A Work Project, presented as part of the requirements for the Award of a Master's degree in Economics from the Nova School of Business and Economics.

THE PRUDENCE BIAS: ASSESSING INSTITUTIONAL FORECASTERS' ASYMMETRIC LOSS FUNCTIONS

Pedro Nuno Garcez Moreira de Sousa

Work project carried out under the supervision of:

Professor João B. Duarte

17-12-2021

The Prudence Bias: Assessing institutional forecasters' asymmetric loss functions

Pedro Nuno Garcez Moreira de Sousa

Abstract

The general posture towards forecasts is often one of scepticism: provided with the same information, predictions differ significantly. But the inefficient use of information might not be the single cause to blame. I assess the potential biased behaviour from fiscal and monetary authorities towards over- and underpredictions of key macroeconomic variables. I propose an estimation routine based on GMM to identify asymmetric loss functions and discuss the plausible underlying economic costs that shape the preferences of institutions and thus respective forecasts. I find that all institutions favour the overprediction of real output growth and monetary authorities prefer to underpredict inflation, whilst fiscal authorities tend to overpredict it.

Keywords: Macroeconometrics, Macroeconomic Forecasting, Loss Function, Asymmetric Preferences, Forecast Rationality.

Acknowledgements

Imagine how much harder Physics would be if electrons had feelings. - Richard Feynman

I thank my family for all the love and Luísa for all the tough love. I've always needed both. I thank Professor João B. Duarte for the guidance in this final chapter of my Master's degree. I thank Professor Filipe Grilo for the mentorship. To all my friends, I will see you in Fontelas.

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

1 Introduction

Economics has come a long way since its genesis. The seminal book "An Inquiry into the Nature and Causes of the Wealth of Nations" is almost two hundred and fifty years old. Economic science has evolved spectacularly since then, keeping up with the pace of technology as society advances, modernizing its tools, broadening its scope to other areas of knowledge while learning from them, "standing on the shoulders of giants". Despite the fact economists have been working on this craft for so long now, some of our predictions still miss spectacularly. Following the postulate of the rational expectations hypothesis, academia has argued for many decades that forecasters are not rational, in the sense they do not use information efficiently. But the most prominent economic institutions, those which perform forecasting exercises using long up-to-date series and state-of-the-art macroeconomic "hybrid"-models, fail predictions nonetheless. While it is true that no one possesses the Economic unified field theory and hence there is room for deviations from the actual values, a fundamental question remains to be further discussed: why do institutions that use virtually equivalent data sets produce significantly different forecasts depending on the nature of the variable under study? One may argue that among the most renowned institutions, there are still those that employ relatively more high-skilled individuals or those that receive relatively more funding and therefore achieve different results. But what if all relevant economic institutions provide rational predictions with the given data and there is still an observed bias? Is there a rational bias that accommodates the skewness of forecasts towards under- and overpredictions?

This study addresses the seemingly paradoxical dichotomy of central banks and executive bodies of the European and American landscapes presenting substantially different forecasts (both cross- and within-bloc), under rationality as the degree of asymmetry of the underlying forecaster's loss function varies. From a policymaker point of view, the following macroeconometric work might suggest that relatively more optimistic or pessimistic forecasts from different institutions do not necessarily mean disagreeing views on the issue at hand but rather that the institutions can share the same underlying forecasted value while attributing different degrees of conservativeness to it. Under- and overpredictions might have different weights on the forecaster's loss function, depending on both the nature of the institution (fiscal or monetary) and the macroeconomic variable under analysis (output growth or inflation). From the perspective of Macroeconomics and Political Economy, the study sheds light on the different aversion degrees institutions might present to similar deviations of the forecast errors from zero. Furthermore, I discuss some intrinsic economic incentives that potentially bias predictions, such as the asymmetric impacts on the real economy of under- and overpredictions. Overall, this study is relevant to better perceive the set of economic objectives driving institutions.

I build on the landmark work of Elliott, Timmermann, and Komunjer 2005 and include recent methodological contributions suggested by the loss function literature, for instance, in what regards to rationality testing, asymmetric loss estimation routine and data treatment. My study shares some of the concerns of the literature regarding potentially problematic assumptions of the benchmark model and scope of the original data and contributes to the ongoing debate on the adequate forecast horizon for such analyses and which vintages should one use as realized values. Concerning application, the requalified model compares the forecast rationality and the loss function parameter estimates of the different fiscal and monetary authorities, so it is possible to compare the results (geography- and objective-wise) and discuss prevailing behaviours observed cross-sample. To the best of my knowledge, there is no published work related to this literature that includes institutions of different geographic locations and economic goals to compare their forecasting behaviours. This comprehensive feature of the study allows us to explore some of the underlying economic mechanisms that otherwise could be overlooked by the literature.

Based on the results of my model, I find evidence for claiming that monetary authorities have incentives to underpredict inflation, whereas fiscal authorities have incentives to overpredict it, although the latter follow a symmetric path more closely. Intuitively, this means that monetary authorities perceive the overprediction of inflation as more costly than its underprediction, whereas fiscal authorities perceive the contrary. All institutions tend to overpredict output growth. I construct a narrative on the potential driving factors behind these phenomena. Importantly, the findings may impact the public policy field. First, to avoid discussing the potential relative optimism or pessimism purely on a subjective basis, but rather on the grounds of empirical- and theoretical-sound evidence; Second, to have a sense of what might be the "true" forecast, after adjusting for the presented bias. Concerning the impact on the macroeconomic field, this study is relevant for the set of theoretical arguments put forth to rationalize the observed biases, since the loss function literature typically focuses its efforts on improving the econometric framework while providing few contributions to the economic fundamentals.

This paper is organized as follows. Section 2 describes the data, advocates for the cross-regional comparability of the institutions, argues which data vintages should one use as realized values and discusses the consequences of considering different forecast horizons. Section 3 explains the benchmark model, the methodological issues raised by the literature and the requalified model. Section 4 presents the empirical results and interprets them. Section 5 discusses the impact of the findings and concludes.

2 Data

2.1 Framework

The datasets include time series concerning four institutions, two macroeconomic variables of interest and two forecast horizons. The institutional setting encompasses two monetary financial institutions (European Central Bank, Federal Reserve System), and two executive bodies (European Commission, Executive Office of the President of the United States). The forecasting exercises are conducted by the respective staffs, namely those of: the European Central Bank (ECB), the Federal Reserve Board (FED), the Directorate-General for Economic and Financial Affairs (ECFIN) and the United States Office of Management and Budget (OMB). The key macroeconomic variables under study are the real gross domestic product (GDP) and the inflation rate, because they are the traditional ones to assess economic development, outlook and risk to price stability. Depending on data availability and region, the inflation variable is either the Consumer Price Index (CPI), Harmonised Index of Consumer Prices (HICP), GNP deflator or GDP deflator (with preference for the latter whenever it is available). I evaluate the bias for two forecast horizons, short run (Figures 1 and 2) and medium run (Figures 3 and 4), due to the potentially different economic costs of failing predictions. However, note that the forecasts lose accuracy as their horizon increases, so the results for the medium run should be analysed with caution, as the deviations from an average forecast error of zero might be caused by both different economic costs and significant accuracy losses. Next, I briefly describe each dataset.

2.2 Data Description

I aggregate all the institutional forecast reports and data files available from 1967 to 2021 to build a thorough dataset. The ECB staff prepares quarterly reports on the Euro Area economy for the Governing Council meetings and shares them with the public through the Macroeconomic Projection Database (the forecasts available are those ranging from 1999:Q1 to 2020:Q3). The FED staff produces quarterly forecasts to present to the Board of Governors, two weeks before the Federal Open Market Committee meetings. These economic forecast reports are usually called Greenbooks and they are made available to the public with a lag of five years, hence the reports from 1967:Q1 to 2015:Q4 are available on the Reserve Bank of Philadelphia's website. To ensure comparability of results, we need to run our model on the same sample period for both monetary authorities, so I impose the samples to begin in 1999:Q1 and end in 2015:Q4.

The European fiscal authority has reports that date back to the time of the Commission of the European Economic Committee, before the Maastricht Treaty (1993) renamed it the European Commission. The forecasts made by the ECFIN present yearly frequency from 1973 to 2014 and quarterly frequency from 2015:Q1 to 2021:Q3. The American fiscal authority was formed in 1970 and has been producing annual forecasts since 1975. The OMB is the agency responsible for supervising the development and execution of the fiscal budget. This time, to ensure comparability of results, not only have I to impose the sample of the institutions to begin in 1975, but also to impose an upper limit in the year 2019 rather than in the most recent forecast report available. Due to the Covid-19 pandemic and the consequent economic and epidemiological uncertainties, it is not reasonable to include the latest set of observations in the analysis. I would be introducing noise in the model because institutional forecasters did not have enough time to incorporate into their forecasting models the cyclical behaviour and probability of emergence of different waves of the SARS-CoV-2 virus. Finally, all institutional projections are subject to judgemental adjustments to the output of both econometric and macroeconomic models and entail specific fiscal and monetary policy assumptions that are presented in the subsection 2.5. Further details concerning the data characteristics can be found in the summary table (Table 3) of the appendix.

2.3 Comparability Between Regions

The member-states of the European Union (EU) and the United States of America (US) are advanced economies that share a common heritage: the Anglo-Saxon institutional background that the US has makes the country comparable with nations from the rest of the Western world in what regards to socio-cultural attributes, maturity of institutions and overall economic stage of development. This common setting allows the EU and the US to have similar fiscal and monetary goals and policy toolkits. Otherwise, one could be discussing forecast biases from mature economies based on the objectives of fiscal and monetary policies from emerging markets' institutions. Moreover, the European and American short- and medium-term business cycles are significantly correlated (Choudhry, Hassan, and Shabi 2020), which facilitates comparisons. It is important to note that, as a consequence of this analysis, regions may face similar economic costs of under- and overpredictions. However, one should also bear in mind that while the US is a federation, the EU is arguably not one, so the conclusions one might draw from the data should not be binding, but rather suggestive, of potentially different economic costs (in case there is indeed any difference).

2.4 Realization Data

All datasets comprise the same elements: the forecast value and the respective realization for each quarter or year (data vintage). Concerning the high-frequency datasets, some authors state that, with a reasonable degree of confidence, trustworthy values are available two quarters after the quarter to which the data refer to (Faust and Wright 2009; Sinclair, Joutz, and Stekler 2010; Tulip 2009). The discussion, therefore, is not about when does one have access to the "true" values but rather whether one wants to use the first release of the data, a later revision of the data or the latest vintage available. Indeed, Melander, Sismanidis, Grenouilleau, et al. 2007 state that the selection of "outturn data" can influence the forecast errors' magnitude and interpretation.

Caunedo et al. 2020 argue that the choice of which vintage one should use depends on whether one perceives the objective of the forecaster as predicting the definitive value of a variable (the case in which we should use the latest vintage available), or as the value that both the forecaster and policymaker observe a few periods after the respective forecast (we should use either the first release of the data or a later revision of it). The authors consider the former as the final goal of the forecaster, even though there might occur some major methodological revisions over the years that change the backcasts, such as the base year used for GDP calculations – a fact that they acknowledge, arguing these data revisions are unpredictable and outside the scope of the forecasting problem of most agents.

However, a relevant feature of my work is precisely the forecaster making projections based on the information it possesses at the moment of each forecasting exercise. The estimates released first attract greater attention than later revisions (Keereman et al. 1999), as proved by the relatively swift reaction of central banks and executive bodies to adverse economic conjunctures. Thus, while I acknowledge that the goal of the forecasters is to predict the actual values, it is also true that they do not wait for the "final say" on some backcast values to perform other forecasts. They do the exercise with the information they have at each moment. Moreover, one should take into account that the forecasts produced do not suffer the same major data revisions the realized values do, hence the adoption of the latest vintage available would have some major repercussions on the empirical results due to the artificial skewness imposed on the forecast errors calculated at the present moment, such as rendering estimates of loss asymmetry significantly exaggerated (Wang and Lee 2014). I use an intermediate vintage: the two-quarters-behind backcasts for quarterly-frequency datasets and one-year-behind backcasts for the yearly-frequency ones.

2.5 Forecast Horizons

The intrinsic economic information the forecasts contain depends on the forecast horizon considered. For instance, the institutional forecaster might follow a New-Keynesian DSGE model with Keynesian features in the short run (due to the short-term stickiness of prices and wages, monetary policy can influence real variables), and Classical features in the long run (neutrality of money). Moreover, the agent can show more or less concern with fulfilling medium-term policy goals. Thus, the choice of which forecast horizon to use is pivotal to this study. All forecasts considered include pre-determined fiscal and monetary paths, such as a package of public investments expected to be implemented soon or a specific interest rate path. Naturally, the forecasts do not include those policy mixes that are designed and implemented within the forecast horizon.

Since comparability is one of the goals of the study, I set a short-run and medium-run forecast horizon of one year and two years respectively, for all institutions under analysis, which enforces coherence regarding the monetary policy. The FED's Board of Governors and the ECB's Governing Council set a medium-term inflation target to be achieved under a two- to three-years horizon and monetary policies are designed considering the position of the expected values relative to the medium-term targets. The time length of the inflation target is such that policymakers can address other pressing issues (*e.g.*, economic output smoothening) in the short run. Regarding fiscal authorities, they differ from monetary ones in several aspects, namely, the policy toolkits, medium-term goals, political meaning attributed to new mandates and respective term length. Nevertheless, the forecast horizons considered make sense on the grounds of the one-year-ahead forecasts allowing authorities to assess the economic state of affairs of the near future, while the two-years-ahead forecasts grossly correspond to half a mandate (the College of Commissioners of the EU offers political leadership for five years and the US Administration for at least four). Since governments tend to value [and potentially depend on] positive results while in office, the forecast horizons considered are legitimate.

3 Methodology

I gather contributions from the loss function literature and create an estimation routine for both the asymmetry loss parameter and the rationality test (leaving the loss parameter unspecified or not). The benchmark model was made by Elliott, Timmermann, and Komunjer 2005 (EKT hereafter). The authors estimate a loss parameter from a family of asymmetric loss functions and create a rationality test based on a Generalized Method of Moments (GMM) framework. The generalized loss function is derived from the interaction of response and utility functions. No assumptions regarding the data generating process are needed. Next, I briefly overview the intuition behind forecast loss functions and present both the EKT method (in a non-exhausting fashion) and the adjustments I introduce into the model.

3.1 Economic Intuition of Loss Functions

We expect the forecaster to minimize a loss function. Thus, the agent cares about the accuracy of the prediction, "internalizes" both the forecast errors and its preferences and adapts the subsequent forecasts accordingly. The reputation or power *w* decreases in the magnitude of the absolute forecast error *e* of a given target variable *y*. We have: $e_{t+1} \equiv y_{t+1} - f_{t+1}$. Note that the economic cost of a higher-than-expected value (*b*) might not equal that of a lower-than-expected value (*c*), hence the authors let w_{t+1} be an asymmetric utility function of the forecast error where *b* and *c* are positive constants corresponding to the respective economic costs. The symmetric scenario corresponds to $b = c$. The response function of the forecaster is the following:

$$w_{t+1} = w(e_{t+1}) \equiv -ln[b|e_{t+1}|\iota(e_{t+1} \geq 0) + c|e_{t+1}|\iota(e_{t+1} < 0)] \quad (1)$$

Where:

ι: indicator function that equals 1 when true and 0 otherwise.

The level of reputation decreases concavely as the size of the forecast error increases; the reputation rate of decay depends on the signal of the forecast error; the intensity of the rate of decay of w_{t+1} depends on the direction of the forecast error. This setting implies that the utility function of the forecaster is:

$$\frac{u'(e_{t+1})}{u'(-e_{t+1})} = (\frac{\alpha}{1-\alpha})^{sgn(e_{t+1})} \quad (2)$$

Where:

$\alpha \equiv \frac{b_p}{b_p+c_p}$;

$0 < \alpha < 1$.

For a given magnitude of the forecast error, the ratio of marginal utilities of underpredicting y_{t+1} (positive forecast error) and overpredicting y_{t+1} (negative forecast error) is $\frac{\alpha}{1-\alpha}$. In other words, when $\alpha < 0.5$, the forecasts are biased towards underprediction, and when $\alpha > 0.5$, the forecasts are biased towards overprediction. As α moves away from 0.5, the loss function becomes more and more asymmetric. If $\alpha = 0.5$, then the loss function is symmetric and the economic costs are equal ($b = c$).

Finally, in order to maximize the expected utility, we have to minimize the expected loss. The authors combine the response and utility functions (Equations 1 and 2) to achieve the flexible class of loss functions defined as the following generalized loss function:

$$L_p(e_{t+1}; \alpha) \equiv [\alpha + (1 - 2\alpha)\iota(e_{t+1} < 0)]|e_{t+1}|^p \tag{3}$$

Where:

α: symmetry parameter;

p: power of the loss function.

The authors take p as given to focus on estimating α. The family of loss functions for which $p = 2$ is called "Quad-Quad" and for $p = 1$ is called "Lin-Lin". The generalized loss function nests the standard Mean Squared Error ($p = 2$, $\alpha = 0.5$) and Mean Absolute Error ($p = 1$, $\alpha = 0.5$) as special cases of loss functions (assuming the forecaster provides the mean of the distribution).

3.2 GMM Estimation

Standard rationality tests estimated by Ordinary Least Squares are not robust to significant asymmetries of the loss function (Elliott, Komunjer, and Timmermann 2008). Moreover, although p is taken as given, α is left unconstrained, because I want to know precisely what is the asymmetry value that each institutional dataset reflects. The optimization problem of the forecaster corresponds to a utility optimization one (in this case, the maximization of a first-order condition). Precisely, Hansen and Singleton 1982 propose a solution to estimate a parameter given optimizing behaviour while optimality restrictions are satisfied: the GMM. Thus, EKT suggest a GMM framework in which the asymmetry parameter α is a by-product of the observation of the sequential forecast errors and respective equation of the linear instrumental variable estimator ($\hat{\alpha}$):

$$\hat{\alpha} \equiv \frac{[\Sigma_{t=\tau}^{T+\tau-1} v_t |e_{t+1}|]' \hat{S}^{-1} [\Sigma_{t=\tau}^{T+\tau-1} v_t |e_{t+1}| \iota(e_{t+1} < 0]}{[\Sigma_{t=\tau}^{T+\tau-1} v_t |e_{t+1}|]' \hat{S}^{-1} [\Sigma_{t=\tau}^{T+\tau-1} v_t |e_{t+1}|]} \quad (4)$$

Where:

v_t: $d \times 1$ vector of instruments;

S: positive definite $d \times d$ weighting matrix.

The matrix S is defined as:

$$S \equiv E\{v_t v_t' [\iota(e_{t+1} < 0 - \alpha_0]^2 |e_{t+1}|^2\} \quad (5)$$

Where:

$p = 2$.

The estimation of α depends on the *S* estimate which in turn depends on the $\hat{\alpha}$, so the estimation is performed iteratively assuming that *S* equals the identity matrix in the first iteration so we can estimate $\hat{\alpha}$ and the model iterates until convergence. Finally, the rationality test takes the form of an overidentification J-test under the null hypothesis that the forecasts are rational.

$$J \equiv \frac{1}{T} \{\Sigma_{t=\tau}^{T+\tau-1} v_t [\hat{\alpha} - \iota(e_{t+1} < 0)] |e_{t+1}|\}' \times \hat{S}^{-1} \{\Sigma_{t=\tau}^{T+\tau-1} v_t [\hat{\alpha} - \iota(e_{t+1} < 0)] |e_{t+1}|\} \quad (6)$$

The intuition of the test presented in Equation 6 is that with each instrument we can obtain an asymmetry parameter that rationalizes the forecast vintages. When the number of instruments is greater than one, the method tests whether there is a common value of α for each moment condition and the J-statistic follows an asymptotic chi-squared distribution of $d - 1$ degrees of freedom. In case no α satisfies all moment conditions simultaneously, the J-statistic becomes large and, under a certain level of confidence, we can reject the null hypothesis of rationality.

In other words, the point of the GMM estimation of the loss parameter is to assess whether the forecast accuracy is improved if we include the value of α as a known parameter, in which case the projections are not rational. If there is not an improvement by incorporating the value of the asymmetry parameter (and respective instruments), then the forecasts are optimal.

3.3 Requalified Model

I use the estimation routine made in MATLAB by Christodoulakis and Mamatzakis 2009 that replicates the EKT method (the code was kindly made available online by the authors). They follow the calibration settings suggested in Elliott, Timmermann, and Komunjer 2005: setting the parameter p to 2 and using an extensive combination of instruments (constant, lagged forecast error and lagged realization).

Demetrescu, Roling, and Titova 2021 raise some concerns about the stationarity assumptions underlying the EKT method (and consequently that of Christodoulakis and Mamatzakis 2009). Particularly, Demetrescu, Roling, and Titova 2021 find that the lagged realization instrument is persistent, rendering the J-statistics inappropriate to use in a standard asymptotic chi-squared distribution. Since the chi-squared critical values lose their adequacy for the J-test, the latter no longer follows the standard limiting distribution, leading to over-rejections of the null hypothesis. In truth, Stock and Watson 2007 prove that certain economic time series (such as inflation) are the sum of “a permanent stochastic trend component and a serially uncorrelated transitory component”, therefore nonstationary. The implication for the rationality testing is that we should limit ourselves to use stationary time series to prevent spurious findings, like the set of instruments “constant and lagged forecast error”, so we get clarifying results for the J-statistics. Thus, I use this combination of instruments to ensure stationarity. Note that although some concerns are raised regarding the rationality testing, the $\hat{\alpha}$ is still consistent, even in the presence of instrument persistence.

There remains to be discussed the number of lags of the instrumental variables and the value of the shape parameter p. In line with the work of Komunjer and Owyang 2012, I use only one lag to prevent size distortions of the rationality test derived from the “many instruments problem” (first raised by Sargan 1958 and later dealt with by Roodman 2009). Finally, I set

p to 2, developing the study on a comparable framework to the tests of forecast rationality derived under the mean squared error loss, although Komunjer and Owyang 2012 assert that all asymmetry parameters are estimated consistently for both $p = 1$ and $p = 2$.

4 Empirical Results

Table 1: Real output growth under asymmetric quadratic loss function

	Real Output Growth									
	Short Run					Medium Run				
	ECB	FED	ECFIN	OMB	Avg.	ECB	FED	ECFIN	OMB	Avg.
$\hat{\alpha}$	0.977	0.781	0.683	0.822	0.816	0.996	0.854	0.785	0.936	0.893
S.E.	0.016	0.063	0.085	0.068		0.007	0.053	0.088	0.035	
J-Statistic	7.798	5.645	0.002	3.002		5.841	7.028	0.007	4.067	
$\alpha = 0.5$ (p-value)	0	0	0.031	0		0	0	0.001	0	

Estimates based on d = 2 instruments (constant and lagged forecast error). J-statistics are distributed as a chi-squared distribution, yielding $d - 1 = 1$ degree of freedom. All institutions produce rational forecasts and the asymmetry parameter is statistically different from 0.5 at the 5% significance level for all institutions.

Table 2: Inflation rate under asymmetric quadratic loss function

Inflation Rate										
	Short Run					Medium Run				
	ECB	FED	ECFIN	OMB	Avg.	ECB	FED	ECFIN	OMB	Avg.
$\hat{\alpha}$	0.176	0.321	0.473	0.828	0.449	0.118	0.345	0.642	0.748	0.463
S.E.	0.057	0.070	0.102	0.066		0.048	0.074	0.112	0.080	
J-Statistic	7.879	5.486	7.139	5.381		9.154	8.772	3.742	4.072	
$\alpha = 0.5$ (p-value)	0	0.011	0.789	0		0	0.036	0.204	0.002	

For more details about the estimates and tests, check the notes of Table 1. All institutions produce rational forecasts. The ECFIN short-run asymmetry parameter is the only one not statistically different from 0.5.

Tables 1 and 2 show that all institutional forecasters are rational and almost all of them operate outside the symmetric-loss scenario. Sample sizes are not big (see Table 3), so one should be cautious interpreting the results, especially those that are more extreme. The null hypothesis of the α estimate being equal to 0.5 is rejected in all cases except for the short-run inflation predictions made by the ECFIN. To get a sense of the relative economic costs of under- and overpredictions, for instance, note that if $\alpha = 0.60$, the response function (Equation 1) shows us that the forecaster puts one and a half times as large a weight on positive forecast errors as on negative ones. The α estimates that shape the institutional forecasters' loss functions suggest there might be significant economic information behind those values. In other words, the institutions find economic arguments to incur in forecast errors that justify the costs of committing them.

The preferences for under- and overprediction in the monetary policy scenario are clear: real output growth forecasts are biased towards overprediction, whereas inflation forecasts are biased towards underprediction. However, the primary analysis of the fiscal policy context is not so straightforward. Although the two fiscal authorities overpredict real output growth, the OMB also tends to overpredict inflation rates and the ECFIN attaches the same weight to positive and negative short-run inflation forecast errors. Thus the stark difference between monetary and fiscal authorities is that the former cross the symmetry threshold to predict the key macroeconomic variables, whereas the latter do not. Notice that the relatively more extreme estimates belong to the ECB (a recent monetary authority still building credibility), and to the OMB (an office of a national government). Generally speaking, there does not seem to exist any remarkable difference between regions. This result is in line with the discussion posed in the Data section (2) about whether the two regions share the same economic costs of under- and overpredictions. Finally, the α estimates of the output growth are systematically less moderate in the medium run. However, the same does not apply to the inflation case: on average, the loss function converges to symmetry, although such statement might not be statistically significant depending on the number of standard deviations from the mean one considers to build a confidence interval to test this hypothesis.

In the first decade of the twenty-first century, most papers published about this topic addressed either the forecasts from the Greenbook, the Survey of Professional Forecasters, the IMF or the OECD. That is, most of the published works addressed neither the EU-bloc as a whole (regardless of the institution), nor American fiscal-side institutions. Note also that the datasets from the beginning of the century were shorter and they encompassed a period when the application of standard monetary policy tools was still the rule. Interestingly, that was the time when the published works on the field used the term “bad outcomes” to refer to lower-than-expected real output growth and higher-than-expected inflation rates, because that was what forecasters were risk-averse to and what they were biased to forecast. According to the literature of that time, they preferred to moderately underpredict real output growth and overpredict inflation rates. Since then, it has become less clear what a “bad outcome” means. Next, I elaborate on potential reasons that justify the observed results.

4.1 Monetary Forecasting Bias

Including data from the first and second decade of the twenty-first century, the datasets encompass the period of unconventional monetary policies that started during the Great Recession. Many advanced economies have hit the “Zero Lower Bound” (ZLB) of short-term nominal interest rates, thus constraining the reference interest rate and therefore the central banks’ ability to stimulate the economy in case a recession hits (Borio and Hofmann 2017). Importantly, this affected the size of the fiscal multiplier (Ji and Xiao 2016), a key indicator to macroeconomic forecasts. In fact, by the time I discuss these results, the ECB has operated under an unconventional set of policy tools for more than half of its lifetime. Consequently, the sample period of monetary authorities has a greater number of observations under unconventional times than under conventional ones. What good and bad outcomes mean under such idiosyncratic circumstances might not be the same as before. Surprisingly at first, if we followed the same line of thought as the previously mentioned works, my results would suggest that monetary authorities are risk-averse to higher-than-expected real output growth and to lower-than-expected inflation rates. Next, I further elaborate on the results obtained under these special conditions.

Independent and impartial policymaking are recognized virtues in both the ECB and the FED. However, the political establishment is sensitive to the social costs of recessions and although these central banks are independent and accountable, they are not completely insensitive to social and political pressures (Aguiar and Martins 2008). This claim gained strength when a former insider from the FED wrote a sharp testimony about the political pressures meddling in the Board of Governors of the central bank: "In most situations the CB will take far more political heat when it tightens pre-emptively to avoid higher inflation than when it eases pre-emptively to avoid higher unemployment." (Blinder 1999). So one should wonder whether the same behaviour applies to the relationship of the monetary policymaker with the forecaster of the same institution. As they both incorporate biases in their respective loss preferences, the functions might be intertwined. For instance, the ECB prefers to build anti-inflationary credibility (Aguiar and Martins 2008) and the forecaster might benefit from aligning its loss function with that of the employer (Ashiya 2009). According to Blinder et al. 2008, communication has become an increasingly important tool of monetary policy. It helps monetary authorities to influence financial markets, to make monetary policy decisions more predictable and to achieve their macroeconomic objectives. The forecasting bias might prove to be another way for policymakers to reveal their preferences to the markets and create the economic conditions to shape business cycles to their goals. In the current macroeconomic setting, expectations play a crucial role. Thus the debate about the underlying reasons for the existence of biased forecasts is intertwined with the debate about the preferences of central banks for deviations of key macroeconomic variables from their targets. That is, the asymmetric loss function of the institutional forecaster might be linked to that of the monetary policymaker.

When an economic downturn is expected, a higher ZLB risk demands an aggressive cut in the reference interest rate (Reifschneider and Williams 2000; Adam and Billi 2007; Nakata 2016), whereas facing an economic expansion after being constrained for a long time by the ZLB (like the Europe and the US cases), a higher ZLB risk (*i.e.*, the possibility that the policy interest rate might be constrained by the ZLB in the future) requires a more gradual increase in the policy rate (Nakov 2018; Evans et al. 2015; Nakata 2017). While the FED has a dual mandate to address full employment and inflation, the ECB focuses on price stability to secure a strong currency, leaving the goals of strong economic growth and employment to the European Commission. If the inflation rate shows signs of acceleration, the monetary authorities might raise the policy rate to ”cool down” the economy. However, raising the reference interest rate decreases the profitability of projects in the economy and therefore said increase might reduce the aggregate investment. Considering this, what central banks want to transmit to the markets during the sample period under analysis is that they expect the economy to grow while not expecting signs of problematic inflation that would make monetary authorities raise their interest rates. The results suggest that monetary institutions try to influence economic agents' expectations. Under unconventional times, central banks act in terms of sending clear signals to the market, expecting agents to incorporate the economic forecasts in their decision-making processes and invest accordingly. In the end, the observed forecast biases show that monetary authorities believe in the accuracy of their policies (overprediction of the real output growth rate), while trying to "tame the animal spirits" of investors (underprediction of inflation, assuring that the reference interest rates will not change in the foreseeable future). The hypothesis of the loss function of the policymaker and that of the institutional forecaster being linked seems plausible.

Looking to the same issue from a different perspective, the results might also suggest that institutions are risk-averse to lower-than-expected inflation outcomes, thus underpredicting the variable and therefore tending to be more pessimistic: the economy faces the risk of deflation while they lack monetary policy tools to tackle it, due to the vicinity of the ZLB. The risk of deflation has been present in OECD countries since late 2002. The Federal Open Market Committee (FOMC) stated in late 2003: "(...) the risk of inflation becoming undesirably low remains the predominant concern for the foreseeable future.". Indeed, the FOMC statement and the overall case for the underprediction of inflation is consistent with both the key interest rates imposed by the central banks (Figures 5a and 5b) and their preferences for underpredictions of inflation (Figures 2a, 2b, 4a and 4b), which approximate the target of "below, but close to, 2%" whenever possible, but overall underpredict inflation since the early 2000s. This hypothesis is consistent with the key interest rates behaviour from this century: since the beginning of the Great Moderation (mid-1980s), reference interest rates have decreased considerably, save occasional exceptions, such as the Dot-com bubble (late 1990s and early 2000s), and the Subprime mortgage crisis that contributed to the previously discussed Great Recession (late 2007). The forecasting bias in monetary policymaking might reflect a risk management, or even a "lesser evil", attitude.

4.2 Fiscal Forecasting Bias

Fiscal authorities do not have a price stability mandate, unlike the monetary authorities. Besides, the appointment process for each institution is not the same. While the nomination for purely political positions is made by the general population, the nomination for central banking positions is not. Thus, their preferences differ. Particularly, governments want to be re-elected, and to achieve that they have to implement policies that somehow satisfy people's preferences. A good example of how this institutional conduct contrasts to the one of the monetary field is a statement by Paul Volcker about the period that became known as the Volcker Disinflation, characterized by high interest rates and unemployment: "I wanted to move the story at least to the front page (...)".

The overprediction of both real output growth and inflation rates in the fiscal context is not as surprising as the observed results concerning monetary authorities. The overprediction of output growth may occur when a national government is optimistic or when a supranational government harmonises the forecasts of all its member-states which also tend to be optimistic.

The posed argument is purely political: government officials seek re-election. Moreover, it is convenient for the EC to be optimistic regarding short- and medium-term predictions, since the forecasts include the economic impact that its funding programmes have on the European community. However, it could be in the best interest of politicians to underpredict output growth during a recession ("Rally 'Round the Flag Effect"). So there may be other reasons for the persistent overpredictions made by the ECFIN and the OMB.

The answer for the overprediction conundrum might be in the public finance field. The expenses of the State Budget are outlined according to the expected fiscal revenue. If it increased, the policymaker could increase the size of the budget. Fiscal revenue increases when the output growth increases (other things held constant). So fiscal authorities might overpredict opportunistically the output growth due to the consequent effect it would have on the forecasted fiscal revenue. Besides, in the European case, the overprediction of output growth helps member-states to signal compliance with the goals of the Stability and Growth Pact. Concerning the overprediction of inflation, Mamatzakis and Tsionas 2020 argue that inflation forecasts influence key macroeconomic variables such as government debt: higher inflation reduces the burden of debt. This implies lower nominal government debt as a percentage of GDP and eases pressure on fiscal consolidation efforts. Finally, if we combined the overprediction of output growth with the potential decrease of nominal government debt (overprediction of inflation), then we would get a substantial decrease of the debt-to-GDP indicator.

5 Conclusion

This study estimates and evaluates the asymmetric loss functions of institutional forecasters through reverse-engineering and GMM estimation using the reported data on forecast and realized values produced by the staffs of the European Central Bank, Directorate-General for Economic and Financial Affairs, Federal Reserve Board and the US Office of Management and Budget. Forecasting requires a loss function that might bias the predictions towards positive or negative forecast errors. Assuming a quadratic symmetric loss is mathematically convenient, as the optimal forecast corresponds to the conditional mean: the economic costs of under- and overpredictions are offsetting. It might be a valid assumption if the forecaster does not have any preference whatsoever over any forecast deviation. Nevertheless, assuming symmetric loss for convenience might lead to invalid inferences about the forecaster's rationality. Besides, it prevents macroeconomists from learning about the substantive implications of forecasting under potentially more realistic functional forms of loss functions.

Relaxing the assumption of quadratic symmetric loss, I find that all institutions produce rational forecasts and almost all of them present forecast biases that are statistically different from the symmetric scenario. Monetary authorities consistently overpredict real output growth and underpredict inflation between 1999:Q1 and 2015:Q4. In general, fiscal authorities overpredict both macroeconomic variables between 1975 and 2019. The results hold for both the short and medium run, except in the case of the short-run inflation forecast made by ECFIN, for which there is statistical evidence of symmetric loss.

The empirical evidence shows that, under the current setting, central banks consider economically more costly to underpredict real output growth and overpredict inflation, whereas governments consider more costly to underpredict both. Regarding monetary authorities, the bias may be linked to sending optimistic signals to the market given the policy toolkit in use. For instance, producing inflation forecasts that are biased towards underprediction could signal that central banks are not planning to raise the reference interest rates in the near- and medium-term. Concerning fiscal authorities, purely political motivations aside, there are political-economic reasons that justify the overpredictions of both macroeconomic variables. For instance, combining the overpredictions of output growth and inflation results in the expectation of a decrease in the debt-to-GDP ratio.

There are two caveats to point out in this study. First, it was only in recent years that the ECB, ECFIN and OMB began to meet consistently. That is, it was relatively frequent to find missing observations before 2008. Importantly, the model I use does not require an uninterruptible flow of observations, since it compares the forecasts and realized values of each data vintage on an individual basis (thus deducting the forecaster's preferences by reverse-engineering). The second caveat is that the model assumes independence across variables, that is, the model does not allow the forecast errors to interact (Komunjer and Owyang 2012).

The results presented in this study have important implications for both revealing the forecast biases of relevant policymakers and dissecting their underlying preferences. Particularly, it is important to note that what a “bad outcome” means nowadays might not correspond to past definitions of the concept. The study is relevant from a policy standpoint since one could wrongly assume that some institutions have a “neutral” forecasting behaviour. It is possible to correct the value of the institutional forecasts by internalizing their degree of ”forecast conservativeness” and therefore adjust the expectations accordingly, allowing for more insightful policy debriefings and policymaking.

References

Adam, Klaus, and Roberto M Billi. 2007. “Discretionary monetary policy and the zero lower bound on nominal interest rates.” *Journal of monetary Economics* 54 (3): 728–752.

Aguiar, Alvaro, and Manuel MF Martins. 2008. “Testing for Asymmetries in the Preferences of the Euro-area Monetary Policymaker.” *Applied Economics* 40 (13): 1651–1667.

Ashiya, Masahiro. 2009. “Strategic bias and professional affiliations of macroeconomic forecasters.” *Journal of Forecasting* 28 (2): 120–130.

Blinder, Alan S. 1999. *Central banking in theory and practice.* Mit press.

Blinder, Alan S, Michael Ehrmann, Marcel Fratzscher, Jakob De Haan, and David-Jan Jansen. 2008. “Central bank communication and monetary policy: A survey of theory and evidence.” *Journal of economic literature* 46 (4): 910–45.

Borio, Claudio EV, and Boris Hofmann. 2017. “Is monetary policy less effective when interest rates are persistently low?”

Caunedo, Julieta, Riccardo DiCecio, Ivana Komunjer, and Michael T Owyang. 2020. “Asymmetry, Complementarities, and State Dependence in Federal Reserve Forecasts.” *Journal of Money, Credit and Banking* 52 (1): 205–228.

Choudhry, Taufiq, Syed S Hassan, and Sarosh Shabi. 2020. “US economic uncertainty, EU business cycles, and the global financial crisis.” *International Journal of Finance & Economics* 25 (1): 28–42.

Christodoulakis, George A, and Emmanuel C Mamatzakis. 2009. “Assessing the prudence of economic forecasts in the EU.” *Journal of Applied Econometrics* 24 (4): 583–606.

Demetrescu, Matei, Christoph Roling, and Anna Titova. 2021. "Reevaluating the prudence of economic forecasts in the EU: The role of instrument persistence." *Journal of Applied Econometrics* 36 (1): 151–161.

Elliott, Graham, Ivana Komunjer, and Allan Timmermann. 2008. "Biases in macroeconomic forecasts: irrationality or asymmetric loss?" *Journal of the European Economic Association* 6 (1): 122–157.

Elliott, Graham, Allan Timmermann, and Ivana Komunjer. 2005. "Estimation and testing of forecast rationality under flexible loss." *The Review of Economic Studies* 72 (4): 1107–1125.

Evans, Charles, Jonas Fisher, François Gourio, and Spencer Krane. 2015. "Risk Management for Monetary Policy at the Zero Lower Bound." *Federal Reserve Bank of Chicago, January.*

Faust, Jon, and Jonathan H Wright. 2009. "Comparing Greenbook and reduced form forecasts using a large realtime dataset." *Journal of Business & Economic Statistics* 27 (4): 468–479.

Hansen, Lars Peter, and Kenneth J Singleton. 1982. "Generalized instrumental variables estimation of nonlinear rational expectations models." *Econometrica: Journal of the Econometric Society,* 1269–1286.

Ji, Yangyang, and Wei Xiao. 2016. "Government spending multipliers and the zero lower bound." *Journal of Macroeconomics* 48:87–100.

Keereman, Filip, et al. 1999. *The track record of the Commission Forecasts.* Technical report. Directorate General Economic and Financial Affairs (DG ECFIN), European . . .

Komunjer, Ivana, and Michael T Owyang. 2012. "Multivariate forecast evaluation and rationality testing." *Review of Economics and Statistics* 94 (4): 1066–1080.

Mamatzakis, Emmanuel C, and Mike G Tsionas. 2020. "Revealing forecaster's preferences: A Bayesian multivariate loss function approach." *Journal of Forecasting* 39 (3): 412–437.

Melander, Annika, George Sismanidis, Daniel Grenouilleau, et al. 2007. *The track record of the Commission's forecasts-an update.* Technical report. Directorate General Economic and Financial Affairs (DG ECFIN), European . . .

Nakata, Taisuke. 2016. "Optimal fiscal and monetary policy with occasionally binding zero bound constraints." *Journal of Economic Dynamics and control* 73:220–240.

———. 2017. "Uncertainty at the zero lower bound." *American Economic Journal: Macroeconomics* 9 (3): 186–221.

Nakov, Anton. 2018. "Optimal and simple monetary policy rules with zero floor on the nominal interest rate." *Thirteenth issue (June 2008) of the International Journal of Central Banking.*

Reifschneider, David, and John C Williams. 2000. "Three lessons for monetary policy in a low-inflation era." *Journal of Money, Credit and Banking,* 936–966.

Roodman, David. 2009. "A note on the theme of too many instruments." *Oxford Bulletin of Economics and statistics* 71 (1): 135–158.

Sargan, John D. 1958. "The estimation of economic relationships using instrumental variables." *Econometrica: Journal of the Econometric Society,* 393–415.

Sinclair, Tara M, Fred Joutz, and Herman O Stekler. 2010. "Can the Fed predict the state of the economy?" *Economics Letters* 108 (1): 28–32.

Stock, James H, and Mark W Watson. 2007. "Why has US inflation become harder to forecast?" *Journal of Money, Credit and banking* 39:3–33.

Tulip, Peter. 2009. "Has the economy become more predictable? Changes in Greenbook forecast accuracy." *Journal of Money, Credit and Banking* 41 (6): 1217–1231.

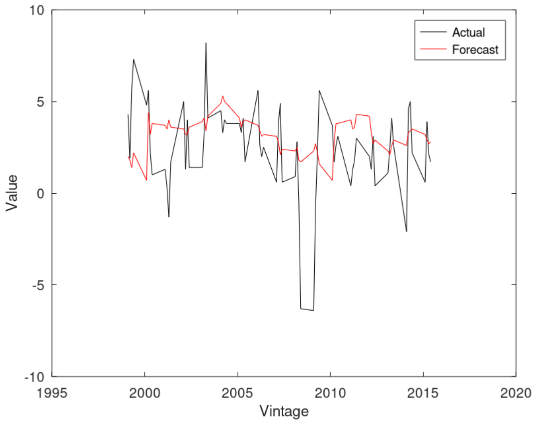
Wang, Yiyao, and Tae-Hwy Lee. 2014. "Asymmetric loss in the Greenbook and the Survey of Professional Forecasters." *International Journal of Forecasting* 30 (2): 235–245.

6 Appendix

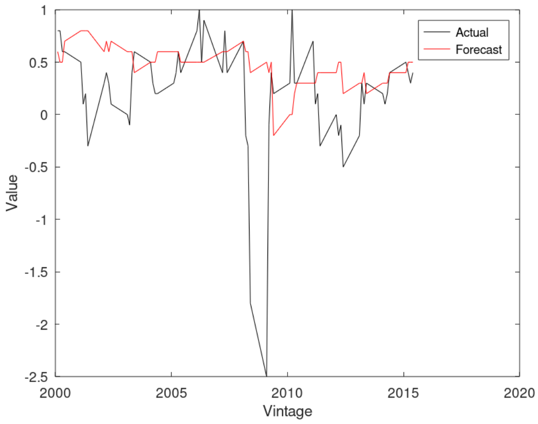
Table 3: Summary table

Institution	Sample Period	Target Variable	Forecast Horizon	Number of Obs.
ECB	1999:Q1-2015:Q4	Real GDP	4 quarters	63
			8 quarters	60
		HICP	4 quarters	60
			8 quarters	53
FED	1999:Q1-2015:Q4	Real GDP	4 quarters	68
			8 quarters	67
		GDP Deflator	4 quarters	68
			8 quarters	67
ECFIN	1975-2019	Real GDP	1 year	51
			2 years	26
		GDP Deflator	1 year	47
			2 years	27
OMB	1975-2019	Real GDP	1 year	45
			2 years	43
		GDP Deflator	1 year	45
			2 years	43

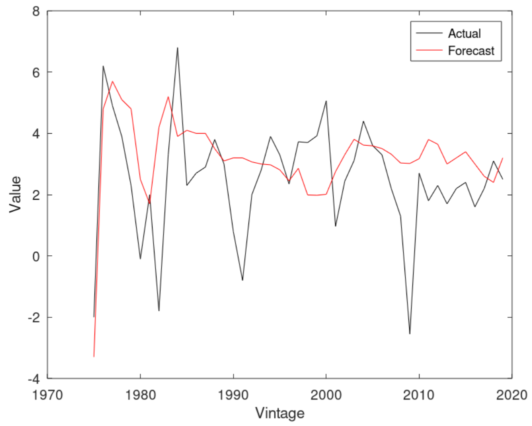
Figure 1: Output Growth Forecasts, Short Run (1 year)



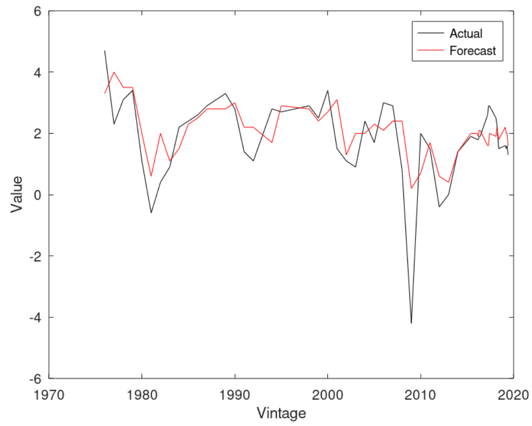
(a) FED



(b) ECB

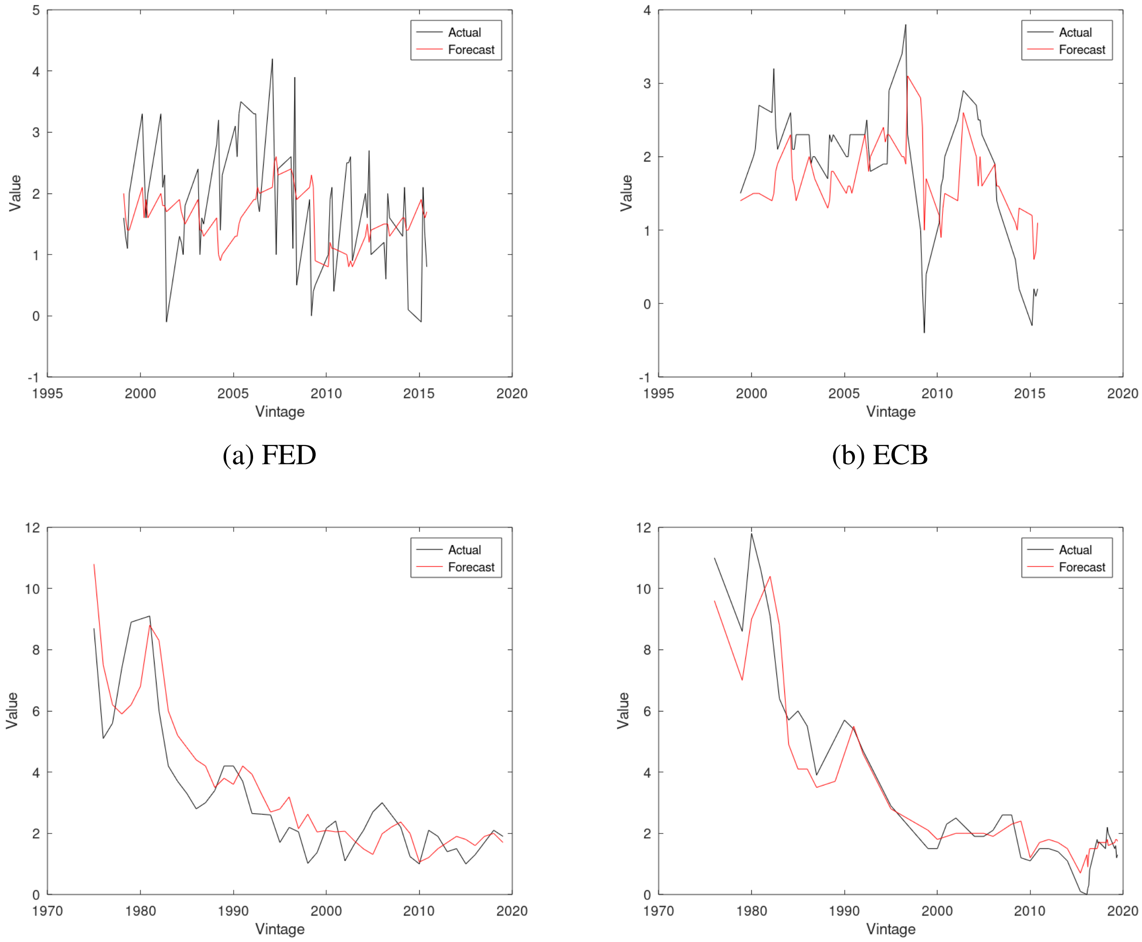


(c) OMB



(d) ECFIN

Figure 2: Inflation Forecasts, Short Run (1 year)



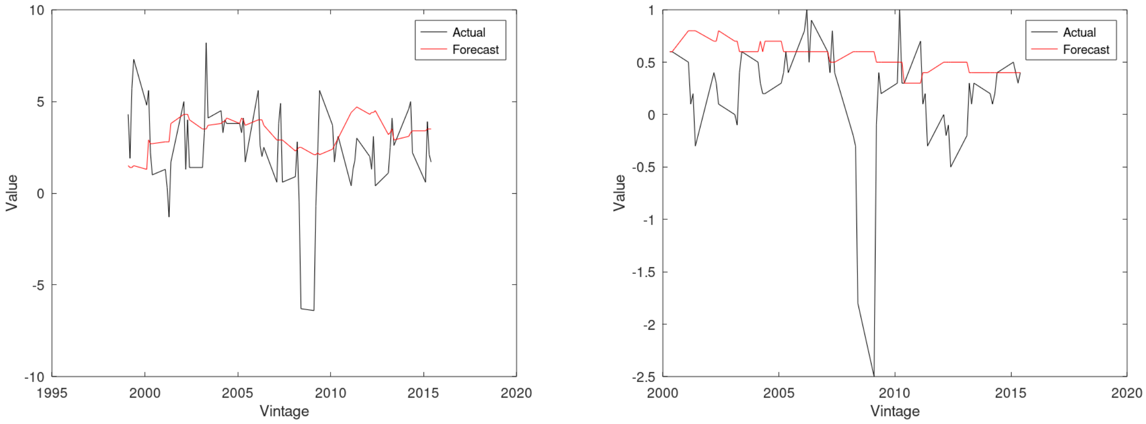
(a) FED

(b) ECB

(c) OMB

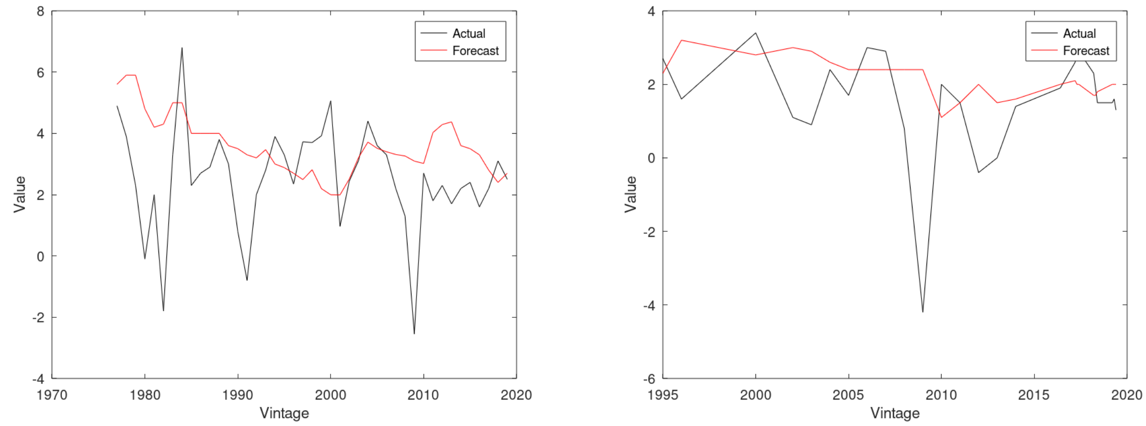
(d) ECFIN

Figure 3: Output Growth Forecasts, Medium Run (2 years)



(a) FED

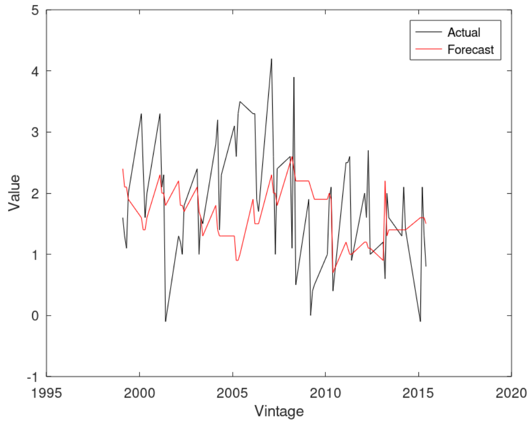
(b) ECB



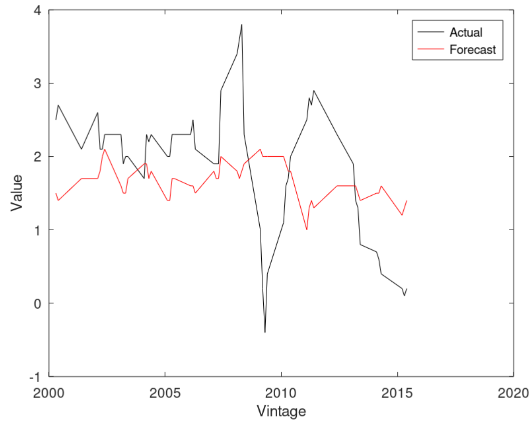
(c) OMB

(d) ECFIN

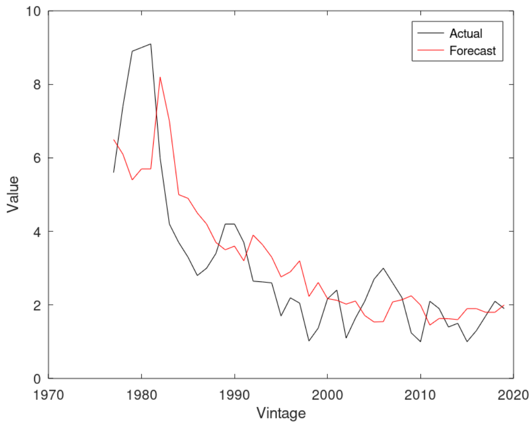
Figure 4: Inflation Forecasts, Medium Run (2 years)



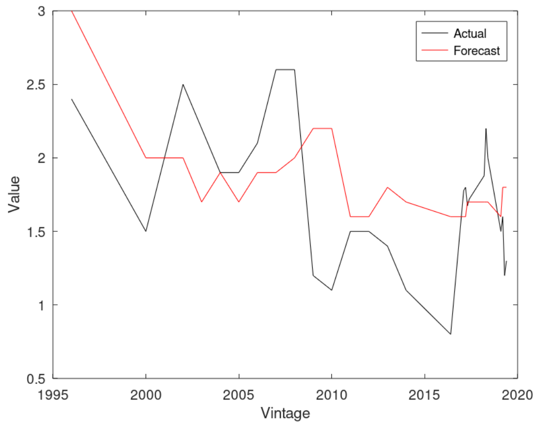
(a) FED



(b) ECB

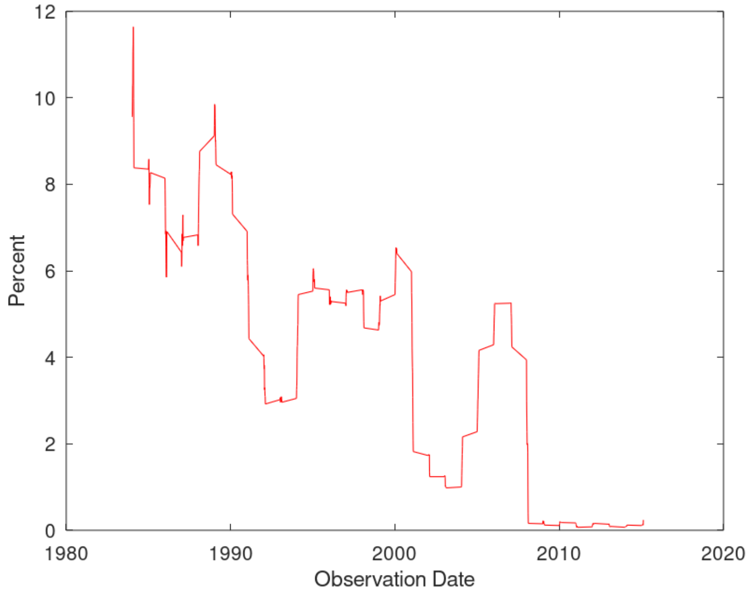


(c) OMB

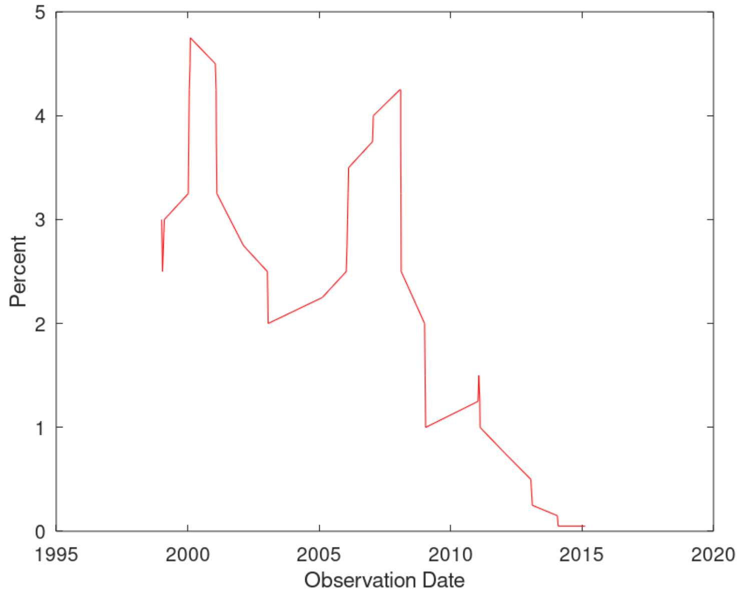


(d) ECFIN

Figure 5: Reference Interest Rates



(a) Federal Funds Effective Rate (the interest rate at which depository institutions trade federal funds with each other overnight); Monthly-frequency (1984-2015); Not seasonally adjusted. Source: FRED.



(b) Interest rate on the Main Refinancing Operations (the interest rate banks pay when they borrow money from the ECB for one week); Daily-frequency (businessweek); Average of observations through undefined period (1999-2015). Source: Statistical Data Warehouse.