upB. ufficio
parlamentare
di bilancio

A retrospective assessment of the macroeconomic forecasts of the MEF and the PBO

1. Introduction and principal findings

The Parliamentary Budget Office (PBO) was established in April 2014, in accordance with the provisions of the Balanced Budget Act (Law 1/2012)[1] and in implementation of the regulations on the new European economic governance arrangements. The "Two-Pack" (Regulation 473/2013) specifies that budgets shall be based on macroeconomic forecasts produced or endorsed by an independent body.[2] The aim is to ensure that official macroeconomic forecasts are not overly optimistic. Excessively high projections of economic activity create fictitious fiscal space, with a consequent imbalancing of the accounts, thus compromising the effectiveness and credibility of budget planning.

In Italy, the PBO is charged, *inter alia*, with endorsing the macroeconomic forecasts prepared by the Government in its policy documents. To perform this function, the PBO constructs its own macroeconomic scenario and coordinates the production of similar scenarios by a panel of independent forecasters (CER, Oxford Economics, Ref.ricerche and Prometeia), which use the same exogenous variables and information on the public finance projections of the Ministry for the Economy and Finance (MEF). The endorsement procedure, which is governed by a memorandum of understanding,[3] involves interaction with the MEF on preliminary versions of the forecasts. In general, the PBO considers acceptable a Government forecast that falls within the range of panel forecasts. This process is not conducted

---

[1] The PBO is governed by Law 243/2012 (Articles 16-19).
[2] See the <u>Hearing</u> of the Chairman of the PBO, Giuseppe Pisauro, on the commencement of the operations of the Parliamentary Budget Office – Joint session of the Budget Committees of the Chamber of

Deputies and the Senate, 24 September 2014 (in Italian).
[3] See https://en.upbilancio.it/wp-content/uploads/2022/10/Protocollo_d_intesa_MEF_UPB_signed-signed.pdf (in Italian).

The Focus Paper was prepared by Cecilia Frale, Libero Monteforte and Carmine Pappalardo

mechanically, but is rather based on general criteria of prudential evaluation criteria. Over the history of the endorsement exercises performed by the PBO, the Government's macroeconomic scenario has been modified repeatedly to reduce the differences with the PBO panel forecasts. In most cases, this has been followed by endorsement of the scenario. The MEF's forecasts have not been endorsed only in two occasions, and in both cases the Government subsequently modified its projections, albeit with different timetables and procedures.[4]

After seven years of experience, a retrospective assessment of the Government's forecasts in the period in which the PBO has performed the endorsement exercises (since mid-2014) could be useful, providing a comparison with both the previous period and with the projections of other forecasters. This Focus therefore presents diagnostics of the quality of forecasts, distinguishing between the period prior to the establishment of the PBO (from 2000 to the spring of 2014) and the subsequent years. The analysis is accompanied by a number of assessments of the sensitivity of the results with respect to various factors, such as the specific error metric selected or the differences in the updating of official national accounts data. A comparison is also conducted with the independent

forecasts of private-sector forecasters and with the experience of other European countries.

The analysis performed shows that the optimism and errors concerning the main macroeconomic aggregates in the Italian policy documents has decreased in the period following the creation of the PBO. The MEF's forecasts from 2014 onwards have on average been fairly close to those of other institutions, albeit slightly more conservative.

The similarity of the macroeconomic scenarios of the Government and the PBO may reflect specific procedural issues. On the one hand, the endorsement process is structured so that the Government's preliminary forecasts can incorporate the observations of the PBO, bringing the MEF scenario closer to that of the PBO panel.[5] On the other hand, the PBO forecasts tend to converge towards those of the MEF, as the main international exogenous variables, and the budget measures, are defined by the MEF and are included by the PBO panel of forecasters. This affects, for example, the official forecasts of price variables in the years in which the safeguard clauses (automatic VAT and excise duty increases) were employed (which incorporated future increases in VAT rates in the current-legislation budget projections).[6] Finally, since the

---

[4] In the autumn of 2016, after endorsing the trend macroeconomic forecasts, the PBO rejected the policy forecasts in the Update of the 2016 Economic and Financial Document. The forecasts were then modified by the MEF shortly thereafter in the Draft Budgetary Plan, thus obtaining endorsement by the PBO. Two years later, the policy macroeconomic scenario in the 2018 Update was again not endorsed and the changes by the Government were made in the final phase of the budget session, in December.

[5] The memorandum of understanding between MEF and PBO provides for the fiscal institution to comment on preliminary versions of Government forecasts and allows the MEF to modify them accordingly. The PBO's requests to the MEF always ask for greater prudence, so this interaction tends to moderate the Government's forecasts.

[6] From the middle of the last decade, most of the forecasters were expecting that the clauses would be deactivated for the following year with the Budget

macroeconomic forecasting calendar used by the MEF and the PBO is the same, the two scenarios incorporate the same economic information, thus diverging from forecasts that draw on more or less timely data. In short, the timing and construction of the macroeconomic scenarios are such as to make the forecasts of the MEF and PBO more uniform than those formulated independently by other bodies.

The following section discusses the indicators used to assess the reliability of the MEF forecasts compared with outturn data, including by comparison those of other institutions. Section 3 presents the findings of sensitivity analyses with regard to the data used as targets and the projections of private forecasters, as well as a comparison with the government forecasts of other large European countries. Section 4 contains brief concluding remarks.

2. Main diagnostics of the forecasts

In order to assess the Government's macroeconomic scenarios, the policy forecasts underlying the budget documents were reconstructed starting from 2000[7] and the sample was divided into two parts: for the period prior to the establishment of the PBO, the scenarios from policy documents between 2000 and the spring of 2014 were considered, while for the subsequent period, with the PBO operational, the forecasts included those between the Update to the 2014 Economic and Financial Document (Update) - the first such scenario endorsed by the PBO - and the 2019 Update, the last prior to the outbreak of the COVID-19 pandemic.[8] For each forecast scenario, the changes in real and nominal GDP over the entire horizon were considered, i.e. for the current year and the next three. For each forecasting step, the errors with respect to National statistical office (Istat) data were calculated and then the measures of the goodness-of-fit of the forecasts, as indicated in Box 1.

With a view to assessing whether the forecasts produced for budget planning are optimistic, an attempt is made to quantify bias, i.e. the tendency to produce errors mainly of the same sign. Accuracy is also analysed, which depends on the size of the errors regardless of their sign. The main indicator selected to measure bias is the arithmetic mean error (ME), which by offsetting negative and positive values measures the optimism (or pessimism) of the forecasts. On the other hand, the amplitude of the error is mainly assessed using the root-mean-square error of the forecast (RMSE, i.e. the standard deviation from historical data), which is normally the most widely adopted diagnostic, both in the academic literature and in practice.

Act, while the MEF and the PBO panel considered the situation on a current-legislation basis, for which they overestimated the deflators for the subsequent years. According to reconstructions published by the PBO of the macroeconomic scenario excluding the safeguard clauses, in the 2017-2019 period, the forecast error of the PBO for the estimate of the GDP deflator in year T+1 virtually disappeared.

[7] The forecasts for nominal GDP were reconstructed as from 2002.
[8] Consistent with the exclusion of 2020 from the sample, in the period preceding 2014 the first recession years (i.e. 2008 and 2012) were also excluded.

Figure 1 shows the mean error (ME) of the policy macroeconomic scenarios of the MEF over two time frames, with a comparison against those of the PBO and the European Commission in the most recent period. The figure shows that the Government's forecasts for both real and nominal GDP were remarkably optimistic up to 2014 (positive values indicate forecasts exceeding the actual figure and vice versa). The mean error was around 1.5 starting from year T+1, i.e. precisely over the time horizon in which the macroeconomic forecasts have the greatest impact on the public finance forecasts (in year T, the forecasts for the general government account benefit from monitoring during the year) and are more relevant for budgetary planning purposes. In the period in which the PBO has operated, the MEF's forecasts became more balanced, especially for real GDP, being marginally pessimistic for the current year. The optimism of the forecast for nominal GDP remains non-negligible, and increasing as the forecast horizon lengthens. This is partly attributable to the assumptions concerning the safeguard clauses. The PBO forecasts were slightly more conservative than those of the MEF and were similar to those of the European Commission (produced in the weeks immediately following the publication of the policy documents),[9] which, however, are characterised by excessive pessimism for the current year.

As a further indicator of forecast bias, we can evaluate the frequency with which forecasts exceed actual figures a posteriori. Figure 2 shows the number of optimistic forecasts produced by the MEF over several periods by comparison with those of the PBO and the European Commission. Consistent with the mean error, in the period following the 2014 Update the indicator improved on average compared with the previous time period, especially for real GDP, while a large share of optimistic forecasts for nominal GDP remained.

---

[9] The comparison with the EFD used the European Commission's spring forecast, while the autumn forecast was used for the Update. The European Commission's forecasts are always published after those of the MEF (and therefore also of the PBO), so the EC forecasts have a slight information advantage.
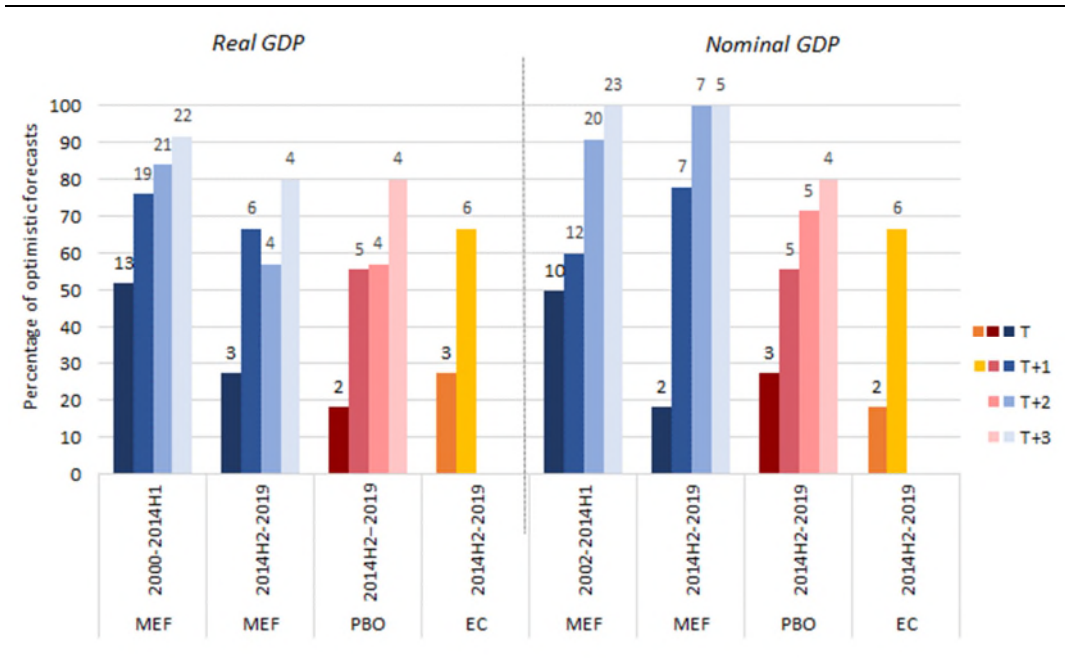
Figure 1 – Mean error (ME) of GDP forecasts (1)
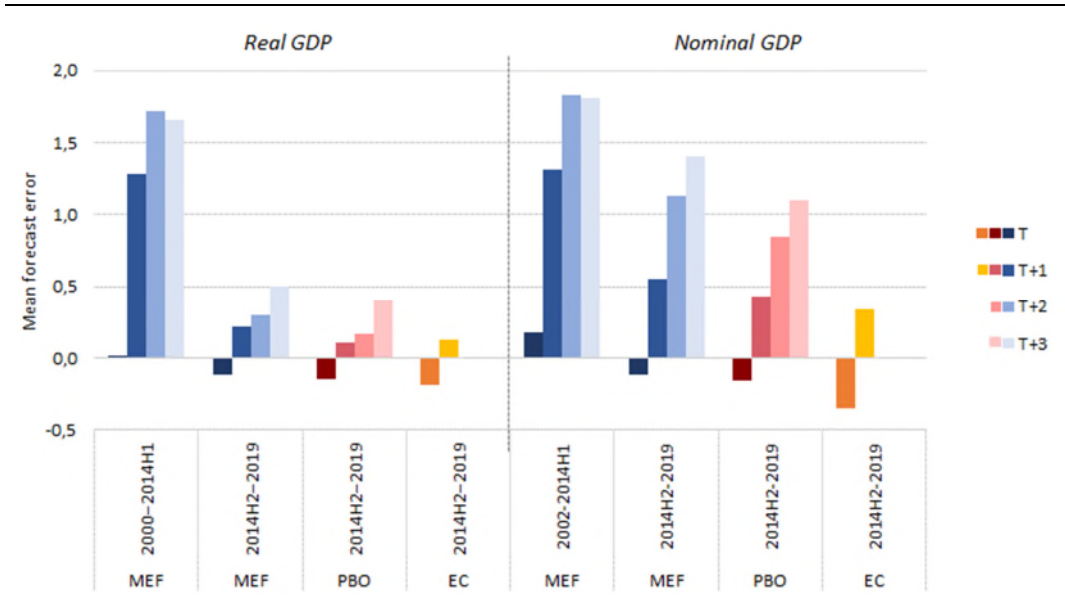
Source: based on MEF, PBO and European Commission forecasts.
(1) 2014H1 refers to the first half of 2014 and therefore includes the 2014 EFD, while the 2014 Update (the first policy document endorsed by the PBO) falls within the subsequent interval (2014H2-2019). The period of analysis preceding 2014 excludes 2008 and 2012, as they were the first years of recession, consistent with the exclusion of the 2020 pandemic year.

Figure 2 – Number of optimistic forecasts (1)

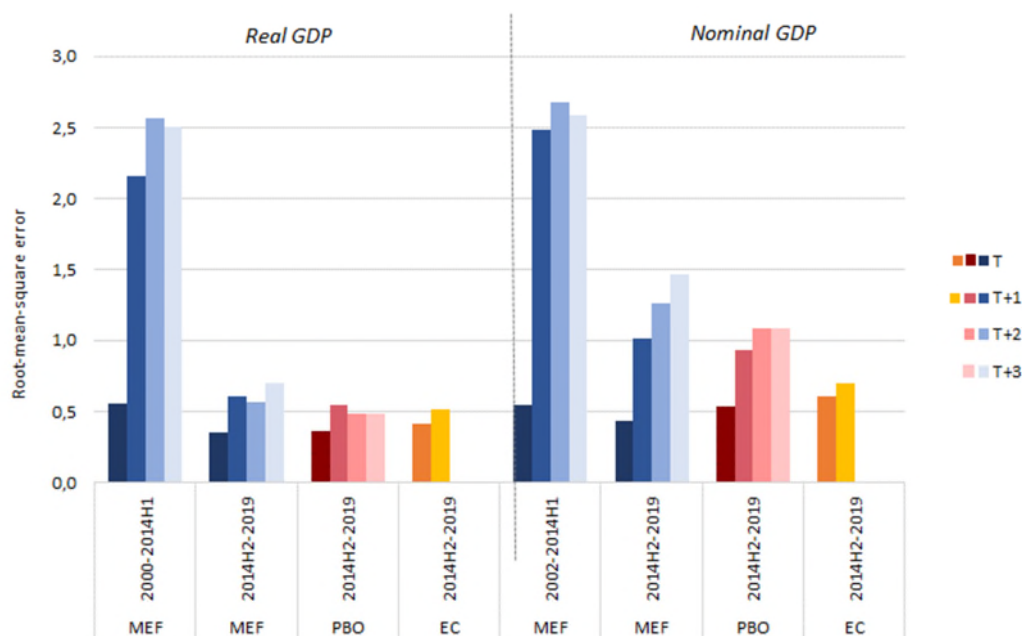Source: based on MEF, PBO and European Commission forecasts.
(1) The histograms represent the share of optimistic forecasts out of total forecasts for the period. The value indicated at the top of the histogram represents the absolute number. 2014H1 refers to the first half of 2014 and therefore includes the 2014 EFD, while the 2014 Update (the first policy document endorsed by the PBO) falls within the subsequent interval (2014H2-2019). The period of analysis preceding 2014 excludes 2008 and 2012, as they were the first years of recession, consistent with the exclusion of the 2020 pandemic year.

For real GDP, the Government's stance is not very different from that of the PBO and the European Commission, but it remains excessively optimistic for nominal GDP over longer horizons. There are no major differences in the comparison between the forecasters, although the PBO tended to formulate non-optimistic forecasts. The absolute number of optimistic forecasts in the post-2014 period is very small, however, so any conclusions drawn from the percentage values must be considered with caution.

With regard to the amplitude of the errors, Figure 3 reports the RMSE, i.e. the square root of the mean squared error. The accuracy of the Government's forecasts has improved significantly in the period following the 2014 Update, especially for the years following the current year. In the previous period, the average RMSE for the MEF forecasts for GDP (real and nominal) was in the order of 2.5 on horizons beyond T, i.e. those that are most relevant for economic policy. For the subsequent period, the diagnostic fell to an average of around 0.5 for real GDP and almost halved for nominal GDP at all horizons. The forecast errors of the MEF are comparable to those of the PBO and the European Commission for real GDP, while they tend to deviate upwards for nominal GDP, especially at more distant horizons. It should be remembered that the policy scenarios of the MEF and the PBO considered the safeguard clauses, which were not incorporated in the forecasts of the European Commission (which assumed their – very probable – deactivation with the subsequent Budget Act).

Figure 3 – Root-mean-square error (RMSE) of GDP forecasts (1)

Source: based on MEF, PBO and European Commission forecasts.
(1) 2014H1 refers to the first half of 2014 and therefore includes the 2014 EFD, while the 2014 Update (the first policy document endorsed by the PBO) falls within the subsequent interval (2014H2-2019). The period of analysis preceding 2014 excludes 2008 and 2012, as they were the first years of recession, consistent with the exclusion of the 2020 pandemic year.

Figure 4 examines the evolution of Government forecasts after the creation of the PBO in greater detail, showing differences between the error metrics on average in the periods preceding and following the 2014 Update. For real GDP, all the diagnostics point to an improvement in forecasts that tends to increase with the forecast horizon, with the exception of the last year, when the gain declines. The same pattern can be found for nominal GDP, although there is less uniformity across the different indicators.

In general, it can be said that after 2014 the accuracy of government forecasts has improved, especially for real GDP and at more distant horizons, regardless of the indicator used to measure the goodness of the forecasts. To get a better view than that provided by the summary metrics and evaluate the form of distribution, the boxplots (described in Box 2) of forecast errors for year T+1, a key period for the drafting of Budget Acts, were also considered. Figures 5 and 6 show that not only have the MEF errors been reduced on average, as already indicated by the diagnostics presented above, but the distributions have also been concentrated, diminishing the uncertainty of the projections. The Government's forecasts remain slightly optimistic (with positive means and medians), but the errors have become more concentrated around the median than in the past, although to a lesser extent than those of the PBO and the European Commission.


Box 2 – Reading boxplots

The "boxplot" provides a statistical representation of the distribution of data. The key elements of a boxplot are the following:

- The central (black) line in the box represents the median. If the data are symmetrical, the median lies in the centre of the box. If the data is asymmetrical, the median will be closer to the top or bottom of the box. Differences between median and mean describe the asymmetry of the distribution.

- The black dot within the box represents the mean.

- The bottom and top of the box correspond to the 25th and 75th quantiles, or percentiles. The height of the box is the difference between these two percentiles and is called the interquartile range (IQR).

- The vertical lines (whiskers) extended from the top and bottom of the box indicate the maximum and minimum values of the data set (within a limit of 1.5 times the IQR in both directions).

- The data outside the span indicated by 1.5 times the IQR above and below the box are given by white dots and are considered outliers.

Figure 4 – Differences between forecasting diagnostics for the MEF before and after the 2014 Update (1)

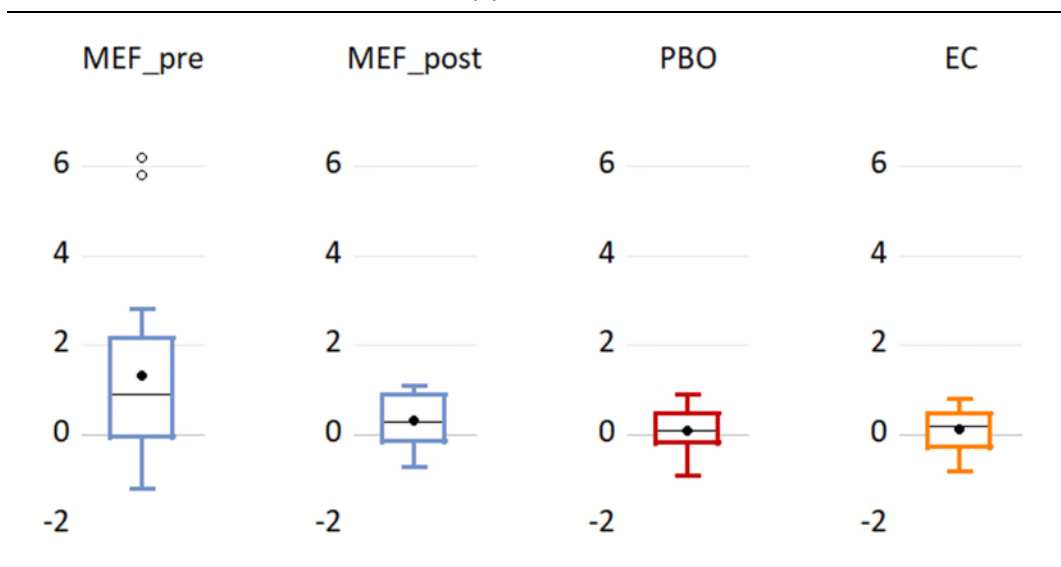*(a) Real GDP*
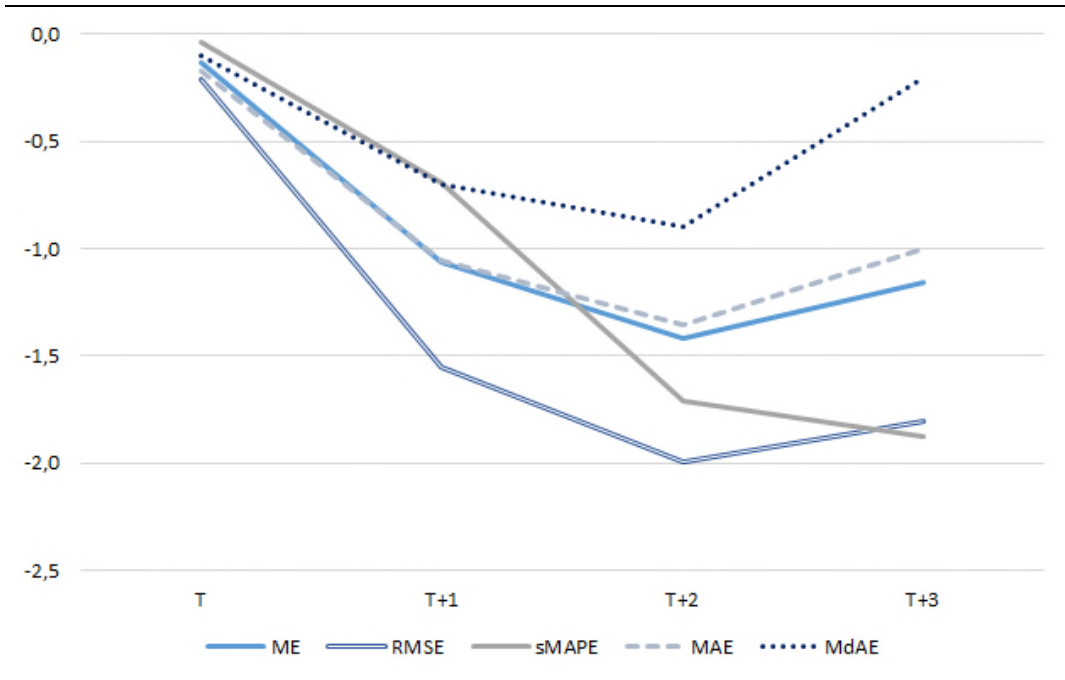
(b) Nominal GDP
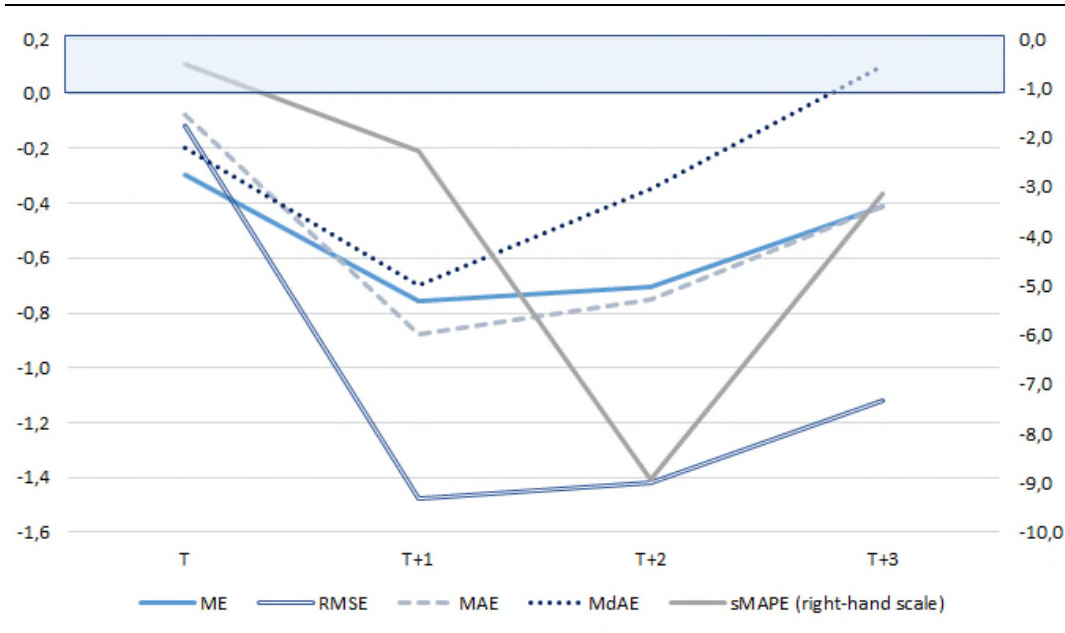
Source: based on MEF forecasts.
(1) The "post" period runs from the 2014 Update to the 2019 Update. The "pre" period comprises forecasting documents from 2000 (2002 for nominal GDP) to the 2014 EFD; excludes 2008 and 2012, as they were the first years of recession, consistent with the exclusion of the 2020 pandemic year.

Figure 5 – Boxplots of forecast errors for T+1 (1)

(a) Real GDP

(b) Nominal GDP

Source: based on MEF, PBO and European Commission forecasts.
(1) The "post" period runs from the 2014 Update to the 2019 Update. The "pre" period comprises forecasting documents from 2000 (2002 for nominal GDP) to the 2014 EFD; excludes 2008 and 2012, as they were the first years of recession, consistent with the exclusion of the 2020 pandemic year.
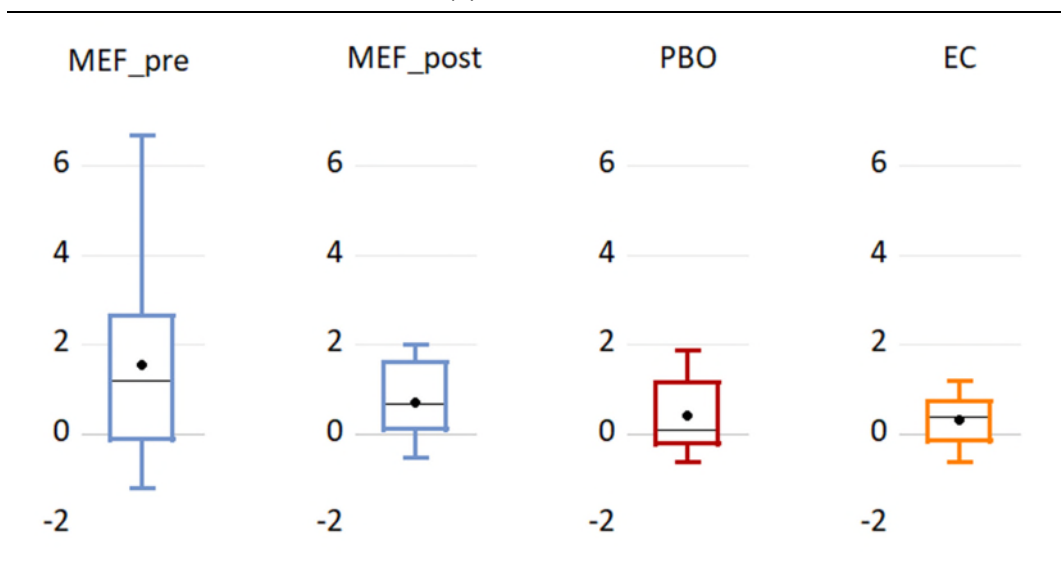
3. Sensitivity analysis and international comparisons

This section presents a number of sensitivity analyses conducted to assess the robustness of the results obtained with respect to a series of influencing factors, such as the data adopted and the forecasters involved.

An initial evaluation concerns the stability of the results with respect to revisions of the national accounts data, which in some cases were significant. The diagnostics presented

previously were calculated using ex post Istat data as a benchmark, while in ordinary forecasting only preliminary data are available, which could subsequently be revised by the statistical institute. To take account of this effect on the validity of the Government's forecasts, the indicators used previously were also calculated on the basis of preliminary data, i.e. the first publication of the annual national accounts. Overall (Figure 6), the bias indicator (ME) would be higher if calculated with respect to the initial Istat data release, rather than with respect to the definitive figures (this implies that in the period considered the data were mainly revised upwards). The effect of the revisions is particularly evident in the ME of year T and tends to attenuate slightly in the subsequent forecasting steps. With regard to the accuracy of the projections, the gap between the indicators constructed with respect to the two different benchmarks is very small for real GDP, while revisions of nominal GDP data improve accuracy more significantly.

In a further sensitivity analysis, the MEF's real GDP forecasts are compared against the virtually contemporary projections formulated by the private forecasters surveyed in *Consensus Forecasts*.[10] In this case, a measure of error is constructed as if the consensus prediction of the external analysts were the historical data in order to directly assess the similarity of the scenarios of the MEF and the analysts. Figure 7 shows that the estimates of GDP for the current year formulated by the MEF do not differ very much on average from the consensus

position, as demonstrated by the low values for ME and RMSE in the period both before and after the establishment of the PBO. The diagnostics for the projections in T+1 diverge more substantially, although the disparity has diminished significantly on average in the period following the formation of the PBO. However, it should be borne in mind that the *Consensus Forecasts* themselves could incorporate information drawn from MEF projections, as they are not exactly contemporaneous, and therefore not completely independent of the Government scenarios.

Finally, we compare the errors for the macroeconomic forecasts of the governments of Germany, France and Spain in the period after the autumn of 2014. With regard to the current year (Figure 8), the bias of the MEF's real GDP forecasts is substantially aligned with that of the other countries, although the macroeconomic scenarios of France and Spain appear more conservative. A similar pattern, with the exception of Spain, is seen in the current-year forecasts for nominal GDP. By contrast, the bias of the MEF scenarios is significantly greater than the other countries for the projections for T+1. The difference is pronounced for nominal GDP, probably due in part to the considerable impact of the safeguard clauses in Italy. The standard deviation of the errors (Figure 9) is similar to or lower than that of the other countries (including Germany) for real GDP, but greater for the nominal GDP projections for the following year.

---

[10] Since *Consensus Forecasts* is published around the middle of the month, we consider the average of the forecast for the month in which the document is published and the following month. The forecasts for the same month reflect the same economic information available to the MEF, while those for the following month may incorporate the budget policy objectives specified in the policy document.

Figure 6 – Difference between forecast diagnostics calculated with ex post data and those calculated with preliminary data (1)

(a) Real GDP
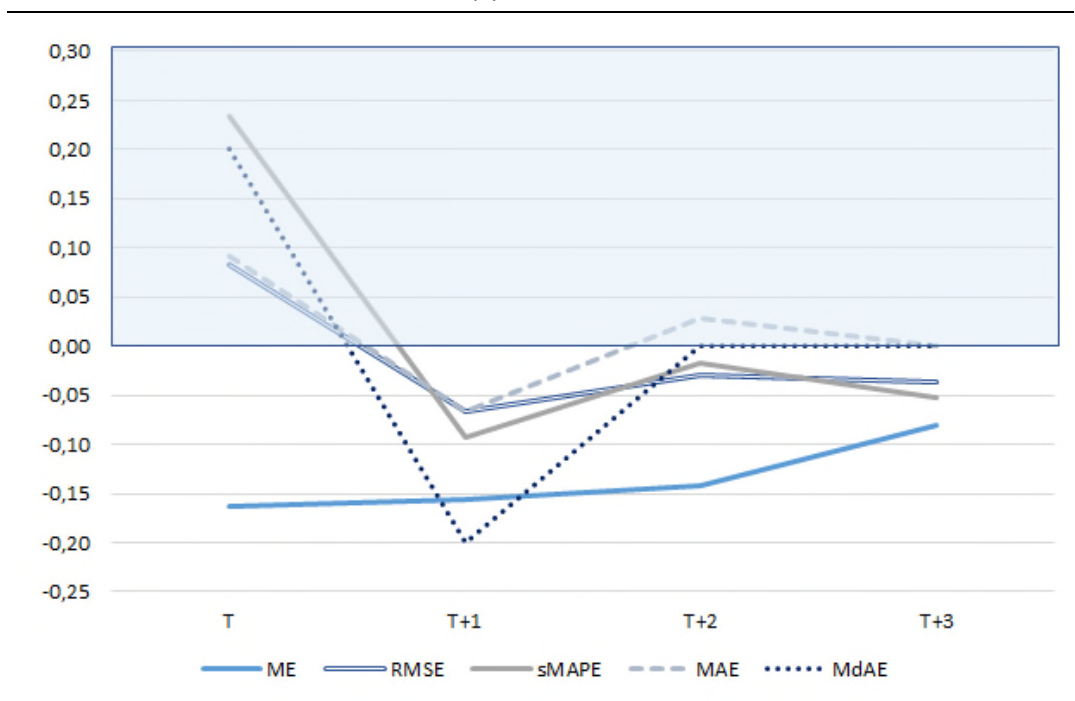
(b) Nominal GDP
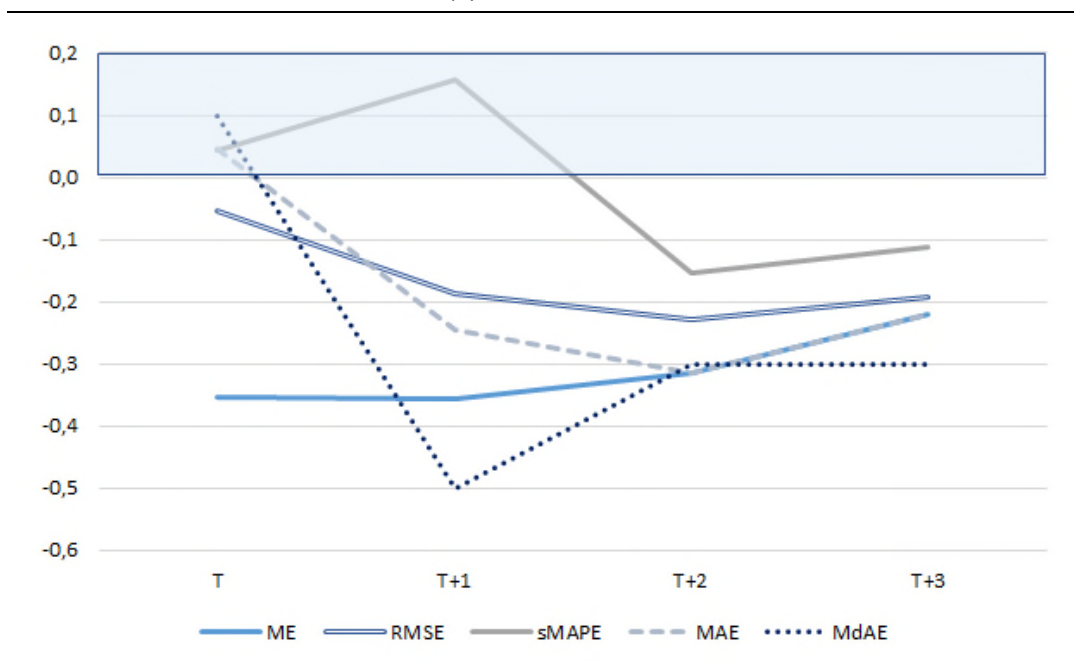
Source: based on MEF forecasts.
(1) The figure indicates the differences for the "post" period (from the 2014 Update to the 2019 Update) between the diagnostics calculated for the forecast error with respect to data available in January 2022 and those with respect to preliminary data (initial publication of annual national accounts data).

Figure 7 – Difference between MEF forecasts for real GDP and those of Consensus Forecasts (1)

(a) Forecasts at time T

(b) Forecasts at time T+1

Legend: ■ 2000-2014H1 ■ 2014H2-2019
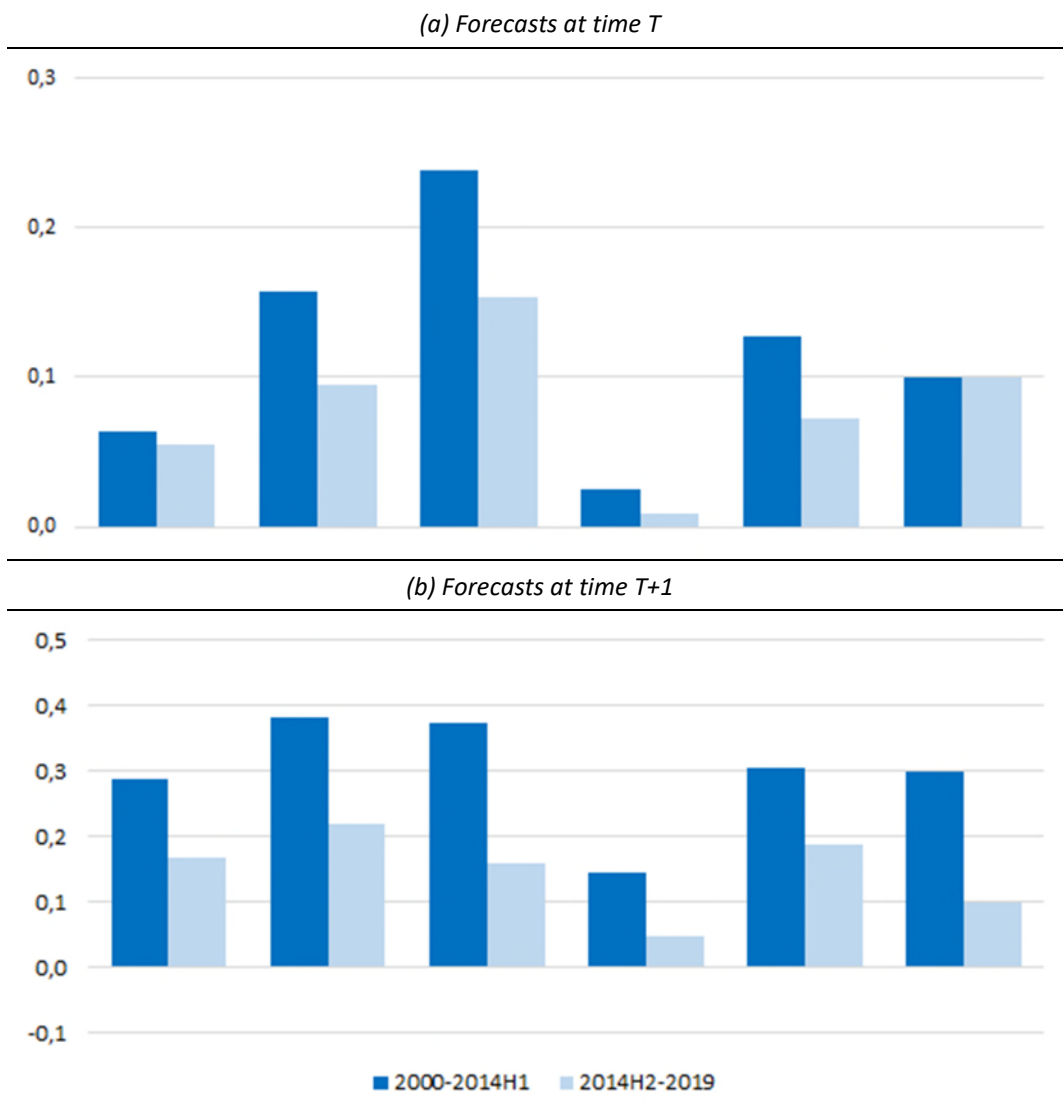
Source: based on MEF and *Consensus Economics* forecasts.

(1) 2014H1 refers to the first half of 2014 and therefore includes the 2014 EFD, while the 2014 Update (the first policy document endorsed by the PBO) falls within the subsequent interval (2014H2-2019). The period of analysis preceding 2014 excludes 2008 and 2012, as they were the first years of recession, consistent with the exclusion of the 2020 pandemic year. In order to ensure consistency with the information (on both the economy and the public finances) available to the Government, we consider the average of the *Consensus Forecasts* projection for the month in which the document is published and that for the following month.

Figure 8 – Official forecasts for selected EU countries – mean error (ME) (1)
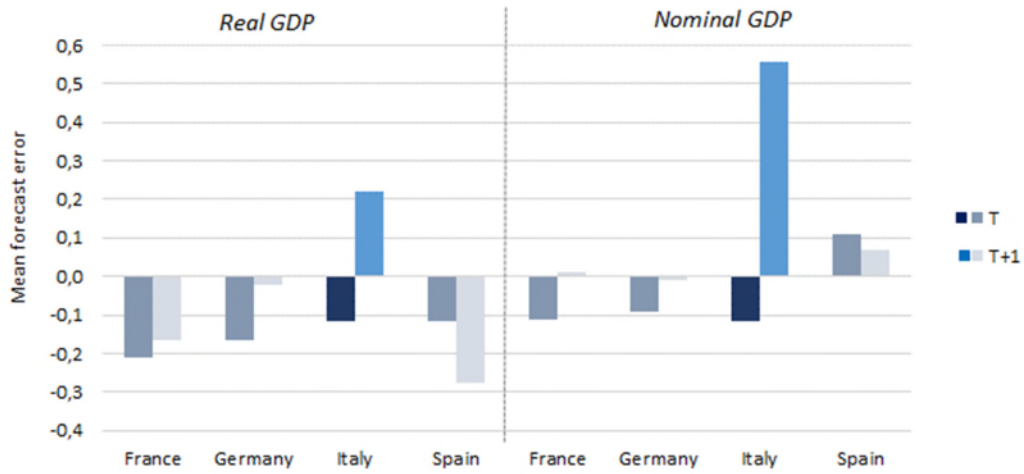
Source: based on forecasts drawn from official budget documents of the individual countries.
(1) The figure provides a comparison between the forecast error diagnostics (between autumn 2014 and 2019) for Italy and the other main EU economies for year T and T+1.

Figure 9 – Official forecasts for selected EU countries – Root-mean-square error (RMSE) (1)
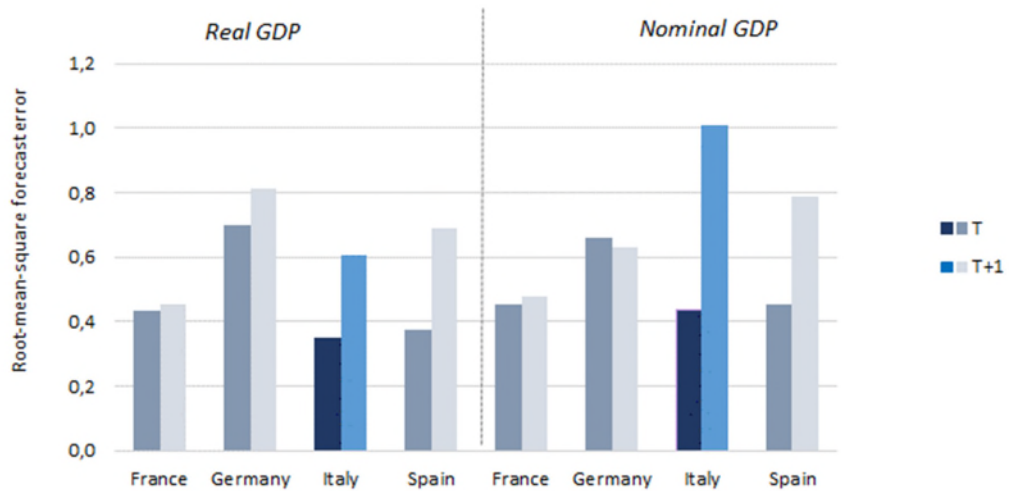
Source: based on forecasts drawn from official budget documents of the individual countries.
(1) The figure provides a comparison between the forecast error diagnostics (between autumn 2014 and 2019) for Italy and the other main EU economies for year T and T+1.

However, bear in mind that these comparisons involved heterogeneous macroeconomic projections, as they regard countries with different economic cycles. Finally, the diagnostics are calculated on the basis of a limited number of observations.

4. Conclusions

The historical reconstruction presented here shows the Italian Government's policy macroeconomic forecasts have improved on average in the period in which the PBO has operated. The upward bias of

forecasts, which signals optimism, has diminished, especially for real GDP, while the reduction has been less marked for nominal GDP, probably due in part to the impact of the safeguard clauses. In the latter half of the last decade the accuracy of the estimates has also improved in terms of the size of the error, although there remains room for further decline in the error for nominal GDP, especially at larger horizons.

By comparison with other institutional forecasters, such as the PBO and the European Commission, the Government's macroeconomic forecasts in 2014-19 were slightly more biased and less accurate. In the same years, the difference between official forecasts and those of private analysts narrowed significantly, signalling the Government's intention to be more consistent with prevailing external expectations. In this Focus we also reconstructed a number of predictive diagnostics for the governmental macroeconomic scenarios of other European countries, although the comparison between different economies is not entirely homogeneous.

In interpreting the descriptive evidence presented, it must be borne in mind that the number of available observations is small, partly reflecting the need to eliminate years affected by anomalous shocks, such as the pandemic, from the sample. This means that more refined inferential statistical assessments are not currently possible. Sounder statistical analysis can only be performed in the future, when longer time series become available, as in the case of independent fiscal institutions that have been operating for many years, such as the Congressional Budget Office in the United States.