ X2(3)	1.80	1.33	1.30
	p = 0.61	p = 0.72	p = 0.73
Statistical inference	H0 cannot be rejected	H0 cannot be rejected	H0 cannot be rejected
Observations	41	41	41

Table 18: Testing unbiasedness of Etla's (*t+1*) and (*t+0*) forecasts with the *t* test

Year ahead (t+1)	*GDP growth (t+1)*	*Unemployment rate (t+1)*	*Inflation (t+1)*
H0: mean error = 0			
Mean (s.e.)	0.49	-0.21	0.12
	(0.42)	(0.23 *modified*)	(0.22)
t value	1.14	-0.91	0.57
Pr(>|t|)	0.26	0.37	0.57
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	41
Current year (t+0)	*GDP growth (t+0)*	*Unemployment rate (t+0)*	*Inflation (t+0)*
H0: mean error = 0			
Mean (s.e.)	-0.03	0.06	0.05
	(0.16)	(0.04)	(0.04)
t value	-0.21	1.65	1.29
Pr(>|t|)	0.83	0.11	0.21
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	41

Modified: standard error taking into account autocorrelation in the t test

Table 19: Testing unbiasedness of Etla's (*t+1*) forecasts with the forecast error regression model, 1976–2016

$e_{t+1,t} = \alpha + \varepsilon_{t+1,t}$	GDP growth (*t+1*)	*Unemployment rate (t+1)*	Inflation (*t+1*)
H0: alpha = 0			
Alpha	0.49	-0.21	0.11
	(0.43)	(0.33 ***HAC***)	(0.22)
t value	1.15	-0.62	0.52
Pr(>|t|)	0.26	0.54	0.61
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	41
$e_{t+0,t} = \alpha + \varepsilon_{t+0,t}$	GDP growth (*t+0*)	*Unemployment rate (t+0)*	Inflation (*t+0*)
H0: alpha = 0			
Alpha	-0.03	0.06	0.34
	(0.16)	(0.038)	(0.04)
t value	-0.22	1.67	0.94
Pr(>|t|)	0.83	0.103	0.35
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	41
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

HAC heteroscedasticity and autocorrelation consistent standard error

Appendix 3: Statistical tests, forecasts by Etla

Table 20: Test on encompassing the naïve forecast, Etla's (*t+1*) forecast

$$y_{t+1} = \alpha + \beta y_{t-1} + \gamma y_{t+0,t} + \delta y_{t+1,t} + \varepsilon_{t+1}$$

	GDP growth(*t+1*)	Unemployment rate (*t+1*)	Inflation (*t+1*)
1) Beta	0.05	-0.66	-0.29*
(standard error)	(0.17)	(0.32 ***HAC***)	(0.12)
H0: beta = 0			
t value	0.30	-2.05*	-2.43
Pr(>|t|)	0.77	0.047*	0.02*
Statistical inference	*H0 cannot be rejected*	***H0 must be rejected***	***H0 must be rejected***
2) Gamma	-0.25	1.7 *	-0.08
(standard error)	(0.25)	(0.81 ***HAC***)	(0.22)
H0: gamma = 0			
t value	-1.00	2.10	-0.38
Pr(>|t|)	0.32	0.04*	0.71
Statistical inference	*H0 cannot be rejected*	***H0 must be rejected***	*H0 cannot be rejected*
3) Delta	1.55**	-0.22	1.46***
(standard error)	(0.49)	(0.60 ***HAC***)	(0.25)
H0: delta ≠ 0, delta > 0			
t value	3.17	-0.37	5.83
Pr(>|t|)	0.003**	0.71	0.000***
Statistical inference	Delta ≠ 0, delta > 0	**Delta = 0** ***H0 must be rejected***	Delta ≠ 0, delta > 0
Observations	39	39	39
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

HAC heteroscedasticity and autocorrelation consistent standard error

Table 21: Test on encompassing the naïve forecast, Etla's (*t+0*) forecast

$$y_t = \alpha + \beta y_{t-1} + \delta y_{t+0,t} + \varepsilon_t$$

	GDP growth(*t+0*)	Unemployment rate (*t+0*)	Inflation (*t+0*)
1) Beta	0.004	-0.13	-0.047*
(standard error)	(0.057)	(0.10)	(0.023)
H0: beta = 0			
t value	0.067	-1.39	-2.1
Pr(>|t|)	0.95	0.17	0.045*
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	***H0 must be rejected***
2) Delta	1.05***	0.99***	1.05***
(standard error)	(0.065)	(0.024)	(0.03)
H0: delta ≠ 0, delta > 0			
t value	16.132	40.66	41.6
Pr(>|t|)	0.000***	0.000***	0.000***
Statistical inference	Delta ≠ 0, delta > 0	Delta ≠ 0, delta > 0	Delta ≠ 0, delta > 0
Observations	40	40	40
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

Appendix 3: Statistical tests, forecasts by Etla

Table 22: Testing information efficiency of Etla's (*t+1*) forecasts with the regression model, 1976–2016

$$y_{t+1} = \alpha + \beta y_{t+1,t} + \varepsilon_{1t+1}$$

	GDP growth *(t+1)*	Unemployment rate *(t+1)*	Inflation *(t+1)*
Alpha	-0.74	0.58	-0.58
	(0.98)	(0.51 ***HAC***)	(0.34)
H0: alpha = 0			
t value	-0.76	1.14	-1.66
Pr(>|t|)	0.45	0.26	0.105
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Beta - 1	0.10	-0.04	0.11
	(0.33)	(0.07 ***HAC***)	(0.07)
H0: (beta-1) = 0			
t value	0.29	-0.64	1.69
Pr(>|t|)	0.77	0.53	0.10
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	41
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

HAC heteroscedasticity and autocorrelation consistent standard error

Table 23: Testing information efficiency of Etla's (*t+0*) forecasts with the regression model, 1976–2016

$$y_t = \alpha + \beta y_{t+0,t} + \varepsilon_{1t}$$

	GDP growth(*t+0*)	Unemployment rate (*t+0*)	Inflation (*t+0*)
Alpha	-0.08	-0.14	-0.047
	(0.19)	(0.0956)	(0.054)
H0: alpha = 0			
t value	*-0.38*	*-1.49*	-0.86
Pr(>|t|)	*0.70*	*0.14*	0.39
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Beta - 1	0.05	0.01***	0.003
	(0.057)	(0.0102)	(0.009)
H0: (beta-1) = 0			
t value	0.88	0.93	0.31
Pr(>|t|)	0.38	0.36	0.76
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	41
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

Appendix 4: Statistical tests, forecasts by the OECD

Appendix 4: Statistical tests, forecasts by the OECD

Table 24: Testing temporal independence of year (*t+1*) and (*t+0*) forecast errors by the OECD, 1976–2016; unemployment rate forecasts 1986–2016 (*t+0*) and 1987–2016 (*t+1*)

Year ahead (t+1)	*GDP growth (t+1)*	*Unemployment rate (t+1)*	*Inflation (t+1)*
H0: autocorrelation coefficients with lags 1–3 = 0			
Q ~ $X^2(3)$	2.61	11.60	2.38
	p = 0.46	*p = 0.008***	*p = 0.50*
Statistical inference	*H0 cannot be rejected*	***H0 must be rejected***	*H0 cannot be rejected*
Observations	41	30	41
Current year (t+0)	*GDP growth (t+0)*	*Unemployment rate (t+0)*	*Inflation (t+0)*
H0: autocorrelation coefficients with lags 1–3 = 0			
Q ~ $X^2(3)$	2.56	6.06	4.99
	p = 0.47	*p = 0.11*	*p = 0.17*
Statistical inference	H0 cannot be rejected	**H0 must be rejected (autocorrelation of lag 1)**	H0 cannot be rejected
Observations	41	31	41

Table 25: Testing unbiasedness of the OECD's (*t+1*) and (*t+0*) forecasts with the *t* test

Year ahead (t+1)	*GDP growth (t+1)*	*Unemployment rate (t+1)*	*Inflation (t+1)*
H0: mean error = 0			
Mean (s.e.)	0.44	-0.34	0.30
	(0.45)	(0.32 modified)	(0.22)
t value	0.97	-1.08	1.36
Pr(>|t|)	0.34	0.30	0.18
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	30	41

Current year (t+0)	*GDP growth (t+0)*	*Unemployment rate (t+0)*	*Inflation (t+0)*
H0: mean error = 0			
Mean (s.e.)	-0.11	-0.04	0.24*
	(0.26)	(0.11 modified)	(0.11)
t value	-0.43	-0.36	2.26
Pr(>|t|)	0.67	0.73	0.03*
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	***H0 must be rejected***
Observations	41	31	41

Modified: standard error taking into account autocorrelation in the t test

Table 26: Testing unbiasedness of the OECD's (*t+1*) forecasts with the forecast error regression model, 1976–2016

$e_{t+1,t} = \alpha + \varepsilon_{t+1,t}$	*GDP growth (t+1)*	*Unemployment rate (t+1)*	*Inflation (t+1)*
H0: alpha = 0			
Alpha	0.43	-0.34	0.29
	(0.45)	(0.53 ***HAC***)	(0.22)
t value	0.97	-0.65	1.33
Pr(>|t|)	0.34	0.52	0.19
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	30	41
$e_{t+0,t} = \alpha + \varepsilon_{t+0,t}$	*GDP growth (t+0)*	*Unemployment rate (t+0)*	*Inflation (t+0)*
H0: alpha = 0			
Alpha	-0.11	-0.039	0.23*
	(0.25)	(0.165 ***HAC***)	(0.11)
t value	-0.43	-0.23	2.14
Pr(>|t|)	0.67	0.81	0.04*
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	***H0 must be rejected***
Observations	41	31	41
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

HAC heteroscedasticity and autocorrelation consistent standard error

Appendix 4: Statistical tests, forecasts by the OECD

Table 27: Test on encompassing the naïve forecast, the OECD's (*t+1*) forecast

$$y_{t+1} = \alpha + \beta y_{t-1} + \gamma y_{t+0,t} + \delta y_{t+1,t} + \varepsilon_{t+1}$$

	GDP growth (t+1)	*Unemployment rate (t+1)*	*Inflation (t+1)*
1) Beta	-0.09	-1.37**	-0.09
(standard error)	(0.20)	(0.40 ***HAC***)	(0.14)
H0: beta = 0			
t value	-0.44	-3.39	-0.64
Pr(>|t|)	0.66	0.002**	0.53
Statistical inference	*H0 cannot be rejected*	***H0 must be rejected***	*H0 cannot be rejected*
2) Gamma	0.05	3.25**	-0.41
(standard error)	(0.36)	(0.77 ***HAC***)	(0.31)
H0: gamma = 0			
t value	0.13	4.25	-1.3
Pr(>|t|)	0.90	0.000***	0.20
Statistical inference	*H0 cannot be rejected*	***H0 must be rejected***	*H0 cannot be rejected*
3) Delta	1.02	-1.17*	1.48***
(standard error)	(0.56)	(0.56 ***HAC***)	(0.28)
H0: delta ≠ 0, delta > 0			
t value	1.82	-2.11	5.22
Pr(>|t|)	0.08	0.04*	0.000***
Statistical inference	**Delta = 0** ***H0 must be rejected***	**Delta = 0,** ***H0 must be rejected***	Delta ≠ 0, delta > 0
Observations	39	28	39

HAC heteroscedasticity and autocorrelation consistent standard error

Table 28: Test on encompassing the naïve forecast, the OECD's (*t+0*) forecast

$$y_t = \alpha + \beta y_{t-1} + \delta y_{t+0,t} + \varepsilon_t$$

	GDP growth (*t+0*)	Unemployment rate (*t+0*)	Inflation (*t+0*)
1) Beta	-0.15	-0.16	-0.06
(standard error)	(0.09)	(0.22 ***HAC***)	(0.08)
H0: beta = 0			
t value	-1.62	-0.72	-0.80
Pr(>|t|)	0.11	0.48	0.43
Statistical inference	*H0 must be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
2) Delta	1.44***	1.25***	1.05***
(standard error)	(0.14)	(0.12 ***HAC***)	(0.09)
H0: delta ≠ 0, delta > 0			
t value	10.41	10.87	12.21
Pr(>|t|)	0.000***	0.000***	0.000***
Statistical inference	Delta ≠ 0, delta > 0	Delta ≠ 0, delta > 0	Delta ≠ 0, delta > 0
Observations	40	29	40
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

HAC heteroscedasticity and autocorrelation consistent standard error

Appendix 4: Statistical tests, forecasts by the OECD

Table 29: Testing information efficiency of the OECD's (*t+1*) forecasts with the regression model, 1976–2016

$$y_{t+1} = \alpha + \beta y_{t+1,t} + \varepsilon_{1t+1}$$

	GDP growth *(t+1)*	Unemployment rate *(t+1)*	Inflation *(t+1)*
Alpha	-0.17 (1.14)	0.76 (0.68 ***HAC***)	-0.40 (0.36)
H0: alpha = 0			
t value	-0.15	1.12	-1.12
Pr(>|t|)	0.88	0.27	0.27
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Beta - 1	-0.10 (0.39)	-0.046 (0.0587 ***HAC***)	0.03 (0.065)
H0: (beta-1) = 0			
t value	-0.26	-0.78	0.39
Pr(>|t|)	0.79	0.44	0.70
Statistical inference	*H0 cannot be rejected*	***H0 must be rejected** – the residual is not random*	*H0 cannot be rejected*
Observations	41	30	41

HAC heteroscedasticity and autocorrelation consistent standard error

Table 30: Testing information efficiency of the OECD's (*t+0*) forecasts with the regression model, 1976–2016

$$y_t = \alpha + \beta y_{t+0,t} + \varepsilon_{1t}$$

	GDP growth *(t+0)*	Unemployment rate *(t+0)*	Inflation *(t+0)*
Alpha	-0.46 (0.34)	-0.19 (0.18 ***HAC***)	-0.23 (0.17)
H0: alpha = 0			
t value	-1.36	-1.04	-1.40
Pr(>|t|)	0.18	0.31	0.17
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Beta – 1	0.27* (0.114)	0.02 (0.024 ***HAC***)	-0.0002 (0.03)
H0: (beta-1) = 0			
t value	2.35	1.028	-0.009
Pr(>|t|)	0.024*	0.31	0.99
Statistical inference	***H0 must be rejected***	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	31	41

HAC heteroscedasticity and autocorrelation consistent standard error

Appendix 5: Statistical tests, forecasts by PT

Appendix 5: Statistical tests, forecasts by PT

Table 31: Testing temporal independence of year (*t+1*) and (*t+0*) forecast errors by PT, (inflation forecasts (*t+0*) 1982–2016, inflation forecasts (*t+1*) 1997–2016)

Year ahead (t+1)	*GDP growth (t+1)*	*Unemployment rate (t+1)*	*Inflation (t+1)*
H0: autocorrelation coefficients with lags 1–3 = 0			
Q ~ $X^2(3)$	6.59	8.73	4.93
	p = 0.086	*p* = 0.03*	*p* = 0.18
Statistical inference	*H0 cannot be rejected*	***H0 must be rejected***	*H0 cannot be rejected*
Observations	41	41	20
Current year (t+0)	*GDP growth (t+0)*	*Unemployment rate (t+0)*	*Inflation (t+0)*
H0: autocorrelation coefficients with lags 1–3 = 0			
Q ~ $X^2(3)$	8.08	10.03	5.03
	p = 0.04*	*p* = 0.018*	*p* = 0.17
Statistical inference	***H0 must be rejected***	***H0 must be rejected***	H0 cannot be rejected
Observations	41	41	41

Table 32: Testing unbiasedness of PT's (*t+1*) and (*t+0*) forecasts with the *t* test

Year ahead (t+1)	*GDP growth (t+1)*	*Unemployment rate (t+1)*	*Inflation (t+1)*
H0: mean error = 0			
Mean (s.e.)	0.46	-0.13	0.28
	(0.46)	(0.27 modified)	(0.23)
t value	1.00	-0.49	1.22
Pr(>|t|)	0.32	0.63	0.23
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	26

Current year (t+0)	*GDP growth (t+0)*	*Unemployment rate (t+0)*	*Inflation (t+0)*
H0: mean error = 0			
Mean (s.e.)	0.03	0.02	0.05
	(0.21 modified)	(0.08 modified)	(0.08)
t value	0.15	0.30	0.59
Pr(>|t|)	0.88	0.77	0.56
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	41

Modified: standard error taking into account autocorrelation in the t test

Table 33: Testing unbiasedness of PT's (*t+1*) forecasts with the forecast error regression model, 1976–2016, (inflation *t+1*: 1993–1995, 1997–2016, inflation *t+0*: 1982–2016)

$e_{t+1,t} = \alpha + \varepsilon_{t+1,t}$	GDP growth (*t+1*)	*Unemployment rate (t+1)*	Inflation (*t+1*)
H0: alpha = 0			
Alpha	0.46	-0.13	0.36
	(0.46)	(0.31 ***HAC***)	(0.24)
t value	0.999	-0.42	1.51
Pr(>|t|)	0.32	0.68	0.15
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	23
$e_{t+0,t} = \alpha + \varepsilon_{t+0,t}$	GDP growth (*t+0*)	*Unemployment rate (t+0)*	Inflation (*t+0*)
H0: alpha = 0			
Alpha	0.03	0.02	0.04
	(0.22 ***HAC***)	(0.084 ***HAC***)	(0.047 ***HAC***)
t value	0.14	0.26	0.90
Pr(>|t|)	0.89	0.795*	0.37
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	41	41	41
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

HAC heteroscedasticity and autocorrelation consistent standard error

Appendix 5: Statistical tests, forecasts by PT

Table 34: Test on encompassing the naïve forecast, PT's (*t+1*) forecast

$$y_{t+1} = \alpha + \beta y_{t-1} + \gamma y_{t+0,t} + \delta y_{t+1,t} + \varepsilon_{t+1}$$

	GDP growth (*t+1*)	Unemployment rate (*t+1*)	Inflation (*t+1*)
1) Beta	-0.12	-0.62	
(standard error)	(0.18)	(0.41 ***HAC***)	
H0: beta = 0			
t value	-0.67	-1.49	
Pr(>|t|)	0.50	0.14	
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	
2) Gamma	-0.01	1.66*	
(standard error)	(0.25)	(0.78 ***HAC***)	
H0: gamma = 0			
t value	-0.025	2.14	
Pr(>|t|)	0.98	0.039*	
Statistical inference	*H0 cannot be rejected*	***H0 must be rejected***	
3) Delta	0.89	-0.23 (*negative*)	
(standard error)	(0.45)	(0.48 ***HAC***)	
H0: delta ≠ 0, delta > 0			
t value	1.975	-0.48	
Pr(>|t|)	0.06	0.64	
Statistical inference	Delta ≠ 0, delta > 0	Delta ≠ 0, delta > 0 ***NO***	
Observations	39	39	n = 23 no testing

HAC heteroscedasticity and autocorrelation consistent standard error

Table 35: Test on encompassing the naïve forecast, PT's (*t+0*) forecast (inflation *t+0*: 1982–2016)

$$y_t = \alpha + \beta y_{t-1} + \delta y_{t+0,t} + \varepsilon_t$$

	GDP growth (*t+0*)	Unemployment rate (*t+0*)	Inflation (*t+0*)
1) Beta	-0.08	0.05	-0.01
(standard error)	(0.06 ***HAC***)	(0.07 ***HAC***)	(0.046 ***HAC***)
H0: beta = 0			
t value	-1.38	0.74	-0.20
Pr(>|t|)	0.18	0.47	0.85
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
2) Delta	1.09***	0.94***	1.02***
(standard error)	(0.12 ***HAC***)	(0.07 ***HAC***)	(0.05 ***HAC***)
H0: delta ≠ 0, delta > 0			
t value	9.33	14.39	19.16
Pr(>|t|)	0.000***	0.000***	0.000***
Statistical inference	Delta ≠ 0, delta > 0	Delta ≠ 0, delta > 0	Delta ≠ 0, delta > 0
Observations	40	40	40
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

HAC heteroscedasticity and autocorrelation consistent standard error

Appendix 5: Statistical tests, forecasts by PT

Table 36: Testing information efficiency of PT's (*t+1*) forecasts with the regression model, 1976–2016

$$y_{t+1} = \alpha + \beta y_{t+1,t} + \varepsilon_{1t+1}$$

	GDP growth *(t+1)*	Unemployment rate *(t+1)*	Inflation *(t+1)*
Alpha	0.21 (1.03)	1.10 (0.60 ***HAC***)	
H0: alpha = 0			
t value	0.20	1.84	
Pr(>|t|)	0.84	0.07	
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	
Beta – 1	-0.25 (0.35)	-0.12 (0.05 ***HAC***)	
H0: (beta-1) = 0			
t value	-0.72	-1.99	
Pr(>|t|)	0.47	0.053	
Statistical inference	*H0 cannot be rejected*	***The test must be rejected (residual non-random)***	
Observations	41	41	n = 23 no testing

HAC heteroscedasticity and autocorrelation consistent standard error

Table 37: Testing information efficiency of PT's (*t+0*) forecasts with the regression model, 1976–2016 (inflation (*t+0*): 1982–2016)

$$y_t = \alpha + \beta y_{t+0,t} + \varepsilon_{1t}$$

	GDP growth *(t+0)*	Unemployment rate *(t+0)*	Inflation *(t+0)*
Alpha	-0.14 (0.42 ***HAC***)	0.07 (0.21 ***HAC***)	-0.06 (0.06 ***HAC***)
H0: alpha = 0			
t value	-0.32	0.35	-1.06
Pr(>|t|)	0.75	0.73	0.29
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Beta – 1	0.05 (0.11 ***HAC***)	-0.01 (0.019 ***HAC***)	0.0075 (0.019 ***HAC***)
H0: (beta-1) = 0			
t value	0.44	-0.60	0.40
Pr(>|t|)	0.67	0.55	0.69
Statistical inference	*H0 cannot be rejected*	***The test must be rejected (residual non-random)***	*H0 cannot be rejected*
Observations	41	31	41

HAC heteroscedasticity and autocorrelation consistent standard error

Appendix 6: Statistical tests, forecasts by PTT

Appendix 6: Statistical tests, forecasts by PTT

Table 38: Testing temporal independence of PTT's (*t+1*) forecast errors in 1982–2016, (*t+0*) forecasts in 1981–2016; ((*t+1*) inflation forecasts (1990, 1992–1993), 1995–2016)

Year ahead (t+1)	*GDP growth (t+1)*	*Unemployment rate (t+1)*	*Inflation (t+1)*
H0: autocorrelation coefficients with lags 1–3 = 0			
Q ~ $X^2(3)$	2.11 *p* = 0.55	2.42 *p* = 0.49	5.66 *p* = 0.13
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	35	35	22
Current year (t+0)	*GDP growth (t+0)*	*Unemployment rate (t+0)*	*Inflation (t+0)*
H0: autocorrelation coefficients with lags 1–3 = 0			
Q ~ $X^2(3)$	1.53 *p* = 0.68	4.31 *p* = 0.23	4.38 *p* = 0.22
Statistical inference	H0 cannot be rejected	H0 cannot be rejected	H0 cannot be rejected
Observations	36	36	36

Table 39: Testing unbiasedness of PTT's (*t+1*) and (*t+0*) forecasts with the *t* test

Year ahead (t+1)	*GDP growth (t+1)*	*Unemployment rate (t+1)*	*Inflation (t+1)*
H0: mean error = 0			
Mean (s.e.)	0.12 (0.45)	0.02 (0.25)	0.21 (0.25)
t value	0.27	0.07	0.85
Pr(>|t|)	0.78	0.94	0.40
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	35	35	26

Current year (t+0)	*GDP growth (t+0)*	*Unemployment rate (t+0)*	*Inflation (t+0)*
H0: mean error = 0			
Mean (s.e.)	-0.15 (0.17)	0.07 (0.052)	0.05 (0.031)
t value	-0.90	1.39	1.62
Pr(>|t|)	0.37	0.17	0.11
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	36	36	36

Table 40: Testing unbiasedness of PTT's (*t+1*) forecasts with the forecast error regression model, 1976–2016

$e_{t+1,t} = \alpha + \varepsilon_{t+1,t}$	GDP growth (*t+1*)	*Unemployment rate (t+1)*	Inflation (*t+1*)
H0: alpha = 0			
Alpha	0.12 (0.44)	0.02 (0.25)	0.28 (0.24)
t value	0.28	0.08	1.18
Pr(>|t|)	0.78	0.94	0.25
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	35	35	22
$e_{t+0,t} = \alpha + \varepsilon_{t+0,t}$	GDP growth (*t+0*)	*Unemployment rate (t+0)*	Inflation (*t+0*)
H0: alpha = 0			
Alpha	-0.15 (0.17)	0.07 (0.0516)	0.04 (0.03)
t value	-0.90	1.346	1.2
Pr(>|t|)	0.38	0.187	0.24
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	36	36	36
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

Appendix 6: Statistical tests, forecasts by PTT

Table 41: Test on encompassing the naïve forecast, PTT's (*t+1*) forecast

$$y_{t+1} = \alpha + \beta y_{t-1} + \gamma y_{t+0,t} + \delta y_{t+1,t} + \varepsilon_{t+1}$$

	GDP growth (*t+1*)	Unemployment rate (*t+1*)	Inflation (*t+1*)
1) Beta	0.08	-0.51	
(standard error)	(0.17)	(0.20)	
H0: beta = 0			
t value	0.47	-2.61	
Pr(>|t|)	0.64	0.01**	
Statistical inference	*H0 cannot be rejected*	***H0 must be rejected***	
2) Gamma	-0.11	1.17*	
(standard error)	(0.23)	(0.53)	
H0: gamma = 0			
t value	-0.49	2.22	
Pr(>|t|)	0.63	0.03*	
Statistical inference	*H0 cannot be rejected*	***H0 must be rejected***	
3) Delta	1.23**	0.18	
(standard error)	(0.38)	(0.42)	
H0: delta ≠ 0, delta > 0			
t value	3.22	0.43	
Pr(>|t|)	0.003**	0.67	
Statistical inference	Delta ≠ 0, delta > 0	**Delta = 0,** ***H0 must be rejected***	
Observations	35	35	n = 22, no testing

Table 42: Test on encompassing the naïve forecast, PTT's (*t+0*) forecast

$$y_t = \alpha + \beta y_{t-1} + \delta y_{t+0,t} + \varepsilon_t$$

	GDP growth (*t+0*)	Unemployment rate (*t+0*)	Inflation (*t+0*)
1) Beta	0.02	0.08**	-0.003
(standard error)	(0.06)	(0.03)	(0.024)
H0: beta = 0			
t value	0.40	2.89	-0.11
Pr(>|t|)	0.70	0.007**	0.91
Statistical inference	*H0 cannot be rejected*	***H0 must be rejected***	*H0 cannot be rejected*
2) Delta	1.07*	0.92***	1.02***
(standard error)	(0.07)	(0.03)	(0.03)
H0: delta ≠ 0, delta > 0			
t value	15.84	33.71	37.98
Pr(>|t|)	0.000*	0.000***	0.000***
Statistical inference	Delta ≠ 0, delta > 0	Delta ≠ 0, delta > 0	Delta ≠ 0, delta > 0
Observations	36	36	36
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

Appendix 6: Statistical tests, forecasts by PTT

Table 43: Testing information efficiency of PTT's (*t+1*) forecasts with the regression model, 1976–2016

$$y_{t+1} = \alpha + \beta y_{t+1,t} + \varepsilon_{1t+1}$$

	GDP growth *(t+1)*	Unemployment rate *(t+1)*	Inflation *(t+1)*
Alpha	-0.37 (0.78)	0.33 (0.68)	
H0: alpha = 0			
t value	-0.47	0.49	
Pr(>|t|)	0.64	0.63	
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	
Beta – 1	0.11 (0.29)	-0.04 (0.07)	
H0: (beta-1) = 0			
t value	0.39	-0.55	
Pr(>|t|)	0.70	0.59	
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	
Observations	35	35	n = 22, no testing

Table 44: Testing information efficiency of PTT's (*t+0*) forecasts with the regression model, 1976–2016

$$y_t = \alpha + \beta y_{t+0,t} + \varepsilon_{1t}$$

	GDP growth *(t+0)*	Unemployment rate *(t+0)*	Inflation *(t+0)*
Alpha	-0.005 (0.202)	0.01 (0.13)	-0.097* (0.048)
H0: alpha = 0			
t value	-0.023	0.07	-2.045
Pr(>|t|)	0.98	0.94	0.0486*
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	***H0 must be rejected***
Beta – 1	0.08 (0.06)	-0.01 (0.134)	0.0187 (0.0114)
H0: (beta-1) = 0			
t value	1.33	-0.065	1.64
Pr(>|t|)	0.19	0.949	0.11
Statistical inference	*H0 cannot be rejected*	*H0 cannot be rejected*	*H0 cannot be rejected*
Observations	36	36	36
	**p < 0.05,*	***p < 0.01,*	****p < 0.001*

References

References

[1] Laki valtiontalouden tarkastusvirastosta, 14 July 2000/676, available at http://www.finlex.fi/fi/
[2] Laki talous- ja rahaliiton vakaudesta, yhteensovittamisesta sekä ohjauksesta ja hallinnasta tehdyn sopimuksen lainsäädännön alaan kuuluvien määräysten voimaansaattamisesta ja sopimuksen soveltamisesta sekä julkisen talouden monivuotisia kehyksiä koskevista vaatimuksista 869/2012, available at http://www.finlex.fi/fi/
[3] *Council Directive 2011/85/EU* of 8 November 2011 on requirements for budgetary frameworks of the Member States, available at http://eur-lex.europa.eu/homepage.html
[4] *Valtioneuvoston asetus julkisen talouden suunnitelmasta* (120/2014) *ja sen muuttamisesta* (601/2017), available at http://www.finlex.fi/fi/
[5] This section is based on the chapter "Aikaisempi tutkimus" (Previous research) of an unpublished audit memorandum of the NAOF, *Tilastoanalyysi makroennusteiden osuvuudesta* (Statistical analysis of the accuracy of macroeconomic forecasts), prepared by Sami Grönberg and Ville Haltia in 2016 (see National Audit Office 2016a).
[6] Konttinen, Jouko (1986). Suomen kansantalouden suhdanne-ennusteista ja niiden osuvuudesta. *Kansantaloudellinen aikakauskirja*, 1986:4
[7] Mutikainen, Tapio and Suvanto, Arvi (1986). Valtiovarainministeriön ennusteiden osuvuus, *Kansantaloudellinen aikakauskirja*, 1986:4
[8] Viren, Matti (1994). Ennustaminen on vaikeaa, *KOP*, *Taloudellinen katsaus*, 2/1994
[9] Valtiontalouden tarkastusvirasto (1999). Valtion talousarvion tulot ja niiden arviointi, *Tarkastuskertomus*, 6/99
[10] Pehkonen, Jaakko (2002). Talousennusteiden osuvuus 1997–2001: valistuneita arvauksia, *Kansantaloudellinen aikakauskirja*, 2002:2
[11] Lahtinen, Markus; Mäki-Fränti, Petri, Määttä, Kalle and Volk, Raija (2009). Valtion talousarvioiden verotuloennusteiden osuvuus, *Eduskunnan tarkastusvaliokunnan julkaisu*, 1/2009
[12] Lanne, Markku (2009). Ennustajien tappiofunktiot ja BKT-ennusteiden rationaalisuus, *Kansantaloudellinen aikakauskirja*, 2009:4
[13] IMF (2015). *Finland, Fiscal Transparency Evaluation*. March 2015.
[14] National Audit Office (2016a). Tilastoanalyysi makroennusteiden osuvuudesta. *Unpublished audit memorandum,* authors Grönberg, Sami and Haltia Ville.
[15] National Audit Office (2016b). Reliability of macroeconomic forecasts. Fiscal policy audit report. *Fiscal policy audit report of the National Audit Office*, 11/2016
[16] Newby, Elisa and Orjasniemi, Seppo (2011). Suomen Pankin ennustevirheet vuosina 2004–2010. *Euro & Talous* 3/2011.
[17] European Commission (1999). The track record of the Commission forecasts. European Economy, *Economic Papers 137*, October 1999
[18] European Commission (2007). The track record of the Commission forecasts – an update. European Economy, *Economic Papers 291*, October 2007
[19] European Commission (2012). The accuracy of the European Commission's forecasts re-examined. European Economy, *Economic Papers 476*, December 2012
[20] European Commission (2016). European Commission's Forecasts Accuracy Revisited. Statistical Properties and Possible Causes of Forecast Errors. *European Economy, Discussion Paper 027*, March 2016
[21] PT forecasts concerning years 1976–1993 refer here to the forecasts of the old Labour Institute for Economic Research. The Finnish name until 1993 was Työväen taloudellinen tutkimuslaitos (TTT) after which the name was changed to Palkansaajien tutkimuslaitos (PT). According to the representatives of PT, before 1993 the institute had more modest targets in its forecast activities and did not necessarily even aim to compete with forecasts of the Finance Mininstry. Therefore, according to the PT representatives, the early forecasts before 1993 may reduce the average accuracy of forecasts by PT in this assessment report.
[22] The presentation of the basic concepts is based on the unpublished audit memorandum of the NAOF, *Tilastoanalyysi makroennusteiden osuvuudesta* prepared by Sami Grönberg and Ville Haltia in 2016 (see National Audit Office 2016a).
[23] See end note 20 above.
[24] Chatfield, C. (2003). *The Analysis of Time Series: An Introduction*. Sixth Edition, Chapman and Hall/CRC, ISBN: 9781584883173

Refrences

[25] Hamilton, James D. (1994). *Time Series Analysis*. Princeton University Press, ISBN: 9780691042893
[26] Ljung, G.M. and Box, G.E.P (1978). On a Measure of a Lack of Fit in Time Series Models, *Biometrika*, 65(2), pp. 297–303
[27] Kokkinen, Arto and Wouters, Hans (2016). EA and EU GDP flash estimates at 30 days, *Eurona,* 1/2016, pp. 71–107, available at http://ec.europa.eu/eurostat/web/products-statistical-books/-/KS-GP-16-001
[28] Newey, W. K., West, K.D. (1987). A Simple, Positive Semi-definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica*, vol. 55, 1987:3, pp. 703–708
[29] Fair, R. C. and Schiller, R.J. (1990). Comparing Information in Forecasts from Econometric Models, *American Economic Review*, 80 (3), pp. 375–389
[30] In these figures, the confidence intervals are centered around the forecast mean error for technical presenting reasons. Another way, which would probably be more consistent in terms of the *t* test, would have been presenting the confidence interval centered around the H0 hypothesis value, which in this case is zero.

NATIONAL AUDIT OFFICE OF FINLAND
PORKKALANKATU 1, P.O.BOX 1119, FI-00101 HELSINKI
TEL. +358 9 4321, WWW.VTV.FI/EN