LUISS

Dipartimento
di Economia e Finanza

Cattedra di Macroeconomics

Are the Italian government forecasts optimistic? A comparative analysis

Prof. Francesco Lippi
RELATORE

Matr. Stefano Carta
CANDIDATO

Anno Accademico 2019/2020

"There are two kinds of forecasters: those who don't know, and those who don't know they don't know."

John Kenneth Galbraith

Acknowledgments

This work marks the end of an important chapter of my academic journey; therefore, I believe it is appropriate to thank those who contributed to this achievement.

First, thanks to my thesis supervisor, Prof. Francesco Lippi, for his precious advice, passionate participation, and overall support to the writing of this thesis.

Thanks to my parents, for everything you have gone through so that I could follow my dreams and passions. You taught me that hard work and dedication can get you everywhere. To you, I dedicate this thesis.

Thanks to my brother for the inspiration he is for me to be a better person and because of the unconditional support and joy he has always given me.

Thanks to my uncles and aunts, they have taught me the valuable lesson to deeply appreciate the small things in life and, above all, to care for each other and be present.

Finally, thanks to all my friends and loved ones who have helped me reach my goals both personally and professionally.

Index

Introduction

Economic and financial forecasting prepared by the government is the basis for policy making and economic surveillance, therefore accurate methods and high quality are essential. This thesis makes an ex-post evaluation of the Italian forecasts which are used for assessing the government's progress compared to the objectives the government itself set in previous years.

The main aim is therefore to collect and evaluate the accuracy of the Italian government forecasts, using statistical tools. Since the future can never be known, forecasts are surrounded by significant uncertainty and will inevitably prove to be wrong in many aspects. Therefore, the thesis will not focus on the errors themselves, but on eventual systematic errors made by the Italian government.

Moreover, the thesis will also compare the accuracy of Italian forecasts with those of the French, German, and Spanish governments to understand if any systematic behaviour is typical of governments in general or if it is country-specific. It will consider forecasts made by the European Commission (EC) and Consensus Economics (CE) for the Italian economy to also understand whether any behaviour belongs only to the Italian government or also to international institutions.

The thesis will analyze forecasts prepared for the years 2001 to 2019, therefore published between 2000 and 2018. It will also study, in isolation, the period which followed the 2008 recession and 2009 European sovereign debt crisis, which understandably confounded many forecasters (Gonzáles L., 2012). In practical terms, two distinct forecasting windows are examined, the full sample (2001 - 2019), and the post-crisis sample (2010 - 2019)[(1)].

For the Italian government, a wide range of variables, both tendency[(2)] and programmatic[(3)], have been considered. For the French, German, and Spanish

[(1)] Predictions for the shorter period are made from 2009 to 2018.
[(2)] Tendency scenario is the hypothetical scenario in which no policies are implemented by the government in charge, its forecasts are therefore more conservative and serve the purpose of creating a benchmark to compare the effectiveness of the new policies with.
[(3)] Programmatic scenario is the actual scenario expected to realize once the government implements the policies proposed. Its forecasts therefore reflect the results of the policy of the current government.

governments and Consensus Economics, comparison is made with real GDP growth while for the European Commission, the analysis is made with real GDP growth and unemployment.

The thesis is structured as follows:

- the first chapter examines the data collected and explains the methods used;
- the second chapter compares the results derived for the Italian government with the performance of the French, German, and Spanish forecasts while also analyzing the Italian forecasts made by the European Commission and Consensus Economics; and,
- Chapter A of the appendix explains the variables being analyzed, the nature of the Italian forecasts and the methodology used to produce them.

The value of these findings lies in the purpose and importance of studying forecasts. Accurate forecasts are crucial for the proper functioning of any economic system.

First, for **transparency and accountability**: finding the reasons that explain the difference between forecast and outcome is key to reassure people that forecasts reflect the professional judgment of the effects of a certain policy rather than politically motivated positivism.

Second, **learning lessons for future forecasts**: an in-depth analysis of how outcomes evolve relative to previous forecasts helps to build a better understanding of the drivers of the economy and public finances, improving therefore our understanding of the forecasts themselves.

Third, **providing information** to use when formulating ex-ante a policy: through the explanation of the causes that create the forecast differences, policymakers can potentially address them.

Fourth, forecasts are the **basis for policy analysis and economic surveillance**, including fiscal surveillance by the European Union (EU) directly.

Finally, any residual forecasting difference which cannot be explained by the previous points must stem from other judgments and assumptions, which will be investigated to derive conclusions that might reveal any systematic error in the forecasts made by the government and institutions that will be evaluated (Office for Budget Responsibility, 2017).

Related literature

To better understand how the thesis reaches its conclusions, it is useful to understand what studies have been made and what their point of view is, related to this topic.

The main paper, on which this thesis is based, was issued in 2016 and revisits the European Commission's forecast accuracy for the period 1969 to 2014 (Fioramanti M., 2016). This paper, which serves the purpose of giving data to compare any results with, illustrates the data which depicts the Italian government as overconfident in its forecasts. According to this research, the main reason for such a result is the misinterpreted influence of external assumptions, such as growth outside the EU, commodity prices, and other financial variables treated as exogenous by the forecasters. Another significant reason that largely explains the error of the one-year ahead forecast, is the assumption of unchanged fiscal policy, a basic assumption that can turn out to be erroneous ex-post.

Other papers, rather than explaining the possible reasons that justify the errors, focus on the incentives that governments and forecasters have, to be overconfident. The working paper of the National Bureau of Economic Research (Frankel J., 2012), describes the idea behind the fiscal caps of the Stability and Growth Pact. It then brings empirical evidence to show that, since 1999, especially for member states in critical condition, how countries were able to postpone painful adjustment by making overly optimistic forecasts arguing that the deficits would fall below the cap within the next years. Harvard Professor Jeffrey Frankel, also argued that the forecasts tend to be more biased in the three-year horizon than at shorter periods and that overly confident official

forecasts can help explain excessive budget deficits, specifically the failure to run surpluses during high output periods (Frankel J., 2012).

Previous studies about the fiscal policy in the EU (Jonung L., 2006), prove over-optimism in one-year forecasts from Italy, 0.5% on average, from 1987 to 2003. The main conclusion points out that an overly optimistic growth assumption is part of a strategy aimed at postponing painful policy decisions by painting a promising picture of the economy to make economic reforms appear less urgent. Such a strategy can also serve the purpose of avoiding politically risky reforms. Even though the study refers to a window of time nearly incompatible with this thesis, similar conclusions will be found in this thesis, therefore it can be still very useful to compare the motifs explaining such results. Jonung's paper explains that the biased nature of Italy depends on the broad set of mutually reinforcing structural problems affecting the labour market, the product market, and the financial market simultaneously. Since addressing only one problem in one area does not necessarily yield relief but rather emphasizes the shortcoming in every other area, and since a thorough reform is more difficult to execute, the government is incentivized to further postpone.

On the other hand, some documents argue that governments are already incentivized enough to properly forecast their growth. For example, Professor Chang (Chang C., 2009), correlates the ability of accurately forecasting, with business confidence, consumer sentiment, and overall trust by the market.

Research carried out on the forecast accuracy of the Japanese government (Ashiya M., 2007), also argues that a government has more incentives to accurately forecast rather than being overconfident. This is, not only to avoid political damage and international criticism but mainly to prevent the risk of carrying out inadequate economic policies based on an illusory picture of the economy.

Main findings

Italian forecast errors, derived from a wide sample of government forecasts, highlight the difference between the values forecast by the Italian government and the outcome at maturity. The variables considered are real GDP growth, passive interest, unemployment rate, primary surplus, net indebtment, public debt (as % of GDP), and nominal GDP. Public debt, among the one-year ahead forecasts, is the variable with the highest average forecast error. Over the 2001-2019 period, it is predicted 3.42% lower than its outcome (Table A1).

The main variable of this study, real GDP growth, on average is predicted 1.38% higher than its outcome for the 2001-2019 period, and 0.86% higher for the shorter window of time, 2010-2019.

For Italy, due to the high volume of data available, it is also possible to evaluate a variety of forecasts. Therefore, different forecast horizons are studied. On average, the predictive accuracy does not change in the two-year ahead forecast, however it worsens significantly in the three-year ahead projection (Table A1).

Moreover, errors for the tendency forecast perform better, with the exception of forecasts for passive interest, unemployment rate, and nominal GDP (Table A1).

In the 2001-2019 window of time under review, when examining other countries, as seen in Table A2, the error in predicting real GDP growth of France is of -1.00%, therefore predictions are, on average, 1.00% higher than realized value. For Germany, the average error is -0.29%.

In the 2010-2019 sample, the French average error for the real GDP growth is -0.52% (over forecasting error), the German one is 0.29% (under forecasting error), and the Spanish one, which is only introduced in this shorter sample, it is of -0.30% (over forecasting error).

When investigating forecasts formulated by other institutions on Italy, over the full period (2001-2019), the European Commission has an error of -1.27% (over forecasting error). Over the shorter sample (2010-2019), the European Commission

reduces its error to -0.75%. Consensus forecast, introduced for the short window, has an error of -0.56%.

Table A1:
One-year forecast average errors for Italy – 2001-2019

2001-2019	GDP growth	Passive interests	Unemployment rate	Primary surplus	Net indebtedness	Public debt	Nominal GDP
Programmatic	-1.38	0.21	0.16	-1.37	-0.99	3.42	9.40
Tendency	-0.92	-0.23	-0.39	-0.72	-0.40	2.40	-69.41

Table A2:
One-year forecast average error of countries and institutions (real GDP growth)

2001-2019	Italy	France	Germany	Spain	Commission	Consensus
Average error	-1.38	-1.00	-0.29	n/a	-1.27	n/a
2010-2019	Italy	France	Germany	Spain	Commission	Consensus
Average error	-0.86	-0.52	0.29	-0.30	-0.75	-0.56

1. A test for the hypothesis of biased forecast

Forecasts from public national institutions are often suspected of being too optimistic. Criticism was voiced against Italy and Europe during the Great Recession of 2008 and the subsequent slower than expected recovery. The matter under consideration is the absence of bias, which implies that on average the forecast error is zero or in other words that there is no systematic under or overestimation.

In practical terms, once the data is collected, analysis looking for any characterizing factor, especially scrutinizing for any positive bias, can be made (OECD, 2016).

This chapter quantifies the forecast error of the Italian government over the forecasting period taken into consideration (2001-2019) so that any conclusions are based on data. However, since forecast accuracy is measured on a spectrum, a bias is not measured in binary terms. To resolve this, the definition made by the Independent Authority of Fiscal Responsibility (AIReF) is used. According to this institution, a bias is considered significant when three conditions are met:

- the bias is large;
- it is unjustified ex-post, for example by being considerably larger than consensus forecasts; and finally,
- it is persistent, denoting at least a four-year occurance.

Nonetheless, when forecasting a government budget, the assessment of biases is far from straightforward. Therefore, even when controlled by the same institutions, different European governments can achieve drastically different results through the same underlying principles (Munoz L. 2018).

1.1 The methodology of the analysis

Related literature over the years has used different indicators for various reasons, however, the common aim is always the same, to evaluate the forecasts and explain the reasons behind the results. In line with such reasoning, for this study, indicators have been chosen based on their effectiveness in describing the aforementioned scope. The description of the methodology used to make the forecasts is found in Annex A.

Number of observations

Forecasts which have a limited amount of observations are less reliable and more susceptible to analysis errors. In this thesis, the variables more prone to such risk are the tendency forecasts. Moreover, the further the time horizon of the forecast, the less the amount of observations available there are.

Serial correlation (or autocorrelation)

It indicates the relationship between a variable and the lagged version of itself. Therefore, this value determines how well the past forecast of a variable predicts the following one, in this case, the study is made on a one-year lagged version of itself. Autocorrelation does not check for a biased or inconsistent estimator but only refers to its efficiency.

Serial correlation is used to evaluate the relationship of each series of forecast errors, the closer the value gets to 1, the more the underlying series examined is serially correlated, the more any given value of the sequence directly predicts the following one. If the serial correlation value is negative, there is an inverse relationship between the value of a series and the following one.

Average

The simple average, when calculated on the forecast error (E_{x_t})[4], is a value which highlights in which direction (positive or negative), and by how much, the forecast differs from the outcome. Although very simple, this measure is very effective due to its ease of understanding and calculation.

An unbiased and accurate estimate would have an average error close to zero, at least over a considerable amount of observations. A negative average value is a result of a forecast, on average, higher than the outcome (overconfidence), vice versa, a positive result of this estimator indicates that on average the forecast was lower than the outcome (under confidence).

Standard deviation (σ)

It is a variable that measures the amount of variation around the mean of the sample. In a normal distribution, 95.4% of all the values of the sample fall within two standard deviations from the mean. While this does not always apply in empirical research, it is reasonable to assume that a very high percentage of the values of any of the forecasts being analyzed will fall within the two σ mark (Agresti, 2013).

MAD – Mean Absolute Deviation

It is the average of the absolute deviations between the outcome and the forecast. It is utilized as a summary statistic of statistical dispersion of variability.

It is the sum of the absolute differences between the outturn and the forecast, divided by the number of observations [5] (Agresti, 2013).

---

[4] Errors are noted as E_{x_t} where x is the year relevant to the error, which is the year the error is calculated for, and t is the time horizon of the forecast used. It is calculated as such: $(O_x - F_{x\prime+t})$ keeping in mind that the x of the outcome and of the error is different from the x' of the forecast by a value of t (e.g. error for 2014 of the three-year forecast made in 2011 will be $E_{14\_3} = O_{14} - F_{11+3}$)

[5] The $\sum_{y=1}^{n}$ part of the formula indicates that the variable n, which indicates the number of observations being axamined, starts from a value equal to 1

As shown in formula:

$$MAD = \frac{\sum_{y=1}^{n} | O_x - F_{x\prime+t}|}{n}$$

MSE – Mean Squared Error

Used directly as an estimator of the accuracy of the forecasts, it is the average of the squared errors, so the average squared difference between the estimated and the actual value. Values closer to zero are signs of the better quality of the estimation process. It incorporates both the variance and the bias of the estimator, respectively how widely spread the estimates are from each sample and how far off the outcome they are (Hyndman, 2006).

It emphasizes larger errors because squaring larger numbers has a greater impact than squaring smaller numbers, therefore the estimator better considers samples with fewer outliers. This must be kept in mind when using it on the full 2001-2019 sample which takes into consideration also the 2009 crisis abnormal error (Agresti, 2013).

$$MSE = \frac{\sum_{y=1}^{n} (O_x - F_{x\prime+t})^2}{n}$$

RMSE – Root mean square error

The Root Mean Square Error (RMSE) is the square root of the MSE and is especially useful when large errors are considered particularly harmful. Also, RMSE is not independent of the number of observations and does not provide information on the direction of the errors (Fioramanti, M. 2016). Although not adding any information other than the one provided by the MSE, it is widely considered the standard accuracy indicator thus, the thesis will base its study on it rather than the MSE (which is still calculated as a step of the RMSE).

As shown in formula:

$$RMSE = \sqrt{\frac{\sum_{y=1}^{n}(O_x - F_{x\prime+t})^{\,2}}{n}}$$

MAPE – Mean absolute percentage error

Mean Absolute Percentage Error is the average of absolute errors divided by actual observation value. It measures the accuracy of the forecasting system as a percentage, as shown in the formula:

$$MAPE = \frac{\sum_{y=1}^{n} \left| \frac{O_x - F_{x\prime+t}}{O_t} \right|}{n} \times 100$$

This estimate has major shortcomings and values can be misleading for one of the following reasons:

A. for forecasts which are too high, there is no upper limit to the percentage error, which explains certain MAPE values which exceed 100% (e.g. outcome is 1, the forecast is 3, so MAPE is 200%;
B. MAPE weighs more negative errors where $O_x < F_{x\prime+t}$ than positive ones. This results in better consideration of series whose forecasts are lower.

The difference in differences ($DD_{m,n}$)

This new variable, $DD_{m,n}$, is used to calculate year by year the difference in the errors of two countries or institutions.

The letters m and n represent the two countries being compared, and the value is calculated by doing:

$$DD_{m,n} = \frac{\sum_{x=01}^{19}(E_{x\_1}^{m} - E_{x\_1}^{n})}{n}$$

Which is a summation of the differences in the forecast errors of the two countries which are then divided by the number of years the errors have been found

for. In other words, $DD_{m,n}$ estimates the average difference of errors. Due to the use of this variable being contingent on the use of two samples, it will be used only starting from Chapter 2. Further economic interpretation is made in Chapter 2.1.1.

The difference in RMSEs of two forecasters ($\boldsymbol{\Delta RMSE_{m,n}}$)

Through computation of the difference of the RMSEs of two countries, this thesis seeks to introduce an estimate which compares the simple difference between two estimates. If the resulting value is greater than 0, the first country considered has a higher RMSE than the second, therefore its forecasts are less accurate around the mean of the error. The following formula is used:

$$\Delta\mathrm{RMSE}_{m,n} = \mathrm{RMSE}_m - \mathrm{RMSE}_n$$

Due to the use of this variable being contingent on the use of two samples, it will be used only starting from Chapter 2.

1.2 The Italian forecasts

This chapter will describe the quantitative results obtained (Table A3) by interpreting the indicators. This analysis is performed for two sample periods, the complete 2001-2019 one and 2010-2019 one, which only considers the post-crisis period. A complete set of summary statistics is provided in Appendix C.

Table A3:
One-year ahead forecasts for Italy – 2001-2019

	GDP growth[6]	Passive interests	Unemployment rate	Primary surplus	Net indebtedness	Public debt	Nominal GDP
Average	-1.38	0.21	0.16	-1.37	-0.99	3.42	9.40
SD	1.94	2.45	1.24	1.40	1.35	4.25	45.05
SE	0.46	0.59	0.29	0.32	0.31	1.00	12.00
Serial Corr.	0.09	-0.19	0.51	0.20	0.19	0.22	0.01
MAD	1.67	0.96	0.92	1.56	1.36	3.95	35.04
RMSE	2.33	2.39	1.21	1.93	1.65	5.36	44.29
MAPE	193.94	22.11	10.27	136.37	47.58	3.24	2.26

Real GDP Growth

Arguably the most important indicator for understanding the behaviour of the government when issuing a forecast, it is the driver of most of the factors discussed in this thesis, thus, any government behavioural trait identified here will also leave traces in the other forecasts.

The **average error** over the full period is -1.38% (for $F_{x'+1}$). This very value high value can be partially explained by the error derived from the forecasts made during the 2008 crisis. For the 2009 year alone, the one-year ahead real GDP growth forecast error was of -6.18% ($E_{09\_1} = -6.18\%$). When observing the two and three-year ahead forecast, the average errors are respectively -1.19% and -1.74%, the

[6] As in the rest of the document, GDP growth is in real terms

noticeable reduction in the two-year ahead forecast can be mostly attributed to the unavailability of the two-year ahead forecast for the year 2009 (E_{09_2}).

By narrowing the forecasting window, so considering the post-crisis period (2010-2019), the **average error** of the one-year ahead predictions ($F_{x\prime+1}$), lowers to -0.86%. Therefore, on average, even without the understandable mistakes of the crisis, the Italian government overestimated its real GDP growth by nearly a full percentage point over the ten years considered.

The **serial correlation** test, applied to the main window of time, 2001-2019, produces a value of 0.09, denoting an absence of any serial correlation. A similar result of 0.14 when computing the average for $F_{x\prime+1}$, $F_{x\prime+2}, F_{x\prime+3}$, reaffirms such a conclusion. The 2010-2019 errors, on the other hand, have a serial correlation value of 0.91.

The average **RMSE** values are of 2.33 for $F_{x\prime+1}$ and 2.19 when averaging $F_{x\prime+1}$, $F_{x\prime+2}, F_{x\prime+3}$ over the full window, reinforcing the thesis brought forward by the study of the European Commission which found an RMSE of 1.84 for the same type of forecast but made between 1969 to 2014 (Fioramanti M., 2016). Not surprisingly, the shorter sample ($F_{x\prime+1}$ over 2010-2019), has a lower RMSE value, 1.87, compared to the sample comprising the crisis.

Although **MAPE** is less reliable for analyses with larger errors and can reach values higher than 100%, which is the case, its use can still highlight the extent to which real GDP growth forecasts are mistaken. MAPE is and 193.94% for the main sample and 134.05% for the shorter window.

Higher values for the full window value are indicative of how much the data set has been influenced by the 2008 and 2009 sovereign debt crisis.

The sharp increase in autocorrelation for the smaller sample is possibly explained by the lower number of observations. Regardless, it is a symptom of an inefficient process, which is also more accurate (see average error data for Italy in Tables 1 and 2).

Passive Interest (as % of GDP)

The prediction of the expenditure on passive interest is a value that gains importance due to its independence from the general methodology used for normal forecasts. Forecasting the expenditure on interest is a process based on the current levels of debt and is expected to have a lower error compared to the rest of the variables.

The **average error** for the full window is 0.21% for the one-year prediction. This value, which indicates consistent underestimation, is in line with the hypothesis that the Italian government has an overconfidence bias, which shows as overestimated positive (such as real GDP growth) forecasts and underestimated negative (such as interest expenditure) forecasts.

More pronounced is the **average error** for this variable calculated over the 2010-2019 window of time equal to 0.69%. The considerable difference between the small and big samples is highlighted also by other estimators. In the first case (2001-2019) the **serial correlation** and the **MAPE** are of -0.19 and 22.11%. In the second one, they are respectively of 0.47 and 140%. This behaviour can be explained by the fact that these indicators, which generally improve during the shorter sample, have the opposite behaviour when the underlying forecast is an expenditure.

Unemployment rate

The forecast **average error** for the unemployment rate is of 0.16% for the one-year forecast, of 0.50% across all forecasts (averaging $F_{x\prime+1}$, $F_{x\prime+2}$ and $F_{x\prime+3}$) and 0.58% for the post-crisis, single-year sample. Although not by a lot compared to other variables, the underestimation of this variable confirms the above-mentioned underestimation bias for variables considered negative, thus an Italian biased model overall.

The **RMSE** of 1.21, grows to 1.66 when also considering the two and three-year ahead forecasts and is of 1.33 for the reduced sample which considers only $F_{x\prime+1}$.

On average, the absolute deviation from the mean (**MAD**), again increases with the elongation of the forecast window, 0.92, 1.19, and 1.53 for $F_{x\prime+1}$, $F_{x\prime+2}$ and $F_{x\prime+3}$ respectively. For the reduced sample, the value is comparable to the longer sample one-year ahead forecast it is only one basis point higher, 0.93.

Autocorrelation goes from 0.51 for $F_{x\prime+1}$, to 0.58 for $F_{x\prime+2}$ up to 0.74 for $F_{x\prime+3}$ (2001-2019) and is of 0.62 for the post-crisis window (2010-2019). These high values signal a high correlation between a single forecasting error and its next value, which is a symptom of an inefficient estimating process.

Primary surplus

The **average error** for the forecast on the primary surplus rises from -1.37% for $F_{x\prime+1}$, to -2.31% for $F_{x\prime+2}$, to -2.98% for $F_{x\prime+3}$. As for most variables, it improves slightly to -1.04 in the minor sample ($F_{x\prime+1}$ for 2010-2019). This variable stands out for the abnormal difference between the average value calculated for $F_{x\prime+1}$ and $F_{x\prime+3}$. A slight divergence is to be expected when describing the three-year ahead forecasts which are more uncertain by nature, but nearly 3 percentage points of average error difference over 19 years (2001-2019) signal a consistent directional shift between the intentions of the government from year to year. This annual change of plans amplifies mainly the error of the three-year forecast. Also, the primary surplus corroborates the hypothesis of overconfidence bias in performing forecasts on positive variables.

Further analyses made on this forecast error highlight a low **serial correlation** of 0.20 ($F_{x\prime+1}$) over 2001-2019 and 0.13 ($F_{x\prime+1}$) over 2010-2019. On the other hand, **MAD** and the **RMSE** change considerably between the two time periods. Indeed, MAD values are 1.56 and 1.13, and values for the RMSE are 1.93 and 1.44 for the normal and short periods, respectively. This dissimilarity is not only explained by the financial crisis, which perplexed forecasters, but also by the volatile nature of this variable in the first years after joining the euro at the beginning of the century.

As expected, the tendency predictions for the whole period are more conservative, consequently, the forecast errors are lower, 2.07 and 1.77 for the RMSE

and MAD of the full sample compared to 2.70 and 2.34 of the programmatic ones (combining $F_{x\prime+1}$, $F_{x\prime+2}$, and $F_{x\prime+3}$).

Net indebtedness

When analyzing this variable, the **average error** increases the further the time horizon of the forecast is, from -0.99% for $F_{x\prime+1}$, to -1.74% for $F_{x\prime+2}$, up to -2.34% for $F_{x\prime+3}$. Considering the shorter sample (2010-2019), the average error noticeably improves, dropping to -0.58%. The negative signs, yet again, mean that the government predicted a value higher than the outturn, supporting the overconfidence bias aforesaid.

The **serial correlation**, on the other hand, remains stable around the value of 0.22 (0.19, 0.22 and 0.24 for $F_{x\prime+1}$, $F_{x\prime+2}$, $F_{x\prime+3}$) in between data sets considering different maturities of the forecast. No major change in autocorrelation for 2001-2019 means that there is no loss or gain in efficiency when performing forecasts. For the post-crisis sample (2010-2019), the serial correlation changes to a -0.17, slightly negatively correlated but still not big enough to prove the inefficiency of this variable.

The **RMSE** is 1.65 for the $F_{x\prime+1}$ normal sample (2001-2019), 1.10 for the small sample ($F_{x\prime+1}$ for 2010-2019), and is 2.18 for the whole sample across maturities ($F_{x\prime+1,2,3}$ for 2001-2019). Such values indicate a high dispersion around the mean of the error with a trend of improving the quality of the forecasts in the post-crisis period.

Public debt over GDP

The **average error** of debt over GDP ratio forecast is of 3.42% for $F_{x\prime+1}$ over the 19 years considered (2001-2019). Comparing different prediction maturities for the long sample (2001-2019) shows values going to 7.28% for $F_{x\prime+2}$ and to 11.92% for $F_{x\prime+3}$. The tendency analysis, supposed to be more conservative on the estimates, provides an average error of 6.64% (for $F_{x\prime+1,2,3}$ over 2001-2019) compared to a 7.54% for the programmatic ones (also $F_{x\prime+1,2,3}$ over 2001-2019).

Having one of the highest debt-over-GDP ratios of the whole Eurozone, Italy has been constantly in the spotlight over the last decades. Consistently, every Italian government in charge has therefore envisaged a reduction of this ratio, not always in the short term $F_{x\prime+1}$, but usually in the medium term $F_{x\prime+2}$ and always in the longer term $F_{x\prime+3}$. Yet, the ratio has been consistently going in the opposite direction, growing at an average rate of 1.31% every year (ISTAT, 2020). This major discrepancy between planning and actual behaviour is reflected in an extremely high forecasting error. In line with other variables, the government underestimated the future debt over GDP ratio in its forecasts, a variable considered negative when too high.

RMSE is 5.36 and the **MAD** is 3.95 for the main window of this study for the one-year forecast (2001-2019) and 9.48 and 7.81 for the average of the different maturities ($F_{x\prime+1,2,3}$). S**erial correlation** 0.22 for $F_{x\prime+1}$, greatly increasing in $F_{x\prime+2}$ to 0.35 and nearly tripling in $F_{x\prime+3}$ up to 0.61. Over the 2010-2019 period, it is of 0.46.

Due to the magnitude of the errors, analyzing the **RMSE** of 9.48 or the **MAD** of 7.81 is trivial (calculated for $F_{x\prime+1,2,3}$ over the 19 years window). However, discussing the serial correlation, which is not directly affected by a biased model, is still viable. A simple interpretation of the $F_{x\prime+1}$ low correlation opposing the $F_{x\prime+3}$ nearly tripled one, maybe that the government is incentivized to spend more than what it can afford and finance its ideas mostly with debt, which in the long run does not pay off. Therefore, even if the expected results are obtained in the short term, they are not when the time horizon is increased. This hypothesis is also supported by the rapid growth of errors in relation to extended maturity of the forecast.

Nominal GDP

These values, which are calculated in thousands of billions, signal an **average** under forecasting **error** of 9.40% for $F_{x\prime+1}$, which then switches in an over forecasting error for $F_{x\prime+2}$ and $F_{x\prime+3}$, becoming -9.55% and -34.06%, respectively. The errors in the main sample, one-year horizon, are not **serially correlated** (0.01). The same cannot be said about $F_{x\prime+2}$ and $F_{x\prime+3}$ over the same horizon, values indeed grow to 0.23 and 0.76. These values are partially explained by the inclination to overestimate the government's potential to affect the nominal GDP value in the medium to long run

RMSE, **MAD,** and **MAPE** (for $F_{x\prime+1}$over 2001-2019) are of 44.29, 35.04, and 2.26% while for the 2010-2019 sample they are 29.55, 21.87, and 21.87%. Such values are hard to compare with others due to the uniqueness of their unit of measure, nominal GDP is not calculated as a percentage of GDP but in absolute terms. However, estimates can be considered relatively close to the mean of the error, thus quite accurate. To point this out further, MAPE values of 2.26%, 2.57%, and 3.29% are useful to understand by which percentage the government has miscalculated this indicator, mistakes for other variables are considerably higher, some by a factor of ten or a hundred.

2. Comparing accuracy across institutions and countries

This chapter compares the accuracy of the real GDP growth forecasts with the German, French, and Spanish governments. Moreover, it analyses the performance of the forecasts for Italy made by the European Commission (EC) and by Consensus Economics (CE). If not specified differently, the paper will focus on the window of time going from 2001 to 2019 for the one-year ahead forecast ($F_{x\prime+1}$). For consistency reasons, a comparison will also be made with the post-crisis period for $F_{x\prime+1}$.

Mostly, other forecasters have published their reports during the same months of the year as the Italian government.

- **France**: since implementing the 2011 stability programme, the French government has published the forecasts in April of each year. Before that, the dates of publication fluctuated between December of the previous year and January;
- **Germany**: evaluations of the German forecasts are also based on the yearly stability programmes, which like France, and starting in 2011, are published in April. Before that, they were usually published in November or December of the previous year with sporadic exceptions in October and January;
- **Spain**: forecasts are published either in March or in April with an exception once in January. However, data is only available from 2009 onwards;
- **European Commission**: The EC publishes its forecasts multiple times during the year, the data retrieved for this thesis has always been published in spring; and finally,
- **Consensus Economics**: data retrieved, has been published every year in June except for 2010, year for which the forecast has been published in March 2009. Data is only available from 2008 onwards.

2.1 Cross-country comparison

Cross-country comparison is made to understand whether any bias is country-specific or, if instead, it is a common practice among sovereign countries to be overconfident when producing macroeconomic forecasts, specifically about real GDP growth.

2.1.1 Italy compared to the French government

The comparison with the French forecasts for the sample period (2001-2019) shows a better performance for the French (Table A4), with the exemption of the serial correlation which is low for both countries. Yet the French government has still high values overall, such as an **average error** of -1.00% (compared to Italian -1.38%), **RMSE** of 1.70 (compared to Italian 2.33), and a **MAPE** of 151.47% (compared to Italian 193.94%).

Table A4:
One-year forecast errors of governments (real GDP growth; 2001-2019)

2001-2019	Average Error	Standard deviation	Serial correlation	MAD	RMSE	MAPE
Italy	-1.38	1.94	0.09	1.67	2.33	193.94
France	-1.00	1.42	0.26	1.11	1.70	151.47
Germany	-0.29	1.83	-0.16	1.31	1.85	132.71
Spain	n/a	n/a	n/a	n/a	n/a	n/a

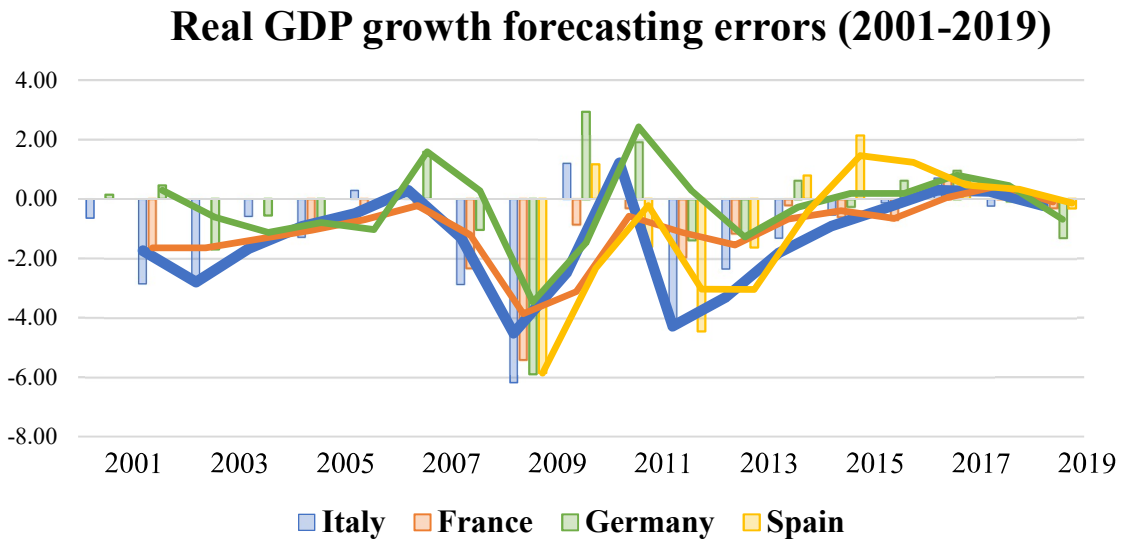
Comparing results for the window 2010-2019 leads to similar results (Table A5). The French standard deviation is less than half, 0.70 compared to 1.77, MAPE improves to 120.04% (193.84% for Italy). Serial correlation drastically changes, 0.21 for the French and 0.91 for the Italian government and RMSE is one point lower, 0.87 compared to the 1.87.

Table A5:
One-year forecast errors of governments (real GDP growth; 2010-2019)

2010-2019	Average Error	Standard deviation	Serial correlation	MAD	RMSE	MAPE
Italy	-0.86	1.77	0.91	1.34	1.87	134.05
France	-0.52	0.74	0.21	0.70	0.87	120.04
Germany	0.29	1.35	0.33	1.13	1.38	105.34
Spain	-0.30	1.87	0.42	1.31	1.80	142.42

As seen in Graph 1, which represents the two-year moving average of the four countries examined, France has consistently overestimated its real GDP growth, the orange line indeed stays mostly under the horizontal axis. However, the French moving average, except for the 2008 financial crisis, stays flat and close to the axis. This graph therefore confirms the data which shows a slight bias but low variance.

Graph 1



Comparing average forecast errors across countries

While evaluating how the French forecast for its GDP growth rate compares with the Italian forecast is very useful, the following paragraph is specifically set out to

determine whether the French and Italian governments have similar behaviours, without dwelling on what the specific behaviours might be.

To do so, real GDP growth forecast errors of the two countries year by year will be compared. Ideally, two countries with the same systematic biases, or lack thereof, will over, under, or accurately estimate real GDP growth in the same way. In this case, the average difference between the errors of the two countries over the period examined should be 0.

This indicator called the **Difference in Differences**, which is now being calculated on Italy and France ($DD_{I,F}$), is of -0.46%. Meaning that on average, regardless of the magnitude of the error of each country, the Italian government had a real GDP forecast error of 0.46% higher than the French one (2001-2019). This process is especially useful to mitigate the hindrance presented by external elements, such as the 2008 crisis, that may have disrupted the conventional forecasting schemes since such factors supposedly puzzled forecasters in the same way.

Likewise, when considering the smaller sample (2010-2019), $DD_{I,F}$ is of -0.29%, meaning that the two governments have marginally converged performances relative to each other after the sovereign debt crisis.

From these analyses, the main takeaway is that, even though the French government has a less than perfect forecasting model, its errors tend to be lower when compared to the Italian ones. Indeed, the indicators applied, indicate an imperfect prediction model for the real GDP growth, yet the French model significantly outperforms the Italian one.

The line of reasoning that governments make similar mistakes cannot be rejected yet, both because the errors made by the French government are still significant and because only one case is not enough proof.

Comparing RMSEs across countries

As a supplement to this study, a comparison of the deviation of the countries, the difference of the **RMSE**, is evaluated (as seen in Table A6).

$\Delta\text{RMSE}_{I,F}$ is 0.64 for the full period (2001-2019) and 1.00 for the post-crisis one (2010-2019). These values denote a high difference in the behaviour of the two institutions which increases even more in the last period. However, the importance of this value lies in the comparison that can be made with other countries ($\Delta\text{RMSE}_{I,F}$ compared with $\Delta\text{RMSE}_{I,G}$ and $\Delta\text{RMSE}_{I,S}$).

Table A6:
Italy vs other countries: comparison of forecast errors (real GDP growth)

2001-2019	$DD_{I,n}$	$\Delta\text{RMSE}_{I,n}$	**2010-2019**	$DD_{I,n}$	$\Delta\text{RMSE}_{I,n}$
France	-0.46	0.64	**France**	-0.29	1.00
Germany	-1.12	0.48	**Germany**	-1.14	0.49
Spain	n/a	n/a	**Spain**	-0.73	0.07

2.1.2 Italy compared to the German government

Concerning the German government, the Italian forecast accuracy for the $F_{x\prime+1}$ for the whole sample period (2001-2019) has an inferior performance. Numerically, the German forecast on real GDP growth has an **average error** of -0.29% over the full sample (compared to -1.38 for Italy). Surprisingly, from 2010 to 2019, Germany has been underpredicting, thus overperforming relative to its forecasts, its real GDP growth by 0.29%, contrary to Italy whose error is -0.86% for the same period. These values alone, confirm the hypothesis that not all forecasters have incentives to systematically overestimate the real GDP growth. Comparing serial correlation (2001-2019), neither government fails the autocorrelation test. Also, German **RMSE** of 1.85 surpasses the 2.33 of the Italian government and **MAPE** is 132.71% (for Italy it is 193.94%). Overall, the German government outperforms Italy's forecasts on real GDP growth.

When observing Graph 1, it appears through the peaks above the horizontal axis, that Germany, more than once, underestimates its real GDP growth. Such behaviour is expected in an unbiased model, that over the long run should have an average error close to 0.

Comparing average forecast errors across countries

When considering the German government, the $\boldsymbol{DD_{I,G}}$ is of -1.12% for the long sample (2001-2019), and to -1.14% for the short one (2010-2019). On average the German government, therefore, made a mistake more than one-percent lower than the Italian government. The hypothesis that governments tend to make similar mistakes is thus partially rejected(7).

Comparing RMSEs across countries

The $\boldsymbol{\Delta RMSE_{I,G}}$ **i**s 0.48 for 2001-2019 and 0.49 for 2010-2019. Therefore, the two governments had a similar performance both before and after the Great Recession. When exclusively considering the variability of the forecast (RMSE), the French had a more distinct forecasting behaviour compared to Germany, whose model is closer to the Italian one.

2.1.3 Italy compared to the Spanish government

Due to a lack of forecasts made by the Spanish government before the year 2009, it is possible to analyze the data only relevant to the post-crisis sample already described in the thesis (2010-2019).

During this period, on average, the government overestimated the real GDP growth for the following year ($F_{x\prime+1}$) by 30 basis points (**forecast error** of -0.30%). In

(7) Only partially since the paper cannot derive definitive conclusions based only on one case, also since the case being considered is Germany, which to date is the strongest European economy

opposition to a low average error, which is considered positive, other values, such as **RMSE** and **MAD** are very high, especially if compared to France and Germany (RMSE of 1.80 and MAD of 1.31).

All the values together can help understand the picture portrayed by the Spanish forecasts. The numbers delineate a government that does not systematically over or underestimate its forecasts but frequently misses by a substantial amount the correct outcome of the prediction.

As seen in Graph 1, the Spanish error is very volatile. However, as seen from the yellow line, representing the two year moving average, due to the few years of underestimation of the real GDP growth starting in 2014, the government was able to considerably reduce the average of the error.

Comparing average forecast errors across countries

When collating the Italian and Spanish samples, the performance is very much alike. The only variable standing out, which is also the most relevant, is the average error. It is therefore expected to have a high $DD_{I,S}$ which is -0.73% (Table A6). This value sits in between the French and the German values and confirms the hypothesis that identifies Italy as the country with the larger systematic bias among those examined.

Comparing RMSEs across countries

Looking at Spain in comparison to Italy, the $\Delta\text{RMSE}_{I,S}$ is 0.07 (Table A6), thus a low value which sets the accuracy of Spain on a similar level to the Italian one. Nonetheless, Italy still has a higher RMSE, therefore confirming the hypothesis that among the countries evaluated, Italy is the less accurate forecaster.

2.2 Cross-institution comparison

Following the example set by previous studies (Frankel J., 2020; OECD, 2014; Fioramanti M., 2016), this paragraph will now compare the results gathered on the Italian forecasts, with the ones of the European Commission and Consensus Economics. The focus of this paragraph will, therefore, be on forecasts made about the Italian government by institutions other than the government itself. The scope of this section is to understand whether the fault for the forecasting errors is to be attributed to a single institution, such as the Italian government, or if it is to be expected by every institution that makes predictions on the government in question.

2.2.1 Italy compared to the European Commission

Already mentioned in the discussion of the methodology of the Italian government, the European Commission can effectively influence the behaviour of countries that do not comply with policies such as the 3% deficit limit by sanctioning, proposing and enforcing legislation, and implementing policies (Hallerberg M., 2009).

The analysis will consider real GDP and unemployment forecasts over the same time windows so far examined. Due to uniform methodology and consistent publications on behalf of the EC, it was possible to compile forecasts for every year considered in this thesis, so 19 values over the 2001-2019 period and 10 values for the 2010-2019 post-crisis one. Bearing in mind that the European Commission only made $F_{x\prime+1}$ forecasts, only $F_{x\prime+1}$ data from the Italian sample will be used for the comparison. Both the data and the analysis follow.

The comparison between the Commission and the Italian government for the whole sample period shows slightly better performance by the EC overall, on average unemployment does not show significant differences with the Italian forecasts. On the other hand, real GDP growth there are some interesting conclusions to be made, as seen in Tables A7 and A8, the **average error** is quite similar to the Italian one, -1.27% (2001-2019) and -0.75% (2010-2019). However, more divergence is found in other

values such as **RMSE** and **MAPE**. The Commission's results, with a focus on real GDP growth, however, did not reveal to be much lower across most indicators analyzed, which partially disproves the superior performance talked about in the latest 2016 study by the European Commission (Fioramanti, M. 2016).

Table A7:
One-year forecast errors of institutions (real GDP growth; 2001-2019)

2001-2019	Average Error	Standard deviation	Serial correlation	MAD	RMSE	MAPE
Italy	-1.38	1.94	0.09	1.67	2.33	193.94
Commission	-1.27	1.84	0.00	1.54	2.20	238.65
Consensus	n/a	n/a	n/a	n/a	n/a	n/a

Table A8:
One-year forecast errors of institutions (real GDP growth; 2010-2019)

2010-2019	Average Error	Standard deviation	Serial correlation	MAD	RMSE	MAPE
Italy	-0.86	1.77	0.91	1.34	1.87	134.05
Commission	-0.75	1.67	0.34	1.20	1.74	73.41
Consensus	-0.56	1.45	0.23	0.99	1.55	57.28

Results on **serial correlation** show a lower value for both variables, respectively by 0.57 and 0.16 for unemployment and GDP growth. For MAPE also, lower results indicate improvement, reducing from 134.05% to 73.41%, and from 93.33% to 8.74%. RMSE values are only better by a small margin (0.17 for GDP growth and 0.10 for unemployment).

Similar results were also found in previous studies, therefore over the full sample period (2001-2019) the hypothesis of a biased estimation process used by Italy cannot be rejected nor confirmed. (Gonzáles, L. 2012).

Every year both the EC and Italy made errors very similar one to the other, proving that the forecasts on average tend to be nearly the same.

Comparing average forecast errors across institutions

The comparison made between institutions is expected to yield more conservative results than the analysis made between countries, the reason being that the former considers forecasts made by different subjects on the same object (every institution estimating Italy's real GDP growth), while the latter is both made by different subjects and on different objects (every country estimating their real GDP growth).

The Commission, over the 19 years covered by this study, mistook on average the **real GDP growth** of the Italian government by 11 basis points. Although not impressive if compared to the rest of the study, it is still a significant improvement in forecast accuracy. Yet, such contrast, especially if taken alone, is not enough to confirm any hypothesis made on the biasedness of the Italian government.

Commission's Methodology

The forecasts are not based on a centralized econometric model but are analyses produced by experts at the Directorate General for Economic and Financial Affairs, using models and field knowledge. Consistency is ensured by several cross-country and cross-variable checks (European Commission, 2020), which may be the reason that the forecasts were so similar to the Italian ones.

2.2.2 Italy compared to Consensus Economics

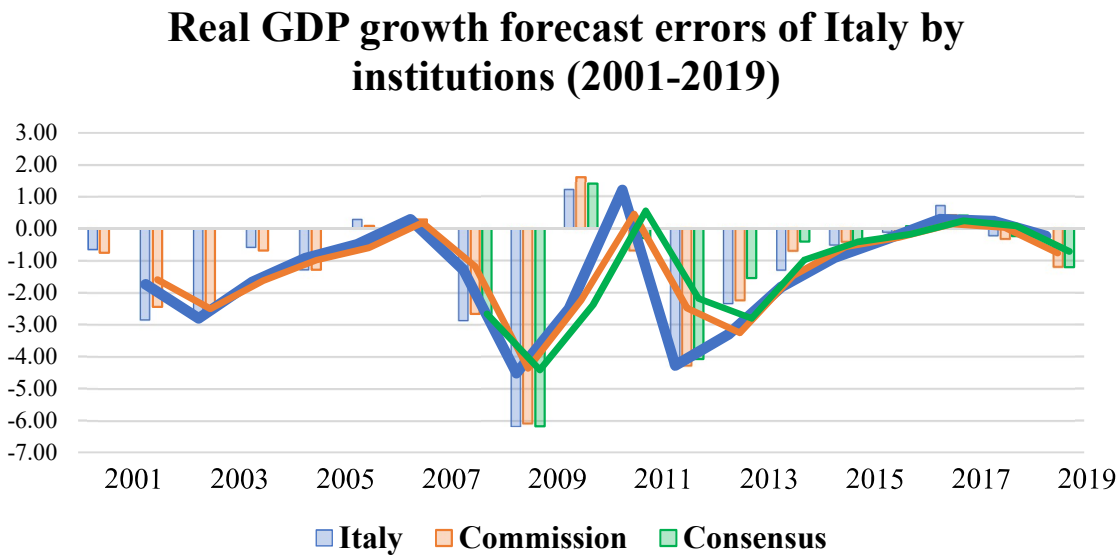
When evaluating the Italian forecasts, comparison with an independent source of forecasting, such as Consensus Economics, will provide a further forecasting institution category to better understand more precisely which type, if any at all, is biased.

Values on the forecasts made by Consensus Economics (CE), which gathers forecasts made by numerous independent forecasters and averages them out, have been retrieved starting from the year 2008. Relevant to this study, therefore, there is a total overlap of data only for the window of time from 2010 to 2019, for this reason, all the indicators in this paragraph refer to the window in consideration.

On average, Consensus forecasts over-estimated the real GDP growth of the Italian economy by more than half percentage point (**average error** = -0.56%). **Serial correlation**, contrary to the Italian forecasts, is very weakly correlated, only 0.22. Moreover, **RMSE** is 1.52 and **MAD** is 1.01, which compared to the forecasts of the Italian government, are respectively 0.35 and 0.33 better. Perhaps the largest improvement is made in the **MAPE**, which is of 57.28% for Consensus, 76.77% better than the Italian counterpart (considering $F_{x\prime+1}$ for 2010-2019).

Moreover, by comparing the lines in Graph 2 which represent the 2 periods moving average of the errors, it appears that the difference across institutions comes down to the magnitude of the forecasts made, which for Italy are always greater. In this study, it is fair to say that these forecasting errors are higher the less independent the institution making them is (as shown also in Tables A7 and A8).

Graph 2



Comparing average forecast errors across institutions

On average, every year of the post-crisis sample, the Italian government forecast the **real GDP** to grow by 0.30% more than what the Consensus Economics prediction, which is a considerable improvement also compared to the Commission's forecasts. The Commission underperformed on every indicator compared to the independent forecast made by CE.

This difference is large enough to partially confirm the hypothesis made at the start of the chapter on the positive bias of the Italian forecasts. Moreover, Consensus Economics, confirms its record of better performance when compared to the European Commission.

Consensus Economics Methodology

Consensus Economics produces its forecasts by surveying the leading independent forecaster around the world. Incoming survey responses are then processed using proprietary software and checked for accuracy, completeness, and integrity. Empirical studies show that pooling the forecasts increases accuracy

(Timmermann A., 2006). The main argument in favour of this type of forecast stems from the contrast with analyses made by individuals that may be subject to some degree of behavioural biases. Therefore, the pooling of data not only can achieve a more accurate result but also invalidates the effect of biases affecting just some of the forecasters.

2.3 The Parliamentary Budget Office

As described in the last paragraph of the Chapter A.2.3 of the appendix, about the methodology of the DEF, the independent organization, the UPB, has the function of observing and improving the forecast throughout every iteration of the draft made by the government.

Since its implementation in 2014, the errors made by the Italian government reduced significantly and aligned with the errors made by both the European Commission and Consensus Economics (as seen in Graph 2).

It would be overly simplistic to assume that all the merits go to this organization, therefore this thesis will just point out that, once an independent institution started reviewing the forecasts being made, the errors also reduced significantly. Such a solution has been proposed by many researchers such as the Doctor of Philosophy in Economics, Carlos Marinheiro, who in 2010 suggested that at least countries with very large forecasting errors, should delegate the forecasting process to independent institutions, which are less likely to be biased, and are generally more accurate (Marinheiro C., 2010).

The UPB, which does not totally control the estimation process but supervises and eventually amends some elements, at least partially adheres to the theory brought forward by the Professor.

Conclusions

This thesis updates the previous research on forecast accuracy until 2014 made by the European Commission (Fioramanti M., 2016) by extending the observation period to 2019. Moreover, it compares the accuracy of the forecast done by Italy, France, Germany, and Spain, as well as the European Commission, and Consensus Economics. Considering a window of time of the first 19 years of the new millennium (2001-2019), allows this document to evaluate the Italian performance before and after the Great Recession of 2008-2009.

The thesis quantifies the mistakes that Italy makes in its forecasts and tries to understand whether they are persistent or not. While mainly focusing on real GDP growth, to better comprehend Italian behaviour also other variables are examined. Indeed, the thesis extends its focus also to passive interest, unemployment rate, primary surplus, net indebtment, public debt (as % of GDP), and nominal GDP. Overall, the forecasting tendency error, which should present a larger error than the programmatic one, on the contrary, is usually lower, confirming previous studies that blamed the Italian government for being overconfident. Furthermore, the fact that growth variables (like real GDP growth) are overestimated, and expenditure ones (like interest expenditure) are underestimated, corroborates the hypothesis of biasedness in the forecasting methods used.

Italian year ahead forecasts are generally found to be directionally accurate, meaning that they successfully predict pick-ups and slowdowns in the variables examined. Pick-ups and slowdowns turn out to be better predicted the shorter the maturity of the prediction is. The longer the time horizon of a forecast, the more legitimate uncertainty is, the more space there is for wishful thinking. Directional accuracy is weakest in longer-term forecasts which are understandably less accurate. Some variables, such as Debt to GDP ratio, even proved to be directionally inaccurate, meaning that forecasts were consistently optimistic and did not reliably anticipate any worsening of the indicator. Indeed, the debt over GDP ratio always anticipated the ratio improving over the next three-years, while the outcome usually progressed in the

opposite direction. This partially explains the positive bias found in these Italian forecasts.

Regarding the different time windows considered, 2001-2019 and 2010-2019, major discrepancies are easily explained by the anomalously large forecast error derived from the 2008-2009 crisis. Such outliers are largely responsible for the difference between time windows, other dissimilarities are due to the natural evolution of the methodology over time and possibly the implementation of the UPB. Overall, Italian data corroborates the hypothesis of a positive bias of the whole model, with a trend of improving behaviour over the past years.

The country-to-country comparison confirms the results already found by previous studies that Italy stands out among European countries for its bias. This thesis studies the difference that sets governments apart when forecasting their real GDP growth. Over the full sample, 2001-2019, the two countries examined, France and Germany, had a considerably improved performance compared to Italy, both regarding accuracy and consistency. The same conclusions can be made about the post-crisis sample, which also considers Spain. Nonetheless, interesting dynamics can be pointed out when evaluating such countries.

In the full sample: Germany, currently the strongest European economy, is also the country whose forecasts are the most accurate. France, based on accuracy, is the second-best forecaster; however, it slightly outperforms the German forecasts when the focus shifts to consistency, embodied by the variance around the mean of the error. Italy comes last for both accuracy and consistency, meaning that both average error and variability measures are worse for this country. In the shorter sample, Spain, whose economy is often compared to the Italian one, also performs better than the French forecasts, proving to be only less accurate than Germany. However, Spanish forecasts are not very consistent. Variability measures, such as standard deviation and RMSE, are indeed not much different from the Italian ones.

The serial correlation test for the full sample across-countries does not return any significant values, on the other hand, for the shorter sample, perhaps due to the lower amount of observations, values increase significantly.

Comparison across-institutions, which seeks to discover whether other institutions can correctly perform forecasts on Italy, considers the European Commission and Consensus Economics. Two key institutions, with the former being the main European monitoring institution and latter being the main independent forecasting organizations. Results for the former, both over the full and shorter samples, are defined by an improvement of the forecasts by more than 10 basis points. The latter institution, however, for which data has only been retrieved for the shorter sample, presented forecasts which on average were better than the Italian ones by 0.30%, a significant improvement also if compared to the European Commission.

Italy is not the first country and will not be the last to be overly optimistic. On the contrary, the overwhelming amount of objective and subjective data analyzed, proves that governments, or at least some of them, have incentives to overestimate their predictions, especially if related to GDP growth (Ashiya M., 2007). If anything, this thesis proves that Italy more than others, believes that the possible benefits of such behaviour outweigh the downfalls.

Over the years, many possible explanations for overconfident behaviour have been found. Historically, in Italy, over-optimistic forecasts worked in favour of the policy-makers, due to the existence of the programmed inflation rate (TIP), which, from 1993 until 2009, was used for adjusting wages to inflation (Ministry of Economy and Finance, 2020).

Harvard Professor, Jeffrey Frankel, argued that overly optimistic official forecasts have facilitated complacency and so have contributed to tax cuts and increases in government spending, and therefore to realized budget deficits (Frankel J., 2012). Over optimism can also explain excessive budget deficit, especially in a country that fails to consistently run surpluses even during periods of high output such as Italy. If a boom is expected to last indefinitely, saving for a rainy day is unnecessary. The Italian government has proven to expect booms to last indefinitely underestimating the persistence of busts.

Furthermore, despite the tendency to exceed the 3% cap on budget deficits, Italy rarely forecast a violation of the limit in the coming years. Officially, if euro area

governments are not in violation of the 3% cap at the time forecast is made, the forecast is no more biased than other countries. On the contrary, deficit caps, such as the one imposed by the Stability and Growth Pact, might have the opposite effect of incentivizing a larger forecasting bias when governments under such restrictions exceed the limits set, they respond by adjusting the forecasts rather than the policies (Frankel J., 2012).

Over-optimism in predicting national growth appears linked to over-optimism in predicting the budget balance and, therefore, in predicting every variable analyzed in this document. This helps to explain why similar behaviour is found in every variable considered in this study.

The current situation, however, will force some changes in the forecasting process. After the Great Recession, research staff was forced to think out of the box and come up with creative new approaches to forecasting (Lucia A., 2014). Similarly, following the current global crisis, countries will implement policies that will play a key role in the process of global recovery. For these policies to be efficient and effective, they will have to be complemented by a forecasting system that aims to achieve the same goals.

Overall, Italy's track record has evolved noticeably. Inefficiencies can be attributed both to political and economic reasons, however, especially since 2014, forecast accuracy has been improving considerably. It will be crucial for future studies, to assess whether recent improvements in the Italian forecasting mechanism are part of a deep reform with long-lasting results or not.

Appendix

A. The data set for the Italian variables

The first step when deriving an ex-post evaluation of forecasts is to check for the availability of data. A complete evaluation must include both qualitative and quantitative data, ranging from a Professors' position paper to a database of outcome data (OECD, 2016). This section of the thesis will collect and show forecasts made by the Italian government over the last nineteen years (2000-2018), and also evaluate the performance of the forecasts of the Italian government made post financial crisis (2009-2018).

To retrieve the forecast data for the Italian government, it is possible to utilize the official annual Economic and Financial Planning Document. This paper, which is now called DEF[8], has changed name multiple times during the last two decades.

Originally called DPEF[9], established in 1988, it was created to define the macroeconomic state of the country and to align economic and financial objectives for institutional organs. Every year, it was presented by the Government to the Parliament before the 30th of June and presented forecasts up to three-years ahead, occasionally extending its time horizon for the whole legislature (Ministry of Economy and Finance, 2020).

In 2009, with the introduction of the Italian law No. 196 (Law No.196 of 2009), the DPEF was substituted with the DFP[10], a document which had to be presented by the 15th of September each year (Ministry of Economy and Finance, 2020).

The DFP however was short-lived, due to an Italian directive enforcing new rules adopted by the European Union regarding economic coordination policies (Law

[8] DEF, which is the Economic and Financial Document, stands for Documento di Economia e Finanza

[9] DPEF, which is the Economic and Financial Planning Document, stands for Documento di Programmazione Economico-Finanziaria

[10] DFP, which is the Public Finance Decision, stands for Decisione di Finanza Pubblica

No.39 of the 7th of April 2011), it was substituted by the DEF. The DFP was published before the 10th of April.

A.1 Explaining the variables analyzed

The variables collected comprise both tendency and programmatic forecasts which, by definition, will substantially differ from each other. The most relevant type is the programmatic forecast, as policymakers aim to reach the values forecast. However, the tendency forecast will be used as a baseline scenario to offer insights on whether the policies achieved the hoped-for-results, or if on the contrary, they did not skew the results away from the spontaneous outturn.

Other than the type of scenario predicted, the thesis will also consider forecasts made for different time horizons, collecting forecasts one-year ($F_{x\prime+1}$), two-year ($F_{x\prime+2}$) and three-year ($F_{x\prime+3}$) ahead of when they are being made. [(11)] Due to different documents being published during different periods of the year, for ease of calculation, the time window indicated by *t* starts at the end of the year in which the prediction was made. Therefore, the so-called one-year forecast is the year ahead forecast, which is on average a twenty-months ahead prediction.

The variables considered in this thesis are:

Real GDP growth: The gross domestic product is the most important indicator of the health of the economy. Being very inclusive by nature, it influences most elements discussed in this study, therefore, its appropriate estimation is of utmost importance for the reputation and reliability of the Italian government on the global markets and in Europe. The GDP growth measures the rate at which the economy is growing. It does so by comparing the Gross Domestic Product (GDP) at the time in

[(11)] $F_{x\prime+t}$: forecast is made in year *x'* for *t* years ahead, where x' stands for the last two digits of the year being considered and *t* is the number of years until the forecasted moment. (e.g. one-year ahead forecast is noted as $F_{x\prime+1}$, forecasts for two-years ahead made in 2016 are noted as F_{16+2})

which the forecast is being made, with one of the years for which the prediction is being made. The four components of the GDP and its growth rate are personal consumption, government spending, business investment, and net trade (Macroeconomics 9th ed., Mankiw). In this thesis, when talking about GDP growth, implicitly all values attached are considered *real*, so already adjusted for inflation.

Passive interests (as % of GDP): This variable measures the expenditure that the government expects to bear due to the interest to be paid on debt as a percentage of GDP.

Unemployment rate: The rate that highlights the percentage of the labour force that is jobless. For the labour force, the population actively seeking a job between the age of 15 and 64 is considered.

Primary surplus: This variable highlights the difference between how much the government collects, for example through taxes, and how much it spends, for example, to finance public services, net of the expenses of the interest on the debt (Net debt – interest expenditures). The primary surplus is considered one of the main variables for the coordination of the balancing policies among all European member states through the Stability and Growth Pact. Specifically, it is used as a measure to monitor the behaviour of a single country and is part of the corrective section of the Pact (European Council, 2020).

The importance of this variable traces back to the policy promoted between 1991 and 1997 by the Prodi government, who increased the primary surplus up to 6% and, during 1997, reduced the deficit from 6.7% to 2.7% to signal good intentions to the EU. This record highlights the relevance of the variable and how important it can be to signal goodwill and foster confidence in the international markets. (Pesole D., 2019).

Net indebtedness: It is the difference between all incomes and all balance expenditures, also calculated by adding the interest expenditures to the primary surplus. Necessary to verify the compliance with the criteria imposed by the Maastricht treaty (State General Accounting Department, 2008), the net debt plays a key role in the yearly economic planning.

Public debt (over GDP): It is considered the key parameter to check for compliance with convergence policies required as part of the European Monetary Union (Bahrke J., 2015), specifically concerning the Stability and Growth Pact (EUR – Lex, 2020).

Nominal GDP: It is the variable that measures the monetary value of the aggregate production of the economy. Being nominal, it considers the current year monetary value.

Usually overlooked, the nominal value is as important as its real counterpart. To do real GDP, and its growth rate, the nominal value must be adjusted to take into account price changes to show whether the output value increases because of higher production or because prices are simply increasing (Callen, 2020). Also, public discussion of economic forecasts tends to focus only on real GDP. But the nominal is more important for the behaviour of the public finances, so it is crucial to understand it to build a picture of the drivers of GDP growth (Office for Budget Responsibility, 2017).

A.2 Methodology used by the Italian treasury

In the 19 years of forecasting considered, the methodology adopted by the Italian government has changed multiple times, this minor aspect adds to the possible reasons for discrepancies and errors. Across time, many econometric models have been used to forecast economic activity. These models differ in structure and the data used, and so their forecasts are not perfectly comparable with each other (Ray C., 1990). At times, forecast values relevant to this study have been omitted illustrating a lack of uniform methodology, for example, tendency estimates have not been consistently illustrated throughout the full-time span under review. This must be kept in mind when evaluating the estimators derived from data which has a lower number of observations. [(12)]

To better comprehend the evolution of the methodology used by the Italian government, the next chapter will describe the main approach used and how it has been revised during the years.

A.2.1 Forecasts of the DPEF

The DPEF did not complement its papers with a specific description of the methods used to create its forecasts, therefore, this section will group the indirect information contained in the DPEF documents. This step is essential to understand how and why the methodology has evolved in more recent years.

DPEF 2002-2004 (Ministry of Economy and Finance, 2001)

Analyzing the methodology of the DPEF, the document itself states how the statistical techniques available are not sufficient to separate effectively the cyclical and structural parts of the various factors. Declaring partial unreliability due to inefficiency, especially for modern standards, of the forecasts made.

[(12)] The number of observations is shown in the annex

However, to mend for it, the document proposes to build evaluation and verification centers[13], offices gathering and dissecting data retrieved, to improve and create a network to share the knowledge about the evaluation, monitoring, and managing of projects and to improve and fasten the cycle of data proofing. When writing about them in 2001, the creation of these groups was already taking place.

DPEF 2003-2005 (Ministry of Economy and Finance, 2002)

The forecasts made during this year, are complemented by the comparison with France and Germany, such juxtaposition, however, is only limited to the general economic condition. Noticeably, in the same 2002 document, the Italian government is said to be on track to improve its AA- rating to AAA rating in 2025, meanwhile France would drop to junk level and Germany to A+. This unjustified forecast may be a symptom of a flawed model overall.

In the paper, forecasts are made with the Multivariate Regression Analysis (VAR) combined with the Hodrick-Prescott model to account also for very short-term components.

DPEF 2006-2008 (Ministry of Economy and Finance, 2005)

In 2005, the Eurostat and the European Commission invited all European member states to be clearer when describing how the national debt is calculated, Italy in response, created a dedicated focus group to define the most efficient and time-saving ways.

A.2.2 Forecasts of the DFP

The DFP, which has only been published in 2010, was complemented by a methodological note, a document describing the methods displayed in the main paper.

[13] Known as *nuclei di valutazione e verifica*

Methodological note DFP 2011-2013 (Ministry of Economy and Finance, 2010)

The methodology described for generating forecasts, which for the most part will be adopted in the following documents, is divided into four main categories:

1- Rudimentary hypotheses for the outline of the macroeconomic framework[14]

The macroeconomic variables are the result of the forecast process utilized by the State General Accounting Department (RGS Department)[15] which utilizes the Judgment Based Model (JBM). This model considers both econometric models and economic and qualitative analysis.

The forecast process comprises three phases: firstly, exogenous variables are defined, both international and public finance ones. Secondly, different models are applied to elaborate an economic overview and lastly, results are compared to yield a single coherent estimate, this variable will then be utilized as an input for the potential GDP and other forecasts.

JBM and another econometric model ITEM separately make their estimates. For the JBM, the cyclical statistical basis, which is the quarterly ISTAT national economic accounts, is the basis for preliminary and short-term estimates. Any available econometric model is also used to complement the information already considered. For medium-term evaluations, historical elasticity is considered. ITEM econometric quarterly model uses a statistical basis in line with the JBM but adds national accounting data regarding institutional sectors, monetary and financial variables. The model publishes quarterly economic accounts while considering high-frequency models such as the Bridge Models and mixed frequency models.

The results of the two models are compared and verified within the working group on forecasts and modeling set up at the Treasury Department. In addition to the comparison between the results of the two models, also a series of internal and

[14] So called macroeconomic framework in Italian is called the *quadro macroeconomico*

[15] Which in Italian is the *Dipartimento della Ragioneria Generale dello Stato*

consensus forecasts are considered. These are formulated by some independent Italian Institutes that carry out research and economic forecasts (CER, RUEF, and Prometeia) and by the major international bodies (OECD, IMF, and European Commission).

2- Criteria adopted to elaborate the forecasting framework for Public Administrations

The Consolidated Account[16] of the Public Administrations (PA) is made on the base of a set of rules called the SEC95, adopted with the European Council Regulation No. 2223/96 on the 25th of June 1996 and implemented for the first time in 1999.

The forecasts for the spending and revenue aggregates, of the net indebtedness and the debt are produced by the Ministry of Economy and Finance (MEF), while the results of the final balance of the PA and the estimates of the debt are made respectively by ISTAT and the Bank of Italy.

The forecasts for public finances are made starting by the macroeconomic framework. Then the Public Finance Model comprises all the information belonging to different qualified offices and elaborates all the cash related forecasts of all the bodies belonging to PA. The forecasts for tax revenues are adjusted according to appropriate coefficients to comply with the Stability Pact.

For the passive interests, the inputs used are the primary surplus of the state sector and the scenarios relative both to the structure of the interest rates and to the composition of portfolios of domestic government bonds.

Expenditures

The second part also explains how the main economic expenditure items are derived. Overall, the estimates are based on historical trends, on macroeconomic variables, and on regulatory mechanisms that manage them. About this, a relevant

[16] Know in Italian as the *Conto Consolidato*

component in making the predictions is the overview of the procedures made in the various sectors, these can cause acceleration or reduction of the related flows, also considered are the evaluations regarding the progress on the main government operations.

Passive interest

The analysis of this variable and its forecast is based on the estimates of future trends of the aggregates of primary cash requirements[17] of the PA, made by the RGS Department and by other bodies that have current accounts open at the Central State Treasury.

Interest forecasts are prepared by combining the cash criterion to estimate the dynamic of the overall need for these administrations over the years and therefore, the evolution of their overall debt stock, with economic competence (SEC95), to guarantee that the information necessary for forecasts of the net debt of the AP, complies with the definitions of the national accounts dictated by the excessive deficit procedure of the Maastricht Treaty.

The overall interest expense on government securities in the DFP estimates is also fueled, among the various items, by interest on securities issued in non-domestic markets, both in public and private placement form, by the hedging swaps of these securities (cross-currency swaps) and other existing positions in derivatives (interest rate swaps).

3- Evaluation of the stock of public debt and the output gap

The third main forecast category when it comes to the methodology used. Debt over GDP ratio is a crucial indicator, especially when evaluating convergence politics of the European Monetary Union.

[17] Which in Italian are called the aggregati di fabbisogno di cassa primario

Output gap and GDP

The methodology utilized for the output gap and structural balances is developed by the European Commission[18] and agreed upon the member states.

In the base hypothesis, the GDP is represented with a production curve with constant returns to scale, such as a Cobb-Douglas. And is analytically represented in the following way:

` (1) $Y_t = \mathrm{L}_t^{\alpha} \times K_t^{1-\alpha} \times TFP_t$

The GDP is expressed in real terms, L is labour, K is capital, and α is labour elasticity. By assuming constant returns to scale, α can be estimated straight from the wage share. Assuming the same Cobb-Douglas characteristics for all European countries, then α is taken as the average value observed in the EU (which is 0.65), estimated on available data ranging from 1960 to 2003.

TFP is the Total Factor Productivity of economic growth. The previously stated assumption includes that the technological process spreads through both factors of production, capital, and labour:

` (2) $TFP_t = (\mathrm{E}_L^{\alpha} E_K^{1-\alpha}) \times (\mathrm{U}_L^{\alpha} U_K^{1-\alpha})$

This formula shows the efficiency of both the input of production labour (E_L), of capital (E_L) and of their degree of use ($U_L U_K$).

To move from the level of real GDP to that of potential GDP[19], it is necessary to obtain one estimate of the potential use or trend level of the individual production

[18] For further insights go to D'Auria F., C. Denis, K. Havik, K. McMorrow, C. Planas, R. Raciboski, W. Roger and A. Rossi, 2010, The production function methodology for calculating potential growth rates and output gaps, European Economy, Economic Paper, n. 420.

[19] The methodology for the potential product and the output gap is based on the estimate of a production function. For more information see Denis, C (2006). "Calculating potential growth rates and output gaps – A revised production function approach". European Economy, European Commission, Directorate-General for Economic and Financial Affairs; Economic Paper No.247

factors (labour, capital, and total productivity). The time series used for the calculation of potential GDP cover, as a rule, the period from 1960 to the last year of the forecast horizon considered. Generally, the trend component of the individual production inputs is obtained through simple filtering procedures using univariate statistical methods such as, for example, the Hodrick and Prescott (HP). However, one of the limitations of the HP filter lies in the so-called endpoint bias, or in the tendency to overestimate the weight of the observations that are at the beginning or end of the sample. Therefore, the methodology of the European Commission provides that the individual series underlying the estimation procedure are extrapolated beyond the time horizon of the forecast for 6 years. Statistical filters are applied to the original series and on the corresponding medium-term extension.

- Changes to methodology made in July 2010

The European Commission recognized some advantages stemming from the implementation of the new methodology based on the bivariate KF when compared to the HP one:

A. a lower number of revisions needed;
B. more realistic results for the short and medium-term. The Kalman filter uses the information on the economic cycle represented by the variable of the capacity used to extract the cyclical component from the TFP;
C. the Kalman filter does not produce biased estimates for the last period of the sample. The KF uses efficiently economic information on the cyclical component of TFP to produce more accurate predictions. The HP filter, on the contrary, does not allow to use of information deriving from other variables and is subject to the end point bias.

Evolution of the public debt stock

The public debt forecasts of the Public Administrations and its subsectors, starting from the data published in the official bulletins of the Bank of Italy regarding the last year, are made by the Treasury Department in the following ways:

A. for Central Administrations, the RGS Department prepares estimates of the central government's cash requirements for each year of the reference period. These are based primarily on the state sector's cash requirements - which also include estimates of the state sector's interest expenditure, but also consider other central government departments other than the state. To estimate the overall annual change in debt, starting with the cash requirements of central governments, the Treasury Department prepares a plan for hedging through debt issues in government securities, also considering cash management needs. Based on the securities that are expected to be issued each year, an estimate of the impact on debt is made. This evaluation considers the spreads of the medium and long-term securities issued (calculated using the difference between the net proceeds from the issue and the nominal value of the debt), and the expected revaluation of the debt due to inflation (deriving from the presence of inflation-linked securities among government bonds).

B. for Social Security Institutions, the estimate is based on the forecasts of their needs in the various years provided by the RGS Department;

C. for Local Administrations, the estimate is based on the forecasts of their needs in the various years provided by the Department of RGS. To arrive at the forecasts of the overall debt of the Public Administrations, an aggregation of the subsectors mentioned above is carried out considering the resulting consolidations.

4. Thematic analysis of insights related to the government budget. The government budget forecasts of the expenditures are analyzed to derive the basic economic tendency forecasts which directly impacts the PA.

Forecast expenditure of the state budget

To formulate the spending forecasts, it is necessary to have the results of the last report at hand. If they are not readily available, an estimate of such results must be made based on the most recent data available on final appropriations, commitments, payments, and residuals[(20)] up to the first of January of the year in which the forecasts are made, while also estimating the left behind to the 31st of December of the year ahead.

Whilst making the forecast, also the historical outcomes of the government budget are considered. It is essential to compile a historical overview. To do so, all subordinate funds are distinguished, taking into consideration the overviews carried out in previous years, the most updated data coming from the current year's monitoring, and the indications coming from the PA.

Consequently, all financial effects which are not included in the initial balance, possibly because they derive from interventions following the approval of the budget law, are considered.

Finally, the accounting framework for entry redeployments[(21)] (D.P.R. No. 469/1999) is considered. When performing the preliminary evaluation, recurrent events that, by nature, are to be accounted for as a statistically significant historical criterion are also accounted for. All considerations and classifications made are then combined with the initial estimates, thus obtaining the definitive forecasts.

Remarks of the DFP

With the revision of the Stability Pact originally made in 2005, the Medium Term Objective (MTO) that every European member state must follow is monitored through the structural budget surplus. [(22)] The Stability and Growth Pact, furthermore,

[(20)] Which in Italian are *stanziamenti definitivi, impegni, pagamenti e residui*

[(21)] Which in Italian is the *quadro contabile delle riassegnazioni di entrata*

[(22)] Which in Italian is the "*saldo strutturale di bilancio"*

says that such an objective must be reached through a reduction of the structural budget balance of at least 0.5% every year.

A.2.3 Forecasts of the DEF

Currently used as the official document of economic and financial forecasting, the DEF was introduced in 2011 due to the necessity to have a more standardized forecasting methodology across all European member states.

DEF 2011 (Ministry of Economy and Finance, 2011)

Despite using similar methodology compared to the DFP, the first document issued following Law No. 39 of the 7th of April 2011, presents some key differences.

The accounts series were adjourned, changing from the ATECO 2002 to the ATECO 2007. The balance indicators were adjusted according to the ECOFIN Council for the potential PIL and the output gap.

The commitment to aim for the medium-term objective (MTO) as designated by the European Commission is reiterated in the DEF. Italy, among the models available, chose a front-loading parameter with a degree of the cost of aging equal to 33%.

Methodological Note DEF 2011 (Ministry of Economy and Finance, 2011)

The changes made to the Law of Accounting and Public Finance (n. 196/09) are aimed at including the new national regulations and rules about the European Semester, adjourning the timing on the forecasting cycle, and introducing novelties on the contents of the financial and economic forecasting.

Comma 4, art. 10, states that the DEF must be complemented by a methodological note describing the methodology of the forecast.

The first quarter of the methodological note is, as for the DFP, about the outline of the macroeconomic overview. The main process applied has been and still is the *Judgment Based Model* (JBM). This process considers both econometric models and qualitative economic analysis, and is organized in three parts:

The first step is still about elaborating exogenous international variables and public finance ones. Updates regarding this step are about supplementary provisions of the Code of Conduct for the definition and contents of the Stability Program. New criteria require member countries to implement a new technical hypothesis for the exogenous variables. The 2011 DEF, about the exogenous variables, utilizes the most adjourned estimates of the main international organizations such as OCSE and FMI.

The second step is about elaborating on the national macroeconomic framework. Using the JBM for the statistical database, represented by the quarterly ISTAT economic accounts, a base for the short-term estimates and the preliminary estimates is made. In the case of structural interventions, the ITEM analysis used is coupled with the ones deriving from the QUEST III model. The QUEST III model, which in 2011 was the most recent version of the DSGE (*Dynamic Stochastic General Equilibrium)*, is made by the European Commission and adopted by the Italian Treasury Department. This analysis constitutes a solid base to understand how the economy answers to shocks and policy interventions.

The final step defines the macroeconomic framework: the results of the models are then reviewed and analyzed. Also, consensus forecasts and independent research are kept into consideration, just like in the DPEF.

Methodological Note DEF 2013 (Ministry of Economy and Finance, 2013)

This document presents the same methodology of one of the previous years, except for the section on the thematic analysis, the fourth quarter of the document. This section bases all its forecasts for the factors impacting the economic accounts of the PA on the analysis of the government balance expenditures forecast.

Also, the methodological note describes two of the key bodies that influence the government forecasting process:

FSC (Fondo per lo Sviluppo e la Coesione), previously known as the FAS, through the legislative decree No. 88 of the 31st of May 2011[(23)], changed its scope to give programmatic and financial consistency to all the additional nationally financed interventions aimed at the economic rebalancing of the country.

CIPE (Comitato Interministeriale per la Programmazione Economica) fosters cooperation on forecasting and national economic policies (legislative decree No. 430/1997). The CIPE defines the guidelines for these areas, framing the primary objectives for economic development and outlining the actions to take to achieve such goals.

Article 61 of Law No. 289/2002 establishes that the FSC must be managed according to specific resolutions of the CIPE, which are submitted for preventive control of the Court of Auditors (Dipartimento per la Programmazione e il Coordinamento della Politica Economica, 2020).

The FSC methodological procedure can be described as follows:

A. Verification of the available resources, both with regards to the supposed residuals of 2013 and concerning the resources entered in the 2013-2015 multi-year budget, considering also the changes laid down in current legislation to take into account the uses of the rules introduced during the course 2012 and the further changes made with the stability law for 2013.

B. Analyzing changes in the budget incurred during the year of forecast formulation. A distinction is made between changes in the reductions of the

[(23)] About provisions for additional resources and special interventions for the removal of economic and social imbalances

FSC, either statutory or due to CIPE resolutions (according to art. 33 paragraphs 2 and 3 of Law no. 183/2011);

C. Assessment of the impacts on the spending trends expected from 2013, both in terms of competence and cash, about the budget changes to be implemented in the application of specific hedging rules and CIPE resolutions. In particular, assessments were carried out on needs of special relevance, concerning their peculiarity and nature (for example, interventions for the post-earthquake reconstructions in Abruzzo, service, and programme contracts for public companies, interventions provided by OPCM and intended for the management of emergencies and environmental remediation, large events, and large public works);

D. Assessment of the additional budget available for national programming, net of regional programming (the uses of which, being bound by compliance with the internal stability pact, do not impact on net debt) and of the aforementioned most relevant needs, taking into account an overall ratio in terms of usage and competence for each year averaged to one third on an annual basis, in line with the estimates made in the context of the technical reports. In addition to this analysis, the residual portions of expenses are included, for which the annuities have not yet been identified by the CIPE, with an impact, on average, lower by a third than the ratio;

E. An economic analysis, by expenditure type, on the foreseeable allocations of resources entered in the residual account, with reference only to the changes selected to provide for the allocation of the Fund, in accordance with CIPE resolutions (unless mentioned in the application guidelines of specific norms).

Methodological Note DEF 2014 (Ministry of Economy and Finance, 2014)

In 2014, the DEF methodological changes reflect the changes happening at a European level. Aiming to strengthen the governance through coordinated policies across all member states, policies started in 2011 with the introduction of the *Six Pack,* which includes the PSC (Stability and Growth Pact), followed by the Treaty of Coordination, followed the *Fiscal Impact* and the *Two-Pack* (No. 472/2013 and No. 473/2013) about the monitoring of balance policies.

Moreover, in addition to the ratification of the Fiscal Compact, in July 2012, the national law complied with the new European law No. 1/2012 and No. 243/2012 enacting the main budget balance.

With regard to the macroeconomic outlook forecast, the DEF 2014 introduces some changes, especially in the second step of the *JBM* assessment process, defining the national macroeconomic framework. In this step, the ITEM model used is complemented not only by the QUEST III model but also by IGEM. This last model is a medium-scale model designed specifically for the Italian economy.

Besides, the Department of Treasury developed two micro-simulation models to be tested and refined, a statistical model called ITAxSIM and a dynamic micro-simulation model called T-Dymm (*Treasury's Dynamic Microsimulation Model).*

Methodological Note DEF 2015 (Ministry of Economy and Finance, 2015)

In the 2015 DEF, some relevant changes are made on the calculation of the structural budget balance. There have been, in fact, adjustments to the Phillips curve introduced by the *Output Gap Working Groups - OGWG*. Changes have been made also to the parameters representing the specific measures of sensitivity. [(24)]

[(24)] Those that measure the responsiveness to the economic cycle for each income and expenditure item, with an impact on the structural entries and outputs

Methodological Note DEF 2016 (Ministry of Economy and Finance, 2016)

Regarding the benchmark macroeconomic forecasts, a whole new section describing micro and macro models is now introduced, the following models, which have already been mentioned, are now going to be explained: (ITEM, IGEM, QUEST III, ORANI-IT&TERMITY, MACGEM-IT)

1) **ITEM**: The Italian Treasury Econometric Model, developed and utilized by the Treasury Department of the Ministry of Economy and Finance, describes the behaviour of the main aggregates of the Italian economy at a macro level. It is a medium-sized model, it includes 371 variables, 247 of which are endogenous. It is based on 36 behavioural equations and 211 identities. It is used both for forecast and for simulations regarding potential new policies. In 2016 an important revision of the model was made to comply with the new European Account system (SED 2010) and to consider a more recent estimate sample.

2) **IGEM** – Italian General Equilibrium Model
   It is a general equilibrium model conceived for the Italian economy, also medium-sized. It is based on an explicit micro foundation and can be used both to study how the economy reacts to temporary shocks and to make long term analyses needed to study structural reforms. IGEM complies with all the requirements for being a neo-Keynesian model, such as the presence of real and nominal rigidities, but it is characterized in particular for the extension and adaptation to the Italian labour market, where strong homogenous entities, both contractual and professional, coexist. The flexibility with which the model has been created grants the ability to simulate a vast spectrum of economic policies.

3) **QUEST III** - Italy
   This general equilibrium model, not only studies how the economy reacts to shocks, but also how it reacts to policy interventions. The version adopted is the one developed by the European Commission (DG ECFIN) for the quantitative evaluation of economic policies updated to consider also endogenous growth.

This model is usually used to analyze policies that foster competition and reforms aimed at improving human capital.

4) **ORANI-IT & TERMITI**

ORAN-IT is a Computable General Equilibrium (CGE) model for the Italian economy. It is a national, multi-sector model with a 62 product and 63 production branches.

The model is composed of a theoretical neoclassical framework, and from a database containing data about the national accounts that emulate the economy and fits into the theoretical structure of the model through equilibrium conditions.

This model comprises also an elaborate fiscal extension, that increases the analytical capacities of the model. This extension includes an *economy-wide framework* that replicates the VAT through a complex system of equations.

This extension makes ORANI-It a powerful policy analysis model that can also extend on a multiregional dynamic level, making it suitable for regional analysis (TERMITI).

5) **MACGEM-IT** – The new CGE model for the Italian economy

Developed by the Directorate I of the Treasury Department. MACGEM-IT is also a CGE model and is used to evaluate the unbundled effects, direct and indirect, of the fiscal policies and the hypothetical reform scenarios.

The 2016 version is static, disaggregated, multi-output, and multi-input. Because of this, the model is best used while estimating sector-specific policies, or policies for which it can be useful to study how the results affect both a specific sector and the economy.

Methodological Note DEF 2017 (Ministry of Economy and Finance, 2017)

In 2017, the methodology regarding the macroeconomic estimates of the DEF changed. The changes, limited to the IGEM model, will now be described:

In the IGEM-PA variation, public capital can increase private sector productivity. The public sector can, in turn, use goods and capital combined with public employment for the production and supply of goods and services. This extension answers the need to quantify the macroeconomic impact of complex scenarios. IGEM-PA also allows us to thoroughly study fiscal multipliers.

fIGEM, another IGEM variation, allows us to consider factors related to the conditions of offers of credit to businesses. This extension permits the evaluation of the impact of structural reforms and fiscal interventions considering different credit alternatives. Thanks to fIGEM it is also possible to study reforms aimed at the reduction of inefficiencies in the credit market.

Methodological Note DEF 2018 (Ministry of Economy and Finance, 2018)

This year, the only change made to the macroeconomic models is the addition of a validation phase of the macroeconomic forecasts by the Parliamentary Budget Office (UPB).

This process is implemented following the EU Regulation n. 473/2013[25] and the Law of the 24th of December 2012 No. 243. 2012 n. 243. [26] The revision process began after the agreement reached on September 15th, 2014 between the Ministry of Economy and Finance and the Parliamentary Budget Office.

The agreement introduced the following steps to be observed for each forecast being made:

A. transmission from the Ministry of Economy and Finance (MEF) to the UPB of the macroeconomic framework; which may propose changes;

[25] Regulation (UE) N. 473/2013 of the European Council and Parliament of the 21st of May 2013 about common provisions for the monitoring an evaluation of programmatic balance documents and for the correction of excessive deficits among the European member states. The regulation requires the validation of the Stability Programme forecasts by an independent national body

[26] Law of the 24th of December 2012, No. 243 about "Disposizioni per l'attuazione del principio del pareggio di bilancio ai sensi dell'articolo 81, sesto comma, della Costituzione"

B. the eventual answer of the new version of the framework by the MEF to the UPB;
C. final validation by the UPB which can be both positive and negative.

B. Hyperlinks to data sets for Italy and other institutions

Outcome data

Italy: Outcome data retrieved by the official Italian Institute of Statistics ISTAT for the exception of the real GDP growth which was retrieved from the World Bank.

France: Data retrieved from Statista for the 2001-2019 period. Retrieved from https://fr.statista.com/statistiques/479446/evolution-annuelle-du-pib-en-volume-france/

Germany: 2001-2018 retrieved from the World Bank, 2019 from the European Commission

Spain: 2009-2018 data retrieved from the World Bank at https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=ES&start=2001 2019 data retrieved from Statista at https://www.statista.com/statistics/263610/gross-domestic-product-gdp-growth-in-spain/

Forecast data

Italy: Data for forecasts made in the years 2000-2009 is retrieved from the DPEF at http://www.dt.mef.gov.it/it/attivita_istituzionali/analisi_programmazione_economico_finanziaria/documenti_programmatici/sezione1/dpef.html

Data for forecasts made in 2010 retrieved from the DFP at http://www.dt.mef.gov.it/modules/documenti_it/analisi_progammazione/documenti_programmatici/DFP_xissn_on-linex_PROTETTO.pdf

Data from 2011 until 2018 is retrieved from the DEF at http://www.mef.gov.it/documenti-pubblicazioni/doc-finanza-pubblica/index.html#cont4

France: Full window of time (2000-2018) retrieved from the French Annual Stability and Convergence Programs, not available values for 2001, 2003, and 2004. Retrieved from: https://www.performance-publique.budget.gouv.fr/ressources-documentaires/textes-documents-reference/programmes-stabilite-francais#.Xs0ZY2gzbcd

Germany: Full forecasting period (2000-2018) retrieved from the German Stability and Convergence Programs, not available value for 2006. Documents retrieved from: https://www.bundesfinanzministerium.de/Web/DE/Themen/Oeffentliche_Finanzen/Stabilitaetspolitik/Stabilitaetsberichte/Deutsche-Stabilitaetsprogramme/deutsche-stabilitaetsprogramme.html;jsessionid=9C940633E2195CCCC1C75B8E7CAE1644.delivery2-master?gts=101278_list%253DdateOfIssue_dt%252Bdesc&gtp=%252681195201-53b3-4359-8158-a29b47dd20d7_list%253D2 .

Spain: All forecasting data retrieved from the Bank of Spain at https://www.bde.es/bde/en/secciones/informes/boletines/relac/Boletin_Economic/Informes_de_proy/index2008.html

European Commission: 2000-2018 data retrieved from the Spring European Economic Forecast. From 2000 to 2015 data retrieved from https://ec.europa.eu/economy_finance/publications/european_economy/forecasts_en.htm

2016 retrieved from https://ec.europa.eu/info/sites/info/files/ip025_en.pdf

2017 retrieved from https://ec.europa.eu/info/sites/info/files/ip053_en_1.pdf

2018 retrieved from https://ec.europa.eu/info/sites/info/files/economy-finance/ip077_en.pdf

Consensus Economics forecasts have been retrieved from the Economic Bulletin published in July by the Bank of Italy

C. Tables[27] [28][29]

Table A1:
One-year forecast average error of countries and institutions (real GDP growth)

2001-2019	Italy	France	Germany	Spain	Commission	Consensus
Average error	-1.38	-1.00	-0.29	n/a	-1.27	n/a
2010-2019	Italy	France	Germany	Spain	Commission	Consensus
Average error	-0.86	-0.52	0.29	-0.30	-0.75	-0.56

Table A2:
One-year forecast average errors for Italy – 2001-2019

2001-2019	GDP growth	Passive interests	Unemployment rate	Primary surplus	Net indebtedness	Public debt	Nominal GDP
Programmatic	-1.38	0.21	0.16	-1.37	-0.99	3.42	9.40
Tendency	-0.92	-0.23	-0.39	-0.72	-0.40	2.40	-69.41

[27] Negative average values are due to average overestimation and positive ones are because of average underestimation

[28] The better the model, the closer the programmatic errors should be closer to zero and the tendency ones equal to the difference between the programmatic and tendency forecasts

[29] Any minor computation discrepancy is to be attributed to approximation errors

Table A3:
One-year ahead forecasts for Italy – 2001-2019

	GDP growth	Passive interests	Unemployment rate	Primary surplus	Net indebtedness	Public debt	Nominal GDP
Average	-1.38	0.21	0.16	-1.37	-0.99	3.42	9.40
SD	1.94	2.45	1.24	1.40	1.35	4.25	45.05
SE	0.46	0.59	0.29	0.32	0.31	1.00	12.00
Serial Corr.	0.09	-0.19	0.51	0.20	0.19	0.22	0.01
MAD	1.67	0.96	0.92	1.56	1.36	3.95	35.04
RMSE	2.33	2.39	1.21	1.93	1.65	5.36	44.29
MAPE	193.94	22.11	10.27	136.37	47.58	3.24	2.26

Table A4:
One-year forecast errors of governments (real GDP growth; 2001-2019)

2001-2019	Average Error	Standard deviation	Serial correlation	MAD	RMSE	MAPE
Italy	-1.38	1.94	0.09	1.67	2.33	193.94
France	-1.00	1.42	0.26	1.11	1.70	151.47
Germany	-0.29	1.83	-0.16	1.31	1.85	132.71
Spain	n/a	n/a	n/a	n/a	n/a	n/a

Table A5:
One-year forecast errors of governments (real GDP growth; 2010-2019)

2010-2019	Average Error	Standard deviation	Serial correlation	MAD	RMSE	MAPE
Italy	-0.86	1.77	0.91	1.34	1.87	134.05
France	-0.52	0.74	0.21	0.70	0.87	120.04
Germany	0.29	1.35	0.33	1.13	1.38	105.34
Spain	-0.30	1.87	0.42	1.31	1.80	142.42

Table A6:
Italy vs other countries: comparison of forecast errors (real GDP growth)

2001-2019	$DD_{I,n}$	$\Delta\text{RMSE}_{I,n}$	2010-2019	$DD_{I,n}$	$\Delta\text{RMSE}_{I,n}$
France	-0.46	0.64	**France**	-0.29	1.00
Germany	-1.12	0.48	**Germany**	-1.14	0.49
Spain	n/a	n/a	**Spain**	-0.73	0.07

Table A7:
One-year forecast errors of institutions (real GDP growth; 2001-2019)

2001-2019	Average Error	Standard deviation	Serial correlation	MAD	RMSE	MAPE
Italy	-1.38	1.94	0.09	1.67	2.33	193.94
Commission	-1.27	1.84	0.00	1.54	2.20	238.65
Consensus	n/a	n/a	n/a	n/a	n/a	n/a

Table A8:
One-year forecast errors of institutions (real GDP growth; 2010-2019)

2010-2019	Average Error	Standard deviation	Serial correlation	MAD	RMSE	MAPE
Italy	-0.86	1.77	0.91	1.34	1.87	134.05
Commission	-0.75	1.67	0.34	1.20	1.74	73.41
Consensus	-0.56	1.45	0.23	0.99	1.55	57.28

Table A9:
Real GDP growth (% of GDP) – Italy – 2001-2019

Prog.	One-year	Two-year	Three-year	Average of the 3 periods
Average	-1.38	-1.19	-1.74	-1.44
SD	1.94	1.20	1.96	1.70
SE	0.46	0.30	0.49	0.41
Serial Corr.	0.09	0.32	0.02	0.14
MAD	1.67	1.29	1.74	1.57
RMSE	2.33	1.66	2.57	2.19
MAPE	193.94	176.30	257.06	209.10
Tend.				
Average	-0.92	-2.06	-1.73	-1.57
SD	1.16	2.46	2.27	1.97
Serial Corr.	0.17	0.53	-0.72	0.00
Obs Prog.	17	15	15	
Obs Tend.	10	9	8	

Table A10:
Passive interests (% of GDP) – Italy – 2001-2019

Prog.	One-year	Two-year	Three-year	Average of the 3 periods
Average	0.21	-0.67	-0.43	-0.30
SD	2.45	0.50	1.58	1.51
SE	0.59	0.13	0.39	0.37
Serial Corr.	-0.19	0.42	0.01	0.08
MAD	0.96	0.69	1.09	0.91
RMSE	2.39	0.83	1.58	1.60
MAPE	22.11	15.88	24.47	20.82
Tend.				
Average	-0.23	-0.29	-0.39	-0.30
SD	0.31	0.44	0.31	0.35
Serial Corr.	0.25	-0.04	0.30	0.17
Obs Prog.	16	15	15	
Obs Tend.	9	8	7	

Table A11:
Unemployment rate – Italy – 2001-2019

Prog.	One-year	Two-year	Three-year	Average of the 3 periods
Average	0.16	0.45	0.89	0.50
SD	1.24	1.74	1.90	1.63
SE	0.29	0.43	0.49	0.41
Serial Corr.	0.51	0.58	0.74	0.61
MAD	0.92	1.19	1.53	1.21
RMSE	1.21	1.74	2.04	1.66
MAPE	10.27	12.92	16.28	13.16
Tend.				
Average	-0.39	-0.13	0.09	-0.14
SD	0.55	1.16	1.89	1.20
Serial Corr.	-0.50	-0.33	0.86	0.01
Obs Prog.	17	15	14	
Obs Tend.	11	10	9	

Table A12:
Primary surplus (% of GDP) – Italy – 2001-2019

Prog.	One-year	Two-year	Three-year	Average of the 3 periods
Average	-1.37	-2.31	-2.98	-2.22
SD	1.40	1.68	1.59	1.55
SE	0.32	0.39	0.39	0.37
Serial Corr.	0.20	0.21	0.26	0.22
MAD	1.56	2.49	2.98	2.34
RMSE	1.93	2.82	3.35	2.70
MAPE	136.37	221.99	303.32	220.56
Tend.				
Average	-0,72	-1.12	-1.43	-1.09
SD	1.41	2.00	2.07	1.83
Serial Corr.	0.30	0.37	0.46	0.37
Obs Prog.	18	17	16	
Obs Tend.	14	13	12	

Table A13:
Net indebtedness (% of GDP) – Italy – 2001-2019

Prog.	One-year	Two-year	Three-year	Average of the 3 periods
Average	-0.99	-1.74	-2.34	-1.69
SD	1.35	1.39	1.36	1.37
SE	0.31	0.33	0.33	0.32
Serial Corr.	0.19	0.22	0.24	0.22
MAD	1.36	1.96	2.39	1.90
RMSE	1.65	2.20	2.69	2.18
MAPE	47.58	66.71	77.20	63.83
Tend.				
Average	-0.40	-0.65	-0.88	-0.64
SD	1.24	1.68	1.73	1.55
Serial Corr.	0.28	0.36	0.42	0.35
Obs Prog.	17	16	15	
Obs Tend.	16	15	15	

Table A14:
Public debt over GDP – Italy – 2001-2019

Prog.	One-year	Two-year	Three-year	Average of the 3 periods
Average	3.42	7.28	11.92	7.54
SD	4.25	6.32	6.74	5.77
SE	1.00	1.53	1.68	1.40
Serial Corr.	0.22	0.35	0.61	0.39
MAD	3.95	7.55	11.92	7.81
RMSE	5.36	9.51	13.58	9.48
MAPE	3.24	6.15	9.73	6.37
Tend.				
Average	2.40	6.19	11.34	6.64
SD	5.18	7.91	8.18	7.09
Serial Corr.	-0.13	-0.35	-0.50	-0.33
Obs Prog.	19	18	17	
Obs Tend.	9	8	7	

Table A15:
Nominal GDP – Italy – 2001-2019

Prog.	One-year	Two-year	Three-year	Average of the 3 periods
Average	9.40	-9.55	-34.06	-11.40
SD	45.05	51.98	57.92	51.65
SE	12.00	14.37	16.65	14.34
Serial Corr.	0.01	0.23	0.76	0.33
MAD	35.04	40.47	53.13	42.88
RMSE	44.29	50.68	64.89	53.28
MAPE	2.26	2.57	3.29	2.71
Tend.				
Average	-69.41	-103.74	-134.56	-102.57
SD	323.66	361.64	366.67	350.66
Serial Corr.	-0.01	-0.11	-0.11	-0.08
Obs Prog.	13	12	11	
Obs Tend.	14	12	12	

Table A16:
One-year ahead programmatic forecasts for Italy – 2010-2019

	GDP growth	Passive interests	Unemployment rate	Primary surplus	Net indebtedness	Public debt	Nominal GDP
Average	-0.86	0.69	0.58	-1.04	-0.58	4.79	0.75
SD	1.77	3.26	1.28	1.05	0.99	3.20	31.90
SE	0.63	1.09	0.43	0.35	0.33	1.13	12.06
Serial Corr.	0.91	0.47	0.62	0.13	-0.17	0.46	0.01
MAD	1.34	1.40	0.93	1.13	0.93	4.79	21,87
RMSE	1.87	3.15	1.33	1.44	1.10	5.65	29.55
MAPE	134.05	140.00	93.33	113.33	93.33	478.75	21.87

Table A17:
One-year forecast errors of European Commission (unemployment)

	Average Error	Standard deviation	Serial correlation	MAD	RMSE	MAPE
2001-2019	-0.17	1.19	0.40	0.96	1.17	10.85
2010-2019	0.14	1.29	0.48	0.94	1.23	8.74

Sources

Accounting and Public Finance Law No.196 of 31 December 2009. Retrieved from: http://www.dt.mef.gov.it/export/sites/sitodt/modules/documenti_it/analisi_progammazione/analisi_programmazione_economico/Documento3.pdf

Agresti, A. (2013). "Statistics the art and science of learning from data". 3rd Edition, Pearson.

Ashiya, M. (2007). "Forecast accuracy of the Japanese government: Its year-ahead GDP forecast is too optimistic". Japan and the World Economy.

Bahrke, J. (2015). "Stability and Growth Pact: Commission issues guidance to encourage structural reforms and investment". Retrieved from: https://ec.europa.eu/commission/presscorner/detail/it/IP_15_3220)

Brunel, C. (2011). "Over- Optimism in Official Budget Agencies' Forecasts". National Bureau of Economics. Retrieved from: https://www.nber.org/digest/nov11/w17239.html

Callen, T. (2020). "Gross Domestic Product: An Economy's All". Retrieved from: https://www.imf.org/external/pubs/ft/fandd/basics/gdp.htm

Chang, C. (2009). "How Accurate are Government Forecast of Economic Fundamentals?". Erasmus University Rotterdam, Econometric Institute, Econometric Institute Report. Retrieved from: https://www.researchgate.net/publication/46434347_How_Accurate_are_Government_Forecast_of_Economic_Fundamentals

Decree of the President of the Republic (D.P.R.) of the 10th of November 1999. Documentazione Economica e Finanziaria. Retrieved from: https://def.finanze.it/DocTribFrontend/getAttoNormativoDetail.do?ACTION=getSommario&id={FAE4F41F-2035-4AD0-B924-4CF6BBBFC770}

Denis, C. (2006). "Calculating potential growth rates and output gaps – A revised production function approach". European Economy, European Commission, Directorate-General for Economic and Financial Affairs; Economic Paper No.247

Diebold, F. (1995). "Comparing Predictive Accuracy. Journal of Business and Economic Statistics". 13, 253-265 Retrieved from: https://www.sas.upenn.edu/~fdiebold/papers/paper68/pa.dm.pdf

Dipartimento per la programmazione e il coordinamento della politica economica, (2020). "Fondo per lo Sviluppo e la Coesione". Retrieved from: http://www.programmazioneeconomica.gov.it/2019/12/02/fondo-per-lo-sviluppo-e-la-coesione-2/

EU regulation of the European Council and Parliament No. 472 of the 21st of May 2013

EU regulation of the European Council and Parliament No. 473 of the 21st of May 2013

EUR - Lex (2020). "L'accesso al diritto dell'Unione europea". Retrieved from: https://eur-lex.europa.eu/summary/glossary/stability_growth_pact.html?locale=it

European Commission (2020). “About the economic forecasts”. Retrieved from: https://ec.europa.eu/info/business-economy-euro/economic-performance-and-forecasts/economic-forecasts/about-economic-forecasts_en

European Council (2020). “Semestre europeo: una guida alle norme e ai documenti principali”. Retrieved from: https://www.consilium.europa.eu/it/policies/european-semester/european-semester-key-rules-and-documents/

Fioramanti, M. (2016). “European Commission’s Forecasts Accuracy Revisited: Statistical Properties and Possible Causes of Forecast Errors. European Commission”. European Commission. Retrieved from: https://ec.europa.eu/info/sites/info/files/dp027_en.pdf

Frankel, J. (2012). “Over-optimism in Forecasts by Official Budget Agencies and Its Implications” (NBER working paper No. 17239). National Bureau of Economic Research. Retrieved from: https://www.nber.org/papers/w17239

Frankel, J. (2012). “Over-optimistic Official Forecasts in the Eurozone and Fiscal Rules” (NBER working paper No. 18283). National Bureau of Economic Research. Retrieved from: https://www.nber.org/papers/w18283.pdf

Gonzáles, L. (2012). “The accuracy of the European Commission’s forecasts re-examined”. Economic papers 476, European Commission. Retrieved from: https://ec.europa.eu/economy_finance/publications/economic_paper/2012/pdf/ecp476_en.pdf

Hallerberg, M. (2009). “Fiscal Governance in Europe”. Cambridge University press.

Hyndman (2006). "Another look at measures of forecast accuracy". International Journal of Forecasting. 22: 679–688

ISTAT (2020). “Unemployment rate”. Retrieved from: http://dati.istat.it/Index.aspx?DataSetCode=DCCV_TAXDISOCCU1#

Jonung, L. (2006). “Fiscal policy in the EU - are official output forecasts biased”. European Commission, Directorate-General for Economic and Financial Affairs

Law No.39 of the 7th of April 2011 Retrieved from: http://www.dt.mef.gov.it/export/sites/sitodt/modules/documenti_it/analisi_progammazione/documenti_programmatici/Legge_7_aprile_2011.pdf

Law No. 183 of the 12th of November 2011

Law No. 243 of the 24th of December 2012

Legislative Decree No. 430 of the 5th of December 1997

Legislative Decree No. 88 of the 31st of May 2011

Lucia A. (2014). “Central Bank Macroeconomic Forecasting during the Global Financial Crisis: The European Central Bank and Federal Reserve Bank of New York Experiences” (Working Paper SerieS No.1688). European Central Bank. Retrieved from: https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1688.pdf

Mankiw, N. (2015). “Macroeconomics” (9th ed.)

Marinheiro, C. (2010). “Fiscal sustainability and the accuracy of macroeconomic forecasts: do supranational forecasts rather than government forecasts make a difference?”. GEMF No.7. Retrieved from: https://pdfs.semanticscholar.org/1d49/271f58a6c178ae523b1208e248029a96725b.pdf

Ministry of Economy and Finance (2000 to 2009). “Economic and Financial Planning Document – DPEF”. Retrieved from: http://www.dt.mef.gov.it/en/attivita_istituzionali/analisi_programmazione_economico_finanziaria/documenti_programmatici/sezione1/dpef.html

Ministry of Economy and Finance (2010). “Methodological note”. About the DFP. Retrieved from: http://www.mef.gov.it/primo-piano/documenti/25275_nota-metodologica-alla-dfp-modellorgs_29sett2010-definitivaxmef.pdf

Ministry of Economy and Finance (2010). “Public Finance Decision – DFP”. Retrieved from: http://www.dt.mef.gov.it/en/attivita_istituzionali/analisi_programmazione_economico_finanziaria/documenti_programmatici/sezione2/dfp.html

Ministry of Economy and Finance, (2011 to 2019). “Economic and Financial Document – DEF”. Retrieved from: http://www.mef.gov.it/documenti-pubblicazioni/doc-finanza-pubblica/index.html#cont4

Ministry of Economy and Finance, (2011 to 2019). “Methodological note”. About the DEF. Retrieved from: http://www.mef.gov.it/documenti-pubblicazioni/doc-finanza-pubblica/index.html#cont4

Ministry of Economy and Finance, (2020). “Planned Inflation Rate”. http://www.dt.mef.gov.it/en/attivita_istituzionali/analisi_programmazione_economico_finanziaria/inflaz_programmata/index.html

Muñoz, L. (2018). “Forecast errors and forecast biases: Fiscal implications”. Autoridad Independiente de Responsabilidad Fiscal. Retrieved from: https://ec.europa.eu/info/sites/info/files/economy-finance/airef_forecast_biases.pdf

OECD (2014), “OECD forecasts during and after the financial crisis: A Post Mortem”, OECD Economics Department Policy Notes, No. 23 February 2014. Retrieved from: https://www.oecd.org/economy/outlook/OECD-Forecast-post-mortem-policy-note.pdf

OECD (2016), “Reference guide on ex post evaluation of competition agencies’ enforcement decisions”. Retrieved from: http://www.oecd.org/daf/competition/Ref-guide-expost-evaluation-2016web.pdf

Office for Budget Responsibility. (2017). “Evaluating forecast accuracy” (Briefing paper No.7). UK Treasury. Retrieved from: https://obr.uk/docs/dlm_uploads/BriefingPaper7WEB.pdf

Pesole, D. (2019). “Rafforzare l’avanzo primario, la possibile strategia del governo”. Il sole 24 ore. Retrieved from: https://www.ilsole24ore.com/art/rafforzare-l-avanzo-primario-strategia-governo-bruxelles-AC2MjNS

Provisions for the formation of the annual and multiannual state budget, Law No. 289 of the 27th of December 2002

Ray C. (1990). “Comparing Information in Forecasts from Econometric Models”. Retrieved from: https://pdfs.semanticscholar.org/7c31/d11844113619103644d878d2e4a43339d164.pdf

State general accounting department (2008). “I principali saldi di finanza pubblica: definizioni, utilizzo, raccordi”. Ministry of economics and finance. Servizio Studi. Retrieved from: http://www.rgs.mef.gov.it/_Documenti/VERSIONE-I/Comunicazione/Eventi/WORKSHOP--1/I-principali-saldi-di-finanza-pubblica-definizioni-utilizzo-raccordi-.pdf

Timmermann, A. (2006). “Chapter 4 Forecast Combinations”. Handbook of Economic Forecasting. 1. pp. 135–196. Retrieved from: https://www.sciencedirect.com/science/article/abs/pii/S1574070605010049?via%3Dihub