Do Expert Experience and Characteristics Affect Inflation Forecasts?*

Jonathan Benchimol,† Makram El-Shagi‡ and Yossi Saadon§

September 2022

Abstract

Each person's characteristics may influence that person's behaviors and outcomes. This study builds and uses a new database to estimate experts' performance and boldness based on their experience and characteristics. Our study classifies experts providing inflation forecasts based on their education, experience, gender, and environment. We provide alternative interpretations of factors affecting experts' inflation forecasting performance, boldness, and pessimism by linking behavioral economics, the economics of education, and forecasting literature. The study finds that an expert with previous experience at a central bank appears to have a lower propensity for predicting deflation.

Keywords: Expert forecast, Behavioral economics, Survival analysis, Panel estimation, Global financial crisis.

JEL Codes: C53, E37, E70.

*This paper does not necessarily reflect the views of the Bank of Israel. The authors thank Philippe Andrade, Lahcen Bounader, Itamar Caspi, Paul Fontanier, Xavier Gabaix, Paul Goldsmith-Pinkham, Ori Heffetz, Ben Schreiber, Michel Strawczynski, Matthias Weber, and Osnat Zohar, participants at the 2021 Asian Meeting of the Econometric Society, the Bank of Israel and Tsinghua University research seminars for their valuable comments.

†Research Department, Bank of Israel, Jerusalem, Israel. Email: jonathan.benchimol@boi.org.il

‡Center for Financial Development and Stability, Henan University, Kaifeng, China. Corresponding author. Email: makram.el-shagi@cfds.henu.education

§Research Department, Bank of Israel, Jerusalem, Israel. Email: yosis@boi.org.il

Please cite this paper as:
Benchimol, J., El-Shagi, M., and Saadon, Y., 2022. Do expert experience and characteristics affect inflation forecasts? Journal of Economic Behavior & Organization, 201, 205-226

1 Introduction

There has been increasing interest in professional economic forecasts in recent decades, partly due to the increasing importance of expectations in modern workhorse macroeconomic models. In a few of the most active fields, authors have compared private-sector and central bank professional forecasters (Romer and Romer, 2000; El-Shagi et al., 2016), and professional forecasts to market-implied forecasts (Adeney et al., 2017; Benchimol and El-Shagi, 2020). Others have assessed whether forecasts (or forecast spreads) incorporate information regarding macroeconomic uncertainty (Bachmann et al., 2013; Bloom, 2014; Rossi and Sekhposyan, 2015).

While there is ample evidence regarding the factors driving the overall (relative) performance of professional forecasts (Giacomini and Rossi, 2010; El-Shagi et al., 2016), there is, to our knowledge, no assessment of the factors driving performance at the individual level. Instead, studies looking at individual-level forecasts have so far mostly looked at general time-series properties of individual forecasts and their interaction with each other (Andrade and Le Bihan, 2013). The main reason seems to be that there is no information regarding individual forecasters linked to the available major forecast databases.

This paper overcomes this limitation. It combines a dataset of US CPI inflation nowcasts and clearly identified forecasters with two new original databases with–mostly web-sourced–detailed information about *forecasters* and *institutions*. Our first database includes the key individual characteristics of experts, such as their previous job experience (location, type, and duration), gender, educational attainment (including the quality of their alma mater), and affiliation (type and place). Our second database includes institutional characteristics, most importantly location and type, which have been found to matter in the previous literature. We merge these databases with CPI nowcast data extracted from the Bloomberg survey of professional forecasters, including the name and institution of each forecaster covered. To our knowledge, this is the first paper to exploit such interconnected data.

The present study contributes to the literature in several ways. Most importantly, it can assess the role of the individual and institutional characteristics in forecast performance to a new degree in the literature. At the same time, the novel dataset allows unique insights into the behavioral aspects of forecasting. We look at boldness in terms of a forecaster's willingness to deviate from the consensus (or, in other words, the herd), which can also be interpreted as overconfidence (Bordalo et al., 2020). Similarly, we assess forecasters' tendency to under or overreact, i.e., a bias towards or away from the long-term mean. In an extension, we consider pessimism, which we define as the tendency to predict deflation. While this list is far from exhaustive regarding possible behavioral traits that might affect forecasts, it provides insights into how factors beyond asymmetric information or disagreement on the underlying data-generating process shape forecast dispersion. While our study is only a first step in that direction, understanding how forecasts are derived is essential to improve forecast pooling approaches and correctly weigh different forecasts.

In addition to being the first to account in such detail for forecaster characteristics, we move away from the previous literature by considering a range of nonlinearities and interactions between relevant indicators, particularly how forecaster characteristics interact with growing experience.

Apart from its central interest, namely macroeconomic forecasts, our study contributes to the exciting debate in cognitive sciences and behavioral economics about the role of education—its level, field, and quality—in expert behaviors. We examine experts' underreactions and the influence of high inflation periods on experts' forecasting performance and boldness. In addition, this paper contributes to the debates in human resources about how experience within or between institution types matters in forecasting performance or boldness.

A deeper understanding of professional forecasts is crucial for three reasons. First, economic agents, including policymakers such as central banks, strongly rely on forecasts (Piotroski and Roulstone, 2004). Since expert forecasts are generally better than market-based forecasts (Adeney et al., 2017; Benchimol and El-Shagi, 2020), the economic agents mostly rely on the average of expert forecasts[1] (Genre et al., 2013; Budescu and Chen, 2015). Usually, the forecasts are aggregated in very simplistic ways, such as the simple average, and policymakers rank them without considering the experts' characteristics (Alessi et al., 2014; Coibion et al., 2020). Nevertheless, as pointed out by Giacomini and Rossi (2010), understanding the (time-varying) conditional relative forecast performance of several forecasts can help to select the appropriate one and generate better forecast combinations.

Forecast combinations are typically superior to individual forecasts because they can overcome model uncertainty and pool private information from various sources. Disentangling dispersion due to behavioral traits rather than actual disagreement is thus essential when not pooling model-based forecasts but forecasts provided by individual forecasters based on unknown models or experience. Sec-

[1]Although it is impossible to certify if the Fed decides according to expert forecasts, it is interesting to know how often the Fed mentions expert forecasts. However, we can reasonably assume the Fed has actively considered expert forecasts in their decision processes since the global financial crisis (GFC). The Federal Open Market Committee (FOMC) meeting minutes detail the record of the committee's policy-setting meetings and offer detailed insights regarding the FOMC's stance on monetary policy. They mention the word "forecaster" only six times over the pre-GFC decade (88 meetings), with more than a fourfold increase over the post-GFC decade (25 times over 82 meetings). The word "survey" does not appear in the interest rate announcements during the pre-GFC decade, while it appears 29 times during the post-GFC decade. The difference in the Chairman's speeches is even more spectacular than the difference in interest rate announcements or monetary policy committee minutes. The word "forecaster" appeared only eight times in 216 speeches by the Chairman during the pre-GFC decade. However, it appeared 58 times in only 160 speeches over the post-GFC decade–almost ten times more than during the pre-GFC decade.

ond, understanding professional forecasts help us to understand the behavioral foundations of expectations. There is a small but growing literature exploring this field. Contrary to our approach, the previous literature has focused on institutional characteristics–particularly the location and type of institution–due to data limitations. For example, regarding location, Bae et al. (2008) show that the earnings forecasts of local financial experts are more precise–the *local analyst advantage*. Berger et al. (2009) demonstrate that institutions based in Frankfurt (or with a subsidiary in Frankfurt) are significantly better at predicting the ECB's interest rate decisions. With respect to the type of institution, Mitchell and Pearce (2007) find evidence that predictions by some economists covered in the biannual Wall Street Journal survey are consistently above the survey mean, while those of others are consistently below, depending on the industry of the economists' employers. In this vein, economists with a public mission–e.g., academics, central bank, and government employees–demonstrate a tendency towards being pessimistic, whereas bankers, in general, are overly optimistic about future stock market developments (Veress and Kaiser, 2017).

Third, inflation expectations from learning-to-forecast experiments[2] are in line with experts, households (Michigan), and industry (Livingston) survey forecasts (Cornand and Hubert, 2020). Understanding the characteristics that influence the outcomes of expert forecasts should contribute to identifying potential factors driving the formation of expectations.

Our study produces a range of results that are novel to the literature. First, we find that experts with central bank experience are less likely to predict deflation. These experts are less pessimistic, but this is mitigated when pessimism turns out to be justified. Second, we highlight the implications and nonlinearities of the role of experience and traits in experts' forecasting performance and boldness. Third, we confirm that the influence of experts' traits on forecasting performance and boldness changed following the GFC. Fourth, we show that underperforming experts are less likely to survive in our expert database, while boldness does not significantly influence this survival rate.

The remainder of the paper is organized as follows. Section 2 outlines literature relevant to our theoretical background. Section 3 discusses our data. Section 4 outlines our empirical methodology and results, including characteristics-based forecasting ability tests and expert characteristics related to pessimism, with their interpretation. Section 5 presents the policy implications of our results and concluding remarks. Appendix A describes a simple theoretical model of expectation formation that demonstrates the importance of accounting for nonlinearities, and Appendix B presents a proportional hazards model to examine the expert's sur-

[2]Experiments are well-incentivized by remuneration. Subjects are asked to submit an inflation forecast and are rewarded solely based on their forecasts ex-post accuracy (Marimon and Sunder, 1993).

vival.

2 Literature Review and Theoretical Background

With forecasts being both the result of forecasters' labor and, to some degree, a measure of their expectations, our research question is at the intersection of labor economics and behavioral economics.

In labor economics and adjacent fields, especially economics of education (Mincer, 1974), there has been extensive discussion regarding what employee traits improve (or limit) his or her labor productivity. This includes innate traits such as gender and origin, as well as acquired permanent traits such as the field of education and experience. It seems plausible that those factors that drive labor productivity also drive forecasters' performance and, thus, the quality of their forecasts. However, as mentioned above, forecasts are not merely a product analytically derived by the forecaster using his human capital, but also a reflection of his expectations. As such, they are subject to a plethora of factors that psychologically and rationally affect the forecaster, including attention to economic variables (Gabaix, 2019, 2020), expert behavioral biases (Thomas, 1999; Davis and Lleo, 2020), and asymmetric information (Keane and Runkle, 1990), among others (Lim, 2001; Coibion and Gorodnichenko, 2015). For example, Malmendier and Nagel (2011) find that people who experienced low stock market returns invest less in the stock market, indicating how personal experience shapes optimism and pessimism. Education not only equips a forecaster with the necessary knowledge to perform his job but also imbues him with a specific world view. Specific life experiences are much more (or less) likely depending on someone's gender, origin, etc., making the role of those factors play in forecasts far less evident than the traditionally estimated effects on productivity.

What complicates matters is that the reported forecasts are not necessarily the experts' true expectations. The objective of the forecasters' employer is typically to obtain the best possible forecasts.[3] However, due to agency problems, the forecaster's objective function does not necessarily mirror the objective function of the employer.[4] While there might be intrinsic motivation to provide good forecasts,

[3]Their specific loss function can differ to some degree, depending on their use of the forecast. For instance, risk-neutral investors might aim to minimize absolute forecast errors rather than squared forecast errors that are more common in the literature. However, all those loss functions that aim for unbiased forecasts with minimal errors are highly correlated, making the differences mostly inconsequential.

[4]The outcomes achieved by any institution depend on its ability to take action today to achieve its objectives tomorrow. Institutions use expert forecasts to shape their economic decisions (Bernanke, 2007; Campbell and Sharpe, 2009). Consequently, expert forecasting accuracy generates opportunity costs (Laster et al., 1999). Fulfilling these objectives depends on the performance of the forecasts they use, including expert forecasts (See Footnote 1).

the main incentive for the forecaster to provide good forecasts is to maximize expected lifetime income.[5]

There has been discussion in the herding literature that failing "alone" is much worse than failing as part of a group. Indeed, it seems likely that an employer reads the failure of its experts as the experts' fault when they alone fail, while he might interpret them as inevitable if everybody fails. Consequently, forecasters who deviate from the herd (following their own beliefs) have a lot to lose and little to gain and might indeed be fired for their supposed incompetence. In this case, their loss function–based on the underlying objective to maximize–would include both the forecast loss and deviation from the herd of other forecasters.[6]

This is why we go beyond merely looking at forecast performance in this study and consider boldness, which we define as the willingness to deviate from the herd. Additionally, further highlighting the aforementioned psychological aspects of forecasts, we consider pessimism.[7]

We extend the literature considerably by adding a battery of traits that might affect either of those aspects of forecasts, including educational attainment (Bachelor, Master, and PhD), field (Economics, Finance, both, or others) and quality (based on the Academic Ranking of World Universities, also known as Shanghai Ranking). As the geography (i.e., location) and the institution type of both the expert and the institution (i.e., its affiliation) matter for forecasting (Batchelor, 2007; Hong and Kacperczyk, 2010), we also contribute to the literature by extending the expert's experience characteristics with institution type (central bank, academic, financial, or other) and location (of both the expert and the institutional affiliation).

In accordance with Malmendier and Nagel (2016) and Malmendier et al. (2021), which show that policymakers' inflation experience affects their inflation forecasts, we test the influence of high inflation periods experienced by experts in their forecasting performance and boldness. We also analyze the determinants of experts' over- and underreaction in predicting inflation along the lines of Barberis et al. (1998) and Daniel et al. (1998).

3 Data

In this study, we combine three unique datasets: individual forecasts for the US CPI (Section 3.1), web-sourced information, and characteristics of the forecaster's CV (Section 3.2), and web-sourced information about institutions where a forecaster worked (Section 3.3). These datasets cover the period from 1997:Q1 to

[5] More precisely, the present value of expected lifetime income.

[6] See Section 3.1, Eq. 1 and 2, for more details.

[7] For details on our definition, refer to Sections A and 4.3.

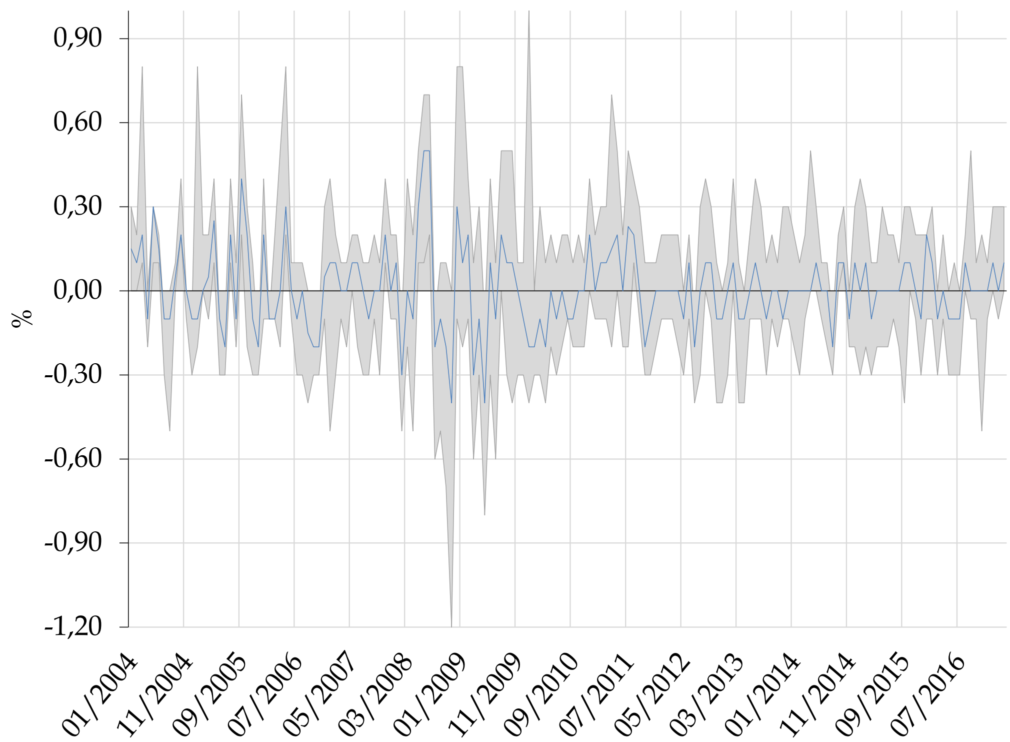
2017:Q4. We focus on inflation forecasts for both data availability in terms of quantity of forecasters and forecasts and the sufficient variability of inflation compared to interest rate forecasts. CPI inflation is also less prone to data revision than GDP forecasts, and forecaster names are unavailable for most GDP forecasts.

3.1 Forecaster Behaviors

Most of the data, including individual point forecasts and the name and affiliations of the experts, come from Bloomberg. Each expert can submit and update their US inflation forecasts until the last day of the month. The publication of the effective US CPI inflation during the corresponding month occurs about fifteen days after this day. Since forgoing the chance to update despite new information being available is irrational, we assume that the final forecasts are considered the best possible forecasts by the submitting experts.

The expert forecast updates are accessible to other experts at any time, allowing forecasters to react to each other. Therefore, any deviation from the *herd* can be considered deliberate. This allows us to capture both the quality (Fig. 1) and boldness (deviating from the herd) of forecasts.

Figure 1: Expert Forecast Performance



Notes: The blue line is the median of expert forecast errors, and the gray area represents the difference between the maximum and minimum forecast errors for each period.

First, we look at *performance*, measured through squared forecast errors as the most commonly used loss function, i.e.,

$$L_{i,t} = \left(\pi_t - \mathbb{E}^i_{t-\tau}\left[\pi_t\right]\right)^2, \tag{1}$$

where i is a specific expert. We use $t - \tau$ to denote a point in time before t but clearly after $t - 1$, since forecasters can update until the last moment.

Second, we assess *boldness*, which we define as deviating from the "herd" of other forecasters. In other words,

$$B_{i,t} = \left(\mathbb{E}^i_{t-\tau}\left[\pi_t\right] - M_{it}\left[\mathbb{E}^i_{t-\tau}\left[\pi_t\right]\right]\right)^2, \tag{2}$$

where M_{it} is the median operator over all expert forecasts i at time t.

It is worth noting that expert forecasts are published after they have been submitted, and that experts can still update until the very last minute until the deadline. Since forecasts are typically not submitted at the last second, we took the last available forecast. This would allow forecasters to adjust to the perceived consensus (median), causing perfect coincidence of forecasts by the time the data is published. Therefore, we can assume that any deviation from this median is an intentional deviation from the group, which corresponds to our model's assumption that each expert is isolated from others (and thus cannot observe other forecasts) and is overconfident in their forecasts since they disregard information from other forecasts.

There is, however, one caveat. Since we use nowcasts, where a detailed information set is already available for all forecasters, the median forecast is typically very close to the actual outcome. The correlation between B and L is 0.573, making it problematic to fully attribute variation in L as pure boldness.

These two measures will mainly be used as the dependent variable in our empirical exercise. However, they will also serve as an explanatory variable in a survival analysis, where we assess the influence of those measures on the probability of submitting another forecast.

3.2 Forecaster Characteristics

Our databases on expert characteristics are based on an automated collection, which is manually augmented where needed. The primary source is LinkedIn.[8] To rule out technical problems, we manually checked LinkedIn itself, where the automated search did not yield results, before looking for alternative sources for a CV,

[8] LinkedIn is a US-based employment-oriented Internet service founded in 2002. It is mainly used for professional networking, with employers posting jobs and job seekers posting their CVs, as well as people seeking to build their professional networks.

such as private and institutional websites. Partial database entries from LinkedIn were completed in the same fashion.

The data we collected about experts encompass 151 experts and 164 institutions that published at least one US CPI inflation forecast, with an average of about 31 forecasts per expert. Since our panel specification includes forecaster fixed-effects, we drop all forecasters who provided less than three forecasts during our sample (1997:Q1–2017:Q4), leaving us with 118 experts who published an average of about 41 US CPI inflation forecasts per expert.

For each forecaster, we collect data about the type and duration of previous job experience and education. The experience type is separated into (a) academic experience, (b) central bank experience, and (c) experience in financial institutions. The expert's experience duration corresponds to the time since the first experience in forecasting according to the expert's CV (resume-based). Our model incorporates forecasting experience (expert experience according to the CV) as a variable expressed in months and experience type (a, b, and c) as dummy variables.

Each forecaster has about 23 years of experience in forecasting on average. We also gather data on crucial expert characteristics such as their location (local or foreign), education (level, type, and quality), gender (male or female), and affiliation (type and location). About 41% of the experts are Americans, 12% are females, and 77% have at least one degree (BA, MA, or PhD) in Economics. Experts have an average of 218 months of experience in financial institutions and only ten months of experience in academia as well as central banks. Their highest degree (Bachelor's, Master's, or PhD) and the corresponding field, where we only distinguish between Economics, Finance, and other fields, are also identified. For about 8%, their highest degree is a BA, 59% have an MA, and 28% hold a Ph.D. The rest have no identified highest degree. Our database also includes the Shanghai Ranking classification for the respective university, following both the Economics and Finance rankings for the highest degree.[9] About 39% of the US CPI inflation experts graduated (BA, MA, or PhD) from a top-ranked university according to the general ranking.

Lastly, we collect additional information, such as citizenship and age. The date of birth is generally not reported on LinkedIn. As such, we scrapped the CVs available on the Internet and Bloomberg to gather the date of birth, which by construction led to less available but more precise data. The measure of age used in the paper is based on the available date of birth and allows us to assume that experts, who faced high inflation periods during their youth or adolescence, were born before 1973.

Table 1 presents the summary statistics of the dataset used in the estimations

[9] We use the Shanghai Ranking's Global Ranking, Academic Subjects, 2017. In addition, we decompose each ranking (Economics and Finance) into four levels: first tier, second tier, third tier, and not ranked.

presented below.

Table 1: Summary Statistics

	Average	Std. Dev.		Average % of the sample
Forecasts per forecaster	43.4	36.6	Local forecaster	44.7
Forecast error	0.14	0.12	Local institution	30.6
Forecast boldness	0.08	0.09	Financial institution	92.9
Experience: Central bank (m)	10	31	Experience: Central bank	16.5
Experience: Academia (m)	10	26	Experience: Academia	20.0
Experience: Total (m)	185	143	Experience: High Inflation	54.1
			Gender	12.9
			Education: MA	58.8
			Education: PhD	28.2
			Education: Economics	71.8
			Education: Finance	14.1
			Ranking: Economics	28.2
			Ranking: Finance	29.4

Notes: (m) indicates the number of months. This table presents a summary of the data at our disposal.

3.3 Institutional Information

The data about experts' affiliations describe the institution type and its primary location (local or foreign headquarters). We classified all forecast providers into several types: retail bank, investment bank, private bank, insurance company, economic and financial analysis firm, fund, investment management firm, brokerage, credit union, savings and loan firm, academia, central bank, and others. Although we built an in-depth database separating several relevant hosting institution types, we rely only upon the simple (and more relevant) difference between private financial, academic, and monetary institutions in our analysis.

4 Results and Interpretation

This section uses the data described in Section 3 to detect characteristics leading to forecasting performance, boldness, and experts' sentiments. We present a random effects analysis to identify the role of inherent traits (Section 4.1), a panel estimation to explore the role of experience (Section 4.2) and over- and under-reaction (Section 4.4), a probit model to assess the expert's pessimistic and optimistic behaviors (Section 4.3), and forecasting ability tests (Section 4.5) to identify the forecasting performance of characteristics-based groups of experts.

We report two sets of results in each of the following two subsections. The first set is based on demeaned variables. Thus, we can interpret the effect of experience as effects at the mean, although it should be noted that the mean forecaster cannot exist (as the dummies take values between 0 and 1 for this hypothetical forecaster). This is particularly helpful in immediately gauging the average effect of experience in a typical case. Essentially, this is a more straightforward way to interpret the significance of individual interaction terms, whereas one would need to utilize joint significance tests to overcome the multicollinearity issues that otherwise arise with interactions.

However, interpreting the results of the interactions is less straightforward since a dummy value of 0 no longer reflects that the dummy variable is *not true*. Therefore, we reestimate the model, but while we still demean experience (our only continuous variables), we do not demean the dummies for location, education, gender, etc. Thus, the baseline for this second estimation is a young, male, foreign forecaster, not from a top university, who holds neither a Master's nor a PhD, and has a background that is neither finance nor economics, has no experience in academia or at a central bank, did not go through high inflation periods, and works at a non-US nonfinancial institution. Consequently, this makes the interpretation of the dummy variable interactions much more straightforward.

Note that the results are mathematically equivalent, and this provides different angles to look at our results and to ease interpretation.

4.1 Random Effects

Many of the forecaster traits we collect in our dataset do not vary over their forecast history and, therefore, cannot be included in the fixed effects panel benchmark models (see Section 4.2). In this preliminary analysis, we use a random-effects estimator–where non-time-varying variables can be used–to have a look at those variables separately. We keep the model as parsimonious as possible and only include time-varying data in the fixed effects model, where we can fully account for unobserved heterogeneity.

To account for the fact that a large part of forecast errors comes from unpredictable shocks to inflation, affecting all forecasts simultaneously, we include time fixed effects.

In addition, we account for both forecaster-specific and institution-specific characteristics, thereby guaranteeing that some institutions with easier access to forecasters with specific traits do not merely drive the effect of forecaster characteristics (e.g., forecasters with a PhD or from the best institutions).

This yields the estimation equation

$$y_{it} = \beta_0 + \beta_1 \ln x_{it} + \Gamma X_{it} + \Lambda Z_{it} + u_i + v_t + \varepsilon_{it}, \tag{3}$$

where X and Z denote forecaster- and institution-specific characteristics, respectively, and Γ and Λ the corresponding vectors of coefficients. u_i, v_t, and ε_{it} represent the forecaster random-effect, time fixed-effect, and idiosyncratic component of the error term, respectively. β_0 and β_1 are the constant and experience regression coefficients, respectively.

We split our sample (1997:Q1-2017:Q4) into two subsamples: the pre-GFC(1997:Q1-2008:Q1) and post-GFC (2008:Q1-2017:Q4). Table 2 presents the random effects estimations for the US CPI inflation expert forecasts over the full sample and the pre- and post-GFC samples, with the correction for time-effects outlined in Eq. 3, for the variables presented in Section 3.

Table 2, given in this subsection, is based on demeaned variables. Results based on the alternative estimation where all dummy variables are set to 0 and only experience is demeaned are reported in Table 9 in the appendix.

Table 2: Panel Estimates with Individual Random Effects - Nonlinear Model

	Performance			Boldness		
	Full sample	Pre-GFC	Post-GFC	Full sample	Pre-GFC	Post-GFC
Experience	-0.0003	-0.0156**	0.0017	-0.0018**	-0.0182***	-0.0004
Age > 60 years old	0.0016	∼0	-0.0015	0.0042	∼0	0.0022
Local forecaster (LF)	0.0037	-0.0145	0.0137	0.0068	-0.0097	0.007
Local institution (LI)	-0.0114	-0.0014	-0.0241**	-0.008*	-0.0113	-0.01**
Financial institution (FI)	-0.0067*	-0.0662***	-0.004	-0.0016	-0.0479***	-0.001
Experience: Central bank (CB)	-0.0044	-0.0167	-0.0036	-0.0016	-0.0203	-0.0001
Experience: Academia (AC)	-0.0117***	-0.057***	-0.0082**	-0.0069***	-0.0512***	-0.0048**
Gender (G)	0.0014	0.0686***	0.0039	0.0023	0.0733***	0.002
Education: MA (MA)	-0.0047	∼0	-0.0128**	-0.0032	∼0	-0.0069**
Education: PhD (PhD)	0.0037	0.0267	0.0031	-0.0023	0.0332	-0.0025
Education: Economics (EC)	0.0147	0.0172	0.0139*	0.0062	0.013	0.0009
Education: Finance (EF)	0.0092	0.0355	0.0156*	-0.0068	0.0479	-0.0049
Ranking: Economics (RE)	0.0024	-0.0062	0.0018	0.0020**	-0.0077	0.0018**
Ranking: Finance (RF)	-0.0007	0.0097	-0.0003	-0.0003	0.0101	-0.0001
Constant	0.0372	0.0329	0.0781***	0.0028	-0.0316	0.0242**
Two-way interactions						
MA × RE	-0.0174***	∼0	-0.0152**	-0.0093***	∼0	-0.0079***
MA × RF	0.0139**	∼0	0.0104**	0.0081**	∼0	0.0058**
PhD × RE	-0.0119*	0.0196	-0.0109*	-0.009**	0.0214	-0.0089***
PhD × RF	0.0075	-0.0298*	0.0077	0.0038	-0.0271*	0.0039
MA × EC	-0.0098	-0.0221	∼0	-0.015	0.0119	∼0
FI × LI	0.0276***	0.131***	0.0224***	0.0094**	0.0955***	0.0061*
G × EF	-0.0434***	-0.184***	∼0	-0.0218**	-0.173***	∼0
MA × LI	0.0497***	0.138***	0.0677***	0.0337***	0.103**	0.0306***
MA × LF	-0.0508***	-0.169***	-0.0769***	-0.036***	-0.165***	-0.0372***
Observations	2829	324	2505	2832	324	2508
R^2	0.422	0.402	0.418	0.167	0.192	0.152

Notes: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Coefficients lower than 0.00005 are reported as ~ 0. Given the scaling of our variables coefficients in this order of magnitude do not have an economically meaningful effect.

Random effects estimates presented in Table 2 show that performance and boldness are influenced differently. We find coefficients of determination around

0.42 for performance and around 0.16 for boldness. This difference can be explained by performance variance, which is greater than boldness variance, and our focus on performance for the modeling strategy benchmark since boldness and performance results are aligned in the literature. Given the generally high precision of nowcasts and the correspondingly relatively low variance of our measures, this level of explanatory power is fairly high, particularly for performance.

As far as forecasting performance is concerned, working at a financial institution (FI) improves the expert's forecasting ability more before the GFC than after. However, experts at local financial institutions (FI $\times$ LI) slightly mitigate this result as this interaction is more significant than FI or LI effects alone. Interestingly, graduating from a top university (Finance ranking) appears to improve expert forecasts before the GFC. However, this is not the case after the GFC or over the full sample. This is mitigated by experts who graduated from a top university (Finance ranking), who achieve lower performance with their forecasts.

Before the GFC, male experts appeared to provide better forecasts than females, but this is not true after the GFC or over the full sample. The fact that male experts were more numerically dominant prior to the GFC, while the percentage of female experts has increased since the GFC, mitigates this result. Our results may also reflect that women benefit less than men from connections in job performance, herding behavior, and subjective evaluation by others (Fang and Huang, 2017).

The situation is different regarding the expert's boldness (herding behavior). First, working at a financial institution does not influence the expert's boldness. Experience in a central bank increases the expert's boldness, but it does not influence the expert's forecasting performance. In addition, graduating in Economics or Finance decreases boldness and increases the performance of expert forecasts after the GFC, while having a PhD from a top-ranked institution in Economics increases the expert's boldness before the GFC. Interestingly, the increase in the boldness of female experts with a degree in Finance was greater than the general decrease in the boldness of females before the GFC. At the same time, the robustness of this result is still questioned due to the paucity of female experts before the GFC. After the GFC, graduating with a Master's degree from a top university in Finance significantly reduces the expert's performance and boldness compared to experts who graduated with a Master's degree in Economics, which contributes significantly to increased forecasting performances and boldness.

In line with Clarke and Subramanian (2006), performance and boldness results in Table 2 are often similar in terms of significance and sign,[10] confirming that significant underperformers are more likely to issue bolder forecasts and vice-versa. Like financial analysts who also tend to exhibit herding behavior, which some-

[10]This is also the case with our next results (Table 3).

times compromises accuracy; our results suggest that social forces (ranking, institution type, and location), education (type and level), and experience (type and duration) influence an expert's rational economic logic and cognitive biases–an interpretation close to Christoffersen and Stæhr (2019) that is presented in our next results (Section 4.2). It should be noted that the highly precise nature of nowcasts partly drives the correlation between accuracy and boldness.

Interestingly, the difference between pre- and post-GFC may reflect the change in the expert's attention or biases induced by the crisis shock on characteristics' effects (Andrade and Le Bihan, 2013; Christoffersen and Stæhr, 2019). Although these results consider a time-fixed-effect, results considering individual fixed-effects like experience are presented in Section 4.2.

A Master's degree in Finance led to greater boldness but more herding behaviors after the GFC and over the full sample than in the pre-GFC period. As a result of cognitive biases and an intuitive reaction to uncertainty and financial instability, experts with lower risk tolerance may herd more (Christoffersen and Stæhr, 2019). Also, being a local forecaster with a Master's degree seems to improve forecasting performances and decrease boldness, which complements Clarke and Subramanian (2006).

Better education quality in Economics[11] improves experts' forecasting performance among those with a Master's or a PhD degree (MA $\times$ RE and PhD $\times$ RE). This was not the case for experts who graduated with a Master's degree from a non-top-ranked university. However, better education quality in Economics (RE) may accentuate experts' herding behaviors, which is not necessarily the case for experts who graduated from a top-ranked university in Finance (RF).

Working at a local institution and having a Master's degree (MA $\times$ LI) increased forecasting performance after the GFC. This result is confirmed for local experts having a Master's degree over the full sample and pre-GFC period (MA $\times$ LF). Our findings that mix geography and education relate to several strands of the literature. Our results demonstrate that the likelihood of herding increases with the expert's forecasting experience and is influenced by institution (Clement and Tse, 2005). Our results also show that experience (Hong et al., 2000; Mikhail et al., 2003) and education (De Franco and Zhou, 2009) influence social interactions, cognitive biases, and intuitive reaction to uncertainty, an interpretation partially shared with Christoffersen and Stæhr (2019).

Female experts underperform men before the GFC, but this result is not significant after the GFC, which may confirm a labor market entry selection bias. However, female experts with education in Finance outperform men (G $\times$ EF) even after the GFC. Female experts in a market segment in which their concentration is

[11] Measured with the ShanghaiRanking's Global Ranking of Academic Subjects in Economics, see Section 3.

lower than in others appear to have better-than-average skills due to self-selection (Kumar, 2010).

While differences in views may persist through time, differences in information sets only cannot explain such differences in opinion. Patton and Timmermann (2010) show they stem from heterogeneity in priors or models and that differences in opinion move countercyclically. Although this heterogeneity is most robust during recessions, our results bring another layer to this conclusion. The GFC not only changed differences in opinion but also modified the influence of experts' characteristics on their forecasting performance and boldness.

4.2 Forecaster Fixed Effects

In our third set of analyses, we assess the effect of experience on the different output measures. Since experience is the only time-varying trait we consider, this essentially boils down to a simple univariate panel regression with time (t) and forecaster (i) specific effects, such as

$$y_{it} = \beta_0 \ln x_{it} + \Gamma \ln x_{it} X_{it} + \Lambda \ln x_{it} Z_{it} + u_i + v_t + \varepsilon_{it}, \tag{4}$$

where y_{it} is one of our two loss functions discussed in Section 3.1 (performance and boldness) and x_{it} is our experience duration measure presented in Section 3.2. β_1 is the regression coefficient on experience. u_i, v_t, and ε_{it} represent the forecaster fixed-effect, time fixed-effect, and idiosyncratic component of the error term, respectively.

There are, however, potential endogeneity issues with this specification since the ability to gain experience (or in other words "to keep your job") might depend on forecast performance. We assess this possibility in detail as outlined in Appendix B. The evidence for the existence of such an effect is mixed, and if it does exist, it seems to be only moderately sized. We correct for it in a robustness test by including a dummy for flagging the last five forecasts submitted by any forecaster. While much simpler, this follows the spirit of selection estimators[12] (Heckman, 1979). The resulting equation is given as

$$y_{it} = \beta_0 \ln x_{it} + \beta_1 \mathbb{1}_{t \in last_5(i)} + \Gamma \ln x_{it} X_{it} + \Lambda \ln x_{it} Z_{it} + u_i + v_t + \varepsilon_{it}, \tag{5}$$

where $last_5(i)$ is the set of the last five periods in which forecaster i submits a forecast to our dataset, X and Z denote forecaster- and institution-specific characteristics, respectively, and Γ and Λ are the corresponding vectors of coefficients. β_0 and β_1 are regression coefficients.

[12]In a full-fledged selection model, one would include a transform of the "inclusion" probability rather than a determinant of the latter. The key problem is that we cannot truly estimate inclusion probabilities here, but merely relative risks of being removed from the sample.

Table 3 presents the panel estimations for the US CPI inflation expert forecasts for the model that controls for individual specific and time fixed-effects. All rows except the first two – reporting the effects of our only two time-varying indicators, namely log experience and the age dummy – are interactions with (log) experience. As the indicators are constant, the coefficients should be interpreted as a response to a change in experience conditional on the dummy (or combination of dummies in the case of three-way interactions).

In both Tables 3 and 10 (Appendix C), some of the coefficients seem relatively large given the low disagreement between nowcasts and the correspondingly low variance of performance and boldness. However, a unit change in log experience is substantial. For the mean forecaster, one additional year of experience corresponds to an increase of 0.1 in log experience.

For both subsamples and the full sample, we find that, on average, experience decreases the squared forecast error and deviation from the consensus forecast. This is very much in line with the previous literature. Experts may underreact less to prior CPI information as experience increases, suggesting one reason why experts, like analysts, become more accurate with experience (Mikhail et al., 2003). This is similar to analysts' firm experience, which is strongly and positively associated with analysts' forecast boldness (Clarke and Subramanian, 2006; Huang et al., 2017). At the mean, a one-year change in experience would lead to a reduction of approximately 0.02 in the squared forecast error (i.e., one-tenth of the coefficient value).

In the full sample and for most of the post-GFC period, domestic forecasters (LF) benefit less from experience than foreign forecasters at the mean, which complements Clarke and Subramanian (2006). This is intuitive since domestic forecasters can be expected to start with a better understanding of the local economy. At the same time, experience is, on average, more helpful at local institutions (LI), which offer better exposure to local data, networking, and news. Generally, people with higher levels of education benefit from experience, which might either reflect that they start with less "practical knowledge" or (hopefully) that their training enables them to utilize information better to improve their performance over time. People with experience in academia benefit less from experience, while a degree in Economics or Finance seems to help in that respect.

In Table 3, the results on expert performance (left panel) are close to the boldness results (right panel) in terms of sign and significance, except for several interesting instances. Experience seems to embolden local forecasters over the full sample and after the GFC, which also complements Clarke and Subramanian (2006). However, a notable difference between performance and boldness results concerning experts who faced high inflation periods shows that their boldness is discouraged by experience without significantly influencing their performance,

Table 3: Panel Estimates with Individual Fixed Effects - Nonlinear Model

	Performance			Boldness		
	Full sample	Pre-GFC	Post-GFC	Full sample	Pre-GFC	Post-GFC
Experience	-0.176***	-2.29	-0.271***	-0.137***	-1.027	-0.127***
Age $>$ 60 years old	-0.03***	$\sim$0	-0.032***	-0.009	$\sim$0	-0.005
$last_5$	0.002	-0.042	-0.005	0.001	-0.011	-0.004
Two-way interactions with experience						
Local forecaster (LF)	3.428***	-1.059	0.281*	1.512***	-0.27	0.188***
Local institution (LI)	-4.409***	-1.634	0.105	-2.009***	-1.309	-0.101
Financial institution (FI)	1.409***	-0.279	0.085	0.662***	0.381	0.09
Experience: Central bank (CB)	0.816***	0.365	0.814***	0.326***	0.061	0.334***
Experience: Academia (AC)	2.351***	-0.006	-1.359***	0.944***	-0.157	-0.6***
Experience: High Inflation (HI)	-0.018	-1.949**	0.034	-0.046**	-1.121	-0.006
Gender (G)	-0.145	-0.625	-0.007	-0.067	-0.76	0.063
Education: MA (MA)	-2.681***	-0.438	-3.147***	-1.327***	-0.195	-1.462***
Education: PhD (PhD)	-3.3***	$\sim$0	-3.29***	-1.552***	$\sim$0	-1.443***
Education: Economics (EC)	0.738***	-1.473*	-0.859***	0.36***	-0.815	-0.331**
Education: Finance (EF)	-3.467***	4.805	0.462	-1.582***	7.649***	0.193
Ranking: Economics (RE)	1.098***	-0.203	0.467***	0.504***	-0.131	0.221***
Ranking: Finance (RF)	-0.041***	0.366	-0.028	0.004	-0.144	0.008
Three-way interactions with experience						
FI $\times$ RE	-0.016	1.19	-0.003	0.013	0.46	0.017
MA $\times$ RE	-7.103***	-0.33	-6.857***	-3.449***	-0.296	-3.151***
PhD $\times$ RE	-7.117***	$\sim$0	-6.927***	-3.462***	$\sim$0	-3.201***
EC $\times$ RE	-3.573***	-0.727	0.262	0.544***	-0.466	0.067
FI $\times$ RF	-0.062	-1.275	-0.013	-0.027	-0.593	-0.028
G $\times$ RF	-0.162***	0.427	-0.041	-0.111***	0.492	0.02
MA $\times$ RF	-0.043**	0.083	-0.093**	-0.031***	0.089	-0.054**
EF $\times$ RF	0.007	-3.304	-0.098	0.154***	-6.123***	-0.034
AC $\times$ LI	-15.052***	2.069	7.791***	-6.39***	1.104	3.248***
MA $\times$ LI	21.214***	-4.6	20.684***	10.349***	-3.378	9.555***
PhD $\times$ LI	21.361***	$\sim$0	20.852***	10.459***	$\sim$0	9.598***
FI $\times$ LF	-6.89***	$\sim$0	-6.908***	-3.278***	$\sim$0	-3.156***
G $\times$ LF	6.441***	4.387	6.725***	2.776***	2.463	2.956***
AC $\times$ LF	14.892***	$\sim$0	-7.965***	6.347***	$\sim$0	-3.207***
MA $\times$ LF	0.023	2.564	$\sim$0	0.06**	2.541	$\sim$0
EC $\times$ LF	7.453***	$\sim$0	-0.48*	3.392***	$\sim$0	-0.014
G $\times$ FI	-15.893***	$\sim$0	-15.794***	-7.647***	$\sim$0	-7.248***
EC $\times$ FI	-0.279	$\sim$0	7.323***	0.167	$\sim$0	3.21***
G $\times$ CB	7.149***	$\sim$0	7.017***	3.322***	$\sim$0	3.019***
EC $\times$ AC	3.749***	$\sim$0	$\sim$0	1.641***	$\sim$0	$\sim$0
EC $\times$ MA	-3.367***	$\sim$0	$\sim$0	-1.357***	$\sim$0	$\sim$0
Observations	2829	324	2505	2832	324	2508
R^2	0.052	0.231	0.041	0.057	0.146	0.035

Notes: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Coefficients lower than 0.00005 are reported as $\sim$ 0. Given the scaling of our variables coefficients in this order of magnitude do not have an economically meaningful effect.

which mitigates Clarke and Subramanian (2006), who linked performance to boldness in terms of sign and significance. High inflation, when the expert was young, may contribute to a specific fear of remaining far from the consensus. Experience, therefore, encourages experts to herd. As a result of cognitive biases and an intuitive reaction to uncertainty and financial instability, experts with lower risk tolerance may herd more (Christoffersen and Stæhr, 2019).

We look at the full set of interactions to fully understand the results, which is more accessible when looking at Table 10 (Appendix C). While the coefficient on experience is positive in the model, where dummies are not demeaned, the vast majority of (single) interactions is negative. Experience has a detrimental but insignificant effect on our baseline forecaster, which is not true for the typical forecaster (who has good education and typically a background in an Economics or Finance-related field). In the full sample and post-GFC, ceteris paribus, most deviations from the benchmark case (i.e., the single interactions) are negative, making experience more helpful. Two notable exceptions are gender (where women seem to benefit less from experience than men or indeed suffer from experience) and financial institutions where forecasters also seem to be deteriorating with growing experience.

The simple (two-way) interactions and the three-way interactions have to be considered jointly. For example, we find that the FI $\times$ LF interaction compensates for the negative effect of financial institutions. That is, local forecasters at financial institutions benefit no less from experience than their counterparts in other institutions that provide forecasts. Similarly, we find a highly negative coefficient for the G $\times$ FI interactions, i.e., while women at nonfinancial institutions benefit less from experience, women at financial institutions benefit more. That said, our gender results should be taken with the caveat that there are relatively few women. In a highly nonlinear model such as ours, this might imply that the results for a female local forecaster in a financial institution (for example) are driven by just one or two women. Nevertheless, the large coefficient values make the combination of the three aforementioned dummies an excellent example for interpretation.

Table 4 shows the marginal effects for all combinations of these three dummies, where all other indicators match the baseline case. It is evidently impossible to look at more than 8000 combinations in the same detail.

However, this way of looking at things sheds more light on seemingly extreme results. For example, a hugely negative coefficient of local institutions is offset mainly by interactions with higher degrees, i.e., the typical highly educated forecaster at those institutions is not that good.

For the modal forecaster in terms of characteristics (except experience), i.e., a male foreign forecaster working at a foreign financial institution, who holds an MA in Economics from a non-ranked university and has some experience with

Table 4: Financial Institutions, Local Forecasters, and Gender

FI	LF	G	Marginal effect of experience
0	0	0	5.551***
0	0	1	9.034***
0	1	0	-0.521
1	0	0	7.298***
0	1	1	6.441***
1	0	1	-15.893***
1	1	0	-6.89***
1	1	1	5.02***

Notes: All other variables except inflation are set to 0, i.e., we are looking at a forecaster from a foreign institution, without a higher degree, who studied neither Economics nor Finance and did not go to an institution that excels at those issues. The forecaster has no experience with high inflation regimes, academia, or central banking.

high inflation (but none in academia or central banking), the total benefit of experience is indistinguishable from zero. For this forecaster, the benefit of experience would be much higher when having a degree in Finance instead of Economics, especially if this degree is from a top institution in Finance, or if he was female. He would benefit much less when having a background in academia. Again, keep in mind that a lower or even detrimental impact of experience does not necessarily imply bad forecasting, but could also reflect the depreciation of relevant knowledge that makes the initial forecasts particularly good.

Three-way interactions show that experts having a Master's degree or a PhD from a top university according to the university rankings in Economics and Finance (MA $\times$ RE, PhD $\times$ RE, and MA $\times$ RF), working at a financial institution and being a local forecaster (FI $\times$ LF), or having experience in academia and working at a local institution (AC $\times$ LI) generally benefit from experience over the full and post-GFC periods. Experience is beneficial to the quality of education in Economics (EC $\times$ RE). For experts who graduated with a degree in Finance from a top-ranked university, this is less true. However, experts who have a Master's or PhD in Economics from a top-ranked university can benefit from experience and the latter seems to make them more confident, emboldening their forecasts.

Unlike local forecasters working at a financial institution (FI $\times$ LF), working at a financial and local institution (FI $\times$ LI) does not influence the contribution of experts' experience to performance or boldness. However, experience for experts at a local institution while having experience in academia (AC $\times$ LI) is beneficial and encourages herding behaviors over the full sample, but experience seems detrimental and emboldens forecasts after the GFC. The total effect of experience on experts at a local institution and a Master's degree or PhD is beneficial and leads to more herding behaviors, which is related to our previous findings for experts in a local institution with a Master's degree (Table 2). The total effect of experience on local forecasters with a degree in Economics (LF $\times$ EC) tends to be detrimental during the full sample but remains neutral in terms of performance after the GFC. Interestingly, experts with a degree in Economics benefit from experience differently if they hold a Master's degree than those with experience in academia. The findings regarding mixed geography and education relate to several strands of the literature. Our results demonstrate that the likelihood of benefiting from experience in terms of boldness increases with the expert's forecasting performance and experience, and is influenced by the institution (Clement and Tse, 2005). Furthermore, the results complement the ones on experience (Hong et al., 2000; Mikhail et al., 2003) and education (De Franco and Zhou, 2009) that influence social interactions, cognitive biases, and intuitive reaction to uncertainty, an interpretation partially shared with Christoffersen and Stæhr (2019).

Local female experts (G $\times$ LF), female experts who graduated from a top-ranked university in Finance (G $\times$ RF), and female experts with central bank experience (G $\times$ CB) benefit more from experience than men with central bank experience, confirming a labor market entry selection bias.

While differences in views may persist through time, differences in information sets cannot explain such differences in opinion. Patton and Timmermann (2010) show that they stem from heterogeneity in priors or models and that differences in opinion move countercyclically. Although this heterogeneity is most robust during recessions, our results bring another layer to this conclusion. The GFC not only changed differences in opinion but also modified the influence of experts' characteristics on their forecasting performance and boldness.

The pre-GFC period significantly differs from the post-GFC one. However, our results show that the GFC contributed significantly to changing the distribution of the effects of characteristics and experience on forecasting performance and boldness, a loosely documented phenomenon.[13] Again, it should be noted that (a) the results are not directly comparable due to different mean forecasters and (b) the lack of significant results for the pre-GFC sample might be driven largely by the low number of observations.

The results for boldness are generally similar to the results for performance. Unsurprisingly, the average forecast is highly accurate at this time horizon. More often than not, deviating from the herd is not beneficial.

Table 3 also reports the coefficients of determination for the fixed effects models. At first glance, these seem relatively low, falling in the range between 0.04

[13] There is a broad literature describing how economic conditions affect (expert) forecasts (Adeney et al., 2017) or inflation perception. However, to our knowledge, there are few sources documenting how economic conditions influence the changing effect of experts' characteristics on forecasting accuracy and boldness.

and 0.15. However, it must be kept in mind that the forecaster fixed effects–i.e., the actual characteristics–are not included in a within-R^2. That is, the coefficient of determination truly only captures the additional effect of experience and its interactions. Also, as mentioned previously, nowcasts are highly precise since much information is already available, limiting the influence of personal belief. Finding that even for nowcasts, experience and characteristics have a statistically significant effect, even though quantitatively small, is a strong result that indicates that personal beliefs and abilities play an important role in understanding expectation formation.

4.3 Pessimism

When looking at the behavioral side of forecasts, it seems evident that one of the most relevant questions is whether forecasts–i.e., expectations–are optimistic or pessimistic. While distinguishing optimism from pessimism is quite straightforward for business cycle forecasts, it is less so for inflation, where "good" and "bad" are less clearly defined. One might look at the deviation from target inflation, but it is hard to argue that 1.9% inflation is worse than 2.1%.

However, our sample includes two brief episodes when the US economy was endangered by–and in some months experiencing–deflation. Unlike low inflation, deflation is almost universally considered highly problematic in Economics, allowing us to use those periods to assess the question of optimism vs. pessimism. Figure 2 shows how in the early period of the GFC (late 2008 to early 2009) and over most of 2015, forecasters disagreed on whether or not there would be deflation.

For those subsamples we estimate a panel probit model explaining the probablity that the forecast $\mathbb{E}_{it}\left(\pi_i t\right)$ would be below zero, taking the form:

$$p\left(\mathbb{E}_{it}\left(\pi_i t\right) < 0\right) = \Phi\left(\psi_0 \ln x_{it} + X_i \Psi\right), \tag{6}$$

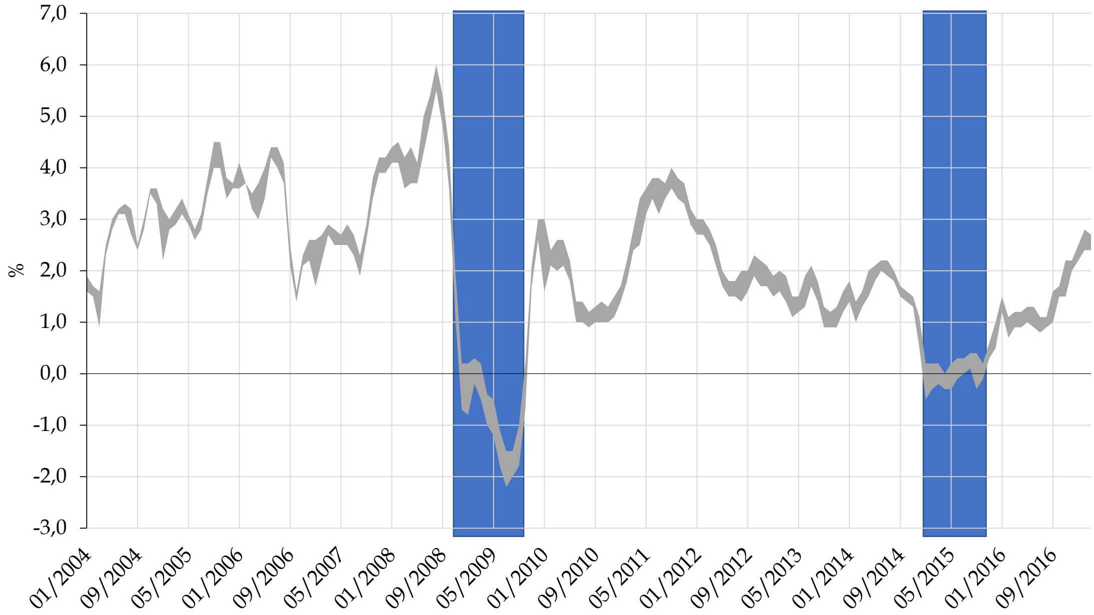
where Φ is the cumulative distribution function of the standard normal distribution and Ψ is a vector of regression coefficients.

We split this sample into the subsamples when deflation was observed (i.e., when pessimism was justified, or–in other terms–the lack of deflation expectation was overly optimistic), and when no deflation was observed (that is, when expecting inflation can genuinely be seen as overly pessimistic).

Table 5 presents the probit panel regressions for a general situation (i.e., during both inflationary and deflationary periods) and during only deflationary periods.

Table 5 shows that experts with more experience, a degree in Economics, or central bank experience are less likely to predict deflation. However, this propensity becomes less significant under deflation for experts with central bank expe-

Figure 2: Forecast Spread - Mixed Inflation and Deflation Expectations Periods



Notes: The gray area represents the spread of forecasts (disagreement). The blue shaded background highlights the situations where both deflation and inflation were considered possible by forecasters.

rience, which strongly influences private information. These experts are less pessimistic, but this is mitigated when pessimism turns out to be justified. These results are not significant under deflation periods and may be explained by the expert's professional experience with few deflation periods and their education curriculum, which may not contain extended attention to deflation or forecasting deflation.

Positive numbers in Table 5 show that experts' characteristics are more likely to lead to expectations of deflation. Experts with central bank experience are optimistic, but people with experience in academia or high inflation are pessimistic. Interestingly, this is mainly driven by actual deflation periods. Experts with central bank experience and those with a degree in Economics do not (want to) see deflation coming, but people with any high inflation experience or experience in academia do.

Table 5 also presents the marginal effects, which translate the probit coefficients into derivatives of the probability with respect to the explanatory variable. While the marginal effects seem large, it has to be noted that we only look at periods with disagreement on whether there will be deflation. Within this subsample, the unconditional probability of predicting deflation exceeds 75%. Moreover, a unit change of log experience at the mean (of this subsample) corresponds to almost 30 years of additional experience (with average log experience being 5.26).

At the mean, a unit change in experience reduces the probability of predicting

Table 5: Probit Estimates - Pessimism

	All		Deflation	
	Estimates	Marginal Effects	Estimates	Marginal Effects
Experience	-0.886***	-0.32	-0.933***	-0.37
Experience: Central bank (CB)	-0.642**	-0.23	-0.61*	-0.24
Experience: Academia (AC)	0.613**	0.22	0.884***	0.35
Education: Economics (EC)	-0.858**	-0.31	-0.697	-0.28
Ranking: Economics (RE)	-2.083	-0.02	-1.444	-0.02
Experience: High Inflation (HI)	0.938**	0.34	1.135***	0.45
Observations	359		257	

Notes: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. *All* stands for both *Inflation* and *Deflation* periods, and *Deflation* stands for only deflation periods. Both samples (*All*) are restricted to periods where there was disagreement, i.e., some experts predicted deflation while others did not. The second sample (*Deflation*) is a subsample of *All*, where the pessimists turned out to be correct.

deflation by 32 percentage points, corresponding to about 1.9 percentage points for one year of experience (a log change of 0.06). This increases to 37 percentage points (or 2.1 percentage points respectively) under deflation.

A degree in Economics reduces the probability of predicting deflation by 31 percentage points under normal times, while it is insignificant under deflation. Central bank experience reduces the probability of predicting deflation by 23 percentage points under normal times, while this is less significant (but still significant at 10 percentage points) under deflation. Having academia or periods of high inflation experience increases the probability of forecasting deflation by 22 percentage points and 34 percentage points, respectively. Under deflation, these responses increase to 35 percentage points and 45 percentage points, respectively.

In line with our theoretical model (Appendix A), these results show that the type and duration of an expert's experience drive his or her sentiment. Their degree of pessimism (optimism) may interact with their private signal processing. According to Manzan (2011), heterogeneous sentiments may influence experts in deciphering newly available information, involving a positive relationship between the interpretation of the mean signal and the prior sentiment (pessimistic or optimistic). Hence, the effect of prior sentiment on an expert's information processing, discussed in Appendix A, may depend on the expert's experience, both duration and type (central bank or not).

We also complement the view stating that disagreement stems from heterogeneity in prior sentiments and moves countercyclically, with heterogeneity being strongest during recessions, when forecasters appear to place greater weight on their prior beliefs (Patton and Timmermann, 2010). Our results may confirm that the weight of experts' prior beliefs, and the beliefs themselves are influenced by both experience and prior experience in a central bank, as well as the current state of inflation (or deflation).

4.4 Over- and under-reaction

In this section, we provide evidence about the influence of cognitive factors on experts' forecast errors. We augment a standard Mincer and Zarnowitz (1969) regression to allow for testing under- and over-reaction for different types of forecasters rather than just looking at the sample as a whole (Barberis et al., 1998; Daniel et al., 1998).

Rather than merely estimating

$$\mathbb{E}^i_{t-1}\left[\pi_t\right] - \pi_t = \beta_0 + \beta_1 \pi_t + \varepsilon_{it}, \tag{7}$$

where β_0 and β_1 are regression coefficients, and ε_{it} is the idiosyncratic component of the error term. We estimate

$$\mathbb{E}^i_{t-1}\left[\pi_t\right] - \pi_t = \alpha_0 + \left(\alpha_1 + \sum_{n=1}^{k} \alpha_{n+1} x_{n,it}\right) \pi_t + u_i + \varepsilon_t, \tag{8}$$

where $x_{n,it}$ is the n^{th} characteristic of forecaster i at time t. That is, through interaction terms with π_t, the effective coefficient β_1 from Equation 7 is allowed to vary depending on forecaster characteristics. At the same time, we account for unobserved heterogeneity through random effects[14] (u_i). We use the version of the test that uses the forecast errors rather than forecasts on the left-hand side of the equation; that is, we can test β_1 or $\alpha_1 + \alpha_n x_{n,it}$ against 0 rather than against 1 to assess over- or under-reaction.

Table 6 presents the directional forecast error regression over the full sample.

Table 6 shows that experts underreact on average. However, experts with central bank experience or more (accumulated) experience in forecasting tend to underreact less than the average. Having central bank experience seems to reduce underreaction behaviors more and more robustly than accumulating experience in forecasting. Moreover, older experts (over 60 years old) underreact more than the average. The underreaction of older experts may reflect a slight misinterpretation of private information and the decrease in forecasting performance with age exhibited in Table 3.

The misinterpretation of genuine new private information affects forecasting performance. Daniel et al. (1998) assume overconfidence about private information involves stronger return predictability in firms with the greatest information asymmetries,[15] and suggest investigating whether the overconfidence of investors and traders can be identified with specific characteristics. As in their model,

[14] A Hausman test prefers random effects to fixed effects in our specification.

[15] This also implies greater inefficiencies in the stock prices of small companies.

Table 6: Random Effects Estimates - Underreaction

	Forecast Error
CPI	-0.0356**
Experience (E)	0.0001*
Experience: Central bank (CB)	0.0079**
Age > 60 years old	-0.0007**
Constant	0.0093*
Observations	2904

Notes: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Unlike our previous exercises, these results are obtained from undemeaned data, and the age variable used here is the age of the expert rather than its dummy. A Hausman test justifies the use of a random effects model.

the uninformed investors of our model could be interpreted as being contrarian-strategy investors (whether institutions or individuals). Identifying the confidence characteristics of different observable experts' categories generate additional implications.

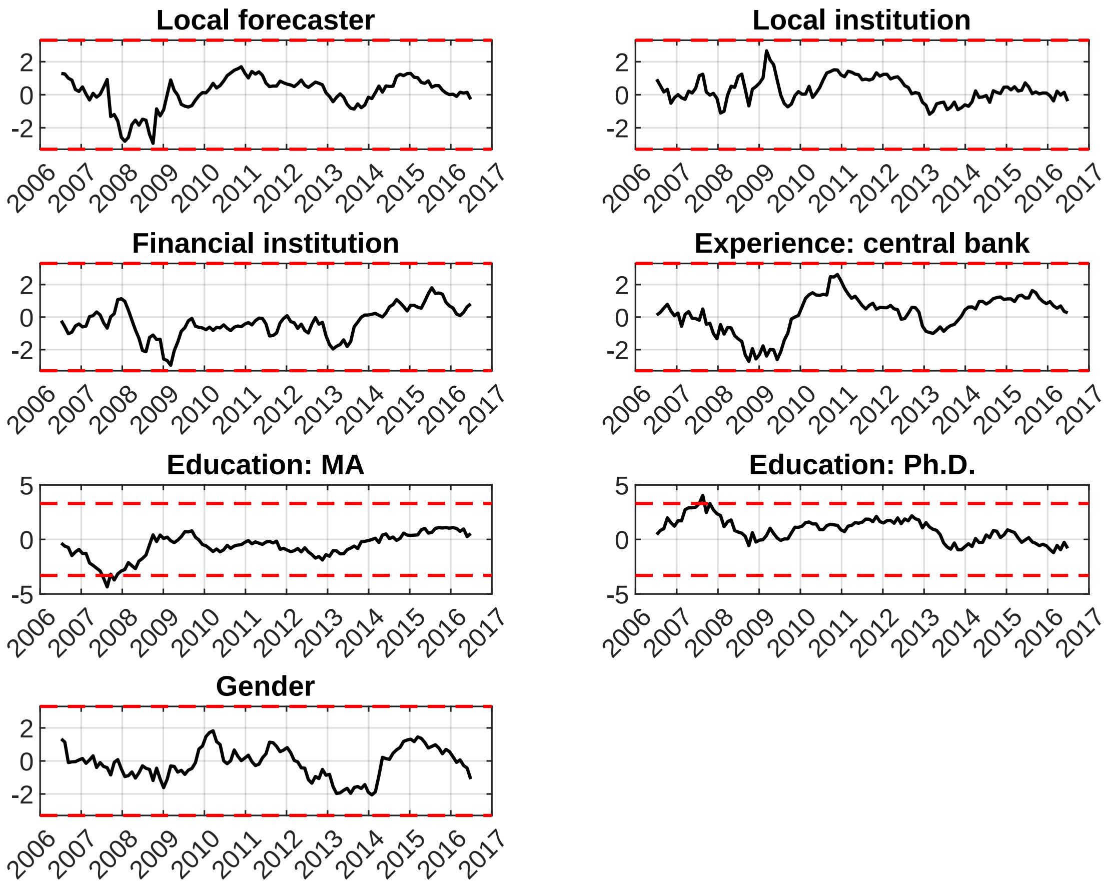
4.5 Forecasting Ability

The characteristics identified in Section 4 should improve the predictive ability of experts. We analyze time variation in the out-of-sample relative forecasting performance to test this hypothesis. More precisely, we test for relative forecast performance of characteristics-based expert groups in unstable environments, as proposed by Giacomini and Rossi (2010). The null hypothesis is that the forecasts under consideration perform equally well at every point in time. Exceeding the critical value does not imply that one group of expert forecasts constantly outperforms the other but that there is a meaningful difference in the predictive ability for a subsample. Fig. 3 presents these fluctuation tests for forecasts grouped by characteristics.

The test rejects only for education level characteristics (MA and PhD, see Fig. 3) that is, experts with a PhD differ from the others significantly at least once, and so do forecasters where the top degree is a Masters's degree. Due to the low number of experts holding neither an MA nor a PhD, those two tests capture almost the same information. Thus, it is unsurprising that the rejection is driven by the same period in the mid-2000s, when PhD educated experts underperformed the competition (mostly MAs), and MAs outperformed their competition (mostly PhDs).

Interestingly, the test does not reject the remaining characteristics, confirming that experts' characteristics may have a significant role in their out-of-sample fore-

Figure 3: Forecasting Ability Tests



Notes: The red dashed lines represent the critical values of the predictive ability test.

casting outcomes even in unstable environments. Our findings may have implications for policymakers. They may decide to select their inputs (multiple expert forecasts) according to the expert's characteristics to maximize their forecasting precision and optimize their decisions.

Our results also provide a possible basis for an alternative to conventional forecast combination methods in the literature. While the bias-adjusted combination method is found to work well in practice (Capistrán and Timmermann, 2009b), we demonstrate that a characteristics-based forecast combination is potentially more desirable than equal-weighted or bias-adjusted forecast combination methods.

Inflation forecasts from the Survey of Professional Forecasters are biased, presenting positive serial correlation in forecast errors, cross-sectional dispersion, and predictability patterns depending on inflation variance. As we control for time fixed-effects, we interpret experts' shifts in forecasting performance not explained by asymmetric loss and rational expectations (Capistrán and Timmermann, 2009a) through their characteristics rather than inflation variance.

All in all, combining forecasts with respect to the expert's characteristics generally improves out-of-sample forecasting performance.

5 Policy Implications and Conclusion

In line with our model, the characteristics we identify are shown to influence experts' performance, boldness, forecasting ability, and sentiment (optimism or pessimism).

Investors and policymakers use forecasts to design or explain their decisions, and sometimes, the efficiency of these decisions depends on forecasts (European Central Bank, 2011, 2014; de Vincent-Humphreys et al., 2019). Identifying the characteristics of the best experts could help firms and policymakers to achieve their objectives efficiently. For instance, Carvalho and Nechio (2014) and Binder (2020) show that households' macroeconomic forecasts–about interest rates, inflation, and unemployment–are not uniform across income and education levels. Forecasts also constitute an essential information channel leading investment portfolio and spending decisions (Duca-Radu et al., 2021). The more accurate the forecast, the less likely a surprise could occur, minimizing the required adjustment costs of the investment portfolio and the corresponding market volatility when the data become publicly available (Laster et al., 1999). We provide an in-depth picture of the personal characteristics of professional forecasters that may affect information rigidity, complementing Coibion and Gorodnichenko (2012, 2015). We also confirm the influence of personal characteristics on forecasting performances of consumers found by Duca-Radu et al. (2021) with professional forecasters.

Policy institutions extensively use expert forecasts for both decision-making and forecasting purposes (Piotroski and Roulstone, 2004; Adeney et al., 2017). While policymakers generally aggregate these forecasts in simplistic ways and rank them without considering the expert characteristics (Alessi et al., 2014; Coibion et al., 2020), the main takeaway from our results is that experts' characteristics drive forecasting outcomes, boldness, and sentiment.

Consequently, policymakers may use our results to group forecasters with respect to some of their characteristics (Section 4.5) to increase the reliability of their inflation forecasts compared to simple averaging, thus improving their policy decision making processes. This should also hamper the spillover effect of pessimistic or optimistic behaviors on inflation forecasts, which is somehow frequent during specific periods such as deflation, if policymakers group expert forecasts according to experience (Section 4.3).

The current state of inflation (or deflation) and of the economy (after or before a crisis like the GFC) influence experts' behaviors and beliefs, and thus the transmission is channeled from their characteristics to their forecasting performance and boldness as demonstrated here.

Underperforming experts are more likely to no longer be part of our expert database, i.e., they are less likely to be in charge of the inflation forecasts contributed to the Bloomberg database, while boldness does not significantly influence the experts' survival rate. Degrees in Finance or Economics do not offer the same protection as career concerns or institutional labor market expectations, while graduating from a top university decreases the expert's survival rate.[16]

Comparing our results from the two subsamples reveals that the GFC changed both financial institutions and the expert's labor market. After the GFC, expert's experience, location, institution type, education field, or quality change their forecasting performance. Our panel estimations also show that before the GFC, only the education field and quality mattered for forecasting performance and boldness under fixed effects. More characteristics play a significant role under random effects.

The expert's location, institution location and type, experience type, and gender affect the expert's forecasting ability. The expert's previous experience or previous experience in a central bank significantly influences the expert's sentiment.

We interpret our results as evidence of the effect of characteristics on experts' inflation forecast outcomes. One implication of our analysis is that experts' characteristics and experience matter for policymakers as long as expert forecasts are considered in their decision-making process.

References

Adeney, R., Arsov, I., Evans, R., 2017. Inflation expectations in advanced economies. RBA Bulletin March Quarter, Reserve Bank of Australia.

Alessi, L., Ghysels, E., Onorante, L., Peach, R., Potter, S., 2014. Central bank macroeconomic forecasting during the global financial crisis: the European Central Bank and Federal Reserve Bank of New York experiences. Journal of Business & Economic Statistics 32 (4), 483–500.

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

Bachmann, R., Elstner, S., Sims, E. R., 2013. Uncertainty and economic activity: evidence from business survey data. American Economic Journal: Macroeconomics 5 (2), 217–249.

[16]See MA $\times$ RF and PhD $\times$ RE in Table 8.

Bae, K.-H., Stulz, R. M., Tan, H., 2008. Do local analysts know more? A cross-country study of the performance of local analysts and foreign analysts. Journal of Financial Economics 88 (3), 581–606.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of Financial Economics 49 (3), 307–343.

Batchelor, R., 2007. Bias in macroeconomic forecasts. International Journal of Forecasting 23 (2), 189–203.

Benchimol, J., El-Shagi, M., 2020. Forecast performance in times of terrorism. Economic Modelling 91, 386–402.

Berger, H., Ehrmann, M., Fratzscher, M., 2009. Forecasting ECB monetary policy: Accuracy is a matter of geography. European Economic Review 53 (8), 1028–1041.

Bernanke, B. S., 2007. Inflation expectations and inflation forecasting. Speech 306, Board of Governors of the Federal Reserve System.

Binder, C., 2020. Coronavirus fears and macroeconomic expectations. Review of Economics and Statistics 102 (4), 721–730.

Bloom, N., 2014. Fluctuations in uncertainty. Journal of Economic Perspectives 28 (2), 153–176.

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

Breslow, N., Chatterjee, N., 1999. Design and analysis of two-phase studies with binary outcome applied to Wilms tumour prognosis. Journal of the Royal Statistical Society Series C 48 (4), 457–468.

Budescu, D. V., Chen, E., 2015. Identifying expertise to extract the wisdom of crowds. Management Science 61.

Campbell, S. D., Sharpe, S. A., 2009. Anchoring bias in consensus forecasts and its effect on market prices. Journal of Financial and Quantitative Analysis 44 (2), 369–390.

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

Capistrán, C., Timmermann, A., 2009b. Forecast combination with entry and exit of experts. Journal of Business & Economic Statistics 27 (4), 428–440.

Carvalho, C., Nechio, F., 2014. Do people understand monetary policy? Journal of Monetary Economics 66 (C), 108–123.

Christoffersen, J., Stæhr, S., 2019. Individual risk tolerance and herding behaviors in financial forecasts. European Financial Management 25 (5), 1348–1377.

Clarke, J., Subramanian, A., 2006. Dynamic forecasting behavior by analysts: Theory and evidence. Journal of Financial Economics 80 (1), 81–113.

Clement, M. B., Tse, S. Y., 2005. Financial analyst caracteristics and herding behavior in forecasting. Journal of Finance 60 (1), 307–341.

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

Coibion, O., 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., Kumar, S., Pedemonte, M., 2020. Inflation expectations as a policy tool? Journal of International Economics 124, 103297.

Cornand, C., 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.

Cox, D. R., 1972. Regression models and life-tables. Journal of the Royal Statistical Society: Series B 34 (2), 187–202.

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. Journal of Finance 53 (6), 1839–1885.

Davis, M., Lleo, S., 2020. Debiased expert forecasts in continuous-time asset allocation. Journal of Banking & Finance 113 (C), 105759.

De Franco, G., Zhou, Y., 2009. The performance of analysts with a CFA designation: The role of human-capital and signaling theories. Accounting Review 84 (2), 383–404.

de Vincent-Humphreys, R., Dimitrova, I., Falck, E., Henkel, L., Meyler, A., 2019. Twenty years of the ECB Survey of Professional Forecasters. Economic Bulletin Articles 1.

Duca-Radu, I., Kenny, G., Reuter, A., 2021. Inflation expectations, consumption and the lower bound: Micro evidence from a large multi-country survey. Journal of Monetary Economics 118 (C), 120–134.

El-Shagi, M., Giesen, S., Jung, A., 2016. Revisiting the relative forecast performances of Fed staff and private forecasters: A dynamic approach. International Journal of Forecasting 32 (2), 313–323.

European Central Bank, 2011. Monthly Bulletin. Executive Board 01/2011, European Central Bank.

European Central Bank, 2014. Monthly Bulletin. Executive Board 01/2014, European Central Bank.

Fang, L., Huang, S., 2017. Gender and connections among Wall Street analysts. Review of Financial Studies 30 (9), 3305–3335.

Gabaix, X., 2019. Behavioral inattention. In: Bernheim, B. D., DellaVigna, S., Laibson, D. (Eds.), Handbook of Behavioral Economics - Foundations and Applications 2. Vol. 2 of Handbook of Behavioral Economics: Applications and Foundations 1. Elsevier, pp. 261–343.

Gabaix, X., 2020. A behavioral New Keynesian model. American Economic Review 110 (8), 2271–2327.

Genre, V., Kenny, G., Meyler, A., Timmermann, A., 2013. Combining expert forecasts: Can anything beat the simple average? International Journal of Forecasting 29 (1), 108–121.

Giacomini, R., Rossi, B., 2010. Forecast comparisons in unstable environments. Journal of Applied Econometrics 25 (4), 595–620.

Heckman, J. J., 1979. Sample selection bias as a specification error. Econometrica 47 (1), 153–161.

Hong, H., Kacperczyk, M., 2010. Competition and bias. Quarterly Journal of Economics 125 (4), 1683–1725.

Hong, H., Kubik, J. D., Solomon, A., 2000. Security analysts' career concerns and herding of earnings forecasts. RAND Journal of Economics 31 (1), 121–144.

Huang, R., Krishnan, M., Shon, J., Zhou, P., 2017. Who herds? Who doesn't? Estimates of analysts' herding propensity in forecasting earnings. Contemporary Accounting Research 34 (1), 374–399.

Jacob, J., Lys, T. Z., Neale, M. A., 1999. Expertise in forecasting performance of security analysts. Journal of Accounting and Economics 28 (1), 51–82.

Kaplan, E. L., Meier, P., 1958. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association 53 (282), 457–481.

Keane, M. P., Runkle, D. E., 1990. Testing the rationality of price forecasts: new evidence from panel data. American Economic Review 80 (4), 714–735.

Kumar, A., 2010. Self-selection and the forecasting abilities of female equity analysts. Journal of Accounting Research 48 (2), 393–435.

Laster, D., Bennett, P., Geoum, I. S., 1999. Rational bias in macroeconomic forecasts. Quarterly Journal of Economics 114 (1), 293–318.

Lim, T., 2001. Rationality and analysts' forecast bias. Journal of Finance 56 (1), 369–385.

Malloy, C. J., 2005. The geography of equity analysis. Journal of Finance 60 (2), 719–755.

Malmendier, U., Nagel, S., 2011. Depression babies: do macroeconomic experiences affect risk taking? Quarterly Journal of Economics 126 (1), 373–416.

Malmendier, U., Nagel, S., 2016. Learning from inflation experiences. Quarterly Journal of Economics 131 (1), 53–87.

Malmendier, U., Nagel, S., Yan, Z., 2021. The making of hawks and doves. Journal of Monetary Economics 117 (C), 19–42.

Manzan, S., 2011. Differential interpretation in the survey of professional forecasters. Journal of Money, Credit and Banking 43 (5), 993–1017.

Marimon, R., Sunder, S., 1993. Indeterminacy of equilibria in a hyperinflationary world: experimental evidence. Econometrica 61 (5), 1073–1107.

Mikhail, M. B., Walther, B. R., Willis, R. H., 2003. The effect of experience on security analyst underreaction. Journal of Accounting and Economics 35 (1), 101–116.

Mincer, J. A., 1974. Schooling, Experience, and Earnings. NBER Books. New York, NY: National Bureau of Economic Research.

Mincer, J. A., Zarnowitz, V., 1969. The evaluation of economic forecasts. In: Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance. NBER Chapters. National Bureau of Economic Research, pp. 3–46.

Mitchell, K., Pearce, D. K., 2007. Professional forecasts of interest rates and exchange rates: Evidence from the Wall Street Journal's panel of economists. Journal of Macroeconomics 29 (4), 840–854.

Patton, A. J., Timmermann, A., 2010. Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion. Journal of Monetary Economics 57 (7), 803–820.

Piotroski, J. D., Roulstone, B. T., 2004. The effects of analyst forecast properties and country-level institutions on the cost of debt. Accounting Review 79 (4), 1119–1151.

Romer, C. D., Romer, D. H., 2000. Federal reserve information and the behavior of interest rates. American Economic Review 90 (3), 429–457.

Rossi, B., Sekhposyan, T., 2015. Macroeconomic uncertainty indices based on nowcast and forecast error distributions. American Economic Review 105 (5), 650–655.

Scharfstein, D. S., Stein, J. C., 1990. Herd behavior and investment. American Economic Review 80 (3), 465–479.

Thomas, L. B., 1999. Survey measures of expected U.S. inflation. Journal of Economic Perspectives 13 (4), 125–144.

Veress, A., Kaiser, L., 2017. Forecasting quality of professionals: Does affiliation matter? Quarterly Review of Economics and Finance 66 (C), 159–168.

Zwiebel, J., 1995. Block investment and partial benefits of corporate control. Review of Economic Studies 62 (2), 161–185.

Appendix

A Model

In this section, we present a stylized model that motivates the inclusion of nonlinearities and interactions. We show that even using a straightforward and standard form of learning–namely Bayesian learning–experience has nonlinear effects that depend strongly on initial conditions. Our model belongs to the "noisy information" model class. As in Coibion and Gorodnichenko (2012, 2015), agents continuously monitor variables and consider the most updated information to formulate their decisions. Different responses may occur in information acquisition when fundamental shocks happen.

Consider a forecaster aiming to forecast inflation π, who is exposed to both public and private signals. We assume both signals are drawn from a normal distribution, they are isolated[17] and mutually independent, and both are unbiased but noisy. In other words:

$$s_t^u \sim \mathcal{N}\left(\pi_t, \tau_u^{-1}\right), \tag{9}$$

and

$$s_t^r \sim \mathcal{N}\left(\pi_t, \tau_r^{-1}\right), \tag{10}$$

where s_t^u and s_t^r are the public and private signals at time t, respectively.

We assume that the precision of the public signal, τ_u, is known and constant. The precision of the private signal, τ_r, is unknown to the forecaster initially. Rather, starting from a prior assumption, he learns about the quality (i.e., the precision) of his private signal over time through Bayesian learning.

For simplicity, we assume that the prior regarding the precision of the private signal is Gamma distributed, where the initial prior has a mean of $\tilde{\tau}_0$ and a variance of $\sigma^2_{\tau,0}$.[18]

Since the Gamma distribution is a conjugate prior to the precision of a normal distribution with a known mean (in this case zero), the Bayesian updating yields a new Gamma distribution with lower variance and a more accurate estimate of the true mean every period.

More precisely, the mean of new distribution is given by

$$\tilde{\tau}_t = \frac{\tilde{\tau}^2_{t-1}/\sigma^2_{\tau,0} + 1/2}{\tilde{\tau}_{t-1}/\sigma^2_{\tau,t-1} + 1/2\left(s^r_{t-1} - \pi_{t-1}\right)^2}, \tag{11}$$

[17]Each expert cannot observe other forecasts when extrapolating this one-agent model to a multiple-agent model, a simplifying assumption corresponding to the findings of Bordalo et al. (2020).

[18]In the more common α, β parameterization, this corresponds to $\alpha_0 = \tilde{\tau}_0^2/\sigma^2_{\tau,0}$ and $\beta_0 = \tilde{\tau}_0/\sigma^2_{\tau,0}$. While this notation makes the updating equations more convoluted, it allows for deriving the interest variable's law of motion more straightforwardly.

and the variance by

$$\sigma_t^2 = \frac{\tilde{\tau}_{t-1}^2/\sigma_{\tau,0}^2 + 1/2}{\left(\tilde{\tau}_{t-1}/\sigma_{\tau,t-1}^2 + 1/2\left(s_{t-1}^r - \pi_{t-1}\right)^2\right)^2}. \tag{12}$$

The Bayesian point estimate for the average forecaster–i.e. the representative forecaster that is repeatedly experiencing errors of a magnitude of exactly one standard deviation–who starts forecasting in $t = 0$ and does so every period at time t is thus given by:

$$\tilde{\tau}_t = \frac{\tilde{\tau}_0^2/\sigma_{\tau,0}^2 + (t-1)/2}{\tilde{\tau}_0/\sigma_{\tau,0}^2 + t/2\tau_r}. \tag{13}$$

A forecaster who aims to maximize the expected precision of his forecast, will then provide a weighted forecast

$$f_t = \frac{\tau_r}{\tau_r + \tilde{\tau}_t} s_t^u + \frac{\tilde{\tau}_t}{\tau_r + \tilde{\tau}_t} s_t^r, \tag{14}$$

implying a precision of

$$\tau_t = \left(\left(\frac{\tau_r}{\tau_r + \tilde{\tau}_t}\right)^2 \tau_u^{-1} + \left(\frac{\tilde{\tau}_t}{\tau_r + \tilde{\tau}_t}\right)^2 \tilde{\tau}_t^{-1}\right)^{-1}. \tag{15}$$

The dynamics implied by this model are quite straightforward:

1. Unless $\tilde{\tau}_0 = \tau_u$, the forecast performance will improve over time, since any misconception–whether it is overly optimistic or pessimistic–regarding the quality of the private signal will lead to a suboptimally weighted forecast.

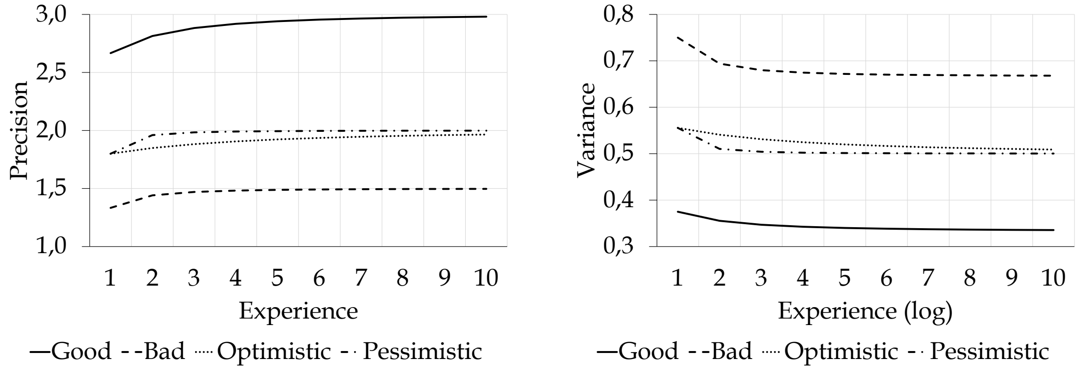
2. Precision converges monotonically to $\tau = \tau_u + \tau_r$.

We can now imagine a range of factors that can potentially drive forecast quality τ_r and the initial optimism (or pessimism) regarding one's own forecast quality $\tilde{\tau}_0$.

Figure 4 shows some core scenarios–namely a good, a bad, and two average forecasters, with the latter two differing in the degree of optimism (or pessimism) regarding the quality of their private signal–linking expected forecast performance to time, i.e., experience.

In the left panel, we show the "raw numbers", i.e., time (experience) t and precision τ_t, obtained from substituting Eq. 13 into Eq. 15. In the right panel, we show the data using the transformations that will be applied in the empirical approach, i.e., the variance (corresponding to the expected values of squared forecast errors) and the natural logarithm of time (experience). The different trajectories of performance in response to time–i.e., experience–are visible. That is,

Figure 4: Evolution of forecast precision and variance over time



Notes: The good forecaster has a precision of two with a prior of one. The bad forecaster has a precision of 0.5 with the same prior. The optimistic forecaster has a precision of one with a prior of two, and the pessimistic forecaster has a precision of one with a prior of 0.5. All priors have a variance of zero, and $\tau_u = 1$.

factors that drive forecast performance–through both actual quality and suboptimal weighting–also affect the performance's response to experience (here shown as time). In other words, it is necessary to include forecast experience and constant factors driving forecast performance in a model, but also to consider possible interactions.

The model predicts that both heterogeneity in prior expectations and understanding of news across forecasters may drive dispersion in forecasts. Heterogeneous interpretation of the incoming information might exacerbate the dispersion of forecasts. Similar results are observed in Coibion and Gorodnichenko (2012, 2015). We assume this interpretation of incoming information is influenced by experts' characteristics such as experience and education, for example.

To some degree, it is always good to include your own information[19] so forecasts do not converge–worse forecasters stay worse. The two experts who have the same "true" quality of the private signal–the pessimist and the optimist–eventually converge as they learn their signal's quality.

B Survival Analysis

In our first set of analyses, we assess forecasters' probability of providing a further forecast in the future, i.e., to survive in the market. Our dataset does not allow us to distinguish the reasons for possible removal from the dataset. Being fired, or at least removed from this particular responsibility, are possibilities, but merely retiring or even being promoted are equally possible. However, the main reason for

[19]Since the signal error is unrelated, even a high variance signal is meaningful.

us to conduct this analysis is that bad forecast performance may lead to exclusion from the sample, thereby creating endogeneity issues as outlined above.

The approach we chose is a proportional hazards model pioneered by Cox (1972). In this model, we estimate

$$h_{it} = h_0\left(t\right) \times e^{\lambda_0 f\left(Y_{it-1}\right) + X_i \Lambda}, \tag{16}$$

where Y_{it-1} is the past history of y_{it}, $h_0\left(t\right)$ corresponds to the time-specific effect, Λ is a vector of regression coefficients, and λ_0 is a regression coefficient. We use three different transformations $f\left(Y_{it-1}\right)$ describing slightly different possibilities of how performance review might be conducted by the hiring institutions: last-period performance or boldness[20] (Raw), five-period backward-looking moving average (MA5), and weighted moving average where the most recent periods have a higher weight (wMA5). First, we simply take the last forecast's value, i.e. $Y_{it-1} = y_{i,t-1}$. Second, we look at a moving average over the past five months, indicating a slightly longer horizon rather than penalizing extremely bad forecasts immediately, i.e. $Y_{it-1} = 1/5 \sum_{s=1}^{5} y_{it-s}$. Finally, we look at a weighted moving average, i.e., $Y_{it-1} = \sum_{s=1}^{5} w_s y_{it-s}$, where the weights decline exponentially by a power of two.

Since we want to stay true to our general multivariate framework that includes continuous time-varying variables, we abstain from traditional Kaplan and Meier (1958) type survival plots, which allow assessing actual expected survival time for distinct subgroups. Therefore, we cannot interpret our results in terms of additional months of survival in the job, but merely interpret the sign and relative magnitude of coefficients.

Before presenting our survival analysis, we present preliminary results in Table 7 to capture our research question's intuition. Tables 7 and 8 present three transformations for each dependent variable. In these tables, positive significant coefficients show that low forecasting performance or boldness increases the probability of removal (i.e., not being a forecaster in the next period).

Table 7 presents the survival estimates for US CPI experts without interactions with few variables.

Table 7 presents the survival estimates for US CPI experts, without interactions, related only to the expert's location and education field (Economics). It shows that these variables matter even in this simplified model. Although the expert's survival depends on his past performance, especially MA5 and wMA5, it also depends on his education in Economics and on his current location.

A local inflation expert may survive longer than a foreign expert, which may reflect employment protection or information advantage effects, in line with the literature on financial analysts (Malloy, 2005). We also find that an expert with a

[20] See Eq. 1 and Eq. 2, respectively, for the definitions of performance and boldness.

Table 7: Survival Estimates - Linear Model

	Performance		
	Raw	MA5	wMA5
Dependent variable	2.435	7.359***	4.699**
Local forecaster (LF)	-0.478**	-0.474**	-0.473**
Education: Economics (EC)	-0.725***	-0.724***	-0.718***

Notes: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Raw stands for the last period dependent variable, MA5 stands for five-period moving average (backward-looking) and wMA5 for weighted moving average where the most recent periods have a higher weight.

degree in Economics may survive longer than those without an Economics degree, a result interpreted below (Table 8).

Although the coefficient looks large in Tables 7 and 8, squared forecast errors for inflation nowcasts are small. Thus, the total impact of forecast quality on survival probability is small, which mitigates endogeneity issues.

Table 7 shows that a low past forecasting performance (MA5 and wMA5) increases the probability of being removed.[21] Providing bolder forecasts decreases expert survival but less significantly. According to these preliminary results, the expert's future seems to be determined by current and past performance.

Table 8 presents the survival estimates for US CPI experts with interactions.[22]

As in Table 7, Table 8 shows that having a degree in Economics immunizes the expert from being removed as an economist. Interestingly, Table 8 shows that a degree in Finance does not offer the same protection as having graduated in Economics. Experts with a degree in Economics may provide more convincing explanations to back their information processing and subsequent forecasts of the US CPI, an economic variable par excellence, than experts with a degree in Finance,[23] increasing their survival rate.[24] Having graduated from a top-ranked university in Economics or Finance[25] seems to improve the expert's survival rate, and having experience in academia seems to impair the expert's survival rate.

The expert's survival rate seems to be determined more by current and past

[21]The presented results use the exact partial likelihood. Under the Breslow and Chatterjee (1999) approximation, low past performance in forecasting inflation (MA5) increases the probability of being removed.

[22]See Section 3 for more details about the variables.

[23]Whatever the expert's outcomes, providing more convincing economic explanations about their inflation forecasts helps.

[24]In a previous version of this paper, an application to Fed fund rates shows that experts with a degree in Finance survive better than those with a degree in Economics. We attributed this result to the fact that the nominal interest rate is both an economic and a financial variable and is thus better explained by experts with an education in Finance or both Economics and Finance.

[25]See Section 3.2 for more details about the ranking variables we use.

Table 8: Survival Estimates - Nonlinear Model

	Performance			Boldness		
	Raw	MA5	wMA5	Raw	MA5	wMA5
Dependent variable	2.509	8.14***	5.173**	3.912	11.612**	7.651*
Age > 60 years old	-0.089	-0.124	-0.103	-0.115	-0.137	-0.13
Local forecaster (LF)	-0.674*	-0.668*	-0.668*	-0.682*	-0.69*	-0.685*
Local institution (LI)	0.173	0.199	0.19	0.196	0.176	0.196
Financial institution (FI)	-0.767*	-0.711*	-0.754*	-0.779**	-0.744*	-0.765**
Experience: Central bank (CB)	0.4	0.425	0.408	0.398	0.4	0.399
Experience: Academia (AC)	0.759*	0.799**	0.791**	0.737*	0.739*	0.747*
Gender (G)	0.319	0.38	0.349	0.335	0.376	0.355
Education: MA (MA)	-0.196	-0.276	-0.231	-0.218	-0.249	-0.236
Education: PhD (PhD)	-0.588	-0.697	-0.641	-0.576	-0.596	-0.591
Education: Economics (EC)	-1.126**	-1.11**	-1.086**	-1.151***	-1.157***	-1.141***
Education: Finance (EF)	-0.528	-0.534	-0.483	-0.565	-0.539	-0.534
Ranking: Economics (RE)	-0.07	-0.076	-0.076	-0.086	-0.094	-0.093
Ranking: Finance (RF)	0.057	0.067	0.057	0.063	0.069	0.064
Interactions						
MA × RF	-0.537*	-0.562**	-0.55*	-0.54*	-0.558**	-0.548**
PhD × RE	-0.598**	-0.602**	-0.596**	-0.58**	-0.578**	-0.571*

Notes: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Raw stands for the last period dependent variable, MA5 stands for five-period moving average (backward-looking) and wMA5 for weighted moving average where the most recent periods have a higher weight.

performance, and to a lesser extent, by only current performance. Low average performance (MA5) significantly decreases the expert's survival rate, while herding influences the experts' survival rate less significantly.[26]

Several theories of reputation and herd behavior indicate that agents' performance and boldness may vary with career concerns (Scharfstein and Stein, 1990; Zwiebel, 1995). Our results suggest that increased reputational capital may increase the labor market attractiveness of top-ranked profiles, thus leading experts originating from or targeting academia to leave the profession more frequently than other experts. This might also mean that these experts are fired or moved to another occupation (in the same or another institution) more often because, for instance, they are hired based on higher or different expectations than others, which they may not meet.

Table 8 provides additional insights about experts' survival. First, it shows that having graduated with a Master's degree from a top Finance institution improves the expert's survival. This result sharply contrasts with an education in Finance that does not significantly influence the expert's survival. Second, an MA from a top institution in Finance and a PhD from a top institution in Economics increase

[26]In general, most experts follow others or are moderately bold. It is evident that, as long as the consensus provides the best forecast on average, an expert always far from the consensus (bold) will have a lower survival rate.

the survival probability of the expert. This increase in the expert's survival rate is slightly more significant for a PhD from a top institution in Economics than for a Master's degree from a top institution in Finance.

It is theoretically possible that the effect we attribute to a PhD degree (and other expert-specific factors) is indeed driven by better institutions hiring based on those factors (whether they matter or not). An expert with a PhD might increase his chances of working at a better institution, with better information-gathering functions, than an expert without a PhD. Financial institutions may have better information-gathering functions, data, and private information access. Also, experts with a PhD may be attracted by these institutions, and these institutions may prefer recruiting people with a PhD, making a case for firm effects that are both an intermediary and a factor. Although we do not have enough companies with more than one forecaster in the sample (otherwise, we would be able to control for firm fixed effects), we believe that it is unlikely that the institution is an intermediary rather than a factor. Indeed, as long as we use Bloomberg forecasts, meaning all our experts have access to at least a Bloomberg terminal, we can reasonably assume that all the forecasters have access to almost the same information and necessary equipment to build their forecasts. Hence, the effect of the PhD is more prevalent than firm effects.

Tables 8 mitigates existing career-concern-motivated herding theories (Hong et al., 2000). While our results show that experts are more likely to lose their jobs after providing inaccurate forecasts, our results confirm that underperformers face higher employment risk than outperformers (Clarke and Subramanian, 2006). The theory linking analysts' boldness with career concerns and ability (Scharfstein and Stein, 1990; Jacob et al., 1999) is partially verified for CPI inflation experts.

C Alternative Dummies

In this section, we present the panel estimates with random and fixed effects with undemeaned dummies. We use the demeaned model for selection, as an interaction matters if it is significant at the mean. Demeaned and not demeaned results will produce different significant coefficients, but the joint significance of the interaction terms that belong together will be the same. However, the interpretation of coefficients of dummy interactions where dummies are not demeaned (Tables 9 and 10) is more straightforward than in Tables 2 and 3.

Table 9: Panel Estimates with Individual Random Effects - Nonlinear Model - Alternative Dummies

	Performance			Boldness		
	Full sample	Pre-GFC	Post-GFC	Full sample	Pre-GFC	Post-GFC
Experience	-0.0003	-0.0156**	0.0017	-0.0018**	-0.0182***	-0.0004
Age $>$ 60 years old	0.0016	~ 0	-0.0015	0.0042	~ 0	0.0022
Local forecaster (LF)	0.0338**	0.0859*	-0.0244***	0.0281***	0.0880**	-0.0043
Local institution (LI)	-0.0648***	-0.197***	~ 0	-0.0361***	-0.155***	~ 0
Financial institution (FI)	-0.0165***	-0.113***	-0.0120**	-0.0049*	-0.0819***	-0.0028
Experience: Central bank (CB)	-0.0044	-0.0167	-0.0036	-0.0016	-0.0203	-0.0001
Experience: Academia (AC)	-0.0117***	-0.0570***	-0.0082**	-0.0069***	-0.0512***	-0.0048**
Gender (G)	0.0061*	0.0883***	0.0039	0.0047**	0.0919***	0.0020
Education: MA (MA)	0.0137	~ 0	-0.0784***	0.0172	~ 0	-0.0308***
Education: PhD (PhD)	0.0037	-0.0356	0.0031	-0.0023	-0.0103	-0.0025
Education: Economics (EC)	0.0205	0.0303	-0.0697***	0.0151	0.0060	-0.0325***
Education: Finance (EF)	0.0143	0.0573	0.0156*	-0.0042	0.0684	-0.0049
Ranking: Economics (RE)	0.0167***	-0.0127	0.0145***	0.0105***	-0.0149*	0.0094***
Ranking: Finance (RF)	-0.0114**	-0.0101	-0.0090*	-0.0063**	-0.0080	-0.0049*
Constant	0.0330	0.141	0.157***	-0.0068	0.0414	0.0256
Two-way interactions						
MA $\times$ RE	-0.0174***	~ 0	-0.0152**	-0.0093***		-0.0079***
MA $\times$ RF	0.0139**	0.0298*	0.0104**	0.0081**	0.0271*	0.0058**
PhD $\times$ RE	-0.0119*	0.0196	-0.0109*	-0.0090**	0.0214	-0.0089***
PhD $\times$ RF	0.0075	~ 0	0.0077	0.0038	~ 0	0.0039
MA $\times$ EC	-0.0098	-0.0221	0.0836***	-0.0150	0.0119	0.0334***
FI $\times$ LI	0.0276***	0.131***	0.0224***	0.0094**	0.0955***	0.0061*
G $\times$ EF	-0.0434***	-0.184***	~ 0	-0.0218**	-0.173***	
MA $\times$ LI	0.0497***	0.138***	-0.0160**	0.0337***	0.103**	-0.0028
MA $\times$ LF	-0.0508***	-0.169***	0.0067	-0.0360***	-0.165***	-0.0038
Observations	2829	324	2505	2832	324	2508
R^2	0.422	0.402	0.418	0.167	0.192	0.152

Notes: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Coefficients lower than 0.00005 are reported as ~ 0. Given the scaling of our variables coefficients in this order of magnitude do not have an economically meaningful effect.

Table 10: Panel Estimates with Individual Fixed Effects - Nonlinear Model - Alternative Dummies

	Performance			Courage		
	Full sample	Pre-GFC	Post-GFC	Full sample	Pre-GFC	Post-GFC
Experience	5.551***	1.884	10.535***	3.008***	0.546	4.777***
Age $>$ 60 years old	-0.03***	$\sim$0	-0.032***	-0.009	$\sim$0	-0.005
$last_5$	0.002	-0.042	-0.005	0.001	-0.011	-0.004
Two-way interactions with experience						
Local forecaster (LF)	-0.521	-3.098	7.423***	-0.089	-2.068	3.208***
Local institution (LI)	-21.209***	0.693	-20.629***	-10.402***	0.48	-9.599***
Financial institution (FI)	7.298***	-0.279	-0.425*	3.221***	0.381	-0.025
Experience: Central bank (CB)	-0.031	0.365	-0.017	-0.067**	0.061	-0.024
Experience: Academia (AC)	-3.598***	-0.742	0.228***	-1.636***	-0.55	-0.001
Experience: High Inflation (HI)	-0.018	-1.949**	0.034	-0.046**	-1.121	-0.006
Gender (G)	9.034***	-3.027	8.95***	4.544***	-2.108	4.276***
Education: MA (MA)	-7.41***	-0.205	-10.509***	-3.902***	-0.384	-4.863***
Education: PhD (PhD)	-10.902***	$\sim$0	-10.711***	-5.274***	$\sim$0	-4.859***
Education: Economics (EC)	-1.829***	-1.473*	-6.951***	-1.155***	-0.815	-3.109***
Education: Finance (EF)	-3.467***	4.805	0.462	-1.582***	7.649***	0.193
Ranking: Economics (RE)	10.706***	-0.428	6.627***	4.993***	0.038	3.087***
Ranking: Finance (RF)	0.057	1.727	0.054	0.043	0.916	0.066*
Three-way interactions with experience						
FI $\times$ RE	-0.016	1.19	-0.003	0.013	0.46	0.017
MA $\times$ RE	-7.103***	-0.33	-6.857***	-3.449***	-0.296	-3.151***
PhD $\times$ RE	-7.117***	$\sim$0	-6.927***	-3.462***	$\sim$0	-3.201***
EC $\times$ RE	-3.573***	-0.727	0.262	-1.544***	-0.466	0.067
FI $\times$ RF	-0.062	-1.275	-0.013	-0.027	-0.593	-0.028
G $\times$ RF	-0.162***	0.427	-0.041	-0.111***	0.492	0.02
MA $\times$ RF	-0.043**	0.083	-0.093**	-0.031***	0.089	-0.054**
EF $\times$ RF	0.007	-3.304	-0.098	0.154***	-6.123***	-0.034
AC $\times$ LI	-15.052***	2.069	7.791***	-6.39***	1.104	3.248***
MA $\times$ LI	21.214***	-4.6	20.684***	10.349***	-3.378	9.555***
PhD $\times$ LI	21.361***	$\sim$0	20.852***	10.459***	$\sim$0	9.598***
FI $\times$ LF	-6.89***	$\sim$0	-6.908***	-3.278***	$\sim$0	-3.156***
G $\times$ LF	6.441***	4.387	6.725***	2.776***	2.463	2.956***
AC $\times$ LF	14.892***	$\sim$0	-7.965***	6.347***	$\sim$0	-3.207***
MA $\times$ LF	0.023	2.564	$\sim$0	0.06**	2.541	$\sim$0
EC $\times$ LF	7.453***	$\sim$0	-0.48*	3.392***	$\sim$0	-0.014
G $\times$ FI	-15.893***	$\sim$0	-15.794***	-7.647***	$\sim$0	-7.248***
EC $\times$ FI	-0.279	$\sim$0	7.323***	0.167	$\sim$0	3.21***
G $\times$ CB	7.149***	$\sim$0	7.017***	3.322***	$\sim$0	3.019***
EC $\times$ AC	3.749***	$\sim$0	$\sim$0	1.641***	$\sim$0	$\sim$0
EC $\times$ MA	-3.367***	$\sim$0	$\sim$0	-1.357***	$\sim$0	$\sim$0
Observations	2829	324	2505	2832	324	2508
R^2	0.052	0.231	0.041	0.057	0.146	0.035

Notes: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Coefficients lower than 0.00005 are reported as ~ 0. Given the scaling of our variables coefficients in this order of magnitude do not have an economically meaningful effect.