EMILIA BELEN
CHOCOBAR

COMILLAS
UNIVERSIDAD PONTIFICIA

Escuela Internacional de Doctorado

Doctorate Program in Business and Regional Competitiveness, Innovation and Sustainability

MEASURING AND EXPLAINING DEVIATIONS FROM RATIONAL EXPECTATIONS:

TIME-VARYING EVIDENCE FROM INFLATION AND FISCAL FORECASTS

MEASURING AND EXPLAINING DEVIATIONS FROM RATIONAL EXPECTATIONSN

Autor: Emilia Belén Chocobar
Director: Peter Claeys

MADRID | Abril 2026

Credit Author Statement

I contributed to this dissertation as follows:
Conceptualization – [in collaboration with Peter Claeys]
Data curation
Formal analysis
Methodology
Software
Validation
Visualization
Writing, first draft, review, and editing - [in collaboration with Peter Claeys]
My contribution is the same for all chapters included in this thesis.

Use of GenAI Statement

During the preparation of this dissertation, I used ChatGPT (OpenAI) in order to review grammar, literary and language proofreader. After using this tool/service, I reviewed and edited the content as needed and take full responsibility for the content of the publication.

ABSTRACT

This thesis examines how and why expert forecasts deviate from the rational expectations benchmark in two central domains of macroeconomic policy: fiscal balances and inflation. Whereas much of the literature isolates behavioral biases, informational frictions, or external shocks, this dissertation integrates them within a single empirical framework. This approach allows for time-varying and expert-level analysis, revealing how these factors interact to shape deviations from rationality across contexts.

The first chapter, "Fiscal Forecasting Rationality among Expert Forecasters," focuses on budget balance forecasts for the U.S., France, Germany, and Italy between 1993 and 2023. The Fluctuation Rationality (FR) test of Rossi and Sekhposyan (2016) evaluates whether expert' forecasters efficiently incorporate economic and political information into their forecasts and identifies the conditions under which rationality weakens or breaks down. The chapter finds that fiscal forecasts respond efficiently under normal conditions but tend to overreact during major structural shifts such as the build-up to euro area fiscal consolidation, the Global Financial Crisis, or the COVID-19 pandemic. Noisy or sticky information, political uncertainty, and macroeconomic changes interact to deviate experts from the rational expectations framework.

The second chapter, "Identifying Drivers of Deviations from Rational Expectations: A New Irrationality Index," extends this approach to monthly U.S. inflation forecasts between 2010 and 2022. This chapter also applies the Rossi and Sekhposyan (2016) to introduce the "Irrationality Index" to jointly model behavioral biases, informational frictions, or external shocks of forecast deviations. This index allows for capturing expert-level, time-varying deviations from rationality. Additionally, it quantifies how behavioral, informational, and external shocks jointly shape the deviations. The results show that external factors dominate once all drivers are considered jointly, while traditional behavioral indicators play only a secondary role.

The third chapter "A Neural Network Approach to Model the Herding Behavior of Inflation Experts' Forecasts" introduces a machine learning approach using Graphical Neural Networks to identify and analyze how hearding behavior evolves over time.

The policy implications of this research are clearly important and relevant for policy makers. Transparent, credible, and consistent communication by central banks and fiscal authorities can help anchor expectations, reduce informational frictions, and avoid overreactions to unexpected shocks. Moreover, the Irrationality Index is a practical tool for policymakers to monitor experts' forecast performance and to anticipate periods of heightened misalignment between expectations and fundamentals.

Keywords: rational expectations, deviations, fiscal forecasts, inflation forecasts, fluctuation rationality test, hearding, neural networks.

RESUMEN

Esta tesis examina cómo y por qué las previsiones de los expertos se desvían del modelo de expectativas racionales en dos ámbitos centrales de la política macroeconómica: los saldos fiscales y la inflación. Mientras que gran parte de la literatura aísla los sesgos conductuales, las fricciones informativas o las perturbaciones externas, esta disertación los integra en un único marco empírico. Este enfoque permite un análisis que varía con el tiempo y a nivel de experto, revelando cómo interactúan estos factores para configurar las desviaciones de la racionalidad en distintos contextos.

El primer capítulo se centra en las previsiones de saldo presupuestario para Estados Unidos, Francia, Alemania e Italia entre 1993 y 2023. Al aplicar el test de Racionalidad de Fluctuación (FR) de Rossi y Sekhposyan (2016) para detectar desviaciones de la racionalidad, se constata que las previsiones fiscales responden de forma eficiente en condiciones normales, pero tienden a reaccionar de forma exagerada durante cambios estructurales importantes. El segundo capítulo extiende este enfoque a las previsiones mensuales de inflación en Estados Unidos entre 2010 y 2022. Este capítulo cuantifica cómo los choques conductuales, informativos y externos influyen conjuntamente en las desviaciones de la racionalidad. Los resultados muestran que los factores externos predominan al considerar todos los determinantes en conjunto, mientras que los indicadores conductuales tradicionales desempeñan un papel secundario. El tercer capítulo introduce un enfoque de aprendizaje automático mediante redes neuronales gráficas para identificar y analizar la evolución del comportamiento de los analistas a lo largo del tiempo.

Las implicaciones políticas de esta investigación son claramente importantes y relevantes para los responsables de la formulación de políticas. Una comunicación transparente, creíble y coherente por parte de los bancos centrales y las autoridades fiscales puede ayudar a afianzar las expectativas, reducir las fricciones informativas y evitar reacciones exageradas ante choques inesperados. Además, el Índice de Irracionalidad es una herramienta práctica para que los responsables de la formulación de políticas supervisen el desempeño de las previsiones de los expertos y anticipen períodos de mayor desalineación entre las expectativas y los fundamentos económicos.

TABLE OF CONTENTS

INTRODUCTION

I. Motivation

How experts form their expectations is at the core of macroeconomic theory and policy advice. Forecasts of inflation, growth, or fiscal balances have an important role. Those numbers inform monetary and fiscal policy decisions, financial markets, and shape public debates. Central banks, fiscal councils, and international organizations rely on the quality of private and public forecasts when setting policy, and the credibility of these experts depends on how accurately such expectations anticipate future conditions.

The expectations formations are framed into the rational expectations theory developed by Muth (1961). This theory assumes that experts make forecasts unbiased, fully informed, and without systematic errors. Nevertheless, both theoretical and empirical studies point to sticky or noisy information (Mankiw and Reis, 2002; Woodford, 2002; Sims, 2003), behavioral biases (Kahneman and Tversky, 1979; Barberis and Thaler, 2003), and external shocks (Coibion and Gorodnichenko, 2012, 2015) as key sources of deviations from rational expectations. The behavioral approach explains the deviations trough cognitive biases such as overconfidence, anchoring, and herding (Kahneman and Tversky, 1979; Barberis and Thaler, 2003; Trueman, 1994). Informational frictions highlight the limits of experts' ability to process data in a frequent way, as in the rational inattention and sticky-information frameworks (Mankiw and Reis, 2002; Sims, 2003; Woodford, 2012). The third explanation focused on the role of external shocks and structural breaks, such as energy price surges, policy regime changes, or geopolitical crises (Coibion and Gorodnichenko, 2015). Although each approach addresses important mechanisms, the literature has studied them in an isolated way. Then, the expectation formation literature leaves unanswered how these mechanisms interact to produce systematic deviations from rationality across different forecasting domains.

An additional limitation of existing work is methodological. Traditional rationality and efficiency tests are usually applied to the entire sample. This approach relies on average coefficients that can ignore heterogeneity among experts and mask time variation in behavior. Such aggregation may hide the fact that individual experts react differently to information depending on their institutional environment, the economic cycle, or the nature of incoming shocks. Furthermore, tests on mean forecasts do not capture the dynamics of dispersion, bias, or herding across forecasters, which are often early indicators of forecast breakdowns.

This dissertation addresses these shortcomings by adopting a unified empirical framework to examine deviations from rational expectations in inflation and fiscal forecasts. Expectations and under what conditions these deviations are most pronounced. This thesis is structured around two stand-alone but closely related papers. Together, they address how and why expert forecasts deviate from the rational expectations benchmark and how these deviations evolve over time in fiscal and macro forecasts.

II. Scope and Outline

This thesis is structured as unified research project to address the limitations in the literature of why experts' forecasters deviate from rational expectations in inflation and fiscal domains. Across all three chapters, I use the Consensus Economics survey data to provide a consistent and comparable environment to study expert forecasting behavior in fiscal and inflation areas. This decision enhances the interpretations of the results across fiscal and inflation forecasts, across econometric and machine-learning methods, and across stable and crisis periods.

The first chapter introduces the paper “Fiscal Forecasting Rationality among Expert Forecasters” and has been published as *Belén Chocobar, Peter Claeys, and Marcos Poplawski-Ribeiro (2025), “Fiscal Forecasting Rationality Among Expert Forecasters” Journal of Forecasting, https://doi.org/10.1002/for.3237*. This chapter answer the question: do fiscal experts forecasters form expectations consistent with rational expectations? How do deviations from rational expectations evolve over time and across countries? This chapter extends the forecast limited literature on the fiscal domain by examining budget balance forecasts for the U.S., France, Germany, and Italy from 1993 to 2023. Additionally, I introduce a novel Irrationality Index based on the Fluctuation Rationality test to capture time-varying deviations from rational expectations at the individual expert level. Particular attention is given to large policy shifts and exogenous shocks-such as the build-up to the euro area’s fiscal consolidation, the Global Financial Crisis, and the COVID-19 pandemic-that may impair experts’ ability to process information accurately. This chapter focuses on how noisy or sticky information, political uncertainty, and macroeconomic volatility interact to shape the accuracy and dispersion of fiscal forecasts over time.

The second chapter presents the paper “Identifying Drivers of Deviations from Rational Expectations” and has been published as *Belén Chocobar, Peter Claeys (2026). “Identifying Drivers of Deviations From Rational Expectations: Using a New Irrational Index for Inflation Forecasts”. Journal of Forecasting 1–13. https://doi.org/10.1002/for.70136*. Here the question and the methodology are the same as before but focused on inflation forecasts. I test whether expert forecasters rationally incorporate economic and political information into their inflation projections and identify the conditions under which forecast rationality weakens or breaks down. The study focusses on inflation forecasts made by experts in the U.S. between 2010 and 2022 because this period of time is marked by both prolonged stability and unprecedented shocks such as the COVID-19 pandemic and the surge in geopolitical risk. The analysis also combines measures of behavioral bias, herding tendencies, informational frictions, and external shocks-such as shifts in interest rate expectations and political risk-to explain why forecasters systematically deviate from rationality in predicting year-ahead inflation. By adopting this unified empirical framework, the chapter goes beyond traditional single-factor explanations and shows how multiple drivers interact to influence inflation forecasts over time. Additionally, this analysis provides a comparison between fiscal and monetary results.

The third chapter, “A Neural Network Approach to Model the Herding Behavior of Inflation Experts’ Forecasts” focus on a different but complementary research question: how does herding among forecasters evolve over time? And how the structure of herding behavior evolves over time? This chapter introduce machine learning techniques—specifically Graph Neural Networks (GNNs)—to study dynamic patterns of herding, clustering, and influence among experts. It demonstrates that during stable periods, experts’ forecasts converge around the mean, while in times of heightened uncertainty (such as the COVID-19 pandemic or geopolitical shocks), the network fragments and experts rely more on private information.

III. Contribution

This dissertation makes both methodological and empirical contributions to the study of expectation formation.

On the methodological side, it advances the literature by adapting and extending the Fluctuation Rationality test developed by Rossi and Sekhposyan (2016) to build new indicators of deviations from rationality at the expert level. Whereas most previous studies evaluate forecast rationality using full-sample averages or mean forecasts, this approach captures time variation and heterogeneity across experts, allowing for the detection of episodes in which rationality breaks down. By incorporating behavioral metrics such as bias, herding and overconfidence alongside macroeconomic and political variables, the methodology integrates insights from three previously separate strands of the literature: behavioral biases (Kahneman and Tversky, 1979; Barberis and Thaler, 2003), informational frictions (Mankiw and Reis, 2002; Sims, 2003; Woodford, 2012), and external shocks (Bloom, 2009; Coibion and Gorodnichenko, 2015; Baker *et al.*, 2016; Caldara and Iacoviello, 2022). Thus, this unified framework moves beyond static or single-factor tests and provides a richer diagnostic of how experts form expectations under uncertainty. Additionally, due to Chapter 3, this thesis also contributes to the literature by using machine learning approaches to study hearding behavior in experts' forecasters.

On the empirical side, the thesis contributes new evidence on the determinants of deviations from rational expectations in both inflation and, particularly in, fiscal forecasts. Expectation formation literature is typically focus on monetary variables leaving fiscal variables unexplored. Fiscal outcomes have a particularly complexity, such as, political constraints, electoral cycles, legislative bargaining, changes in fiscal rules, etc. which increase the scope for structural breaks and make information harder to process in real time. The firsts chapter demonstrates that fiscal forecasters in the U.S., France, Germany, and Italy systematically incorporate economic and political news under normal conditions but tend to overreact during large policy shifts (Strauch *et al.*, 2004; Jalles *et al.*, 2015; An and Jalles, 2021). The second chapter shows that deviations from rationality in U.S. inflation forecasts between 2010 and 2022 are not random but strongly associated with external information shocks, long-term interest rate expectations, and geopolitical risk (Afrouzi *et al.*, 2021; Bordalo *et al.*, 2020; Caldara and Iacoviello, 2018). By applying the same empirical strategy across inflation and fiscal forecasts, the thesis provides a comparative view of how uncertainty and policy changes jointly shape the accuracy and dispersion of expert expectations.

Finally, it provides policy-relevant insights on mitigating deviations from rationality through institutional design and communication strategies. By combining cross-country fiscal evidence with individual-level inflation forecasts, the thesis shows that transparent and credible policy frameworks-such as fiscal rules or consistent central bank communication-can help anchor expectations and reduce overreactions to shocks (Gennaioli *et al.*, 2015; Angeletos and Lian, 2018). The results imply that policymakers should pay attention to periods of heightened geopolitical or economic uncertainty, when experts are most prone to deviate from rationality. Additionally, the Irrationality Index offers policymakers a practical tool for monitoring experts' performance and to anticipate periods when market expectations may become misaligned with fundamentals.

CHAPTER 1: Fiscal Forecasting Rationality among Expert Forecasters

I. Introduction

Understanding how agents form expectations holds significant importance in macroeconomics. Two predominant theories hold opposing views: Mankiw and Reis (2002) argue that forecasters infrequently update their information sets due to fixed acquisition costs; Woodford (2002) and Sims (2003) in turn argue that agents continuously receive updates about the true state of the economy but must extract relevant changes from noisy signals. Both theories predict information rigidities leading to deviations from full information rational expectations.

A recent literature has focused on the drivers underlying the dispersion among experts' forecasts, notably on growth and inflation, and the influence of macroeconomic or political factors on those deviations (Coibion and Gorodnichenko, 2012). This literature has not yet come to robust evidence about the quantitative importance and nature of information rigidities faced by different types of economic agents, as it is hard to get detailed data on expectations from different sources, either in surveys or laboratory experiments (Cornand and Hubert, 2020).

This paper innovates by testing the drivers underlying the dispersion of experts' forecasts on fiscal policy and by employing new methods to test the rationality of those expectations. We rely on recent advances in testing forecaster rationality over time-the Fluctuation Rationality test (Rossi and Sekhposyan, 2016)-to examine the role of economic or political information in updating forecasts. We apply the test to a quarterly panel dataset with survey forecasts from four countries (France, Germany, Italy and the US) and with a maximum of 94 (current-year) and 101 (year-ahead) forecasters on a long sample with quite some fiscal shifts between 1993 and 2023[1]. After applying a battery of standard tests for absolute and relative performance of budget balance forecasts, we then test for the presence of information rigidities on those forecasts, and for the impact of external information on these rigidities. Next, we apply the Fluctuation Rationality Test ("FR test") to detect when expert forecasters deviate from rationality, followed by an exploration of the role of external information-in particular, policy frameworks, global risk factors, or macroeconomic factors-on these deviations from rationality.

Our findings indicate that the forecaster behavior for budget balances is best modelled as a mixture of sticky and noisy information. The reason is that traditional tests using full-sample averages iron out any changes in forecaster behavior. In fact, we confirm that expert forecasters are not efficient in revising forecasts over the period 1993-2023. For European countries, the large-scale consolidation leading up to EMU entry in 1999, the repercussions of the 2008 Global Financial Crisis culminating in the early 2010s Debt Sovereign Crisis, and the disruptions created by the Pandemic were responsible to major sudden changes in the experts' fiscal forecasts. In the case of the US, the common informational shock is captured by large swings in tax policies, and increased outlays owing to both the Global Financial Crisis and military events. In the presence of such large shifts, the capacity of each expert forecaster to understand the budgetary implications is compromised, and forecasters overreact, especially in countries where fiscal policy lacks a stable framework.

Nevertheless, a lack of rationality is concentrated only in those episodes. When explaining the time-varying changes in rationality, we find that expert forecasters do adjust their forecast of the budget balance in reaction to specific news. In normal times, forecasters do systematically

[1] The difference between the current-year forecasts (94) and the year-ahead forecasts (101) is driven by the data availability and not by GDP data.

incorporate economic and political news in their budget forecast revisions, and seem to become more cautious during episodes of political and economic uncertainty.

Our paper contributes to the literature in two ways. First, we extend the tests developed by Dovern *et al.* (2015) to test forecaster efficiency, and shed a new light on testing information rigidities. We do so first by augmenting the Dovern *et al*. (2015) test with news on economic or political variables, and second by using the Fluctuation Rationality test to examine time-variation in forecast efficiency. Second, from a practical point of view, our paper further extends another branch investigating absolute and relative performance of fiscal forecasting (Leal *et al*., 2008; Jalles *et al*., 2015). We do so by detailing how expert forecasters deviate from rationality, and shed a new light on the quality of fiscal forecasts (Leal *et al.*, 2008). Expert budget forecasters seem to be overall performing rational projections, but overreact to structural breaks. Forecasters have a hard time filtering out the future tendencies after these large policy shifts. Such a view is in line with models of noisy information.

The remainder of the paper is structured as follows. Section II describes the dataset; Section III first tests absolute forecasting performance of budget forecasts, then checks for information rigidities, to apply next tests of forecast rationality. Section IV then explains deviations from rationality with a combination of economic and political variables. Finally, Section V presents conclusions and policy implications.

II. Expert forecasts on fiscal data

We use *Consensus Economics (CE)* forecasts data to investigate how experts form fiscal expectations. CE conducts a monthly survey in up to 100 countries and queries respondents every first week of each month about current and future developments for a number of macroeconomic and financial variables. These forecasts are then published early in the second week of the same month. The fiscal variable surveyed by CE is the overall budget balance, which a significantly lower coverage than other macroeconomic variables, such as GDP growth and inflation, for example. This variable is reported in nominal local currency terms on an annual basis, both for the current and the next fiscal year.[2]

Unlike other surveys, individual forecasts in CE should not suffer a bias owing to the release of strategic forecasts, as often happens for official projections released by governmental agencies (Ottaviani and Sorensen, 2006; D'Agostino and Ehrmann, 2014). In addition, evidence shows that CE forecasts are less biased and more accurate than forecasts of some international institutions. CE data is public, which helps to prevent a participant from reproducing others' forecasts and also limits the possibility of herding (Trueman, 1994)[3]. Moreover, forecasters are bound in their survey answers by their recommendations to their clients, and discrepancies between the survey and their private recommendation would be hard to justify (Keane and Runkle, 1990). Overall, we can reasonably argue that the CE survey data broadly reflects the spectrum of expectations of market experts.

Our analysis focus on France, Germany, Italy, and the US with data covering a sample of forecasts surveyed between January 1993 to December 2022. Fiscal forecasts have not always received the same attention by forecasters over time. Varying data availability for different

[2] For a few countries and time frames, additional fiscal variables include government debt, revenues, and expenditures, yet its coverage in the survey is patchy, even in the main advanced economies.
[3] The Consensus Economics access requires a paid subscription.

months implies that in some months several fiscal forecasts were available, while others had none. To ensure a consistent and reliable series of forecasts over a longer period, we follow Heppke-Falk and Hüfner (2004) and Cimadomo *et al.* (2016) and proceed in various steps.[4]

Firstly, we aggregate monthly data on a quarterly basis. For each quarter, we select the most recent monthly forecast available within that quarter. This means that if there were forecasts available in January, February, and March, we chose the March forecast as representative value for quarter 1, but if there are no data for March, we select the February forecast. Secondly, CE reports the budget balance in nominal local currency terms only, yet the literature has typically compared absolute and relative forecasting performance of the budget balance to GDP ratio (Artis and Marcellino, 2001; Leal *et al.,* 2008; Frankel, 2011; Merola and Perez, 2013; Jalles *et al*., 2015). We therefore transform each forecast to a ratio of GDP by dividing the fixed-horizon current (or one-year-ahead) forecast of the nominal budget balance in a certain month *m* by the GDP forecast for the same year. As in Cimadomo *et al.* (2016), given that the CE dataset only provides forecasts of GDP growth rates, we compute the year-ahead nominal GDP level forecast by applying the CE growth rate to the latest available estimate for the same year GDP level. The latter is taken from IMF WEO (see Appendix A for more details).

CE conducts surveys among professional economists working for commercial or investment banks, government agencies, research centers and university departments[5]. We distinguish four categories of forecasters: banks, financial services, consultants, and research departments. Banks include domestic or international banks that primarily engage in providing financial services to individuals and businesses. The group of experts from financial services includes insurance companies, investment firms, or hedge funds. Consultants refer to consulting companies and other specialized organizations that provide macroeconomic forecasts and analyzes. Research departments are specialized units within universities, think tanks, or major corporations focused on producing in-depth economic research. A detailed list of the forecasters in the sample is available in Appendixes B (per country) and C (per subgroup).

Table 1 shows some descriptive statistics of the projected budget balances for the current-year and year-ahead in our sample, reporting their overall mean as well as the mean (and other statistics) for each group of experts.[6] The mean forecast is computed using information available for the expert forecasters in each quarter.

Panel (a) of Figure 1 compares the mean forecast across experts for the current-year to the realized budget balance ratio to GDP, while panel (b) does the same for the year-ahead forecast. As realized values, we take the different issues of the IMF World Economic Outlook's budget balance for European countries. For the U.S., we use the outturn data reported by the US Congressional Budget Office.

---

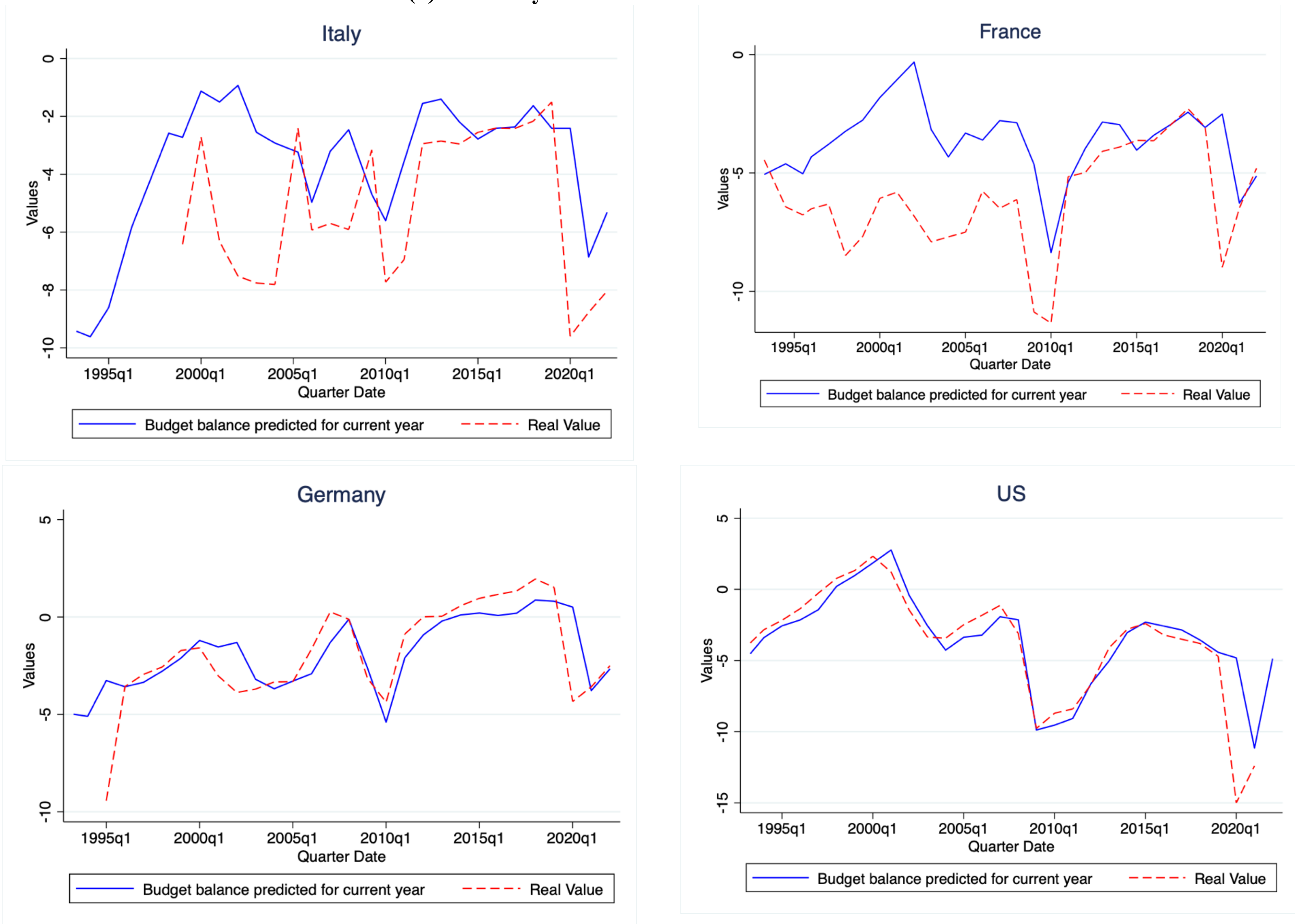
[4] Although the survey is conducted at a monthly frequency, the forecast target is annual.

[5] All surveyed experts provide forecasts for their own country only. Some international financial institutions or research institutes are included in the sample in several countries, but it is the national representative of that institution who provides the domestic forecast. Further information on how the survey is conducted is available at www.consensuseconomics.com.

[6] Italy has the smallest sample of expert forecasters, and as a result, only banks, financial services and consultants are included in the subgroups.

Figure 1. Mean of budget balance forecasts.

(a) current year forecast versus realized value



(b) year-ahead forecast versus realized value

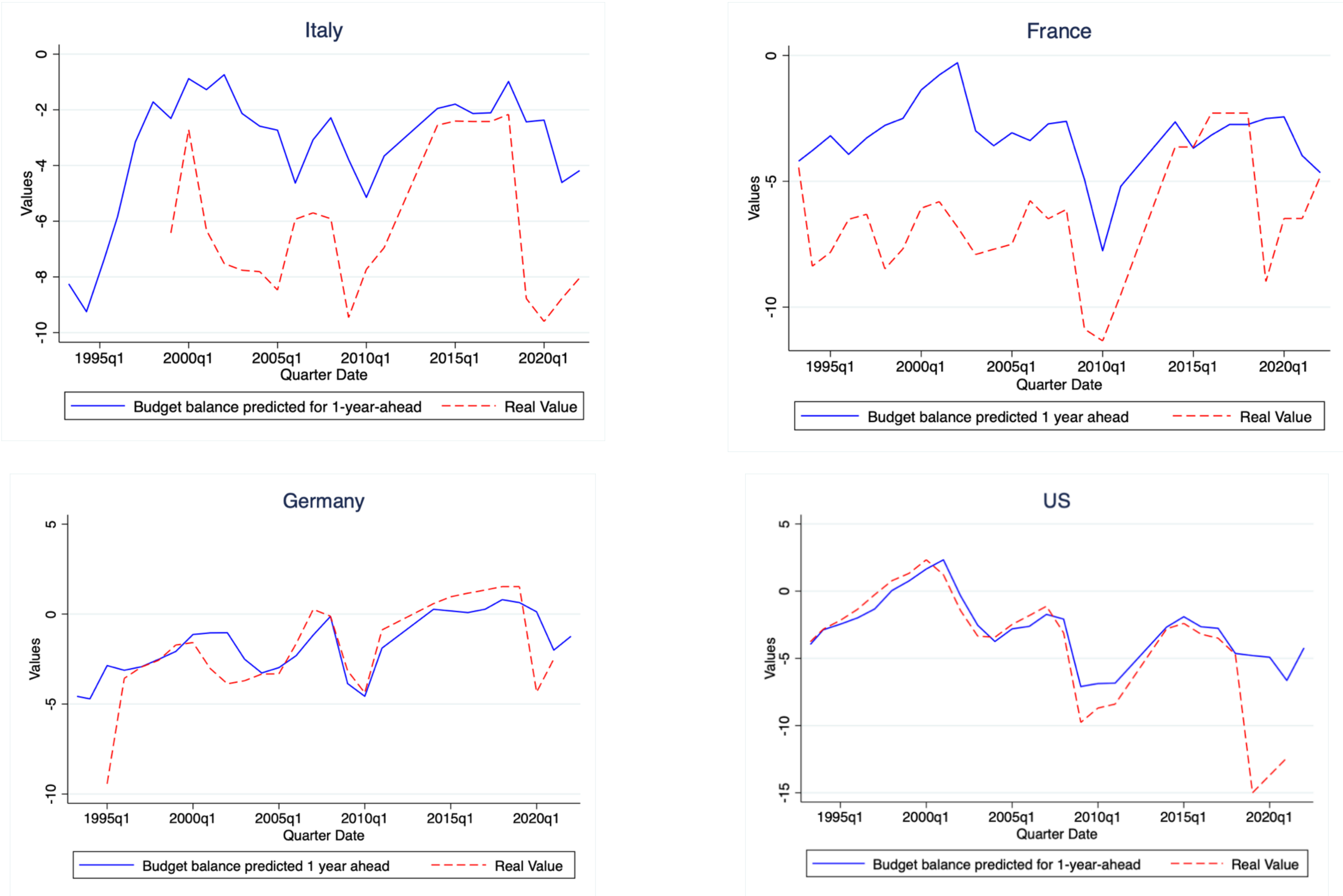


Table 1. Descriptive statistics for budget balance forecasts, overall and per group of expert forecasters.

		(a) current year						(b) year-ahead					
	Group	mean	s.e	min	max	Obs.	N of experts	mean	s.e	min	max	Obs.	N of experts
Italy	All	-4.31	(3.11)	-14.19	-0.88	422	17	-3.40	(1.73)	-10.44	-0.22	415	21
	Banks	-4.16	(3.23)	-14.19	-0.98	195	8	-3.11	(1.53)	-7.50	-0.67	125	8
	Consultants	-4.61	(2.97)	-13.67	-1.17	167	7	-3.76	(1.90)	-10.44	-0.71	155	8
	Financial services	-3.93	(2.64)	-13.06	-0.88	60	2	-3.26	(1.63)	-9.33	-0.22	135	5
	Research departments	-	-	-	-	-	-	-	-	-	-	-	-
France	All	-4.14	(2.11)	-12.78	-0.21	596	24	-3.49	(1.54)	-9.23	-0.19	436	23
	Banks	-4.13	(2.27)	-12.55	-0.24	209	8	-3.55	(1.51)	-8.53	-0.19	102	8
	Consultants	-3.90	(1.32)	-7.87	-2.67	357	13	-3.67	(1.67)	-9.23	-0.22	228	11
	Financial services	-3.90	(1.32)	-7.86	-2.67	15	1	-3.26	(1.11)	-5.75	-0.97	47	2
	Research departments	-3.59	(1.45)	-7.17	-2.34	15	1	-2.89	(1.18)	-5.80	-0.27	59	2
Germany	All	-1.78	(2.20)	-11.46	2.04	1,005	27	-1.76	(1.77)	-7.65	2.03	973	27
	Banks	-0.74	(2.31)	-10.10	2.04	94	5	-0.72	(1.82)	-5.42	1.45	49	5
	Consultants	-2.05	(2.21)	-11.46	1.16	663	14	-1.94	(1.73)	-7.27	1.56	662	14
	Financial services	-0.78	(1.70)	-4.01	1.25	35	1	-1.15	(1.51)	-4.59	1.05	35	1
	Research departments	-1.56	(1.97)	-5.62	1.75	213	7	-1.55	(1.81)	-7.65	2.03	227	7
US	All	-3.98	(3.40)	-20.26	3.15	626	26	-2.82	(2.61)	-12.61	3.42	580	30
	Banks	-4.32	(3.03)	-17.00	2.65	144	7	-2.96	(2.44)	-9.55	2.67	171	8
	Consultants	-4.53	(3.63)	-20.26	2.45	239	10	-3.62	(2.81)	-11.53	2.70	128	9
	Financial services	-3.10	(3.02)	-17.18	3.15	106	4	-2.29	(2.41)	-12.61	3.42	118	7
	Research departments	-3.36	(3.43)	-19.34	2.48	137	5	-2.44	(2.61)	-11.42	2.80	163	6

Sources: Consensus Economic Forecasts and authors' calculations.

III. An evaluation of expert forecasters' performance

i. Understanding bias in expert forecasters' budget forecast

The accuracy of government budget forecasts is critical for effective fiscal policy, for accountability in using public funds, and oversight from fiscal councils (Auerbach, 1999; Jonung and Larch, 2006). Nevertheless, evidence of systematic bias in fiscal forecasts (An and Jalles, 2021), and in particular the persistence of overly optimistic forecasts in public forecasts, has been a continuous concern (Estefania-Flores *et al.*, 2023; Perrelli *et al.,* 2024). On the one hand, bias might create the perception that budget forecasts are central to the political process, and there is pressure on forecasters to present favorable outcomes, potentially leading to biased predictions (Strauch *et al.*, 2004; Moulin and Wierts, 2006). On the other hand, major economic shocks like the Global Financial Crisis, the Pandemic or international conflict, might create unforeseen budgetary pressures. Assessing the performance of budget balance forecasts, and refining forecasting techniques should lead to better informed policy decisions (Leal *et al.*, 2008).

A first test for absolute forecasting performance is on bias. The standard test for bias by Holden and Peel (1990) examines whether the forecast error is zero on average in-sample. We define the forecast error as follows,

$$e_{i,t}^{h} = R_t - F_{i,t}^{h}, \qquad (1)$$

in which R_t is the realized value of the budget balance-to-GDP ratio for period t, and $F_{i,t}^{h}$ is the forecast budget balance ratio by an expert forecaster i at time t. The forecast error $e_{i,t}^{h}$ for each expert forecaster i can be produced for different horizons h, in our case either the current-year or at one-year-ahead.

A simple way to test bias is to regress the forecast error on a constant as follows[7]:

$$e_{i,t}^{h} = \alpha + v_{i,t}^{h} \qquad (2)$$

Forecasts are unbiased if we cannot reject the null hypothesis that $\alpha = 0$. A negative α indicates an optimistic bias. In this case, forecasters systematically underestimate the actual budget deficit (or overestimate the surplus). By contrast, a positive α suggests a pessimistic bias, where forecasts tend to overstate deficits or understate surpluses.

In order to compare our results to previous studies, we test (2) on the mean forecast error across all forecasters, for both the current-year and the year-ahead forecast, in all four countries. We additionally test the mean forecast for bias in each of the four subgroups of forecasters. We estimate model (2) with an OLS estimator correcting standard errors as in Driscoll and Kraay (1998) as there is a potential correlation structure.

The results in Table 2, panel (a) demonstrate notable differences in forecast bias of budget balances across countries and groups of experts. Forecasts for the Italian budget, for example, consistently shows large underestimations, especially in one-year-ahead forecasts, where the bias reaches -1.79 percent of GDP. France also displays a negative bias, with a drop in

[7] All tests are applied to forecasts of the budget balance as percent of GDP. Tests on an alternative measure as the real local currency budget balance led to qualitatively similar results. Results are available upon request.

performance when looking at the year-ahead forecasts. By contrast, forecasts for Germany shows an overall slightly positive bias for current-year forecasts (+0.19 percent), but it turns negative for the one-year-ahead projections. The US has a similar bias in current-year predictions as in Germany. Bias is much more pronounced for one-year-ahead forecasts overall, which is to be expected given the larger uncertainty that surrounds longer-term projections.

Table 2. Testing bias budget balance forecast, overall and per group of expert forecasters.

(a) in the mean

		current year		year-ahead	
Country	Groups	α	s.e.	α	s.e.
Italy	All	-0.46***	(0.10)	-1.79***	(0.14)
	Banks	-0.29**	(0.14)	-1.86***	(0.25)
	Consultants	-0.45***	(0.15)	-1.90***	(0.24)
	Financial services	-1.06***	(0.36)	-1.60***	(0.28)
	Research departments”	-	-	-	-
France	All	-0.21**	(0.10)	-0.74***	(0.14)
	Banks	-0.19	(0.17)	-0.86***	(0.28)
	Consultants	-0.17	(0.11)	-0.72***	(0.20)
	Financial services	-0.83	(0.48)	-0.54	(0.41)
	Research departments	-0.79*	(0.43)	-0.76*	(0.39)
Germany	All	0.19***	(0.04)	-0.45***	(0.05)
	Banks	0.45***	(0.11)	0.01	(0.29)
	Consultants	0.24***	(0.06)	-0.52***	(0.07)
	Financial services	0.36**	(0.16)	-0.14	(0.28)
	Research departments	-0.10	(0.10)	-0.41***	(0.13)
US	All	0.04	(0.05)	-0.76***	(0.10)
	Banks	-0.20*	(0.12)	-0.68***	(0.17)
	Consultants	0.12	(0.08)	-1.43***	(0.27)
	Financial services	0.01	(0.05)	-0.62***	(0.19)
	Research departments	0.16**	(0.71)	-0.45***	(0.16)

Sources: Consensus Economic Forecasts and authors' estimations.

Notes: OLS estimation. Standard robust errors in parentheses *** p<0.01, ** p<0.05, * p<0.1; ” no research institutes in Italy over full sample.

(b) in panel model

		current year		year-ahead	
Country	Groups	α	s.e.	α	s.e.
Italy	All	-2.11***	(0.52)	-3.56***	(0.48)
	Banks	-2.01***	(0.52)	-3.30***	(0.57)
	Consultants	-1.82**	(0.49)	-3.15***	(0.60)
	Financial services	-2.89	(0.84)	-4.24***	(0.55)
	Research departments"	-	-	-	
France	All	-2.02***	(0.49)	-3.69***	(0.46)
	Banks	-2.37***	(0.58)	-3.98***	(0.46)
	Consultants	-1.86***	(0.45)	-3.79***	(0.52)
	Financial services	0.00	(0.00)	-2.59	(0.87)
	Research departments	0.00	(0.00)	-3.77	(0.63)
Germany	All	0.02	(0.31)	-0.49	(0.38)
	Banks	0.14	(0.52)	-0.05	(0.51)
	Consultants	0.07	(0.29)	-0.53	(0.39)
	Financial services	0.03	(0.69)	0.02	(0.86)
	Research departments	-0.18	(0.36)	-0.55	(0.40)
US	All	-0.08	(0.24)	-0.75**	(0.34)
	Banks	-0.70	(0.61)	-0.90**	(0.38)
	Consultants	0.09	(0.26)	-1.23**	(0.53)
	Financial services	-0.03	(0.16)	-0.61*	(0.31)
	Research departments	0.24	(0.14)	-0.30	(0.35)

Sources: Consensus Economic Forecasts and authors’ estimations.
Notes: Fixed effects panel model. Standard robust errors in parentheses *** p<0.01, ** p<0.05, * p<0.1; fixed effects panel including time dummies.

Across the subgroups of experts, some differences in bias become clear, reflecting the varying forecasting abilities of different institutions. Financial institutions in Italy show the largest negative bias, particularly for current-year forecasts (-1.06), and so they are in France. By contrast, Germany's international banks stand out with a strong positive bias in current-year forecasts (+0.45). But in the US, international banks tend to be very negatively biased, while research departments tend to be overly pessimistic. The year-ahead forecasts are more strongly biased, and forecasters in all subgroups tend to expect a smaller deficit than what actually happens. Only the German international banks tend to get the forecast about right, in stark contrast to the mostly negative results for groups.

In general, results in panel (a) of Table 2 align with previous research on budget forecasts. Strauch *et al.* (2004) show a statistically significant optimism bias in EU countries; Jalles *et al.* (2015) find statistically significant biases in fiscal forecasts on advanced economies using the same data but in a shorter period of time (from 1993-2009). Yet, when these authors investigate forecasts by country they find a positive bias for current forecast in US, Germany and Italy, and a negative bias in France, while for year-ahead forecast all negative bias.

Most of the literature on fiscal forecasts evaluates bias for specific institutions' forecasts, such as the IMF or the EC, as in Artis and Marcellino (2001) or Frankel (2011). When multiple forecasts from the same source are tested – as in Jalles *et al.* (2015) or An and Jalles (2020) – often only the mean forecast is tested for bias, rather than the entire distribution.[8]

Testing the bias on the mean forecast has a drawback, though. Carabotta and Claeys (2024) show that different forecast combinations-of which the mean is just one-smooths out differences across forecasters, but also their different biases. We therefore use the entire panel of expert forecasters to test for bias, and estimate (2) on the panel of all expert forecasters using fixed effects and correcting standard errors as in Driscoll and Kraay (1998). This correction does not require strong assumptions on the form of the cross-sectional and temporal correlation in the error terms (Dovern *et al.*, 2015). We also include time dummies to account for potential common shocks or events that jointly shift the budget forecasts across experts. We report in Table 2 results for both the full sample of expert forecasters, and for the subgroup samples.

Results in panel (b) of Table 2 show a complementary picture of the mean results in panel (a). The fixed-effect panel data shows significantly higher negative biases (both for the current-year and year-ahead forecast) in Italy and France. Even if we look at specific groups like banks or consultants, we observe this increase in bias. By contrast to the forecasts for France or Italy, the German expert forecasters never display a bias (at both horizons): none of the coefficients is statistically significant. Something similar is true for the US budget balance forecasts in the current-year. Yet, for the year-ahead forecast an optimistic bias is obtained. These results add insight into mean forecasts, as it shows that taking into account individual forecaster and time fixed effects, leaves us with a similar insight. Expert forecasters-on average, and on a full sample-display a bias even if there is a wide dispersion across them. This shows that averaging out over different information sets across forecasters can be useful to improve forecasting performance, as Carabotta and Claeys (2024) show. However, it does not reflect how expert forecasters update their information, or what type of information is taken into account.

[8] Jalles *et al.* (2015) test bias on the mean budget balance-to GDP ratio in advanced and emerging countries over the period 1993-2009. An and Jalles (2020) evaluate the performance of individual private and public forecasters in the US. Carabotta and Claeys (2024) test absolute and relative forecasting performance of private and public expert forecasters for the Italian budget.

ii. Analysis of information rigidities in budget forecasts

Under rational expectations theory, professional forecasters are expected to use all available information and process it correctly and without any systematic bias. The theory implies that revisions to forecasts should be uncorrelated because if forecasters are rational, they would have already incorporated all relevant information into their forecasts (Nordhaus, 1987). This scenario assumes no informational frictions.

A large literature in the macroeconomics field shows that, contrary to theory, in practice, forecast revisions tend to be positively correlated, such as shown in Nordhaus (1987) or Coibion and Gorodnichenko (2012) who show that revisions to forecasts often contain a predictable component, suggesting that forecasters face frictions in processing information (Cornand and Hubert, 2020). Such frictions lead to stickiness in forecasts. The source of this stickiness is not clear, however. On the one hand, the processing of information itself may be sticky: Mankiw and Reis (2002) posit that forecasters update their information sets infrequently, either due to the costs of acquiring and processing new information or due to cognitive biases, leading to a lag in the full assimilation of new data. On the other hand, agents have to extract from a large set of information those changes that could modify economic developments in the future. Noisy information could make agents only slowly adapt their forecasts as they have to learn what information is relevant and distinguish it from irrelevant noise.

In both cases, information processing makes forecasters, even expert ones, slow to adjust. One test for measuring stickiness in information is the one developed by Dovern *et al.* (2015). They argue that as forecasters receive new information, they should revise their new forecast. The test suggests testing forecast smoothing by regressing forecast revisions on past forecast revisions, as in (3). Revisions are calculated as follows: $r_{i,t}^h = F_{i,t}^h - F_{i,t-1}^h$, where F_t^h is the forecast made by expert forecaster i on horizon h (current-year or one-year-ahead) at time t, and F_{t-1}^h was the forecast made in the previous period.

$$r_{i,t}^h = \lambda r_{i,t-1}^h + \upsilon_{i,t-1}^h. \qquad (3)$$

In (3), the parameter λ is the estimated autocorrelation on the forecast revision, and $\upsilon_{i,t-1}^{C,h}$ an independent and identically distributed error term. As Dovern *et al.* (2015) argue, when λ=0, forecasts are (weakly) efficient-i.e., there is no stickiness and revisions are updated.

We first estimate equation (3) on the mean expert forecast (for each country, and across subgroups), and then estimate a panel fixed effects model on the sample of all expert forecasters to account for their singular forecast behavior. The mean test for forecast efficiency in (3) is estimated with OLS, whereas the panel version is estimated with a fixed-effects model. In the panel, we also include time dummies to account for common shocks that affect all forecasters in every quarter.

Table 3. Standard test of forecast stickiness.

(a) in the mean

	current-year				year-ahead			
	Italy	France	Germany	US	Italy	France	Germany	US
λ	-0.22***	-0.14***	-0.17***	-0.16***	-0.30***	-0.13**	-0.14***	-0.18***
	(0.06)	(0.05)	(0.03)	(0.04)	(0.06)	(0.06)	(0.04)	(0.05)
Obs.	311	522	857	522	310	361	819	461
R^2	0.04	0.02	0.03	0.02	0.1	0.01	0.02	0.02

Sources: Consensus Economic Forecasts and authors' estimations.
Notes: OLS estimation. Standard robust errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

(b) in panel model

	current-year				year-ahead			
	Italy	France	Germany	US	Italy	France	Germany	US
λ	-0.18*	-0.15*	-0.18**	-0.18	-0.29***	-0.14	-0.16	-0.21**
	(0.09)	(0.08)	(0.07)	(0.05)	(0.05)	(0.13)	(0.12)	(0.16)
Obs.	311	522	857	522	310	361	819	461
N of experts	17	24	27	26	21	23	38	30
Within R^2	0.09	0.02	0.04	0.02	0.12	0.04	0.04	0.05

Sources: Consensus Economic Forecasts and authors' estimations.
Notes: Fixed effects panel model. Standard robust errors in parentheses *** p<0.01, ** p<0.05, * p<0.1; fixed effects panel including time dummies.

Panel (a) of Table 3 shows the results of testing stickiness in the mean. We observe a clear pattern of stickiness on both horizons. In all cases, the coefficient λ is found to be negative and statistically significant, indicating that forecasters tend to over-adjust their predictions when faced with new information, which leads to persistent forecast errors. There is no evidence that this pattern is different for longer term projections. The stickiness coefficient is not significantly different from the current-year measure of expectations.

Dovern *et al*. (2015) argue that he degree of information rigidity in the mean forecasts is substantially higher than that in individual forecasts. A comparison of the mean versus the panel version of the efficiency test in panel (b) of Table 3 confirms the findings on the mean forecast though. The stickiness coefficient is quite similar, even if the response is not always significant for all countries. This result indicates that budget forecasts behave differently than GDP or inflation forecasts.

These results might be surprising given that most studies on GDP or inflation forecasts find a substantial amount of rigidity (Coibion and Gorodnichenko, 2012). By contrast, the results indicates that forecasters tend to overreact to new information. This would induce a mean-reverting behavior in the forecasts. If forecasters make an upward revision in one period, the next period might see a downward revision, moving back toward the mean or "true" value. This behavior suggests that forecasts are not smoothly adjusting to the correct values but rather oscillating. Results are not a specific outcome of the panel model; even in a model without time dummies, the same degree of stickiness is found.[9] Jalles *et al.* (2015) examine rigidities in mean fiscal forecast (revisions) on the CE data for the first half of our sample, and in all cases find a mean strong positive correlation among revisions.

iii. Is stickiness influenced by information?

The drivers of stickiness of forecasts could be several. As theory postulates, it might reflect genuine rigidities in information processing, or efforts to understand new information. In the latter case, lack of efficiency in fiscal forecasts could be explained by economic and political conditions, and not only by autocorrelation in the forecast revisions.

The kind of economic or political information that expert forecasters digest is varied. On the one hand, changes in GDP growth, inflation rates, or policy decisions directly affect the fiscal outlook. For example, during an economic downturn, governments run up deficits due to lower tax revenues while simultaneously facing higher spending on social welfare programs. Higher inflation or rising interest rates can lead to increased debt servicing costs. On the other hand, political conditions also shape expert forecasters expectations. Governments often adjust their fiscal policies in response to political pressures or upcoming elections. Beetsma *et al.* (2013) study budget forecast errors from national Stability and Convergence Programs and find that political factors, specifically upcoming elections, are import factors that explain optimism in expectations. Pina and Venes (2011) also highlight the importance of upcoming elections in optimistic fiscal forecasts.

Some papers analyze the disagreement between market experts, such as Dovern *et al.* (2012), which document multivariate forecast disagreement about GDP growth among expert forecasters of the Euro area economy and discuss implications for models of heterogeneous expectation formation. Poplawski-Ribeiro and Ruelke (2011) also analyze the dispersion of financial market forecasts on government budget deficits in France, Germany, Italy, and the UK and how the Stability and Growth Pact (SGP) fiscal rule changed those. They find that accuracy of financial expert deficit forecasts significantly increased in France with the introduction of the SGP. Yet, the heightened monitoring of fiscal figures on account of the SGP rules were not statistically significant in explaining accuracy of fiscal forecasts from market experts in other countries analyzed.

We suggest adding to model (3) additional control variables for all countries that contain news on economic and political developments. We add three political variables-the economic policy uncertainty index by Baker *et al.* (2016), a geopolitical risk index by Caldara and Iacoviello (2022), and a dummy series for elections[10]-and two macroeconomic variables, the output gap and realized interest rate of governments bonds[11]. We again look at the impact on the mean forecast (per country, for the current-year and the year-ahead), and at the panel of expert forecasters (per country and per group of experts). Panel (a) of Table 4 reports an OLS regression, while panel (b) shows a fixed effect panel model including time dummies.

[9] Results are available upon request. The only difference is that significance is reduced.
[10] Presidential elections in the US and France, federal Bundestag elections in Germany, and parliamentary general elections in Italy.
[11] The output gap and real interest rate are collected from World Bank database

Table 4. Augmented test of forecast stickiness.

(a) in the mean

	current-year				year-ahead			
	Italy	France	Germany	US	Italy	France	Germany	US
λ	-0.27***	-0.24***	-0.28***	-0.19***	-0.35***	-0.20***	-0.26***	-0.34***
	(0.05)	(0.04)	(0.03)	(0.05)	(0.06)	(0.06)	(0.04)	(0.05)
Interest rate	-0.05	0.02	0.09***	-0.03	0.01	0.06*	0.03*	0.13***
	(0.09)	(0.06)	(0.02)	(0.09)	(0.05)	(0.03)	(0.02)	(0.05)
Output gap	0.42***	0.36***	0.42***	0.12**	0.29***	0.17***	0.26***	0.07**
	(0.07)	(0.06)	(0.03)	(0.05)	(0.04)	(0.04)	(0.02)	(0.03)
Policy uncertainty	-0.00	0.00**	-0.00	-0.01***	-0.00	0.00	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Global political risk	2.78***	0.25	0.57***	0.44***	1.92***	0.10	0.51***	0.11**
	(1.05)	(0.20)	(0.21)	(0.12)	(0.63)	(0.16)	(0.17)	(0.05)
Elections	0.67**	0.30*	0.33***	-0.88***	-0.17	-0.44***	0.15*	-0.12
	(0.30)	(0.18)	(0.10)	(0.28)	(0.21)	(0.13)	(0.08)	(0.15)
Obs.	311	522	857	522	310	361	819	461
R^2	0.22	0.14	0.20	0.12	0.25	0.11	0.17	0.17

Sources: Consensus Economic Forecasts and authors' estimations.
Notes: OLS estimation. Standard robust errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

(b) in panel model

	current-year				year-ahead			
	Italy	France	Germany	US	Italy	France	Germany	US
λ	-0.26***	-0.24***	-0.29***	-0.22**	-0.34***	-0.23	-0.26*	-0.36**
	(0.04)	(0.08)	(0.10)	(0.09)	(0.08)	(0.16)	(0.14)	(0.14)
Interest rate	-0.06	0.08	0.10	0.05	0.01	0.04	0.01	0.16***
	(0.14)	(0.10)	(0.07)	(0.16)	(0.09)	(0.07)	(0.05)	(0.06)
Output gap	0.45**	0.41*	0.43**	0.14	0.31**	0.21**	0.28***	0.09
	(0.19)	(0.22)	(0.20)	(0.14)	(0.13)	(0.10)	(0.10)	(0.07)
Policy uncertainty	-0.00	0.00*	-0.00	-0.01*	-0.00	0.00	-0.00***	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Global political risk	1.87	0.19	0.55	0.49	1.46**	0.05	0.51	0.10
	(1.74)	(0.24)	(0.33)	(0.34)	(0.57)	(0.23)	(0.33)	(0.10)
Elections	0.78*	0.29	0.34	-0.86	-0.18	-0.44	0.20	-0.13
	(0.37)	(0.29)	(0.23)	(0.61)	(0.25)	(0.29)	(0.20)	(0.28)
Obs.	311	522	857	522	310	361	819	461
N of experts	17	24	27	26	21	23	38	30
R^2	0.302	0.186	0.215	0.149	0.319	0.138	0.187	0.198

Sources: Consensus Economic Forecasts and authors' estimations.
Notes: Fixed effects panel model. Standard robust errors in parentheses *** p<0.01, ** p<0.05, * p<0.1; fixed effects panel including time dummies.

If we first turn to panel (a), the stickiness coefficient remains negative and statistically significant for all countries, and there is also no appreciable difference between current-year and year-ahead forecasts again. Results are not greatly affected by the control for political and macroeconomic variables, yet their impact on the revision is significant (and raises explanatory power of the forecast revisions significantly). An economic boom (a higher output gap), elections and global political risk tend to let forecasters update more strongly their forecast, which should not come as a surprise for budget forecasts.

Panel (b) shows a slightly different picture. The adjustment in forecasts is still negative, yet the impact of economic or political variables is much more limited. The impact of the economic climate, particularly the output gap, is incorporated by forecasters made year-ahead. Global political risk has positive impact on revisions in all countries, except in France. This result suggests that information rigidities for budget forecasts are mostly due to inherent rigidities in the way expert forecasters process information, and implies that different forecasters do not incorporate all the news they could in their set of information.

It is possible that banks, financial services, consultants, or research departments have different incentives to digest economic or political information more or less rapidly. Some are focused on economic monitoring, and producing forecasts is not just a by-product of their main business. We saw in Table 2 that bias differs across expert forecasters; and so may information rigidities. We therefore test model (3) – adding the control variables – with a panel fixed effects model on each group of forecasters to examine their exposure to political and economic information.

Table 5, in panels (a) and (b), shows the results for an OLS on the mean budget balance forecast for the current year and for the year ahead, respectively. Panel (a) shows that overall stickiness is smallest for research departments. This is to be expected given that they are usually focused (or obliged) to analyze the budget. There are few significant differences across the other groups. We also observe that additional economic or political information does not substantially modify the adjustment coefficient, nor are they often significant. So, the results suggest that expert forecasters are behaving rather similarly. As per the previous set of results in panel (b) of Table 4, panel (b) of Table 5 shows that accounting with a panel model for different types of expert forecasters does not substantially modify the outcomes. While the parameter λ is not always significant, it is in the same range for all forecasters, with the exception of research departments, and economic or political variables matter little. Panels (c) and (d) in Table 5 analyze the presence of stickiness in panel datasets for current-year and year-ahead forecasts, with a fixed effect model, including time dummies, and broadly confirm the previous findings.

Table 5. Augment test of stickiness per group of expert forecasters

(a) in the mean – current year

	Italy			France				Germany				US			
	Banks	Consultants	Financial Services	Banks	Consultants	Financial Services	Research Departments	Banks	Consultants	Financial Services	Research Departments	Banks	Consultants	Financial Services	Research Departments
λ	-0.31***	-0.26**	-0.28***	-0.25***	-0.05	-0.20	-0.17	-0.18	-0.22***	-0.15	-0.30***	-0.35***	-0.15	-0.39***	-0.31***
	(0.08)	(0.10)	(0.10)	(0.07)	(0.09)	(0.15)	(0.17)	(0.13)	(0.05)	(0.21)	(0.08)	(0.08)	(0.13)	(0.12)	(0.10)
Interest rate	-0.30**	-0.01	0.09	-0.11	0.00	0.06	-0.07	-0.05	0.04**	-0.22	0.03	0.00	0.16*	0.23***	0.07
	(0.15)	(0.10)	(0.06)	(0.10)	(0.05)	(0.09)	(0.06)	(0.36)	(0.02)	(0.21)	(0.04)	(0.22)	(0.09)	(0.08)	(0.11)
Output gap	0.61***	0.40***	0.12**	0.41***	0.19***	0.19	-0.05	1.00***	0.24***	0.25	0.30***	-0.02	-0.07	0.07	0.25***
	(0.11)	(0.09)	(0.05)	(0.10)	(0.05)	(0.12)	(0.07)	(0.24)	(0.03)	(0.17)	(0.05)	(0.09)	(0.05)	(0.08)	(0.08)
Policy uncertainty	0.00	-0.00	-0.00	0.00	0.00	0.00	-0.00	0.00	-0.00***	-0.00	-0.00***	-0.01*	-0.00	-0.00	-0.00
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Global political risk	3.10*	2.29*	0.85	0.25	-0.08	0.78*	-0.15	0.49	0.55***	-0.72	0.76**	0.46**	0.02	0.03	0.20
	(1.61)	(1.22)	(0.82)	(0.32)	(0.27)	(0.40)	(0.20)	(0.89)	(0.19)	(1.46)	(0.37)	(0.23)	(0.09)	(0.12)	(0.13)
Elections	0.90*	-0.40	-0.05	0.28	-0.26	-0.55	-0.83***	-0.37	0.18**	-0.55	0.25	-0.98*	0.10	0.31	-0.13
	(0.46)	(0.52)	(0.28)	(0.29)	(0.18)	(0.39)	(0.27)	(0.76)	(0.09)	(1.14)	(0.18)	(0.52)	(0.31)	(0.28)	(0.35)
Obs.	148	118	43	182	275	15	16	63	555	30	202	138	187	91	128
R^2	0.30	0.32	0.15	0.14	0.08	0.21	0.43	0.31	0.17	0.21	0.22	0.19	0.10	0.21	0.19

Sources: Consensus Economic Forecasts and authors' estimations.
Notes: OLS estimation. Standard robust errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

(b) in the mean – year ahead

	Italy			France				Germany				US			
	Banks	Consultants	Financial Services	Banks	Consultants	Financial Services	Research Departments	Banks	Consultants	Financial Services	Research Departments	Banks	Consultants	Financial Services	Research Departments
λ	-0.48***	-0.26**	-0.28***	-0.46***	-0.05	-0.20	-0.17	-0.47***	-0.22***	-0.15	-0.30***	-0.50***	-0.15	-0.39***	-0.31***
	(0.09)	(0.10)	(0.10)	(0.11)	(0.09)	(0.15)	(0.17)	(0.16)	(0.05)	(0.21)	(0.08)	(0.09)	(0.13)	(0.12)	(0.10)
Interest rate	-0.20*	-0.01	0.09	0.16**	0.00	0.06	-0.07	0.18	0.04**	-0.22	0.03	0.10	0.16*	0.23***	0.07
	(0.11)	(0.10)	(0.06)	(0.07)	(0.05)	(0.09)	(0.06)	(0.83)	(0.02)	(0.21)	(0.04)	(0.09)	(0.09)	(0.08)	(0.11)
Output gap	0.48***	0.40***	0.12**	0.35***	0.19***	0.19	-0.05	0.62**	0.24***	0.25	0.30***	0.07	-0.07	0.07	0.25***
	(0.09)	(0.09)	(0.05)	(0.09)	(0.05)	(0.12)	(0.07)	(0.29)	(0.03)	(0.17)	(0.05)	(0.05)	(0.05)	(0.08)	(0.08)
Policy uncertainty	-0.00	-0.00	-0.00	0.00**	0.00	0.00	-0.00	0.00	-0.00***	-0.00	-0.00***	-0.01***	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Global political risk	2.70**	2.29*	0.85	0.25	-0.08	0.78*	-0.15	-0.90	0.55***	-0.72	0.76**	0.14	0.02	0.03	0.20
	(1.29)	(1.22)	(0.82)	(0.44)	(0.27)	(0.40)	(0.20)	(1.92)	(0.19)	(1.46)	(0.37)	(0.09)	(0.09)	(0.12)	(0.13)
Elections	-0.28	-0.40	-0.05	-0.76**	-0.26	-0.55	-0.83***	-0.74	0.18**	-0.55	0.25	-0.40	0.10	0.31	-0.13
	(0.39)	(0.52)	(0.28)	(0.31)	(0.18)	(0.39)	(0.27)	(1.04)	(0.09)	(1.14)	(0.18)	(0.26)	(0.31)	(0.28)	(0.35)
Obs.	105	90	115	88	181	43	49	31	569	30	184	149	88	94	125
R^2	0.36	0.32	0.15	0.29	0.08	0.21	0.43	0.35	0.17	0.21	0.22	0.32	0.10	0.21	0.19

Sources: Consensus Economic Forecasts and authors' estimations.
Notes: OLS estimation. Standard robust errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

(c) in panel model – current year

	Italy			France				Germany				US			
	Banks	Consultants	Financial Services	Banks	Consultants	Financial Services	Research Departments	Banks	Consultants	Financial Services	Research Departments	Banks	Consultants	Financial Services	Research Departments
λ	-0.27***	-0.16**	-0.55	-0.24***	-0.27**	-0.08	0.16	-0.10	-0.34**	-0.12	-0.20	-0.37**	-0.17	-0.51	-0.03
	(0.05)	(0.06)	(0.09)	(0.06)	(0.09)	(0.09)	(0.14)	(0.08)	(0.11)	(0.13)	(0.11)	(0.11)	(0.12)	(0.26)	(0.11)
Interest rate	-0.28	0.03	0.26	-0.10	0.13	1.34	1.07	-0.11	0.10	-0.11	0.10	0.20	0.17	-0.04	-0.07
	(0.16)	(0.18)	(0.26)	(0.15)	(0.10)	(0.50)	(0.12)	(0.39)	(0.08)	(0.10)	(0.05)	(0.29)	(0.36)	(0.23)	(0.20)
Output gap	0.61**	0.33**	0.16	0.48*	0.43*	0.24	0.31	0.90*	0.44*	0.36	0.31**	-0.03	0.18	0.14	0.14
	(0.25)	(0.13)	(0.12)	(0.22)	(0.24)	(0.18)	(0.12)	(0.34)	(0.23)	(0.14)	(0.12)	(0.15)	(0.17)	(0.25)	(0.22)
Policy uncertainty	0.00	-0.00	-0.02	0.00	0.00	0.01	-0.00	0.00	-0.00	-0.00	0.00	-0.01	-0.01	-0.01*	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)
Global political risk	1.78	1.95	3.13	0.15	0.25	0.84	-0.60	-0.15	0.45	0.74	0.83**	0.56	0.67	0.02	0.19
	(1.97)	(1.72)	(1.27)	(0.26)	(0.29)	(0.42)	(0.31)	(0.73)	(0.40)	(0.82)	(0.31)	(0.35)	(0.41)	(0.29)	(0.27)
Elections	0.89	0.74**	0.97	0.19	0.25	-1.51	-0.33	-0.08	0.42	-0.16	0.13	-0.76	-1.60*	-0.17	0.10
	(0.49)	(0.30)	(0.57)	(0.27)	(0.34)	(0.86)	(0.35)	(0.58)	(0.31)	(0.47)	(0.13)	(0.73)	(0.86)	(0.49)	(0.28)
Obs.	148	118	43	182	275	15	16	63	555	30	202	138	187	91	128
N of experts	8	7	2	8	13	1	1	5	14	1	7	7	10	4	5
Within R^2	0.360	0.308	0.455	0.196	0.199	0.557	0.927	0.424	0.225	0.448	0.225	0.219	0.248	0.264	0.0522

Sources: Consensus Economic Forecasts and authors' estimations.
Notes: Fixed effects panel model and OLS where N of experts is one. Standard robust errors in parentheses *** p<0.01, ** p<0.05, * p<0.1; fixed effects panel including time dummies.

(d) in panel model – year-ahead

	Italy			France				Germany				US			
	Banks	Consultants	Financial Services	Banks	Consultants	Financial Services	Research Departments	Banks	Consultants	Financial Services	Research Departments	Banks	Consultants	Financial Services	Research Departments
λ	-0.51***	-0.28***	-0.18	-0.50**	-0.05	-0.20	-0.19	-0.52***	-0.22	-0.09	-0.31	-0.53**	-0.14	-0.52	-0.30*
	(0.12)	(0.08)	(0.09)	(0.20)	(0.11)	(0.12)	(0.15)	(0.06)	(0.16)	(0.15)	(0.20)	(0.16)	(0.09)	(0.28)	(0.13)
Interest rate	-0.20	0.05	0.09	-0.08	-0.01	0.04	-0.05	0.39	0.02	-0.27	-0.00	0.23**	0.17**	0.20**	0.12
	(0.11)	(0.13)	(0.09)	(0.10)	(0.10)	(0.07)	(0.08)	(0.41)	(0.05)	(0.18)	(0.05)	(0.08)	(0.06)	(0.05)	(0.10)
Output gap	0.57*	0.42***	0.13	0.46**	0.19*	0.17	-0.03	0.71**	0.26**	0.24	0.30**	0.07	-0.05	0.14	0.22
	(0.25)	(0.11)	(0.06)	(0.17)	(0.09)	(0.08)	(0.05)	(0.18)	(0.10)	(0.11)	(0.11)	(0.08)	(0.04)	(0.08)	(0.16)
Policy uncertainty	-0.00	-0.00	-0.00	0.00*	0.00	-0.00	-0.00	0.00	-0.00***	-0.00	-0.00***	-0.01**	-0.00*	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Global political risk	2.60*	2.54**	-0.28	-0.12	0.04	0.89	-0.07	-0.37	0.55*	-0.64	0.75	0.17	0.03	-0.07	0.23
	(1.20)	(0.99)	(0.46)	(0.36)	(0.29)	(0.23)	(0.12)	(1.22)	(0.29)	(0.94)	(0.42)	(0.10)	(0.08)	(0.10)	(0.15)
Elections	-0.52	-0.61	0.11	-0.62**	-0.27	-0.58	-0.87	-1.42	0.22	-0.74	0.27	-0.34	0.10	0.26	-0.12
	(0.46)	(0.45)	(0.15)	(0.22)	(0.27)	(0.47)	(0.29)	(0.90)	(0.19)	(0.41)	(0.25)	(0.32)	(0.32)	(0.25)	(0.29)
Obs.	105	90	115	88	181	43	49	31	569	30	184	149	88	94	125
N of experts	8	8	5	8	11	2	2	5	14	1	7	8	9	7	6
Within R^2	0.433	0.375	0.324	0.408	0.106	0.296	0.482	0.477	0.192	0.238	0.220	0.364	0.148	0.282	0.218

Sources: Consensus Economic Forecasts and authors' estimations.
Notes: Fixed effects panel model and OLS where N of experts is one. Standard robust errors in parentheses *** p<0.01, ** p<0.05, * p<0.1; fixed effects panel including time dummies.

IV. Rationality of forecasters: a time-varying test

i. Time-variation in rationality applying the FR test

Economic forecasts are inherently dynamic. Rational expectations mean that agents update their expectations based on all available information, and take optimal decisions without bias and systematic errors. As we just showed, expert forecasts do not produce forecasts in a rational way for budget balance forecast in our sample as relevant macroeconomic or political factors are ignored. We would therefore endorse macroeconomic theories that argue agents may receive information only with a delay, so that forecasts adapt only slowly to incoming information. This sticky information makes agents only slowly adapt their forecasts.

Fiscal policy forecasts are particularly interesting in this regard. The lack of success in producing forecasts for budget variables is not only the result of an optimistic bias from public forecasts. Structural changes, both in the budget and the economy, can come from various sources, and complicate budget forecasts (Leal *et al.*, 2008).

We now go one step further and innovate on this field by adapting a test by Rossi and Sekhposyan's (2016) for time variation in forecast rationality. Their Fluctuation Rationality test allows us to detect periods where deviations from rationality breaks down, even if the overall sample average suggests otherwise. By analyzing the time variation in forecasting performance, we can detect and test over- or underestimations during specific sub-periods, or specific adaptations in the forecaster's behavior.

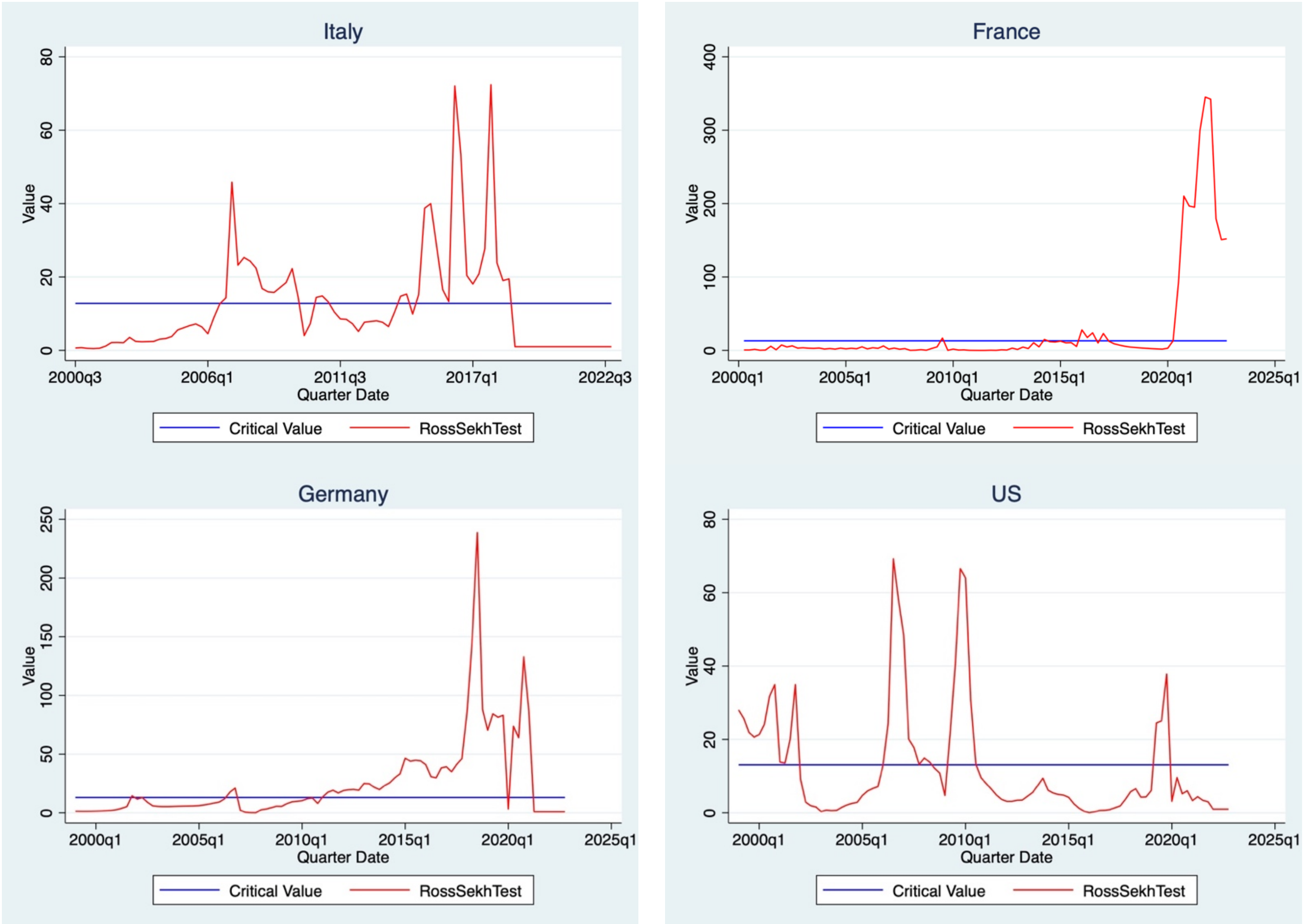
The FR test evaluates rationality over time using rolling windows estimates of the forecast values (current year and year-ahead) and comparing it with the real value of budget balance. As in similar parametric tests for structural instability, the test statistic is based on a Wald test, but must be adjusted for the fact that we test forecasts. Under the null of rationality (i.e., that forecast errors are unbiased and serially uncorrelated), the test statistic follows a χ^2 distribution (with degrees of freedom equal to the number of restrictions tested). The null hypothesis of stability is discarded if the supremum value of the Wald test statistic series $\theta_{i,t}$ exceeds the critical threshold. In the presence of parameter estimation error, the variance of the test statistic needs to be corrected using Newey-West estimators to account for heteroskedasticity and autocorrelation. The variance correction is applied based on the length of the rolling window and the number of lags (Rossi and Sekhposyan, 2016).[12]

To examine time-varying forecast rationality, we run FR test with a window of 24 for both current year and year-ahead forecasts. Figure 2 shows the FR test values and the critical value: if the FR test value (red line) exceeds the critical value (blue line) the forecast is deviating from rationality. Panel (a) shows test values for the mean current-year forecast, and panel (b) the one-year-ahead forecast. The interesting finding that emerges from these graphs is that forecasters do indeed change their behavior significantly over time.

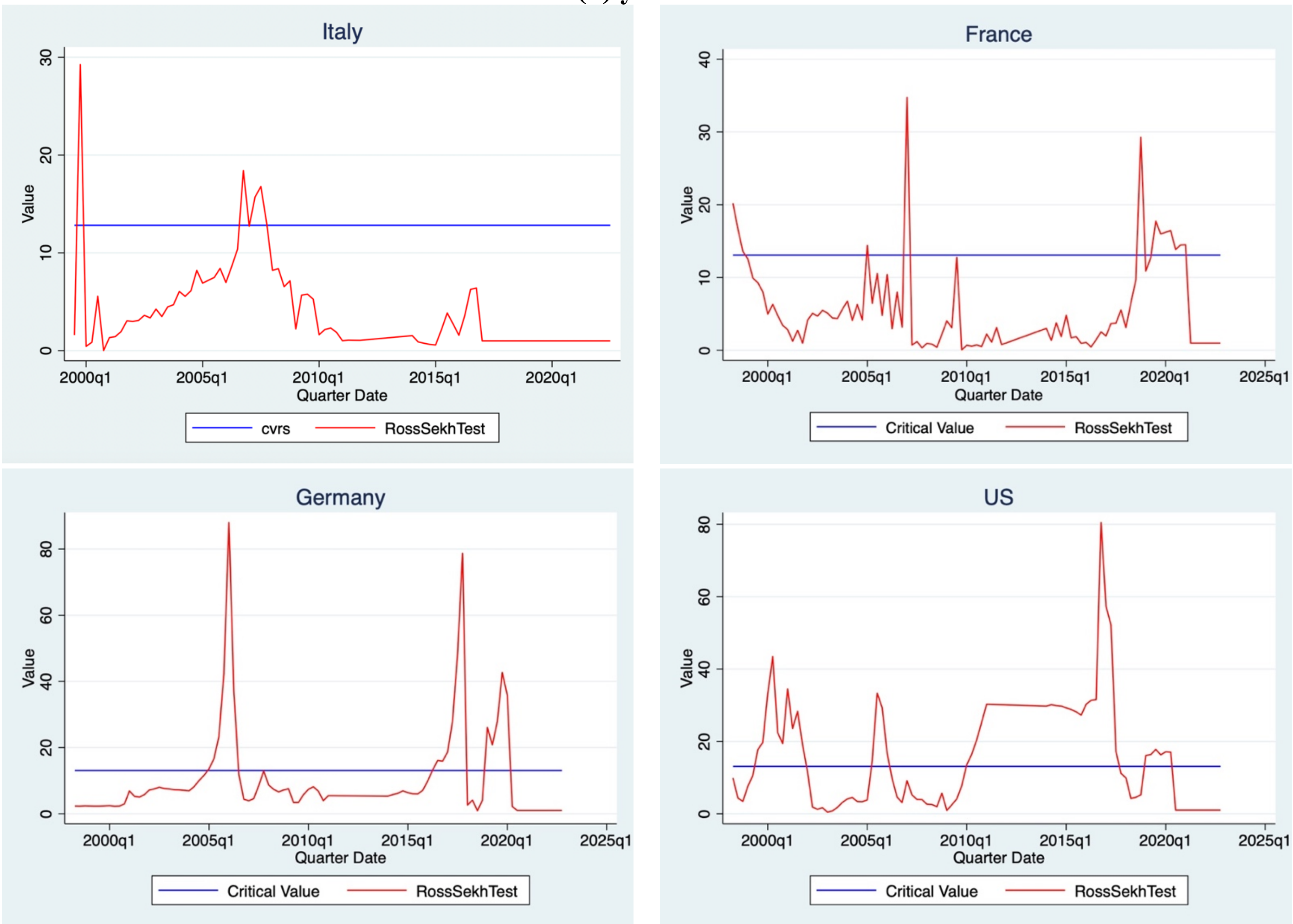
[12] The FR test has so far been applied only for very few variables. Rossi and Sekhposyan (2016) apply their test to inflation forecasts. Tsuchiya (2024) analyzes government construction forecasts.

Figure 2. Fluctuation rationality test for mean budget balance forecast.

(a) current-year



(b) year-ahead



Italy
Value
Quarter Date
2000q1
2005q1
2010q1
2015q1
2020q1
cvrs
RossSekhTest
France
Value
Quarter Date
2000q1
2005q1
2010q1
2015q1
2020q1
2025q1
Critical Value
RossSekhTest
Germany
Value
Quarter Date
2000q1
2005q1
2010q1
2015q1
2020q1
2025q1
Critical Value
RossSekhTest
US
Value
Quarter Date
2000q1
2005q1
2010q1
2015q1
2020q1
2025q1
Critical Value
RossSekhTest

Most of these changes for the current-year forecast seem to be associated with either the Global Financial Crisis, or the Pandemic, both periods in which the budget balance deteriorated quickly. Such deviations in the Global Financial Crisis were incorporated by expert forecasters in France of Germany, though. Changes in forecasting performance can be gradual-as the slow decline in performance in Germany shows-or rather quick-as the spiky behavior of the FR test in the US demonstrates.[13] What the FR test suggests is that large shifts in the economic and political situation make expert forecasters stray away from rationality, yet they do adjust to this new information. Such shifts need to be big before forecasters incorporate them, which seems to endorse once more the 'sticky information' view in which fixed acquisition costs make forecasters update their information sets only intermittently.

Panel (b) of Figure 2 also sheds an interesting light on how expert forecasters see longer term budgetary developments. The results so far seemed to suggest that forecasters did not adjust the year-ahead forecast in a substantially different way than current-year forecasts. But this hides the fact that expert forecasters revise more quickly the year-ahead forecast. This can be seen from the same periods in which the Global Financial Crisis or the Pandemic made expert forecasters deviate, but with a much faster adjustment. In addition, forecasters appear to experience more easily deviations, but then return more quickly to efficient forecasts. The sometimes-erratic jumps in the FR test in France over different periods illustrate this.[14] This result seems to suggest that expert forecasters incorporate the news they receive for current-year forecasters, and in that case also adjust their longer-term forecast simultaneously. This explains both the smaller deviations from rationality, and the same adjustment behavior we found in Table 4.

ii. Explaining deviations from rationality

Periods of deviations from rationality can be assessed also in a more formal way. We can consider the FR test statistics as a measure of the deviations from rationality over time. Periods of behavior are then classified as irrational when the test statistic exceeds the critical value, and as a rational if it stays under it. Hence, we regress this binary measure against the set of macroeconomic and political variables we used in Section III, testing equation (5) below with a logit model, where 1 indicates if exceeds the threshold and 0 if not.

$$\hat{\theta}_t = \alpha + \beta Z_t + \varepsilon_t \qquad (5)$$

Panel (a) of Table 6 indicates a few interesting patterns in the behavior of expert forecasters. Firstly, higher interest rates make it less likely that forecasters deviate, albeit in Italy, the effect is insignificant. The inverse occurs in the US current year, though. Stronger economic growth also raises probability of deviations, but just in Germany. This growth effect also carries over to the next budget year.

[13] Results are very similar for the different groups of forecasters. As the number of observations is much smaller, only the Pandemic seems to lead to large shifts in rationality. Figures are available upon request.
[14] Very similar results hold for the different groups of forecasters. Figures are available upon request.

Table 6. Explaining the deviation of rationality by country.

	(a) current year				(b) year-ahead			
	Italy	France	Germany	US	Italy	France	Germany	US
Interest rate	2.05	-0.61*	-3.00***	1.57***	-1.69*	-1.69*	-0.49**	-0.73***
	(2.07)	(0.31)	(0.89)	(0.39)	(0.96)	(0.96)	(0.25)	(0.27)
Output gap	1.42	-1.11**	1.49**	-0.04	0.12	0.12	0.45*	0.01
	(1.06)	(0.43)	(0.59)	(0.12)	(0.61)	(0.61)	(0.26)	(0.12)
Policy uncertainty	-0.03	0.00	-0.01*	0.01*	-0.04**	-0.04**	-0.01	-0.00
	(0.04)	(0.00)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.00)
Global political risk	1.95	0.49	5.69*	-0.54**	-0.74	-0.74	0.05	0.20
	(8.45)	(0.94)	(2.93)	(0.25)	(3.11)	(3.11)	(2.07)	(0.20)
Elections	2.19	0.59	-0.09	-1.07	-	-	0.95	0.38
	(2.15)	(0.85)	(1.24)	(0.81)	-	-	(0.82)	(0.63)
Constant	-9.46	-3.09*	7.72***	-6.15***	8.17	8.17	0.83	2.20
	(11.23)	(1.73)	(2.56)	(1.95)	(5.70)	(5.70)	(1.77)	(1.47)
Obs.	71	91	89	92	72	78	84	85
McFadden Pseudo R^2	0.26	0.41	0.14	0.27	0.23	0.02	0.17	0.11

Sources: Consensus Economic Forecasts and authors' estimations.
Notes: Logit estimation. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Secondly, heightened uncertainty about economic policy (France current year, and Italy and France, year-ahead) reduces the probability of making large mistakes, underscoring our previous finding that forecasters become more cautious in the revisions of their forecast when faced wirh uncertainty on economic policies. This result is in line with most of the literature on the effects of uncertainty (Bekaert *et al*., 2013; Kang *et al*., 2014; and Baker *et al*., 2016). Uncertainty is also more important than just election news. Elections actually have no impact on the rationality of forecasters. Geopolitical risk only impact significantly German and US current year forecasts. Nevertheless, both results are completely different in magnitude and direction. Higher geopolitical risk increases the probability of deviations in Germany, while in the US it slightly decreases the chance of deviations. This is a surprising finding given that global political tensions are more likely to lead to conflict and military expenditure in the US. These effects do not carry over to the next budget year.

Finally, the explanatory power of the logit model seems to indicate that expert forecasters update their information, mostly overreacting to big shifts, as we argued before, and that such deviations are coming from relevant economic or political news, to which they more slowly (in the current year) or more quickly (for the coming year) adjust.

V. Conclusion

Macroeconomic theories attribute expectations formation to sticky or noisy information. Information processing under both types is characterized by rigidities. The degree of stickiness of forecasts can be examined with standard rationality tests. Recent advances in these tests allow looking into time variation in forecasting performance. In particular, the Fluctuation Rationality test by Rossi and Sekhposyan (2016) allows detecting departures from forecaster rationality. We apply the exercise to fiscal forecasts from Consensus Economics, on a large panel of individual expert forecasters in four major OECD countries between 1993 to 2023.

The first contribution of this paper is to demonstrate that forecast rationality breaks down in specific episodes. We build on previous papers that examine forecast bias and rationality over full samples, but such averaging ignores the large swings that characterize fiscal policy changes over time, and that can be easily missed even by expert forecasters. Forecasters subsequently overreact to big changes.

The second contribution of this paper is to relate individual forecaster behavior to economic or political factors. We find evidence that is in line with the noisy information view: specific economic and political news is systematically incorporated in forecast revisions. It is hard though for forecasters to understand future policy tracks after sudden big shifts, like the Global Financial Crisis or the Pandemic.[15]

A lesson for policymakers is that the dispersion of experts' expectations about the future fiscal and economic outlook is directly related to anchoring market expectations on fiscal policies. In this context, better public financial management systems that improve fiscal reporting, including through the use of fiscal rules or national fiscal councils, could improve budget transparency (Beetsma *et al*., 2022) and the implementation of fiscal policy (Lledó and Poplawski-Ribeiro, 2013), affecting how frequently expert forecasts update their information sets and forecasts of the budget balance. Credibly anchoring market expectations can further significantly improve the effectiveness of macroeconomic policies by alleviating uncertainty, and therefore, risk premia (Cimadomo *et al*., 2016).

Next to the theoretical implications, the methodology of this paper also carries practical policy relevance. Fiscal forecasting is considered as a complicated exercise (Leal *et al.*, 2008). Detecting time variation in forecast dispersion can be useful for upgrading budget forecasts.

[15] In a similar manner, Kontogeorgos and Lambrias (2021) evaluate the accuracy of inflation forecasts in Europe and find how the persistent and significant bias changes over time in the EU sample. Before the global financial crisis, those authors find that inflation was persistently underpredicted, while in post-2013, the bias reverses into overprediction.

CHAPTER 2: Identifying drivers of deviations from rational expectations

I. Introduction

Understanding how experts form expectations holds significant importance in macroeconomics due to its influence on policy decisions and market behavior. Rational expectation theory (Muth, 1961) is built on the assumption that experts produce unbiased forecasts using all available information, and without systematic errors. However, experts form expectations in highly unstable ways, often shifting due to behavioral biases or external factors (Maćkowiak *et al.*, 2023). A key challenge has been explaining why deviations from rationality still persist, even when agents have full access to information.

A large body of literature has looked into behavioral reasons, informational problems, and external shocks as reasons for these deviations in rationality by experts. These studies have measured each of these deviations separately, analyzing only one aspect. For example, Kahneman and Tversky (1979) and Barberis and Thaler (2003) develop a behavioral perspective which indicates that cognitive biases such as overconfidence, anchoring, and herding can distort forecast accuracy. Another line explains the deviations from rationality as a consequence of informational frictions. Sims' (2003) rational inattention framework argues that experts face cognitive constraints in processing huge amounts of information. Recent empirical studies, including Coibion and Gorodnichenko (2012, 2015), provide evidence of inattentive forecasting behavior and highlight that most agents' deviations are sizable and consistent with moderate attention to macroeconomic data.[16] Macroeconomic shocks and structural changes also impact forecast deviations (Ascari *et al.*, 2023). These shocks can include cost-push disturbances such as energy price spikes (Coibion and Gorodnichenko, 2015); monetary policy shocks - especially when central bank communication is ambiguous or lacks credibility (see Romer and Romer, 2004; Gürkaynak *et al.*, 2005) -, or fiscal shocks (Bloom, 2009). The literature also highlights the role of institutional characteristics in shaping experts' expectations. Experts are often influenced by political and organizational decisions (Sanders and Manrodt, 2003). Fildes and Goodwin (2007) empirically confirm that institutional forecasts often deviate due to organizational constraints.

While these approaches highlight important mechanisms, they usually treat each channel in isolation without considering how these multiple factors may simultaneously interact to drive deviations from rational expectations. Dovern and Weisser (2011) attempt to move beyond this single-explanation approach, demonstrating that inflation forecasts systematically violate the assumptions of rational expectations. In their definition, this violation means that forecast errors are not white noise but exhibit bias, autocorrelation, and heterogeneity across the expert forecasters. By decomposing forecast errors Dovern and Weisster (2011) demonstrate that deviations cannot be explained solely by macro-level shocks but also stem from persistent individual-level factors. However, Dovern and Weister's (2011) decomposition measure bias and efficiency as time-invariant parameters and estimate them separately. Their results remain static because they do not model the dynamic processes of bias or efficiency, nor do they empirically identify the specific behavioral mechanisms behind forecast deviations, or how these evolve over time or differ systematically across individuals.

Building on this literature, and using a similar dataset as Dovern and Weisser (2011), we focus on monthly year-ahead forecasts of US inflation in the period of January 2010 to April 2022

[16] See Maćkowiak *et al.*, 2023, for a more detailed rational inattention literature review.

from the Consensus Economics survey to evaluate evolution of deviations from rationality and its drivers.

Our study makes two main contributions. First, we introduce a novel Irrationality Index that aggregates how each expert deviates from the rational expectations assumptions over time. We use the Fluctuation Rationality (FR) test developed by Rossi and Sekhposyan's (2016) to construct this Irrationality index. The null hypothesis of this test is that forecast errors are rational, implying unbiasedness and no systematic pattern in the forecast errors. The FR Test applies rolling-window regressions to detect time-specific deviations in forecast rationality, and compares forecast errors against realized outcomes. We compute the irrationality index as the difference between the time varying FR test statistical test and the critical value. A negative value of the Index indicates a rejection of rational forecaster behavior, while a positive value indicates consistency with rational expectations assumptions.

The Irrationality Index for each expert allows us to then to explain jointly the behavioral and informational drivers of each expert's deviations from rationality. Because we can trace how deviations differ across experts and time, we use a fixed-effects panel model to explain deviations from rationality by different drivers –behavioral reasons, informational problems, external shocks– as explanatory variables. Using panel models allows us to move away from the traditional approach of estimating each phenomenon separately. Instead, it provides an integrated framework to identify how multiple behavioral and informational drivers interact and shape deviations from rational expectations.

Our main findings show that deviations from rational expectations are important, and occur regularly across experts. However, these deviations are not random, but mostly driven by behavioral reasons, informational problems, and external shocks. For example, long-term interest rate expectations are also associated with greater deviations from rationality. This could indicate overconfidence or misjudgment in interpreting monetary policy signals. In addition, external information shocks increase deviations from rationality, experts may overreact to new data or struggle to distinguish relevant information. These results are consistent with evidence from Bordalo *et al.* (2020) and Afrouzi *et al.* (2021). Geopolitical risk is a key determinant of irrationality across all model specifications, indicating that rising global risk significantly weakens experts' forecasting rationality.

To ensure robustness of our results, we additionally conduct three complementary checks. We include a dynamic Arellano–Bond GMM panel model to account for the persistence of irrationality and address endogeneity concerns, a Common Correlated Effects Mean Group (CCEMG) estimator to control for unobserved global shocks and cross-sectional dependence across experts (Pesaran, 2006), and a conditional fixed-effects logit models to treat deviation from rationality as a binary outcome. We confirm that geopolitical risk and interest rate expectations increase deviations from rationality as found on the fixed-effects specification.

The remainder of the paper is structured as follows. Section II describes the dataset, and derives the baseline metrics to measure the deviations of rationality across experts. Section III first derives the Irrationality Index, and then examines the drivers of irrationality. Section IV presents the main robustness checks. Finally, Section V concludes.

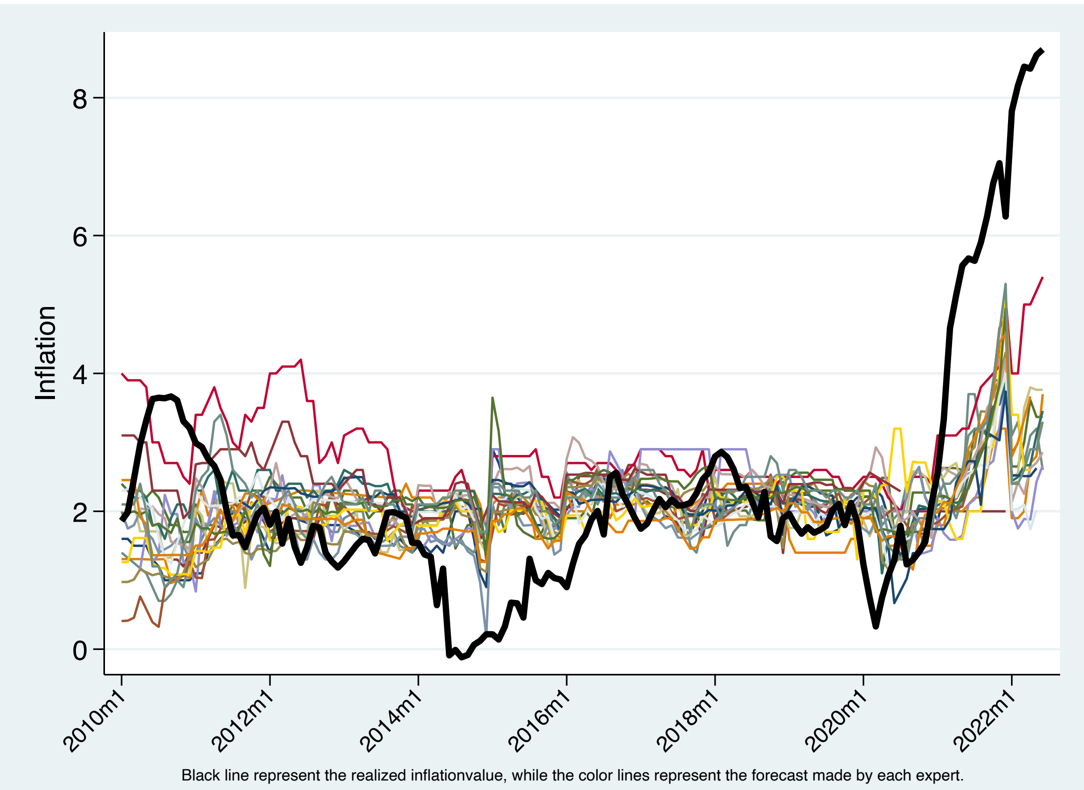
II. Data

We use Consensus Economics (CE) forecasts data to investigate how experts form their macro-economic expectations. This survey provides inflation, GDP, and interest rate forecasts from a wide range of professional economists (investment firms, banks, and research organizations) across more than 100 countries (Cimadomo et al., 2016). These forecasts are then published early in the second week of the same month, reflecting real-time expert expectations under current market and policy conditions.

Unlike other surveys, experts forecast in CE do not suffer a bias, as often happens for official government official projections (Ottaviani and Sorensen, 2006; D'Agostino and Ehrmann, 2014). CE data is public, which should help to prevent the herding behavior (Trueman, 1994) as experts should not show discrepancies between the survey and their private recommendation to their clients. Overall, the CE survey data broadly reflects the spectrum of expectations of market experts.[17]

We focus on monthly year-ahead US inflation forecasts between January 2010 and April 2022. Given the data availability across time, we retained only those experts who submitted forecasts in at least 70% of the months in the sample period. The final dataset results in a panel of 19 experts. Figure 3 shows the comparison between realized inflation and the inflation forecast one year-ahead made by all experts at each moment of time.

Figure 3. Realized and forecast inflation.



[17] Despite the methodological advantages of the Consensus Economics (CE) dataset, some authors have pointed out limitations in the way forecasts are aggregated. For example, Crowe (2010) demonstrates that Consensus Economics forecasts tend to overreact to the prior and underutilize new information. This inefficiency is not due to individual errors, it is due to the information aggregation process itself.

Figure 3 suggests that experts tend to systematically under- or overestimate year-ahead inflation, often deviating from the realized values. This graph motivates the construction of several behavioral and informational metrics-bias, herding, response to external shocks, overconfidence, and dispersion-to systematically analyze the dispersion among experts and the deviations from rationality. These metrics are computed at the individual level and shown in Figure 4, enabling us to track expert-specific behavioral patterns over time.

Firstly, forecast bias is calculated as the difference between forecast and realized inflation values, averaged across experts in each period. A positive bias indicates overestimation, while a negative value indicates underestimation. In Figure 4(a), we observe that although bias fluctuates modestly over time and is consistent with the concept of adaptive learning (Coibion *et al.*, 2019), where experts update their inflation forecast in response to past forecast errors rather than remaining passive. This persistent pattern is consistent with Coibion and Gorodnichenko (2012) and Andrade and Le Bihan (2013) who also find biases in expert's year ahead US inflation forecasts, including anchoring on past inflation rates and underreacting to new information.

Secondly, herding behavior captures experts' tendency to align closely with peers' experts rather than independently processing information. As in Lamont (2002) and Ehrbeck and Waldmann (1996), we measure herding by the absolute deviation of each expert forecast from the mean forecast as shown in Equation 6, in which $F_{i,t}$ is the forecast by expert i at time t, and N_t is the number of experts reporting forecasts in a period:

$$H_{i,t} = \frac{1}{N_{i,t}} \left| F_{i,t} - \overline{F}_t \right| \quad (6)$$

Lower values of the herding index implies that experts are aligning their forecast more closely with peers rather than relying on independent assessments, whereas higher values reflect less herding behavior. Until around 2021 most inflation forecasters were clustered around 2 to 3%, showing high herding behavior. This can also be observed in Figure 4(b) with values between 0.2 and 0.5 where experts align with the consensus. Spikes in the herding index (which indicates more disagreement between experts) are especially visible during periods of uncertainty and geopolitical conflicts, for example during COVID-19, and the Russia-Ukraine war, where the herding index exceeds 2. These spikes occur when inflation diverges from prior patterns and experts break from the mean. These results match with previous research by Glick and Kouchekinia (2021) who find higher inflation forecast disagreement between expert in the US since 2021. Kahneman and Tversky (1979) and Barberis and Thaler (2003) show that individuals become overconfident, anchor their predictions based on past values, or follow their peers rather than processing information by their own. These mechanisms explain why we often observe persistent forecast bias and herding behavior on Figure 4(a) and 4(b).

Thirdly, there are informational constraints that make experts stray away from rationality. Sims (2003) and Woodford (2012) argue that forecasters face cognitive limits in processing information so they may react only to new data or with delay. We measure this with an external information shocks index (in Equation 7 and in Figure 4(c)). This metric shows how deviations from rational expectations can arise by external shocks, which complicate the interpretation of macroeconomic signals and increase deviations from rational behavior as demonstrated in Bekaert *et al.*, (2013), Kang *et al.*, (2014), and Baker *et al.*, (2016).

In order to measure information shocks Davies and Lahiri (1995) propose to measure it as the difference between an expert's forecast revision and the average revision across all other experts, at each period of time. Formally, the shock is calculated as:

$$S_{i,t} = (F_{i,t} - F_{i,t-1}) - \frac{1}{N_t}\sum_{j=1}^{N_t}(F_{j,t} - F_{j,t-1}) \quad (7)$$

A value close to zero indicates that the expert is not affected by shocks of external information, while large positive values indicates that the expert overreacts to new information. In Figure 4c, we can see that between 2012 and 2014 the shock of external information remains tightly clustered around zero, reflecting anchored inflation expectations. Nevertheless, there are several points-early 2012, during the COVID-19, and again in 2022–2023 with post-pandemic inflation, the Russia–Ukraine war and energy price shocks-when the index increases sharply. These match with the periods in Figure 3 when realized inflation either diverged from or surged beyond pre-pandemic norms, especially after 2021 when inflation increase but experts stayed clustered.
Fourthly, we include a dispersion index that captures the heterogeneity of experts at each point in the time in Equation 8, $F_{i,t}$ is the forecast of expert i at moment t, and F_t the mean across all experts (N_t).

$$D_{i,t} = \sqrt{\frac{1}{N_t}(F_{i,t} - \overline{F}_t)^2} \quad (8)$$

The dispersion index captures the degree of heterogeneity among experts. Low dispersion values show tightly clustered forecasts reflecting consensus or potential herding, whereas high values signal greater dispersion and heightened uncertainty (Mankiw *et al.*, 2003; Carroll, 2003). Results on Figure 4(d) follow a similar pattern as for the herding metrics. There is less dispersion from 2010 to around 2016, reaching a trough near 0.2. Starting in 2020, dispersion increases sharply, peaking close to 0.8 during 2022. This period matches with the post-COVID inflation surge, supply-chain disruptions, and the Russia–Ukraine war, all of which introduced considerable uncertainty into inflation dynamics (as seen in Figure 3). The rising dispersion therefore signals greater disagreement among experts precisely when inflation became harder to predict.

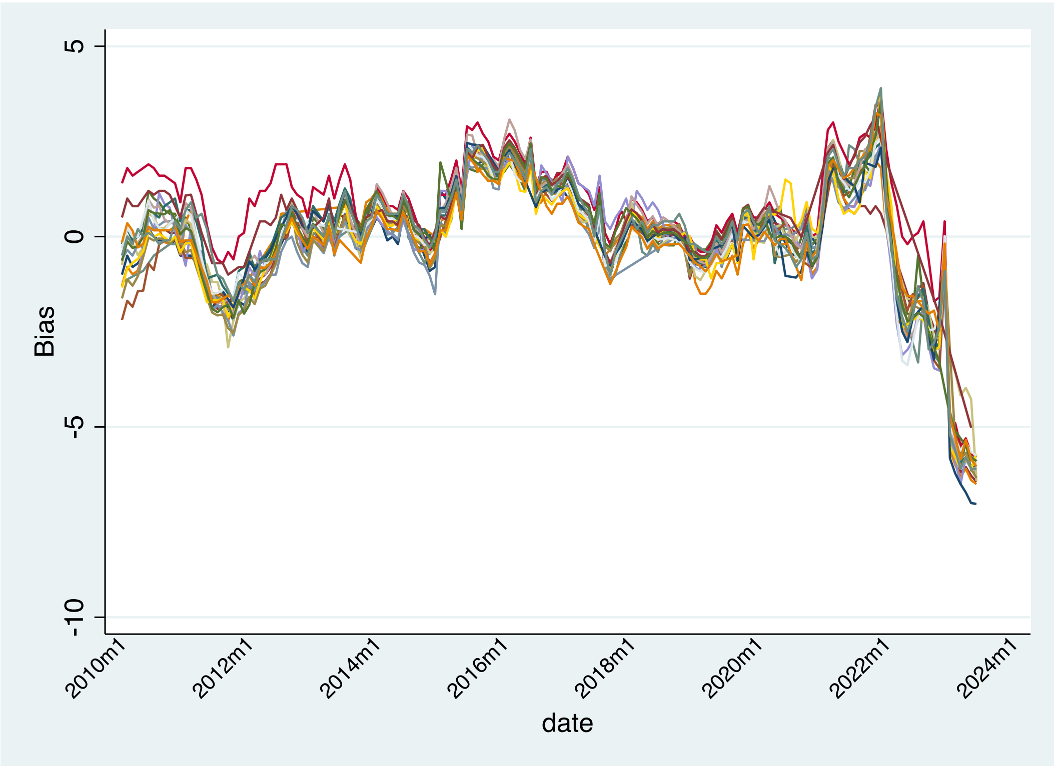
Finally, we also compute some standard statistics of forecasting performance. On Figure 4(e) we show the metrics for RMSE to summarize each expert's average forecasting accuracy. Experts with lower values are more accurate rather those ones with higher RMSE values. Since this measure averages over time, the metrics on Figure 4(e) shows the long-run forecasting accuracy rather than time variation.

We also calculate the skewness of forecast errors, computed over a 24-month rolling window and 0.05 as significance level for each institution. A positive skewness indicates underestimation of inflation (errors skewed to the right), while negative skewness means overestimation. This approach aligns with Adrian *et al.* (2020), who model time-varying macroeconomic risk by estimating the asymmetry in the conditional distribution of forecast outcomes (real GDP growth, unemployment, and inflation). They find that these distributions are often skewed, and that this skewness contains valuable information about risks and economic uncertainty. In our context, applying skewness to forecast errors provides a simple way to capture how individual experts perceive inflation risks. In Figure 4(f), experts show persistent positive skew between 2016-2017.

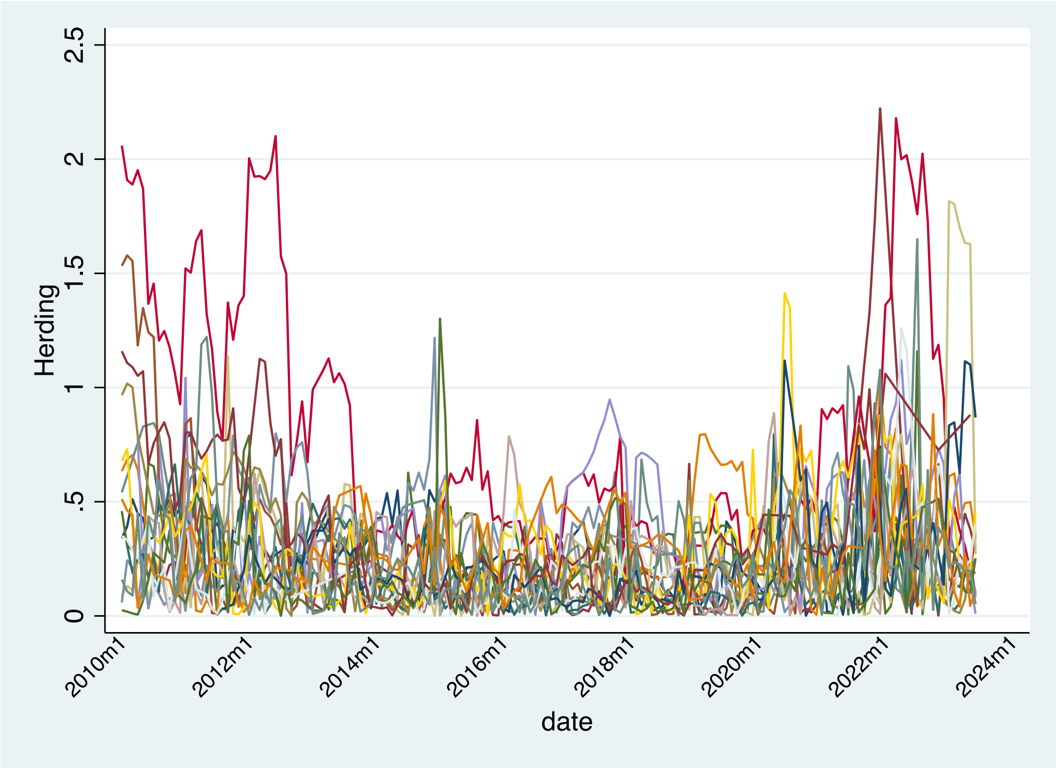
Taken together, these patterns illustrate that deviations from rationality are not driven by a unique explanation. Rather different theoretical explanations seem to interact and explain the deviations. For example, under conditions of stability, forecasts are tightly clustered and revisions are smooth, but in periods of uncertainty (COVID-19, Ukraine-Russia war) herding, delayed information processing, and asymmetric risk perceptions seems to reinforce one another. This interaction is consistent with models of sticky expectations (Mankiw and Reis, 2002), rational inattention (Sims, 2003), and adaptive learning (Coibion *et al*., 2019), in which forecasters update gradually, rely on past anchors, and interpret new information through heterogeneous perspectives.

Figure 4. Behavioral and informational metrics of year-ahead inflation forecasts at expert level

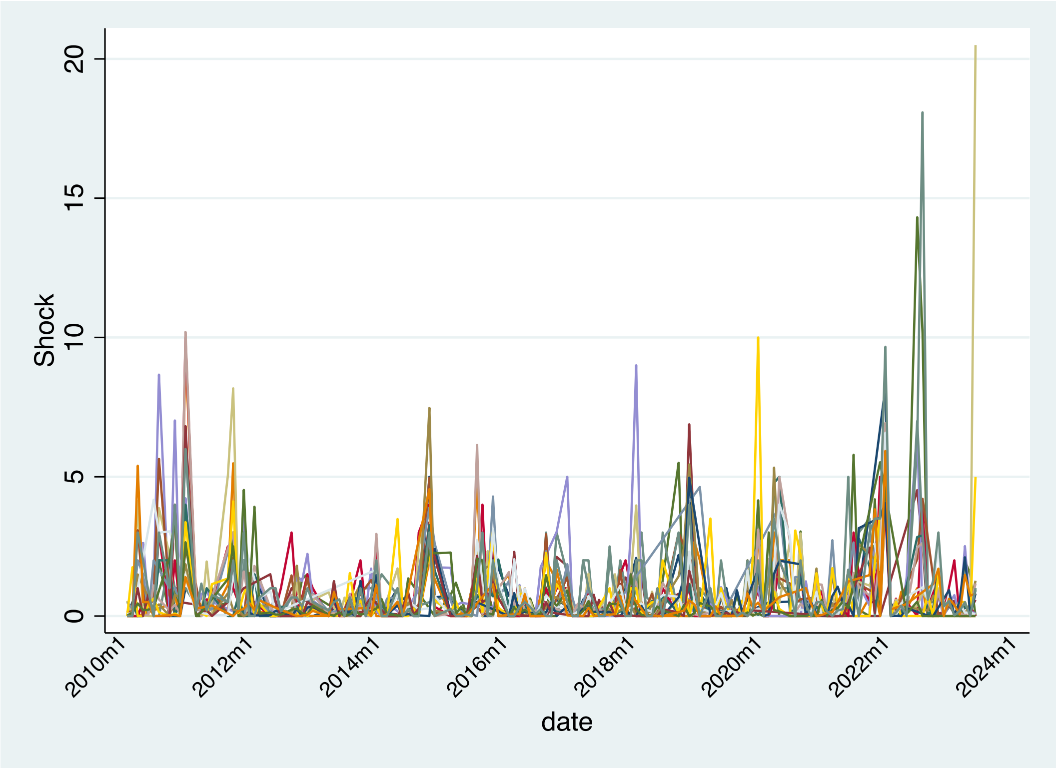
(a) Bias



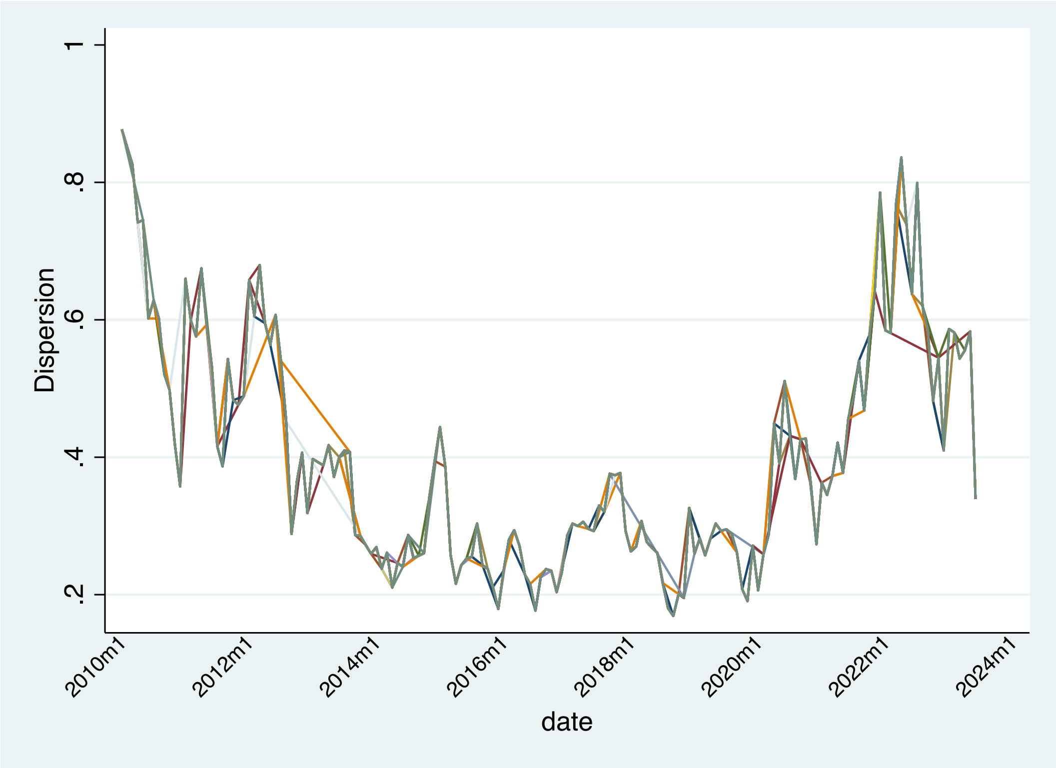
(b) Herding



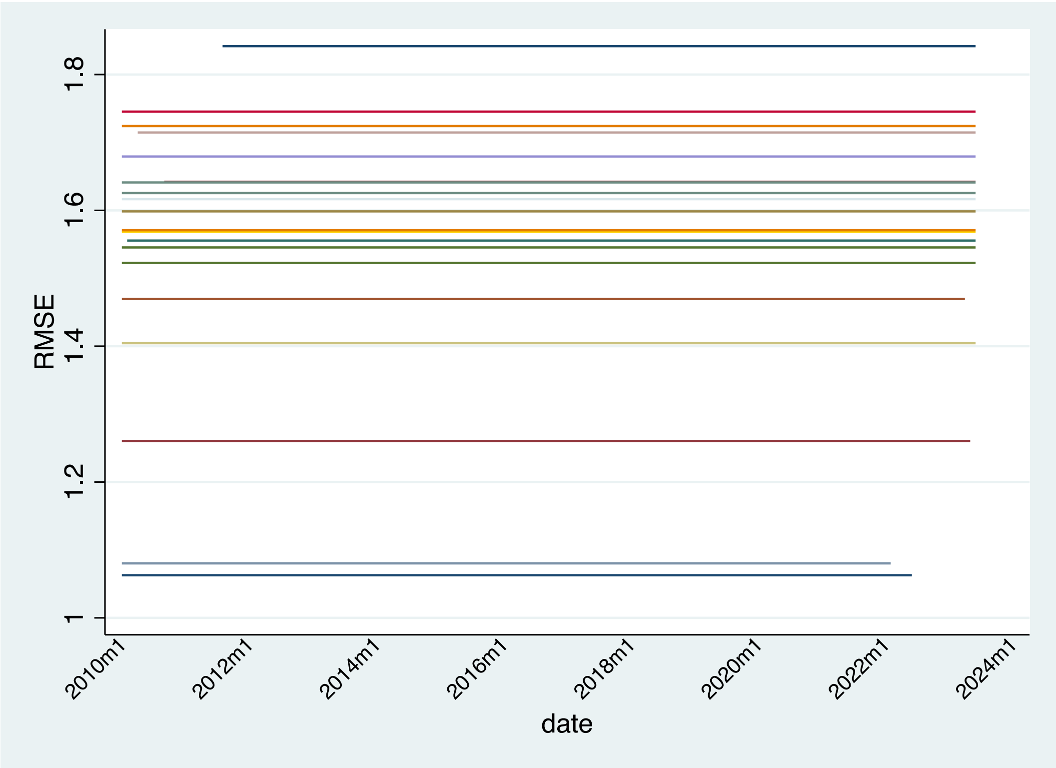
(c) Shock of external information



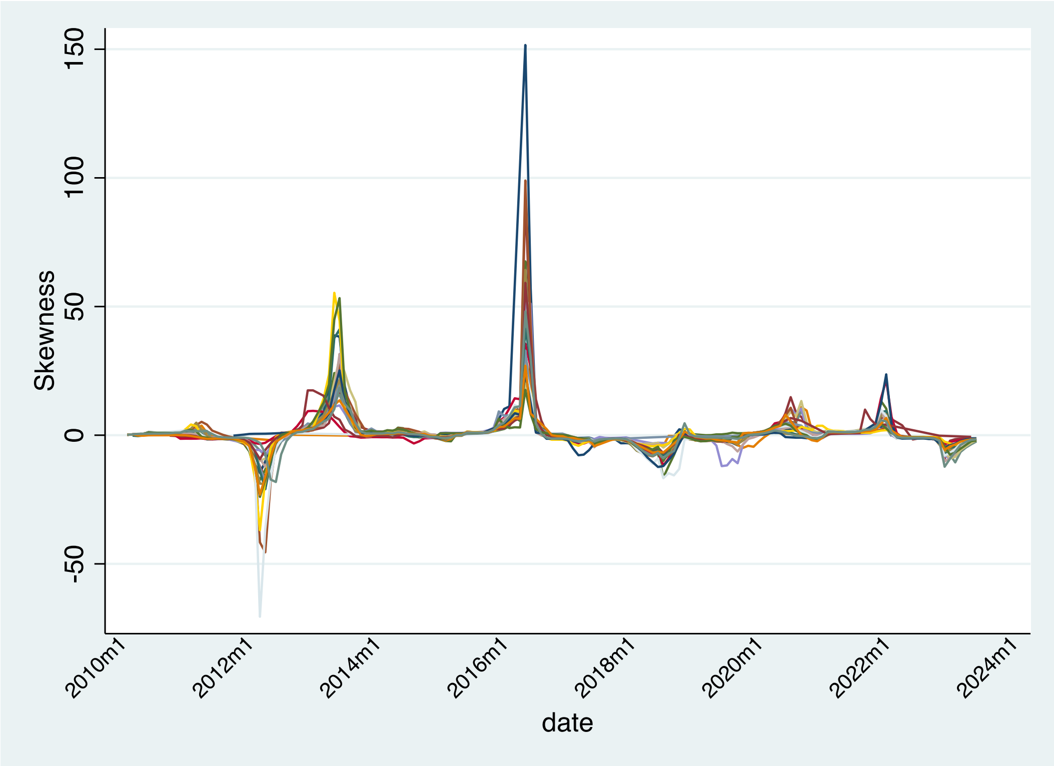
(d) Dispersion



(e) RMSE



(f) Skewness



III. Identifying the drivers of irrationality

Most studies typically focus on measuring why experts deviate from rationality by analyzing one source in isolation without considering how these multiple factors may simultaneously interact. Dovern and Weisser (2011) attempt to move beyond this single explanation. The authors decompose inflation forecast errors to assess the accuracy, bias, and weak efficiency of macroeconomic forecasts. Their decomposition allows the separation of macro-level shocks from individual-specific components, but the framework remains estimating the effect of each source in isolation. They did not empirically identify the specific behavioral or informational drivers of forecast deviations, nor does it capture how the deviations from rationality evolves over time or differ across individuals.

Building on this literature, and based on the descriptive statistics just shown, our hypothesis is that behavioral reasons, informational problems, external shocks at the same time influence deviations from rational expectations. Cognitive biases and herding are expected to produce autocorrelated errors, whereas limited attention and processing constraints may generate persistent heterogeneity across forecasters. Macroeconomic shocks and high uncertainty are likely to amplify these deviations, especially during periods of policy changes, or global risk (Coibion and Gorodnichenko, 2012, 2015; Caldara and Iacoviello, 2022).

To address these limitations, we introduce two key methodological innovations. First, we construct a time-varying, expert-level irrationality index that captures when and how experts deviate from rational expectations. Second, we then relate this index to jointly identify how different drivers -bias, herding, dispersion, overconfidence, political risk, economic uncertainty, etc.- impact on deviations from rationality. We use a fixed-effects panel regressions, which allow us to control for unobserved heterogeneity across experts and time. This approach allows us to move from the traditional approach of estimating separate regressions to, instead provide an integrated framework to identify how multiple behavioral and informational drivers interact and shape deviations from rational expectations.

i. The new irrationality index

The concept of rationality is fundamental in economic theory, particularly in the area of expectations formation and forecasts. The concept of rationality implies three key conditions: unbiadnesses, efficiency and errors do not follow any systematic pattern over time (Muth, 1961; Fama, 1970). These conditions ensure that deviations from actual outcomes are random and reflect noise rather than systematic distortions. However, as the literature has shown, forecasts often fail to meet these criteria. Behavioral biases (Kahneman and Tversky, 1979; Barberis and Thaler, 2003), informational frictions (Sims, 2003; Woodford, 2012), and external shocks (Coibion and Gorodnichenko, 2015; Gürkaynak *et al.*, 2005) are the three main explanations of drivers of deviation from rationality. The statistics shown in Figure 4 confirms that our data sample also exhibit these deviations from rationality.

We construct a novel Irrationality Index that allows us to identify when experts are deviating from rational expectations assumptions. This index is based on the Fluctuation Rationality (FR) Test developed by Rossi and Sekhposyan (2016) and the null hypothesis is that forecast errors are rational, implying unbiasedness and forecast errors do not follow any systematic pattern. The FR Test applies rolling-window regressions to detect time-specific deviations in forecast rationality, it compares forecast errors against realized outcomes using a Wald statistic corrected for heteroskedasticity and serial correlation (Newey-West adjustment).

Formally, the FR Test builds upon the econometric framework originally proposed by West and McCracken (1998):

$$v_{t+h}(\hat{\gamma}_t, R) = \hat{g}'_t \cdot \theta + \eta_{t+h}, \qquad t = R, \dots T \qquad (9)$$

Where $\hat{g}'_t \equiv g_t(\hat{\gamma}_t, R)$ is an $(lx1)$ vector function of period t. θ is an $lx1$ parameter vector and $v_{t+h}(\hat{\gamma}_t, R)$ is the forecast error defines as the difference between forecast and realized value.

In this equation (9), the forecast error $v_{t+h}(\hat{\gamma}_t, R)$ (defined as the difference between forecasted and actual values) is explained by the vector function $\hat{g}'_t \equiv g_t(\hat{\gamma}_t, R)$, which is an $(lx1)$ vector function of period t. The parameter vector θ captures deviations from rationality. Under the null hypothesis of rational forecasts, these parameters equal zero:

$$H_0\colon \theta = \theta_0 \ vs. H_A\colon \theta \neq \theta_0 \ \ where\ \theta_0 = 0 \ \ (10)$$

The Wald statistic computed is:

$$w_P = P\left(\hat{\theta}_P - \theta_0\right)' \hat{V}_{\theta,P}^{-1}\left(\hat{\theta}_P - \theta_0\right) \qquad (11)$$

Where $\hat{\theta}_P$ represents the estimated parameters, and $\hat{V}_{\theta,P}^{-1}$ is the variance-covariance matrix robustly corrected for heteroskedasticity and autocorrelation (Newey-West adjustments). A rejection of the null hypothesis occurs when the maximum observed value of the Wald statistic series exceeds the critical threshold, indicating systematic deviations from rationality.

Summarizing, the Fluctuation Rationality Test covers the following cases:

1. forecast unbiasedness tests, where $\hat{g}_t = 1$;
2. forecast efficiency, where $\hat{g}_t = y_{t+h|t}$;
3. forecast rationality (Mincer and Zarnowitz, 1969), where $\hat{g}_t = [1 \; y_{t+h|t}]$;
4. forecast encompassing tests, where $\hat{g}_t$ is the forecast of the encompassed model;
5. serial uncorrelation tests, where $\hat{g}_t = v_t(\hat{y}_{t-h}, R)$.

We compute the FR Test at individual expert level using a 12-months rolling window and significance level of 0.05. The Irrationality Index is then constructed as the difference between the FR test statistic and its critical value. A negative value of the Index indicates a rejection of rational behavior, while a positive value signals consistency with rational expectations.

Figure 5 presents the Irrationality Index for each expert over time, while Table 7 show some summary statistics.[18] In general, most of the values in the index reject the hypothesis of rational behavior. Between mid-2021 and early 2022 there are peaks of irrationality which aligns with the post Covid-19 period. In Figure 3 we saw how realized inflation rapidly increased while experts' forecasts remained anchored at lower levels, generally around 2 to 3%. During this time, experts underestimated the inflationary consequences of reopening the economy. Figure 4 supports this interpretation by showing an increase in dispersion, bias, and an elevated herding behavior during this time, as experts clustered around the mean forecast rather than responding independently to new signals.

Another episode of deviations from rationality appears at the beginning of 2015. Figure 3 clearly shows that many experts' inflation forecasts were significantly higher than the realized values. The realized inflation was close to or below 1% in early 2015, while some experts exceeded 3% or even 4%. This difference highlights a systematic overestimation of inflation, that could possibly be driven by expectations of monetary policy normalization or just misinterpretation of macro signals. Figure 4 reinforces this interpretation with elevated levels of bias and herding in 2015, suggesting that forecasters failed to adjust adequately.

Similar patterns reappear during later crises. In 2020, COVID-19, and in early 2022 (post-pandemic reopening and the Russia–Ukraine war), Figures 4 again show sharp spikes in dispersion, herding, and shocks, aligned with deep drops in the irrationality index. These episodes confirm that deviations from rational assumptions are recurrent and heterogeneous across experts, with some adapting faster to new conditions while others misinterpret signals.

[18] This index follows a similar interpretative logic to other composite indicators widely used in macroeconomic research-such as the Economic Policy Uncertainty Index (Baker *et al.*, 2016), the Geopolitical Risk Index (Caldara & Iacoviello, 2022), the Financial Stress Index (e.g., Federal Reserve Bank of Kansas City), the VIX Volatility Index, the Michigan Consumer Sentiment Index, the Index of Leading Economic Indicators (The Conference Board), and the FCI (Financial Conditions Index)-where the emphasis is placed on tracking the direction and relative magnitude of changes over time, rather than providing a direct economic meaning or absolute thresholds. These indices are primarily designed to facilitate comparative analysis across time periods, countries, or agents, without requiring a fixed benchmark for interpretation.

Figure 5. Irrationality index by expert.

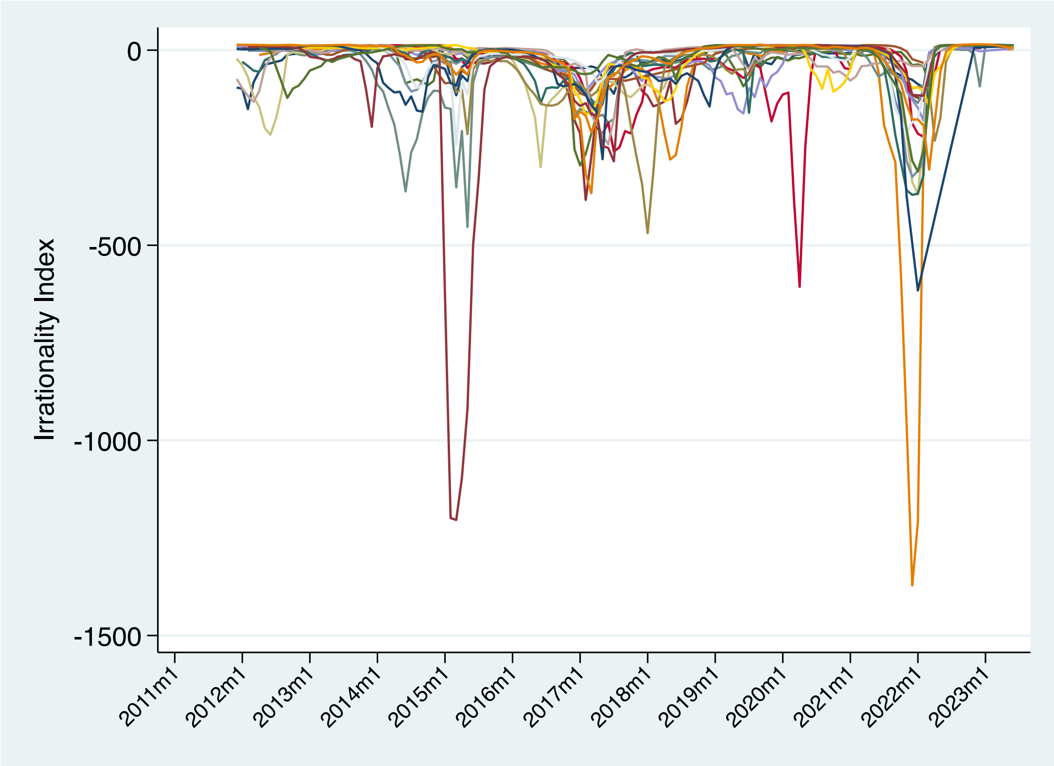


Table 7. Descriptive statistics of irrationality index by expert

Experts	Obs	Mean	Standard deviation	Min.	Max.
A	118	-24.09	33	-151.2	12.86
B	145	-26.18	62.7	-383.93	13
C	129	-24.2	51.61	-295.49	13.06
D	138	-67.53	198.75	-1371.77	13.02
E	150	-52.02	80.48	-452.86	13.92
F	149	-43.88	90.33	-606.46	12.91
G	133	-24.49	42.04	-180.88	9.94
H	73	-40.6	77.12	-369.72	13.08
I	84	-18.21	33.53	-116.63	12.82
J	44	-22.99	60.42	-324.73	13.03
K	126	-36.61	70.25	-370.72	13.07
L	140	-43.93	76.31	-468.83	13.06
M	143	-18.68	43.78	-241.45	12.98
N	147	-24.98	45.64	-166.31	13.07
O	129	-25.13	48.65	-242.38	13.02
P	113	-46.48	84.22	-615.9	12.71
Q	120	-68.14	208.93	-1203.87	13.07
R	140	-23.29	57.15	-311.03	13.04
S	117	-20.24	56.41	-306.34	14.9

ii. Analysis of drivers of deviations from rationality

After identifying periods of deviations from rational expectations, we now explore the underlying mechanisms driving these deviations. The literature generally points to three drivers: behavioral biases, informational frictions, and external shocks. Moreover, as already demonstrated on Figure 4 and in Dovern and Weisser (2011), deviations from rationality are not driven by a unique explanation. Rather, several theoretical explanations seem to interact and explain these deviations.

We therefore include in our empirical model a set of variables that align with these drivers. Forecast bias, herding, overconfidence, and forecast dispersion capture behavioral mechanisms. External information shocks are proxied by the metric in (7), but also using the geopolitical risk index (Caldara and Iacoviello, 2022), economic policy uncertainty (Baker *et al.*, 2016), and expert's macro expectations such GDP growth and interest rates collected from Consensus Economic dataset. On Appendix D we report diagnostic tests-including a full correlation matrix and VIFs-to confirm that multicollinearity and overfitting are not driving our results.

We use a fixed-effects model to control for unobserved, time-invariant heterogeneity across experts, capturing within-expert variation over time rather than differences across them.[19] Due to potential heteroscedasticity and autocorrelation in the panel, we cluster standard errors by experts.

[19] A Hausman test rejects the null hypothesis of no systematic difference between a fixed and random effects model ($\chi^2 = 27.4$).

Table 8. Drivers of irrationality on inflation expectations

	Fixed-effects
Bias	-12.49
	(7.54)
Herding index	20.41
	(20.78)
Shock of external information	-2.31***
	(1.42)
Forecast dispersion	-117.27
	(71.22)
Overconfidence	0.43
	(0.26)
Economic Political Uncertainty index	0.03
	(0.04)
Global Political Risk	-16.82***
	(4.81)
% Change of expected current-year GDP	-0.67
	(1.75)
% Change of expected year-ahead GDP	-8.10
	(5.13)
Interest rate 12-month's forecast	-37.91**
	(17.46)
Interest rate 3-month's forecast	32.14
	(19.64)
Constant	75.31*
	(42.28)
Observations	1,496
R-within	0.12
R-between	0.11
R^2	0.12
No. of experts	19
F test	19.27***

Sources: Consensus Economic Forecasts and authors' estimations.
Notes: Clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

The results on Table 8 demonstrate that deviations from forecast rationality are not random, instead, they are influenced primary by shock of external information, global political risk and interest rate forecast. Other drivers, such as bias, herding, GDP expectations are not significant in our estimation.

The external information shocks have a significant and negative effect (−2.31) implying that greater responsiveness to new information is associated with higher deviations from rationality. This finding suggests that experts may overreact to new data or misinterpret signals. It is also consistent with recent literature of systematic overreaction in expert forecasts (Bordalo *et al*., 2020; Afrouzi *et al*., 2021; Ba *et al.,* 2024). Similarly, the Geopolitical Risk Index has a significant negative effect (−16.82), which indicates that periods of high uncertainty increase deviations from rational expectations, likely due to increased ambiguity and complexity in interpreting macroeconomic signals. Additionally, the 12-month interest rate forecast is negatively associated with the index (−37.91), suggesting that long-horizon monetary policy expectations may introduce distortions, either through overconfidence or misjudgment, thus reducing forecast rationality.

The other variables do not have a significant impact on deviations from rationality. Nevertheless, the results show interesting insights. For instance, the negative coefficient on forecast bias (−12.49) implies that greater forecast bias is associated with more negative values of the irrationality index. Similarly, the large negative coefficient on forecast dispersion (−117.27) suggests that higher disagreement among experts coincides with more deviations from rational expectations. This could indicate that dispersion, rather than reflecting diversity of independent thought, may sometimes signal confusion or lack of consensus in interpreting economic signals. The herding index, on the other hand, shows a positive coefficient (20.41), which would imply that more herding is associated with more positive index values.

The coefficient on skewness of forecast errors is positive, meaning greater asymmetry in forecast errors is weakly associated with more rational forecasts, though this effect is small (0.43) and not significant.

Likewise, the Economic Policy Uncertainty Index appears to be positively associated with the Irrationality Index, implying slightly more rational forecasts under higher policy uncertainty. This result is in line with most of the literature on the effects of uncertainty (Bekaert *et al*., 2013; Kang *et al*., 2014; and Baker *et al*., 2016) which posit that experts become more cautious in the revisions of their forecasts when faced with uncertainty on economic policies.

For the macroeconomic controls, the expected change in current-year GDP and year-ahead GDP both have negative signs, indicating that more optimistic growth forecasts correlate with more irrational forecasts. Meanwhile, the forecast for short-term interest rates is positive, suggesting that expectations of near-term policy tightening may improve rationality, though the result is not significant.

In sum, this integrated approach shows that deviations from rational expectations are not explained by a single channel but by the combined influence of shock of external information, global political risk, and interest rate expectations, while alternative explanations related to behavioral bias are less relevant.

IV. Robustness checks

To validate the empirical findings, we include some robustness checks using three complementary econometric approaches. Firstly, we use a dynamic Arellano–Bond GMM model to account for the temporal persistence of deviations from rationality and address potential endogeneity through lagged instruments (Arellano and Bond, 1991; Roodman, 2009)[20]. specification additionally accounts for temporal dependence and potential endogeneity.

Secondly, we apply the Common Correlated Effects Mean Group (CCEMG) estimator to control for unobserved common factors that may influence experts simultaneously over time. In addition, CCEMG allows for heterogeneous slope coefficients across individuals and

[20] We conduct standard diagnostic checks for the GMM specifications. These include tests such as AR(1) and AR(2) and over-identification, Hansen and Sargan tests. All models pass the AR(2) requirement (p-values > 0.1), and most demonstrate acceptable Hansen p-values (above 0.25), indicating no major concerns regarding instrument validity or overfitting. Moreover, to mitigate potential endogeneity between forecast irrationality and behavioral indicators-such as bias, herding, and forecast dispersion-we instrument endogenous regressors with their own lagged levels (lags 2 to 4). As instruments, we also include predetermined external variables, such as the Economic Policy Uncertainty (EPU) Index and Global Political Risk (GPR) Index.

corrects for cross-sectional dependence, providing more robust estimates in the presence of unobserved global shocks or correlated errors (Pesaran, 2006).

Finally, as it may be hard to interpret deviations from rationality in absolute terms, one can consider any significant deviation from rationality by an expert as a period of irrational behavior. We thus classify all months with a negative irrationality index as 0, and all other 'rational' months as a one. To model the probability of irrationality, we estimate conditional fixed-effects logit models (Greene, 2012) that account for within-expert variation in binary outcomes.

Table 9. Robustness checks.

	Dynamic GGM	CCEMG	Logit Model
Lag of irrationality	0.87***		
	(0.15)		
Bias	-54.46	-9.95*	-0.49***
	(113.55)	(5.62)	(0.09)
Herding index	60.78	-2.55	0.23
	(264.35)	(15.91)	(0.31)
Shock of external information	0.90	-2.85	0.11
	(18.17)	(2.78)	(0.06)
Forecast dispersion	-225.37	-77.73	0.59
	(612.87)	(51.89)	(0.64)
Overconfidence	-7.11	0.78	0.01*
	(10.03)	(0.40)	(0.01)
Economic Political Uncertainty index	19.43	-0.02	0.00*
	(67.73)	(0.04)	(0.00)
Global Political Risk	-43.19	-15.04***	-0.63***
	(208.73)	(4.30)	(0.13)
% Change of expected current-year GDP	50.71	3.11	0.03
	(76.20)	(3.21)	(0.04)
% Change of expected year-ahead GDP	-170.96	-5.06	-0.09
	(372.62)	(8.16)	(0.09)
Interest rate 12-months forecast	-1,962.11	-38.74***	-2.35***
	(7,468.21)	(12.10)	(0.23)
Interest rate 3-months forecast	0.87***	30.60*	2.29***
	(0.15)	(16.31)	(0.25)
Constant	-54.46		
	(113.55)		
Observations	1496	1,496	1496
R-within		0.64	
R-between		0.60	
R^2		0.61	
No of experts	19	19	19
F test	149687***		475,5***
AR(1) p-val	0.108		
AR(2) p-val	0.766		
Hansen p-val	0.287		
Sargan p-val	1.19e-05		
Instruments	19		

Sources: Consensus Economic Forecasts and authors' estimations.

Notes: Logit estimation. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 9 presents the results for the three complementary regressions. First, the dynamic GMM model confirms the persistence of irrationality over time, as indicated by the highly significant lag coefficient.[21] This finding complements the panel FE estimates by showing that deviations from rationality today tend to carry over into subsequent periods. While the baseline FE model captures the structural drivers of irrationality after controlling for unobserved heterogeneity, the dynamic GMM

We then show the CCEMG estimates. As for the panel results, the significance of geopolitical risk and long-term interest rate expectations underlines the role of externals shocks. This finding aligns with earlier studies showing that geopolitical events and policy signals shape forecast behavior (Caldara and Iacoviello, 2022; Baker *et al*., 2016). The behavioral elements driving experts' deviations from rationality are not important at all.

Finally, the conditional fixed-effects logit model further supports the significant impact of geopolitical risk and interest expectations as both significantly increase the likelihood of deviations from rationality. In contrast to previous results, short-term interest rate expectations are positively and significantly associated with rationality. This distinction suggests that while long-horizon monetary policy expectations may raise overreaction, shorter-term expectations can anchor forecasts more effectively which is consistent with Bordalo *et al*., (2020) as they find an overreaction to distant signals.

Across all three approaches, behavioral variables such as herding, dispersion, and policy uncertainty remain statistically insignificant, as in the panel FE model. These results confirm that these factors play only a secondary role relative to external shocks and interest rate expectations.

V. Conclusion

This study contributes to understanding the deviations from rationality in expert forecasts by developing a novel "irrationality index" based on Rossi and Sekhposyan's (2016) Fluctuation Rationality Test. Unlike previous literature, which tends either to analyze isolated dimensions of forecast deviations-such as behavioral biases (Kahneman and Tversky, 1979), informational frictions (Sims, 2003; Woodford, 2012)-or to interpret forecast deviations as random (Blanchard and Watson, 1982; Orphanides and Williams, 2005), our index integrates temporal fluctuations and expert-level heterogeneity into a unified empirical framework.

The two key contributions are the irrationality index and the identification of the drivers of it. First, the index allows us to identify periods of heightened deviations from rational expectations across experts and time. Second, this study advances by jointly identifying the drivers of deviation from rationality: bias, herding behavior, forecast dispersion, uncertainty, global political risk, and macro expectations.

Our results show that deviations from rationality are certainly not random but instead are influenced by different factors. Yet, among these factors it is interest rate expectations, external information, and in particular geopolitical risk that is important. This result reinforces the idea that uncertainty complicates the interpretation of macroeconomic signals. Similarly, long-term interest rate expectations is associated with increased deviations from rational expectations. Experts may misinterpret or overreact to monetary policy outlooks news (Gennaioli *et al*.,

[21] We conduct standard disgnostic tests for the GMM model. The AR(1) and AR(2) tests do not reject the absence of second-order serial correlation, and the Hansen test supports the validity of the instruments (p = 0.287). The Sargan test, however, is significant, suggesting some degree of instrument proliferation, a common issue in GMM applications that we attempt by restricting lag value.

2015; Angeletos and Lian, 2018). Furthermore, the persistence of deviations over time, supported by GMM results, indicates that experts may have learning lags or cognitive inertia.

Other variables, such as forecast bias, dispersion, and information shocks, also contribute to irrational behavior, though their significance varies. For instance, high dispersion-often interpreted as diversity of opinion-may sometimes signal confusion rather than disagreement, while herding behavior, surprisingly, does not increase deviations in our models. This could reflect strategic alignment rather than a mimicry (Trueman, 1994).

These findings have direct policy implications. For example, central banks and economic institutions must consider that experts may be systematically deviating from rationality, especially during high-risk or uncertain periods such as the post-Covid-19 recovery or geopolitical crises. Transparent, consistent, and forward-looking communication can help to reduce these informational problems, avoid overreactions to external shocks, and realign expectations. In addition, monitoring the irrationality index can support early detection of expectation misalignments and inform more responsive policy design.

CHAPTER 3: A Neural Network approach to model the herding behavior of inflation experts' forecasts

I. Introduction

Forecasters do not form their expectations completely independently; instead, they are often influenced by the work or opinions made by others experts in the community. In other words, rather than formulating forecasts in isolation, experts tend to adjust their expectations based on those of their colleagues (Banerjee, 1992).

Experts herd for several reasons, not only due to reputational concerns (Scharfstein and Stein, 1990), but also because of uncertainty aversion (Banerjee, 1992), or shared information sources (Bikhchandani, Hirshleifer, and Welch, 1992). These theories suggest that experts may ignore their private signals and access to information, and imitate others' behavior, especially when they believe other experts have "better" information. Building on this perspective, Devenow and Welch (1996) introduced the idea of “rational herding” by which agents in financial markets may herd not only for informational reasons but also strategically to preserve or improve their reputations. These theories conclude that herding arises endogenously from information asymmetries and strategic interactions.

Several empirical studies have tested whether expert forecasters display herding behavior or not. One test measures deviation from the consensus, for example Rülke, Silgoner, and Wörz (2016), utilize individual Consensus Economics data for various countries and apply tests of herding based on how each GDP forecast deviates from the consensus. They measure the frequency and values of such deviations and find systematic anti-herding, especially at the year-ahead horizon. Clements (2018) demonstrates that standard consensus-clustering tests for inflation and output growth can yield false positives in the presence of noise or informational rigidities. They use U.S. SPF quarterly forecasts for inflation and output (1981–2013) and test the robustness to these issues; nevertheless, they do not find evidence of herding. Another test bases on the frequency in which experts coincide with the mean, Clements (2015) applies the Bernhardt, Campello, and Kutsoati (2006) frequency test to U.S. inflation and growth forecasts, documenting that forecasters tend to exaggerate differences, except at the shortest horizon, where some herding behavior is observed. Capistrán and Timmermann (2009) apply dispersion approaches to measure dispersion, bias, and the serial correlation of forecast errors, showing that disagreement rises with the level and variance of inflation. They also find that many forecasters exhibit bias, including a shift around the early 1980s, patterns consistent with incentives that can mimic herding without requiring it. Outside the U.S., a South Africa study applies the Bernhardt–Campello–Kutsoati test to Bloomberg short-horizon inflation forecasts (2000–2014). This test measures the frequency with which individual forecasts align with the consensus and identifies herding behavior when forecasters primarily observe past consensus, which is more pronounced under high inflation volatility, but exhibits anti-herding behavior when volatility is low.

A small but growing body of literature has begun to apply machine learning techniques to the study of herding behavior, but it has only been applied to financial studies. For example, Rique, Hosein, and Arjoon (2019) analyze herding patterns in the Singapore stock exchange using regression-based machine learning methods and find that herding is particularly pronounced among high-volume stock price returns. Similarly, Asim, Khan, and Shafi (2024) use support VAR models to study herding in the UK stock market, focusing on periods of high uncertainty.

However, both traditional econometric and emerging ML-based studies tend to only detect the herding behavior rather than focusing on how the herding expert’s network occurs and evolves. An initial approach in this direction is Bales (2022), who applies wavelet-based network analysis to capture time–frequency contagion patterns in financial systems. This framework captures dynamic spillovers among experts over multiple time horizons; nevertheless, it does not cover learning-based representations to detect how relationships and connections change over time. Another relevant contribution is by Hommes et al. (2008), who study expectation formation in experimental asset markets and show that even in a simplified environment, participants tend to coordinate on common prediction strategies, leading to speculative bubbles through trend-following behavior. This work highlights the endogenous nature of coordination and herding dynamics, but does not model the evolving network structure of influence among forecasters over time.

To address how herding expert networks occur and evolve, we introduce a novel model based on Graph Neural Networks (GNNs) in U.S. experts' inflation year-ahead forecasts from the Consensus Economics dataset from January 2010 to June 2022. GNNs provide a graphical representation of how relationships among experts evolve. I capture herding behavior by isolating the expert-specific expectation of each forecast, excluding the influence of external factors, using a residual-based filtering approach (Ang, Bekaert, and Wei, 2007). This method regresses individual forecasts on external variables (GDP, realized inflation, interest rates, Economic Policy Uncertainty, and Geopolitical Risk indexes). By keeping only, the residuals, I isolate the expert's inherent expectation, thereby eliminating the influence of external factors.

Based on these expert-specific expectations, I apply GNN using cosine similarity (Jin et al., 2021) to model the dynamic interactions between experts at each time period. From the graphical representation of the network, I capture the structure and evolution of herding behavior using different network metrics (network density, clustering, and centrality). Density and clustering directly measure the intensity and formation of herding behavior, as these metrics indicate how closely experts align or group their expectations. Others, like centrality measures, capture experts' influence and structural roles in shaping consensus.

The GNN results indicate that experts primarily cluster into two groups. One group of experts concentrates around the mean of forecast values, and the other around realized inflation values. The network metrics analysis indicates that herding behavior changes over time and increases during periods of macroeconomic stability, while weakening in moments of elevated uncertainty. In particular, we observe reduced forecast dispersion and more substantial alignment among experts prior to 2020, followed by a notable breakdown in herding during the COVID-19 crisis. These findings suggest that in uncertain environments, forecasters rely more on private signals and deviate from the consensus—consistent with the insights of Andrade and Le Bihan (2013) and Coibion, Gorodnichenko, and Kamdar (2018).

Overall, this study detects herding behavior and shows how the dynamic interactions between experts change over time. The GNN approach provides a richer understanding of experts' forecasting dynamics since we can track how interactions between experts change over time. This study contributes to methodological advancements in the literature by leveraging state-of-the-art ML tools to capture complex behavioral patterns, while also offering new insights into how and why herding behavior unfolds over time, with implications for economic forecasting, policy design, and the stability of informational networks.

This chapter is structured as follows. Section II presents the dataset used in the analysis. The following sections explain the methodological approach, beginning with the traditional measure of herding behavior, followed by the network-based methodology, and compare its results with the traditional approach. Finally, I discuss the findings and conclude.

II. Data

We use Consensus Economics (CE) forecast data to study the herding behavior of expert forecasters. We focus on U.S. year-ahead inflation expectations from January 2010 to June 2022 on a monthly basis. Since not all experts report their expectations every month, we decided to retain only experts who report 85% of the total 149 months in our timeframe, resulting in 22 experts.

Nevertheless, there are still missing values, so we employ linear interpolation to complete the database. The interpolation method estimates missing values based on the closest data points, preserving the original dataset structure, trends, and cycles (Lepot, Aubin, and Clemens 2017). The linear interpolation follows equation (12), where missing values are $\pi_{i,t}$.

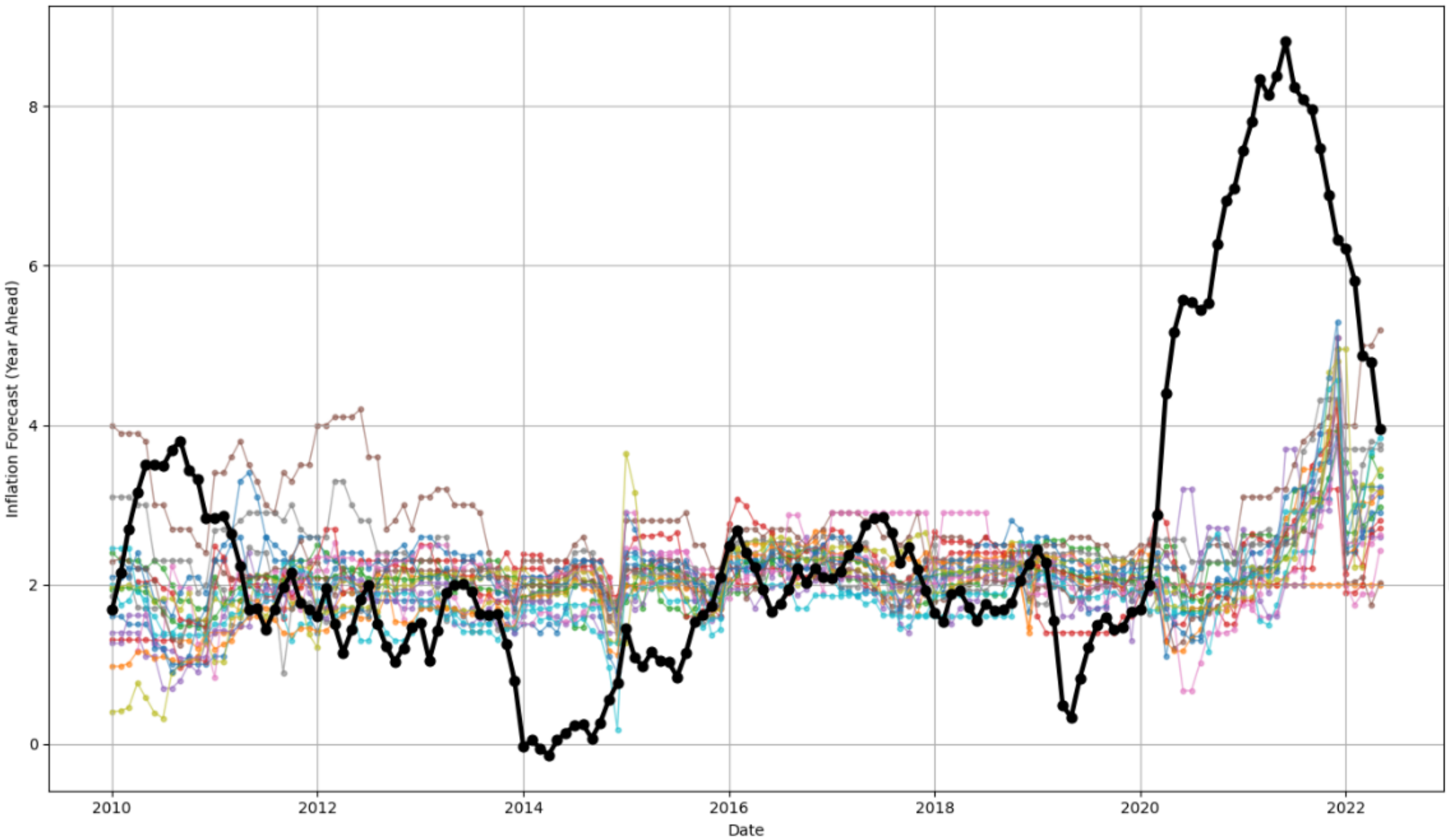
$$\pi_{i,t} = \frac{(\pi_{i,t+1} - \pi_{i,t-1})}{2} \qquad (12)$$

$\pi_{i,t-1}$ and $\pi_{i,t+1}$ represent the closest available data points before and after the missing observation.

The method has several advantages (Lepot, et al., 2017): (i) it is non-intrusive because it does not artificially smooth fluctuations in the data, (ii) it preserves the time-series structure, ensuring that trends and cycles remain present, and (iii) it avoids distortions introduced by alternative methods, such as mean imputation, which can reduce variability and bias results.

Figure 6 illustrates the final dataset of all experts, with the black lines representing the realized inflation values. The first four years remain relatively stable around 2%, followed by a dispersion between 2014 and 2015, where experts' expectations did not align with the realized values. This scenario repeats between 2020 and 2022.

Figure 6. Experts' inflation year-ahead forecast and realized values



III. Herding behavior

The herding behavior concept, introduced by Banerjee (1992), explains how experts adjust their expectations based on the forecasts of others rather than their own internal expectations. To quantify the herding effect, we follow Lamont (2002) and Ehrbeck and Waldmann (1996), who measure it as the absolute deviation of each institution's forecast from the consensus (mean forecast):

$$H_t = \frac{1}{N}\sum_{i=1}^{N}(F_{i,t} - \bar{F}_t)^2 \qquad (13)$$

Where $F_{i,t}$ represents the standard deviation of inflation forecasts for expert i in month t, and $\bar{F}_t$ is the average forecast across all N experts in that month. A higher H_t suggests higher dispersion in forecasts, implying less herding, while a lower value suggests forecasts are more similar, implying stronger herding behavior.

Figure 7. The evolution of herding behavior

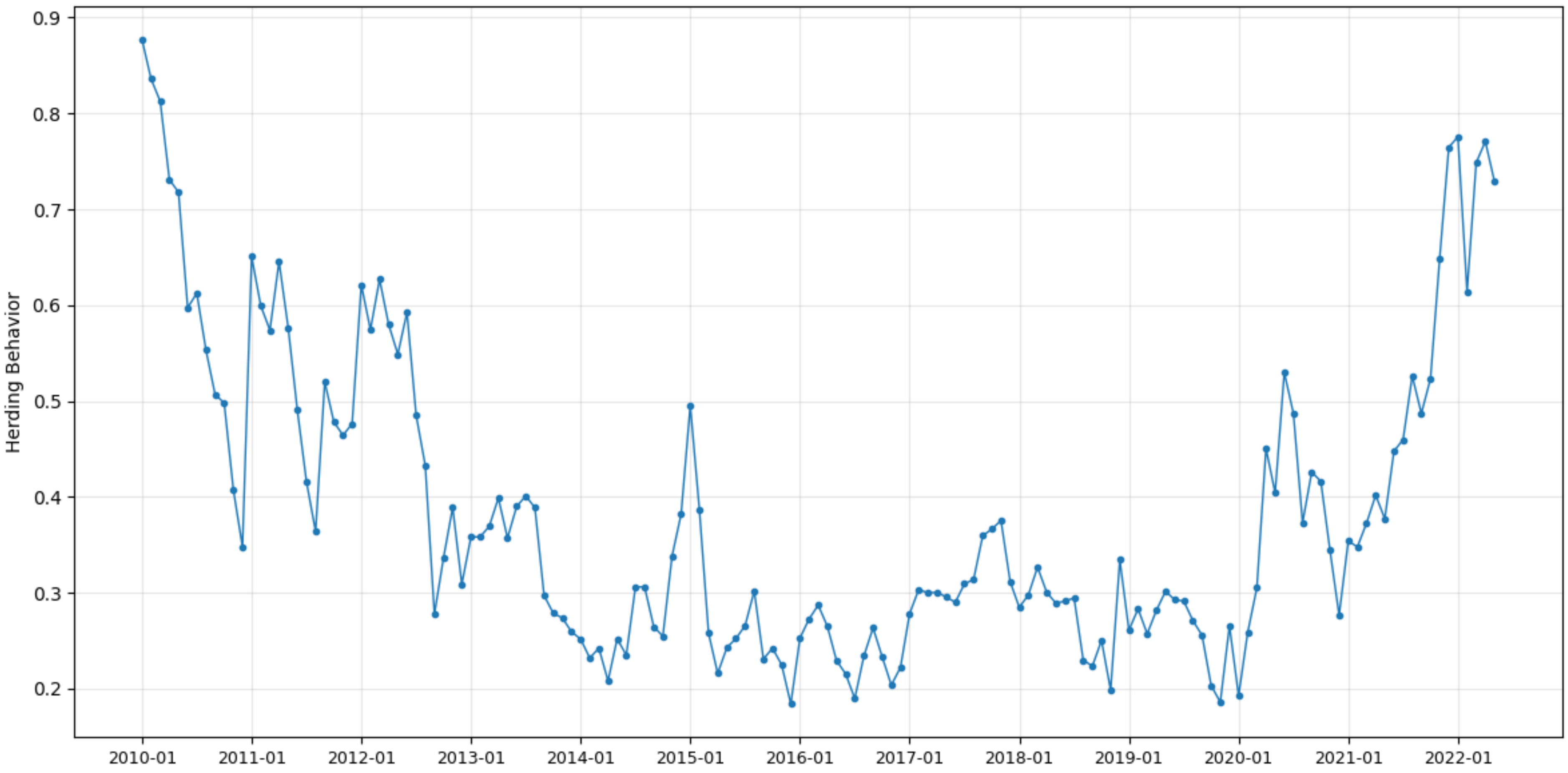


Figure 7 illustrates the herding behavior among U.S. inflation experts from 2010 to 2022, on a monthly basis. From 2010 to around 2019, herding appears to intensify, which is consistent with the findings of Mankiw, Reis, and Wolfers (2003) and Capistrán and Ramos-Francia (2010), who show that in periods of low inflation volatility and credible monetary policy, forecast disagreement tends to decline.

Starting in 2020, however, herding behavior declines sharply, with dispersion in forecasts rising significantly. This shift happens in the COVID-19 pandemic, followed by an inflationary surge in 2021–2022. The spike in forecast dispersion during this period suggests that forecasters relied more on their private information and expectations—possibly due to elevated uncertainty, conflicting signals from global supply shocks, and shifts in monetary and fiscal policy. These results align with Andrade and Le Bihan (2013) and Coibion, Gorodnichenko, and Kamdar (2018), who show that forecast disagreement tends to increase during periods of heightened macroeconomic uncertainty.

Overall, the observed trend also confirms the theoretical insights of Ehrbeck and Waldmann (1996) and Lamont (2002), which suggest that when the cost of deviating from the consensus is low and uncertainty is high, professional forecasters tend to vary more in their expectations, resulting in lower herding behavior.

IV. A new herding approach: Dynamic behavior Network

The traditional way of measuring herding focuses only on its detection; literature does not analyze how experts' relationships form, dissolve, and rewire as conditions change over time. We therefore model herding as a dynamic behavior network that explicitly tracks the evolution of connections and relationships among experts over time.

Graph Neural Networks (GNNs) offer a more flexible approach, enabling the identification of clustering patterns across experts and detecting shifts over time. Unlike traditional econometric techniques that only test the presence or intensity of herding behavior, GNNs can capture the how the networks form, the connections inside the networks, and how the relationship evolve (Scarselli et al., 2008; Kipf and Welling, 2017), tracking how these relationships form and evolve over time. Using a GNN with a cosine similarity metric, we can measure how closely experts' forecasts align or not over time in their inflation forecasts.

The first step in constructing the GNN is reshaping the dataset into a structured format, where each month data is arranged into a matrix:

$$M = \begin{bmatrix} \varepsilon_{1,t} & \varepsilon_{1,t+1} & \cdots & \varepsilon_{1,T} \\ \varepsilon_{2,t} & \varepsilon_{2,t} & \cdots & \varepsilon_{2,T} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{N,t} & \varepsilon_{N,t+1} & \cdots & \varepsilon_{N,T} \end{bmatrix} \qquad (14)$$

Each row represents an expert, and each column represents inflation expectation over time. This matrix structure enables us to systematically measure forecasting similarity across experts.

I use cosine similarity to quantify the degree of similarity between experts' inflation expectations. This metric quantifies the relationship of inflation expectations made by different experts, and it is a widely used metric in machine learning and network science (Blasques et al., 2021)[22]. For example, in Equation 4, A and B are vectors of inflation expectations provided by two different experts.

$$\text{Cosine similarity}(A, B) = \frac{A \cdot B}{\|A\| \cdot \|B\|} \qquad (15)$$

Cosine similarity normalizes differences, captures directional alignment rather than absolute magnitude variations, and it is more robust to outliers than euclidean distance metrics (Aggarwal, 2001). The output of equation 4 is a symmetric expert-by-expert similarity matrix, where each element reflects the degree of alignment, which we consider as herding. We apply a similarity threshold to ensure that the network reflects only meaningful relationships of herding, i.e., retaining only edges where the cosine similarity score exceeds 0.75.

$$S_{i,j} = \begin{cases} \text{cosine similarity}(i, j), & \text{andif cosine similarity}(i, j) > 0.75 \\ 0, & \text{andotherwise} \end{cases}$$

Where $S_{i,j}$ is the similarity matrix element capturing the forecasting similarity between expert i, j.

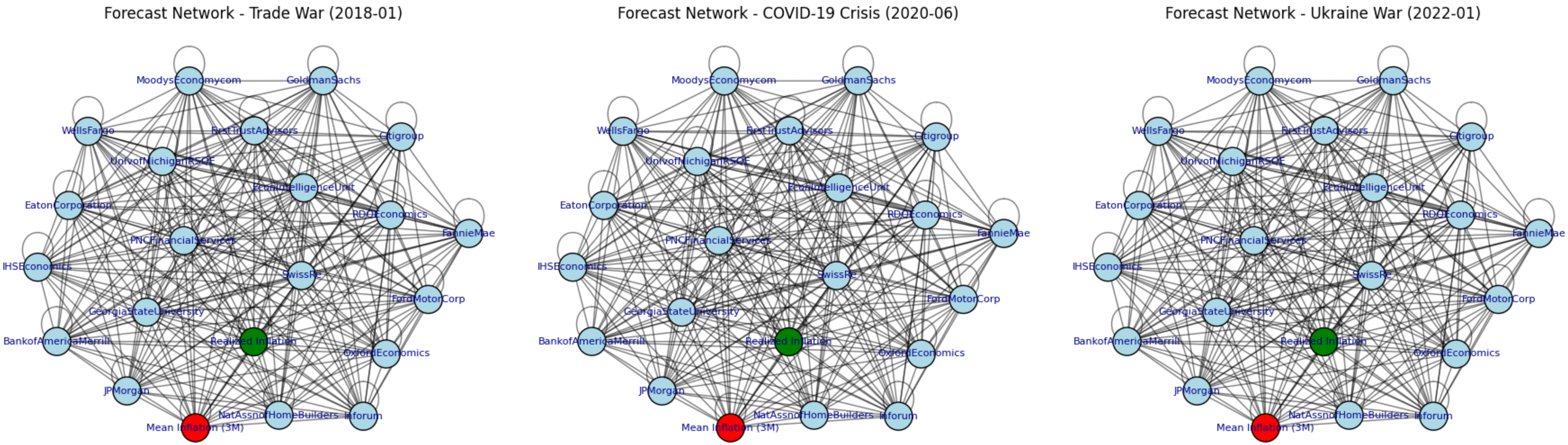
Rather than merely testing for the presence or intensity of herding, GNNs extract latent representations of experts and their ties that evolve jointly with the network by learning from the dynamic structure created in the similarity matrix. GNN and the similarity matrix enable to capture who experts becomes more aligned with whom, and when clusters cohere or disperse (Scarselli et al., 2009; Kipf and Welling, 2017). The GNN is trained in the following way: given graphs up to time t, the model predicts the future edge weight between experts i and j; parameters are learned by minimizing the error between predicted and observed connections.

The graphical representation of the GNN and the similarity matrix shows the connections between experts, where each node represents one of them, with links indicating strong alignment in forecast patterns (based on a cosine similarity threshold).

In Figure 8, I present the output of the GNN for three significant and recent macroeconomic events, as they have a notable impact on expert expectations and how they process the new information. These events are the U.S.-China Trade War, which has important implications in long policy decisions (Caldara et al., 2020); the COVID-19 pandemic for its impact on supply and demand shocks (Baker et al., 2020; Altig et al., 2020); and finally, the Ukraine War for its impact on energy prices and global inflation (Götz and Stjepanovic, 2023). These moments of time align with the sharp fluctuations in forecast dispersion observed in Figure 6 and herding picks in Figure 7. These results also align with findings from Bordalo et al. (2020) and Angeletos et al. (2020), which show that experts tend to differ significantly during periods of geopolitical and macroeconomic stress. To further contextualize these networks, we also include two synthetic nodes: the mean forecast (in red) and realized inflation (in green). These two nodes enable the comparison of realized inflation values, collective expectations, and individual expert behavior.

[22] See, for example, Cai (2022), who uses cosine similarity across banks' balance sheets to assess how similar their positions are. The author considers higher cosine similarity as higher herding behavior.

Figure 8. Forecast Neural Network at an expert level without a filtering approach



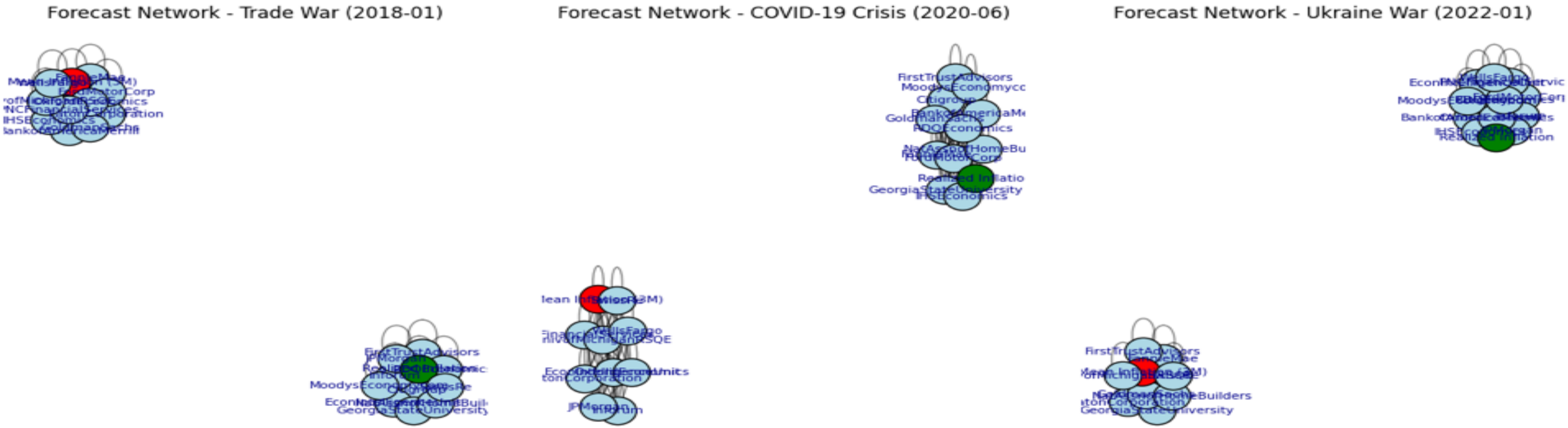
In Figure 8, we can see how all experts are interconnected, suggesting a high level of herding behavior. Nevertheless, these connections can also arise due to external factors that all experts face simultaneously. Since this research aims to understand pure herding behavior, we decided to isolate experts' specific inflation expectations from macroeconomic and political shocks. By isolating the residual, we effectively remove all external influences and retain only the expert-specific component of the forecast — allowing us to analyze pure herding behavior and forecast heterogeneity among forecasters. We use a residual-based filtering approach (Ang et al., 2007) where we estimate an OLS regression between the forecast made by each expert against a macro-political variable (see equation 16) in order to calculate the residual term.

$$\pi_{i,t+1} = \beta_0 + \beta_1 EPU_{i,t} + \beta_2 GPR_{i,t} + \beta_3 GDP_{i,t} + \beta_4 RI_{i,t} + \beta_5 IR_{i,t} + \varepsilon_{1,t} \quad (16)$$

Where $\pi_{i,t+1}$ is the one-year-ahead inflation forecast for expert *i* in month *t*, $EPU_{i,t}$, is the Economic Policy Uncertainty index, $GPR_{i,t}$ is the Geopolitical Risk index, $GDP_{i,t}$ represents GDP growth forecasts, $RI_{i,t}$, is realized inflation, and $IR_{i,t}$ is the realized interest rate. The residual term $\varepsilon_{1,t}$ captures expert-specific inflation expectations, removing the impact of external macroeconomic uncertainty and geopolitical shocks.

Figure 9 presents the GNN networks for this expert-specific component of the experts. Contrary to Figure 8, where all experts were interconnected, in Figure 9 we can identify two main groups of experts: one centered around the mean of forecast values and the other centered around the realized value. During the Trade War and the beginning of the Ukraine War, the expert network shows a moderate density and clustering, suggesting partial convergence but still allowing for diverse interpretations. However, during the COVID-19 crisis, the network fragments dramatically, indicating reduced herding and heightened dispersion as forecasters relied more on their information. This change is also shown in Figure 7 by a notable spike in the herding index, consistent with literature such as Andrade and Le Bihan (2013) and Coibion, Gorodnichenko, and Kamdar (2018), which document increased disagreement during uncertainty shocks. Additionally, the weakening of links to the mean forecast and the realized inflation value during these turbulent episodes reinforces the idea that forecast accuracy and convergence deteriorate under stress.

Figure 9. Forecast Neural Network at expert level



V. Dynamics of the Network metrics over time

The graphical representation of experts' neural networks allows us to see how forecasters connect with each other and how those relationships evolve over time. From the GNN, we can also capture the structure and evolution of herding behavior by using different network metrics, such as network density, clustering, centrality, betweenness, and eigenvector measures (D'Arcangelis, and Rotundo, 2021). Some of these metrics—particularly density and clustering—directly reflect the intensity and formation of herding behavior, indicating how closely forecasters align or group together in their expectations. Others, such as centrality measures, capture the influence and structural roles that experts play in shaping consensus.

(i) Network density (D) measures the proportion of realized connections relative to all possible connections. A higher density indicates greater consensus among forecasters.

$$D = \frac{2E}{N(N-1)} \qquad (17)$$

where E is the number of edges, and N is the number of forecasting experts.

(ii) Clustering coefficient (C) show the degree in which experts form tightly connected clusters. A higher clustering coefficient implies the presence of forecasting groups that share similar expectations.

$$C = \frac{1}{N}\sum_{i=1}^{N}\frac{2P_i}{k_i(k_i-1)} \qquad (18)$$

Where T_i is the number of triangles expert i is part of, and k_i represents its degree (number of direct connections).

(iii) Degree Centrality (DC) measures the relative importance of each expert by counting the number of direct connections it has.

$$DC(i) = \frac{k_i}{N-1} \qquad (19)$$

(iv) Betweenness Centrality (BC) identifies experts who act as bridges in the network. σ_{st} is the total number of shortest paths between expert s and t; while $\sigma_{st}(i)$ is the number of shortest paths that go through node i.

$$BC = \sum_{s \neq i \neq t}\frac{\sigma_{st}(i)}{\sigma_{st}} \qquad (20)$$

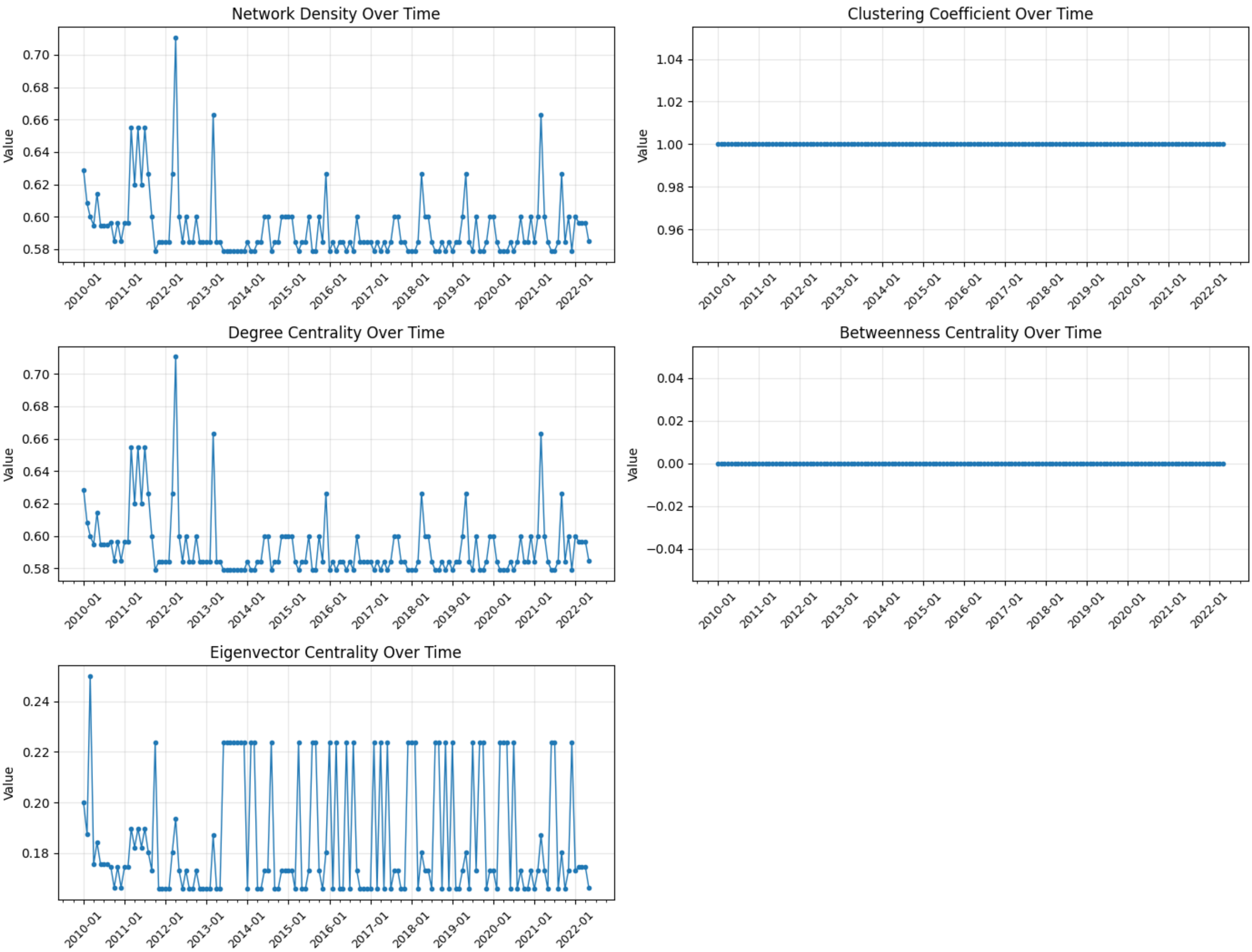
(v) Eigenvector Centrality (EC) measures the influence of experts based on the importance on the neighbor. λ is the largest eigenvalue of the similarity matrix.

$$EC(i) = \frac{1}{\lambda}\sum_{j \in N} S_{i,j} EC(j) \qquad (21)$$

Figure 10 shows how these network metrics evolve. Density measures the number of connections among all possible ones, which reflects the degree of alignment in inflation expectations among forecasters. From 2010 to around 2012, density gradually increases, suggesting stronger consensus or herding among experts—a trend that directly aligns with the declining dispersion shown in Figure 6. However, during the COVID-19 crisis, density drops, indicating that herding behavior reduces, similar to what is shown in Figure 6. The clustering coefficient remains high over time, close to 1, indicating the presence of tightly knit groups of forecasters who remain closely aligned with each other. This suggests that while herding may weaken at the broader network level during periods of macroeconomic stress (e.g., COVID-19), local or subgroup persistence may occur. Although it is more dispersed and without a high level of herding.

On the other hand, centrality metrics— degree, betweenness, and eigenvector centrality—offer deeper insight into the structural roles and influence of experts within the forecasting network. DC captures how connected an expert is by counting their direct links to others; high DC values indicate consensus between experts. Figure 10 shows that DC metric increases during stable periods (e.g., pre-COVID-19) and then fall due to uncertainty. These changes suggest that during stable periods there is high level of herding, while in periods with uncertainty there is less herding behavior. BC identifies how experts act as bridges. The BC metric also becomes more volatile during crises like COVID-19 or the Ukraine War. EC measures the influence based on the importance of an expert's neighbors. This metric remains stable in calm periods but declines during shocks, as the network becomes more decentralized. All these shifts mirror findings from Hommes et al. (2008), who show that in experimental asset markets, forecasters often coordinate on shared heuristics even without direct communication—leading to bubbles driven by trend-chasing and feedback loops. Just as high centrality nodes in our network shape consensus, participants in Hommes' experiments developed synchronized expectations and errors, reinforcing the idea that structural influence and collective alignment can emerge endogenously—even under bounded rationality and minimal external shocks.

Figure 10. Network metrics



VI. Conclusion

This study presents a new approach to study herding behavior in experts' inflation expectations by integrating GNN into macroeconomic forecasting analysis. While previous studies focused on identifying the presence of herding, this research takes it a step further by modeling how experts' relationships and interactions evolve. I use a GNN with a residual-based filtering approach to isolate expert-specific expectations, to enable a more accurate representation of herding behavior.

By applying GNNs to inflation forecasting data, I show how machine learning tools can uncover hidden structures and dynamic relationships that traditional econometric techniques did not cover yet. The GNN approach allows to identify herding forms and how the structure of expert networks evolves. The use of network metrics, such as density, clustering, and centrality, gives a deeper characterization of the network environment. The metrics reveal patterns of alignment, influence, and fragmentation among experts. For example, during periods of uncertainty, such as the COVID-19, the networks become more fragmented, while in more stable times, connections between experts strengthen, forming more cohesive clusters. Another important result is that experts tend to organize into two main clusters: one centered around the mean forecast and the other around the realized inflation value. Given these advantages and results, applying GNNs represents a significant step forward in understanding expert networks and herding behavior.

This new methodological approach and findings also have important policy implications. For example, central banks who rely on experts to assess inflation expectations and inform their policy decisions they can monitor the structure and evolution of expert networks and have a valuable real-time indicator of expectation coordination.

In conclusion, this study presents a new approach of GNN in the analysis of herding behavior in expert inflation forecasts. It demonstrates that experts do not work independently but instead are influenced by networks well-defined around mean forecasts and realized values. The proposed methodology enables the detection of common patterns and the identification of influential experts, as well as shifts in forecast dispersion, during periods of uncertainty. Furthermore, this study contributes to a recent line of research by bridging the fields of network science, behavioral economics, and machine learning.

CONCLUSION

This thesis set out to understand how and why experts deviate from rational expectations in two core areas of macroeconomics: inflation and public finances. Starting from the point that expert forecasts are crucial for central banks, fiscal authorities, and international organizations, this thesis develops a unified empirical framework to identify and explain deviations from rational expectations. Unlike previous literature, which has typically examined biases, informational frictions, or external shocks in isolation, this thesis integrates them within a single framework and observes how they evolve over time and across forecasters. The result is a richer and more dynamic account of expectation formation that combines methodological innovation with empirical evidence.

The first chapter, "Fiscal Forecasting Rationality among Expert Forecasters" analyzes budget balance forecasts for the U.S., France, Germany, and Italy between 1993 and 2023. Applying the rationality test and a panel methodology to examine under what conditions expert forecasters rationally incorporate economic and political information into their expectations and when this rationality breaks down. This chapter showed that experts generally react efficiently under normal conditions, systematically integrating economic and political news, but tend to overreact during major structural shifts, such as the build-up to euro area fiscal consolidation, the Global Financial Crisis, or the COVID-19 pandemic. These findings support the view that noisy or sticky information, combined with political uncertainty and macroeconomic volatility, impairs experts' ability to process fiscal data accurately and generates biases and dispersions that do not appear in full-sample averages.

The second chapter, "Identifying Drivers of Deviations from Rational Expectations: A New Irrationality Index," focused on U.S. inflation forecasts between 2010 and 2022, a period characterized by prolonged stability as well as extraordinary shocks such as the COVID-19 pandemic and the surge in geopolitical risk. This chapter develops a new Irrationality Index based on the Fluctuation Rationality (FR) test of Rossi and Sekhposyan (2016), enabling individual and time-varying measurement of deviations from rational expectations. By combining metrics of bias, herding behavior, informational frictions, and external shocks, the chapter shows that deviations from rationality are not random but respond to identifiable factors. The main finding is that external information shocks, long-term interest rate expectations, and geopolitical risk explain a large share of the deviations. Nevertheless, traditional behavioral indicators, such as bias or herding, play a secondary role once all factors are considered jointly.

The third chapter "A Neural Network Approach to Model the Herding Behavior of Inflation Experts' Forecasts" introduce a Graphical Neural Networks to study how experts hearding relationships. This new approach provides a dynamic visualization of how experts align in two main hearding groups, one around the mean of expectations, and the other, around the realized value of inflation.

Taken together, the three chapters offer a multi-layered view of expectation formation: (i) fiscal rationality and structural shifts; (ii) behavioral and informational deviations through the Irrationality Index; and (iii) collective dynamics and herding networks. The integration of econometric and AI tools contributes to a unified framework for understanding deviations from rational expectations across domains and time.

Another key point is the thesis reliance on a single panel of expert's forecasters: Consensus Economics, across all three chapters. By using the same data source, the thesis ensure consistency in definitions, timings, and institutional coverage, so the differences in behavior forecast can be attributed to economic mechanisms rather than data heterogeneity.

The three chapters demonstrate a complete understanding of expectation formation. First, deviations from rational expectations are recurrent and predictable and driven more by structural factors-external shocks, regime changes, interest rate expectations, and geopolitical risk-than by isolated behavioral biases. Second, the thesis offers a comparative perspective by applying a single methodological framework to two fiscal and monetary areas. It establishes a basis for understanding how uncertainty and policy changes jointly shape the accuracy and dispersion of expert expectations. Third, the integration of machine learning techniques to study how evolves hearding relationships among time.

Finally, these findings have direct policy implications. Transparent, credible, and consistent communication by central banks and fiscal authorities can help anchor expectations, reduce informational frictions, and avoid overreactions to unexpected shocks. Moreover, the Irrationality Index could be extended to other macroeconomic variables, integrated with higher-frequency data, and combined with machine-learning techniques to assess how experts process qualitative and quantitative information in real time.

REFERENCES

Adrian, T., Boyarchenko, N., and Giannone, D. (2020). Vulnerable growth. American Economic Review, 110(4), 1436–1471.

Afrouzi, H., Bordalo, P., Gennaioli, N., and Shleifer, A. (2021). Overreaction in macroeconomic expectations. American Economic Journal: Macroeconomics, 13(3), 132–160.

Aggarwal, C. (2001). Re-designing distance functions and distance-based applications for high dimensional data. ACM Sigmod Record, 30(1), 13-18.

Altig, D., Baker, S., Barrero, J., Bloom, N., Bunn, P., Chen, S., and Mizen, P. (2020). Economic uncertainty before and during the COVID-19 pandemic. Journal of Public Economics, 191, 104274.

An, Z., and Jalles, J. (2021). On the performance of US fiscal forecasts: Government vs. private information. Journal of Economic Studies, 48(2), 367-391.

Andrade, P., and Le Bihan, H. (2013). Inattentive professional forecasters. Journal of Monetary Economics, 60(8), 967-982.

Angeletos, G., Collard, F., and Dellas, H. (2020). Quantifying confidence. Econometrica, 88(5), 1769–1808.

Angeletos, G., and Lian, C. (2018). Forward guidance without common knowledge. American Economic Review, 108(9), 2477–2512.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies, 58(2), 277–297.

Artis, M. and Marcellino, M. (2001). Fiscal forecasting: The track record of the IMF, OECD and EC. The Econometrics Journal, 4(1), 20-36.

Asim, M., Khan, M., and Shafi, K. (2024). Investigation of herding behavior using machine learning models. Review of Behavioral Finance, 16(3), 424-438.

Auerbach, A., and Gale, W. (1999). Does the budget surplus justify large-scale tax cuts? Updates and extensions. Tax Notes, 85, 369-376.

Baker, S., Bloom, N., and Davis, S. (2016). Measuring economic policy uncertainty. Quarterly Journal of Economics, 131(4), 1593-1636.

Baker, S., Bloom, N., Davis, S., and Terry, S. (2020). COVID-induced economic uncertainty. NBER Working Paper No. 26983.

Bales, S. (2022). Sovereign and bank contagion in the Eurozone: A multi-scale approach using wavelet-network analysis. International Review of Financial Analysis, 83.

Banerjee, A. (1992). A simple model of herd behavior. The quarterly journal of economics, 107(3), 797-817.

Barberis, N., and Thaler, R. (2003). A survey of behavioral finance. Handbook of the Economics of Finance (pp. 1053–1123). Elsevier.

Beetsma, R., Debrun, X., and Sloof, R. (2022). The political economy of fiscal transparency and independent fiscal councils. European Economic Review, 145, 104118.

Beetsma, R., Giuliodori, M., De Jong, F., and Widijanto, D. (2013). Spread the news: The impact of news on the European sovereign bond markets during the crisis. Journal of International Money and Finance, 34, 83-101.

Bekaert, G., Hoerova, M., and Lo Duca, M. (2013). Risk, uncertainty, and monetary policy. Journal of Monetary Economics, 60(7), 771–788.

Berg, R., Kipf, T., and Welling, M. (2017). Graph convolutional matrix completion. arXiv preprint arXiv:1706.02263.

Bernhardt, D., Campello, M., and Kutsoati, E. (2006). Who herds? Journal of Financial Economics, 80(3), 657–675.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of political Economy, 100(5), 992-1026.

Blanchard, O., and Watson, M. (1982). Bubbles, rational expectations and financial markets. In P. Wachtel (Ed.), Crises in the Economic and Financial Structure (pp. 295–316). Lexington Books.

Bloom, N. (2009). The impact of uncertainty shocks. Econometrica, 77(3), 623–685.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2020). Overreaction in macroeconomic expectations. American Economic Review, 110(9), 2748–2782.

Caldara, D., and Iacoviello, M. (2022). Measuring geopolitical risk. American Economic Review, 112(4), 1194-1225.

Caldara, D., Iacoviello, M., Molligo, P., Prestipino, A., and Raffo, A. (2020). The economic effects of trade policy uncertainty. Journal of Monetary Economics, 109, 38–59.

Capistrán, C., and Ramos-Francia, M. (2010). Does inflation targeting affect the dispersion of inflation expectations? Journal of Money, Credit and Banking, 42(1), 113-134.

Capistrán, C., and Timmermann, A. (2009). Disagreement and biases in inflation expectations. Journal of Money, Credit and Banking, 41(2–3), 365–396.

Carabotta, L., and Claeys, P. (2024). Combine to compete: Improving fiscal forecast accuracy over time. Journal of Forecasting, 43(4), 948-982.

Carroll, C. (2003). Macroeconomic expectations of households and professional forecasters. Quarterly Journal of Economics, 118(1), 269–298.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. Review of Economic Studies, 47(1), 225–238.

Cimadomo, J., Claeys, P., and Poplawski-Ribeiro, M. (2016). How do experts forecast sovereign spreads? European Economic Review, 87, 216-235.

Clements, M. (2015). Do US macroeconomic forecasters exaggerate their differences? Journal of Forecasting, 34(8), 649–660.

Clements, M. (2018). Do macroforecasters herd? Journal of Money, Credit and Banking, 50(2–3), 265–292.

Coibion, O., and Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities? Journal of Political Economy, 120(1), 116–159.

Coibion, O., and Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. American Economic Review, 105(8), 2644–2678.

Coibion, O., Gorodnichenko, Y., and Kamdar, R. (2019). The formation of expectations, inflation, and the Phillips curve. Journal of Economic Literature, 57(4), 135–188.

Cornand, C., and Hubert, P. (2020). On the external validity of experimental inflation forecasts: A comparison with five categories of field expectations. Journal of Economic Dynamics and Control, 110, 103746.

D’Agostino, A., and Ehrmann, M. (2014). The pricing of G7 sovereign bond spreads. Journal of Banking and Finance, 47, 155-176.

D'Arcangelis, A. M., and Rotundo, G. (2021). Herding in mutual funds: A complex network approach. Journal of Business Research, 129, 679-686.

Davies, A., and Lahiri, K. (1995). A new framework for analyzing survey forecasts using three-dimensional panel data. Journal of Econometrics, 68(1), 205–227.

Devenow, A., and Welch, I. (1996). Rational herding in financial economics. European economic review, 40(3-5), 603-615.

Dovern, J., and Weisser, J. (2011). Accuracy, unbiasedness and efficiency of professional macroeconomic forecasts: An empirical comparison for the G7. International Journal of Forecasting, 27(2), 452–465.

Dovern, J., Fritsche, U., Loungani, P., and Tamirisa, N. (2015). Information rigidities: Comparing average and individual forecasts for a large international panel. International Journal of Forecasting, 31(1), 144-154.

Ehrbeck, T., and Waldmann, R. (1996). Why are professional forecasters biased? Agency versus behavioral explanations. Quarterly Journal of Economics, 111(1), 21–40.

Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. Journal of Finance, 25(2), 383–417.

Fildes, R., and Goodwin, P. (2007). Against your better judgment? How organizations can improve their use of management judgment in forecasting. Interfaces, 37(6), 570–576.

Frankel, J., and Schreger, J. (2016). Bias in official fiscal forecasts: Can private forecasts help? (No. w22349). National Bureau of Economic Research.

Gennaioli, N., and Shleifer, A. (2015). Expectations and investment. In Handbook of Field Experiments (Vol. 2, pp. 379–402).

Glick, R., and Kouchekinia, N. (2021). Disagreement about US and Euro-area inflation forecasts. FRBSF Economic Letter, 8, 1-5.

Götz, T., and Stjepanovic, A. (2023). War in Ukraine and global inflation: Evidence from survey expectations. European Central Bank Working Paper Series, No. 2807.

Greene, W. (2012). Econometric analysis (7th ed.). Prentice Hall.

Gürkaynak, R., Sack, B., and Swanson, E. (2005). Do actions speak louder than words? The response of asset prices to monetary policy actions and statements. International Journal of Central Banking, 1(1), 55–93.

Heppe-Falk, K., and Hufner, F. (2004). Expected budget deficits and interest rate swap spreads: Evidence for France, Germany, and Italy. Deutsche Bundesbank Discussion Paper, 40/2004.

Holden, K., and Peel, D. (1990). On testing for unbiasedness and efficiency of forecasts. The Manchester School of Economic and Social Studies, 58(2), 120-127.

Hommes, C., Sonnemans, J., Tuinstra, J., and Van de Velden, H. (2008). Expectations and bubbles in asset pricing experiments. Journal of Economic Behavior and Organization, 67(1), 116-133.

Jalles, J. (2015). How quickly is news incorporated in fiscal forecasts? Economics Bulletin, 35(4), 2802-2812.

Jalles, J., Karibzhanov, I., and Loungani, P. (2015). Cross-country evidence on the quality of private sector fiscal forecasts. Journal of Macroeconomics, 45, 186-201.

Jin, W., Derr, T., Wang, Y., Y., Liu, Z., and Tang, J. (2021). Node Similarity Preserving Graph Convolutional Networks (SimP-GCN). WSDM '21, 148–156

Jonung, L., and Larch, M. (2006). Improving fiscal policy in the EU: The case for independent forecasts. Economic Policy, 21(47), 492-534.

Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263–291.

Kang, W., Lee, K., and Ratti, R. (2014). Economic policy uncertainty and firm-level investment. Journal of Macroeconomics, 39, 42–53.

Kang, W., Ratti, R., and Vespignani, J. (2017). Oil price shocks and policy uncertainty: New evidence on the effects of US and non-US oil production. Energy Economics, 66, 536-546.

Keane, M., and Runkle, D. (1990). Testing the rationality of price forecasts: New evidence from panel data. American Economic Review, 80, 714-735.

Lamont, O. (2002). Macroeconomic forecasts and microeconomic forecasters. Journal of Economic Behavior and Organization, 48(3), 265–280.

Leal, T., Pérez, J., Tujula, M., and Vidal, J. (2008). Fiscal forecasting: Lessons from the literature and challenges. Fiscal Studies, 29(3), 347-386.

Lepot, M., Aubin, J., and Clemens, F. (2017). Interpolation in time series: An introductive overview of existing methods, their performance criteria and uncertainty assessment. Water, 9(10), 796.

Maćkowiak, B., Wiederholt, M., and Matejka, F. (2023). Rational inattention: A review. Journal of Economic Literature, 61(1), 1–52.

Mankiw, N., and Reis, R. (2002). Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips curve. The Quarterly Journal of Economics, 117(4), 1295-1328.

Mankiw, N., Reis, R., and Wolfers, J. (2003). Disagreement about inflation expectations. In NBER Macroeconomics Annual 2003, Volume 18 (pp. 209–248). MIT Press.

Merola, R., and Pérez, J. (2013). Fiscal forecast errors: Governments versus independent agencies? European Journal of Political Economy, 32, 285-299.

Mincer, J., and Zarnowitz, V. (1969). The evaluation of economic forecasts. In Economic Forecasts and Expectations (pp. 3–46). NBER.

Moulin, L., and Wierts, P. (2006). How credible are multiannual budgetary plans in the EU? Available at SSRN 2005228.

Muth, J. (1961). Rational expectations and the theory of price movements. Econometrica, 29(3), 315–335.

Newey, W., and West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55(3), 703–708.

Nordhaus, W. (1987). Forecasting efficiency: Concepts and applications. The Review of Economics and Statistics, 667-674.

Orphanides, A., and Williams, J. (2004). Imperfect knowledge, inflation expectations, and monetary policy. In The inflation-targeting debate (pp. 201-246). University of Chicago Press.

Ottaviani, M., and Sorensen, P. (2006). The strategy of professional forecasting. Journal of Financial Economics, 81, 441-466.

Pesaran, M. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica, 74(4), 967–1012.

Pina, Á, and Venes, N. (2011). The political economy of EDP fiscal forecasts: An empirical assessment. European Journal of Political Economy, 27(3), 534-546.

Poplawski-Ribeiro, M., and Rülke, J. (2011). Fiscal expectations under the stability and growth pact: Evidence from survey data. International Monetary Fund.

Rique, G., Hosein, P., and Arjoon, V. (2019). Analysis of herd behavior in stock prices using machine learning. In International Conference on Internet Science (pp. 326-331). Cham: Springer International Publishing.

Romer, C., and Romer, D. (2004). A new measure of monetary shocks: Derivation and implications. American Economic Review, 94(4), 1055–1084.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. Stata Journal, 9(1), 86–136.

Rossi, B., and Sekhposyan, T. (2016). Forecast rationality tests in the presence of instabilities, with applications to federal reserve and survey forecasts. Journal of Applied Econometrics, 31(3), 507-532.

Rülke, J., Silgoner, M., and Wörz, J. (2016). Herding behavior of business cycle forecasters. International Journal of Forecasting, 32(1), 23-33.

Sanders, N., and Manrodt, K. (2003). The efficacy of using judgmental versus quantitative forecasting methods in practice. International Journal of Forecasting, 19(3), 415–425.

Scarselli, F., Gori, M., Tsoi, A., Hagenbuchner, M., and Monfardini, G. (2008). The graph neural network model. IEEE transactions on neural networks, 20(1), 61-80.

Scharfstein, D., and Stein, J. (1990). Herd behavior and investment. The American economic review, 465-479.

Sims, C. (2003). Implications of rational inattention. Journal of Monetary Economics, 50(3), 665-690.

Strauch, R., Hallerberg, M., and von Hagen, J. (2004). Budgetary forecasts in Europe: The track record of stability and convergence programmes. Working Paper Series, 307, European Central Bank.

Trueman, B. (1994). Analyst forecasts and herding behavior. Review of Financial Studies, 7(1), 97–124.

West, K., and McCracken, M. (1998). Regression-based tests of predictive ability. International Economic Review, 39(4), 817–840.

Woodford, M. (2012). Imperfect common knowledge and the effects of monetary policy. In P. Aghion and S. N. Durlauf (Eds.), Handbook of Monetary Economics (Vol. 3A, pp. 61–142). Elsevier.

Wooldridge, J. (2010). Econometric analysis of cross section and panel data (2nd ed.). MIT Press.

APPENDIX

A. Calculation of the forecasted budget balance (as a ratio of GDP)

The CE provides forecasts for the total deficit only in nominal values (local currency). Hence, we follow Heppke-Falk and Hüfner (2004) and Poplawski-Ribeiro and Rülke (2011) to construct a forecast measure of deficit ratio to GDP (percentage of GDP). For that, we cannot simply scale the nominal value deficit forecast by the GDP forecast, since the CEF surveys for growth rates only, and not for the GDP in nominal value.

We construct a measure of the expected nominal year-ahead GDP forecast of forecaster i at month m and year t as follows. In the first step, we take a real-time measure of real GDP in levels for a particular year t. We use the real-time forecast of the same-year real GDP (in levels) coming from the most recent IMF World Economic Outlook (WEO) vintage available at any particular month m of year t. The IMF WEOs are published either in April or October, hence from May to October we use the April issue, and the October issue in the other months.

The second step is to compute the year-ahead GDP forecast in nominal value. We multiply the real-time (WEO) measure of same-year real GDP (in levels), $E_{WEO,t}[y_t]$, by the year-ahead market (Consensus) forecasts for GDP $growth$, $E_{i,t,m}[\Delta y_{t+1}]$, and inflation, $E_{i,t,m}[\pi_{t+1}]$, for each forecaster i at a particular month m of year t. The expected year-ahead nominal GDP value for each country is then

$$E_{i,t,m}[\Delta y_{t+1}] = E_{WEO,t}[y_t] \times (1 + E_{i,t,m}[\Delta y_{t+1}] + E_{i,t,m}[\pi_{t+1}]) \qquad \text{(A.1)}$$

The year-ahead expected budget balance for each country is then:

$$E_{i,t,m}[b_{t+1}] = \frac{E_{i,t,m}[b_{t+1}^{nom}]}{E_{i,t,m}[y_{t+1}]} \qquad \text{(A.2)}$$

where $E_{i,t,m}[b_{t+1}^{nom}]$ is the (CE) forecast of the nominal budget balance by forecaster i in month m of year t for one-year-ahead $t+1$.

B. List of expert forecasters from CE, by country

Italy	France	Germany	US
ABI	AXAInvestmentManagers	Allianz	Action Economics
Allianz	Allianz	BHF Bank	Allianz
Bank of America Merrill	BIPE	Bank of America Merrill	AmericanIntlGroup
Capital Economics	BNP Paribas	BayernLB	BBVA
CentroEuropaRicerche	Bank of America Merrill	Berliner Sparkasse	BMO Capital Markets
Confindustria	Barclays	Capital Economics	Barclays
EIU	Capital Economics	Citigroup	CIBCCapitalMarkets
FitchRatings	CentrePrevlExpansion	Commerzbank	DuPont
IHS Economics	Citigroup	DIW Berlin	EYParthenon
ING	Coe Rexecode	DZ Bank	EIU
IntesaSanpaolo	Crédit Agricole	DekaBank	Fannie Mae
LCMacroAdvisors	EIU	Deutsche Bank	FedExCorporation
MoodysAnalytics	Euler Hermes	FERI	First Trust Advisors
Natixis	FitchRatings	Goldman Sachs	Ford Motor Company
Oxford Economics	IHS Economics	HSBC Trinkaus	Georgia State University
Prometeia	ING	HWWI)	Goldman Sachs
Ref	MoodysAnalytics	Helaba Frankfurt	ICIS
SPGlobalMarketIntel	Morgan Stanley	IFO Munich Institute	MacroeconomicAdvisers
Société Générale	OFCE	IHS Economics	MoodysAnalytics
UBS	Oddo BHF	IW Cologne Institute	Morgan Stanley
UniCredit	Oxford Economics	IWH Halle Institute	Nomura
	Rexecode	IfW Kiel Institute	NorthernTrust
	UniCredit	MM Warburg	Oxford Economics
		Morgan Stanley	Robert Fry Economics
		Oxford Economics	RoubiniGlobalEcon
		RWI Essen	SPGlobalMktIntelligence
		UniCredit	Standard & Poor's
			Swiss Re
			TheConferenceBoard
			Wells Capital Management

C.List of expert forecasters from CE, by subgroup

Banks	Consultants	Research departments	Financial services
BNP Paribas	BIPE	OFCE	Euler Hermes
Bank of America Merrill	Capital Economics	DIW Berlin	Allianz
Citigroup	Coe Rexecode	HWWI	Confindustria
Crédit Agricole	EIU	IFO Munich Institute	IntesaSanpaolo
Exane	GAMA	IW Cologne Institute	Fannie Mae
Goldman Sachs	IHS Economics	IWH Halle Institute	First Trust Advisors
HSBC	Oxford Economics	IfW Kiel Institute	Swiss Re
La Banque Postale	PAIR Conseil	RWI Essen	Wells Capital Management
Morgan Stanley	Rexecode	DuPont	
Natixis	FERI	Eaton Corporation	
Oddo BHF	Kiel Economics	Ford Motor Company	
Société Générale	CapitalEconomics	General Motors	
UniCredit	CentroEuropaRicerche	Georgia State University	
BHF Bank	IHSEconomics		
BayernLB	LCMacroAdvisors		
Berliner Sparkasse	OxfordEconomics		
Commerzbank	Prometeia		
DZ Bank	Ref		
DekaBank	Action Economics		
Deutsche Bank	National Association of Home Builders		
HSBC Trinkaus	RDQ Economics		
Helaba Frankfurt	Robert Fry Economics		
MM Warburg			
ABI			
BankofAmericaMerrill			
GoldmanSachs			
ING			
Barclays			
Credit Suisse			
UBS			
BMO Capital Markets			
JPMorgan			
Wells Fargo			

D.Model diagnostics

To ensure the robustness of our regression specifications and address potential multicollinearity concerns, we conducted two key diagnostic checks: a correlation matrix between all explanatory variables and Variance Inflation Factors (VIF). Table D1 shows the correlation matrices for the full set of explanatory variables demonstrating no correlations exceeding 0.8, supporting the conclusion that the included controls measure distinct dimensions of irrational forecasting behavior. Second, on Table D2, we compute Variance Inflation Factors (VIFs) for each country using pooled OLS models. While the short- and long-term interest rate forecasts exhibited higher VIFs (Germany: up to 35), all other regressors-such as bias, herding, forecast dispersion, RMSE, skewness, and risk indicators-had VIFs below standard thresholds (VIF < 5), suggesting no severe multicollinearity among behavioral or macroeconomic variables. These results confirm that the model does not suffer from specification problems due to multicollinearity or overfitting.

Table D1. Correlations.

		Bias	Herding index	Shock of external information	Forecast dispersion	Forecast RMSE	Skewness on errors	Economic Political Uncertainty index	Global Political Risk	% Change of expected current-year GDP	% Change of expected year-ahead GDP	Interest rate 12-months forecast	Interest rate 3-months forecast
U.S.	Bias	1.00	-0.09	-0.02	-0.40	-0.05	0.28	-0.16	-0.20	0.27	0.54	-0.58	-0.72
	Herding index	-0.09	1.00	0.11	0.39	0.08	-0.03	0.07	0.08	0.01	-0.01	0.05	0.08
	Shock of external information	-0.02	0.11	1.00	0.19	0.02	0.01	0.08	0.03	0.01	-0.01	0.03	0.04
	Forecast dispersion	-0.40	0.39	0.19	1.00	0.08	-0.11	0.19	0.17	0.05	-0.07	0.10	0.17
	Forecast RMSE	-0.05	0.08	0.02	0.08	1.00	-0.07	0.12	0.04	-0.01	0.02	0.10	0.12
	Skewness on errors	0.28	-0.03	0.01	-0.11	-0.07	1.00	-0.04	-0.15	-0.05	0.12	-0.26	-0.27
	Economic Political Uncertainty index	-0.16	0.07	0.08	0.19	0.12	-0.04	1.00	-0.14	-0.51	0.05	0.01	0.11
	Global Political Risk	-0.20	0.08	0.03	0.17	0.04	-0.15	-0.14	1.00	0.18	-0.27	0.43	0.33
	% Change of expected current-year GDP	0.27	0.01	0.01	0.05	-0.01	-0.05	-0.51	0.18	1.00	0.04	0.04	-0.02
	% Change of expected year-ahead GDP	0.54	-0.01	-0.01	-0.07	0.02	0.12	0.05	-0.27	0.04	1.00	-0.62	-0.66
	Interest rate 12-months forecast	-0.58	0.05	0.03	0.10	0.10	-0.26	0.01	0.43	0.04	-0.62	1.00	0.94
	Interest rate 3-months forecast	-0.72	0.08	0.04	0.17	0.12	-0.27	0.11	0.33	-0.02	-0.66	0.94	1.00

Table D2.VIF results

	U.S.
Interest rate 12-month’s forecast	18
Interest rate 3-month’s forecast	12.95
Forecast dispersion	4.01
Bias	1.95
Economic Political Uncertainty index	1.88
Herding index	1.83
% Change of expected current-year GDP	1.71
% Change of expected year-ahead GDP	1.42
Global Political Risk	1.23
Shock of external information	1.14
Skewness on errors	1.06
Forecast RMSE	1.05