Do Expert Experience and Characteristics Affect Inflation Forecasts?*

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

November 1, 2020

Abstract

Each person's characteristics may influence that person's behaviors and their outcomes. We build and use a new database to estimate experts' performance and boldness based on their experience and characteristics. We classify 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. 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 Classification: 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, Osnat Zohar, and participants at the Bank of Israel and Tsinghua University research seminars for their useful comments.

†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

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

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. We combine a nowcast dataset for US CPI inflation with clearly identified forecasters with two new original databases with–mostly web-sourced–detailed information about *forecasters* and *institutions*. We contribute to the literature in several ways. First, we show 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 Global Financial Crisis (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.

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

[1] Although it is impossible to certify if the Fed decides according to expert forecasts, it is interesting to know how often the Fed often mentions expert forecasts. However, we can reasonably assume the Fed actively considers expert forecasts since the GFC, among other indicators, in their decision and communication processes. 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"

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). However, 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/or generate better forecast combinations.

Second, understanding professional forecasts helps 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–in particular 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.

We construct two original databases on expert and institutional characteristics. Our first database includes the key individual characteristics of experts, such as their past job experience (location, type and duration), gender, educational attainment (including the quality of their alma mater), and affiliation (type and place).

only six times over the pre-GFC decade (88 meetings) with a more than fourfold increase over the post-GFC decade (25 times over 82 meetings). The word "survey" does not appear in the pre-GFC decade interest rate announcements, 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 the 216 governor speeches during the pre-GFC decade. However, it appeared 58 times in only 160 governor speeches over the post-GFC decade–almost ten times more than during the pre-GFC decade.

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

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, which include both the name and institution of each forecaster covered. To our knowledge, this is the first paper exploiting such interconnected data.

Linking this data to the forecast data allows us a unique insight into the behavioral aspects of forecasting. We can assess the role of the individual and institutional characteristics in forecast performance, and boldness at the individual level. We understand boldness as deviation from the consensus, which can also be interpreted as overconfidence (Bordalo et al., 2020). In an extension, we consider pessimism, which we define as the tendency to predict deflation.

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

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. It also contributes to the debates in human resources about how experience within or between institution types matters in terms of forecasting performance, or boldness. We find several traits that influence forecasting performance, herding behaviors (opposite of boldness), and expert survival.

The remainder of the paper is organized as follows. Section 2 outlines literature relevant to our theoretical background. Section 3 describes a simplistic theoretical model of expectation formation that demonstrates the importance of accounting for nonlinarities. Section 4 discusses our data. Section 5 outlines our empirical methodology and empirical results, including characteristics-based forecasting ability tests and expert characteristics related to pessimism, with their interpretation. Section 6 presents the policy implications of our results and concluding remarks.

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, in particular 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, acquired permanent traits such as the field of education and, last but not least, 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 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, 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 might indeed be fired for their supposed incompetence, and little to gain. In this case, their loss function–based on the underlying objective to maximize–would include

[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).

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

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 paper 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 other) 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).

3 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. 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[8] and mutually independent, and both are unbiased but noisy. In other words:

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

and

$$s_t^r \sim \mathcal{N}\left(\pi_t, \tau_r^{-1}\right), \tag{2}$$

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.

[6] See Section 4.1, Eq. 8 and 9, for more details.

[7] For details on our definition, refer to Sections 3 and 5.4.

[8] 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).

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 $\sigma^2_{\tau,0}$.[9]

Since the Gamma distribution is a conjugate prior for the precision of a normal distribution with a known mean (in this case 0), 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:

$$\tilde{\tau}_t = \frac{\tilde{\tau}^2_{t-1}/\sigma^2_{\tau,0} + 1/2}{\tilde{\tau}_{t-1}/\sigma^2_{\tau,t-1} + 1/2\left(s^r_{t-1} - \pi_{t-1}\right)^2}, \tag{3}$$

and the variance by

$$\sigma^2_t = \frac{\tilde{\tau}^2_{t-1}/\sigma^2_{\tau,0} + 1/2}{\left(\tilde{\tau}_{t-1}/\sigma^2_{\tau,t-1} + 1/2\left(s^r_{t-1} - \pi_{t-1}\right)^2\right)^2}. \tag{4}$$

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 0 and does so every period at time t is thus given by:

$$\tilde{\tau}_t = \frac{\tilde{\tau}^2_0/\sigma^2_{\tau,0} + (t-1)/2}{\tilde{\tau}_0/\sigma^2_{\tau,0} + t/2\tau_r}. \tag{5}$$

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

$$f_t = \frac{\tau_r}{\tau_r + \tilde{\tau}_t} s^u_t + \frac{\tilde{\tau}_t}{\tau_r + \tilde{\tau}_t} s^r_t, \tag{6}$$

implying a precision of

$$\tau_t = \left(\left(\frac{\tau_r}{\tau_r + \tilde{\tau}_t}\right)^2 \tau_u^{-1} + \left(\frac{\tilde{\tau}_t}{\tau_r + \tilde{\tau}_t}\right)^2 \tilde{\tau}_t^{-1}\right)^{-1}. \tag{7}$$

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.

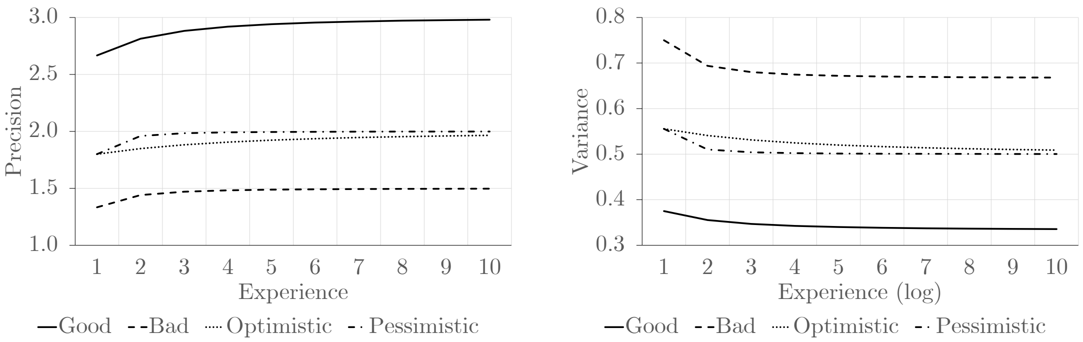
[9] In the more common α, β parameterization, this corresponds to $\alpha_0 = \tilde{\tau}^2_0/\sigma^2_{\tau,0}$ and $\beta_0 = \tilde{\tau}_0/\sigma^2_{\tau,0}$. While this notation makes the updating equations more convoluted, it allows for deriving the interest variable's law of motion more straightforwardly.

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 1 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.

Figure 1: Evolution of forecast precision and variance over time



Note: 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 $\tau_u = 1$.

In the left panel, we show the "raw numbers", i.e., time (experience) t and precision τ_t, obtained from substituting Eq. 5 into Eq. 7. 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, 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. 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[10] 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.

4 Data

In this paper, we combine three unique datasets: individual forecasts for the US CPI (Section 4.1), web-sourced information and characteristics about the forecaster's CV (Section 4.2), and web-sourced information about institutions where a forecaster worked (Section 4.3). These datasets cover the period from 1997:Q1 to 2017:Q4.

4.1 Forecaster Behaviors

Most of the data, including individual point forecasts and the name and affiliation of the experts, come from Bloomberg. Each expert can submit and update US inflation forecasts until the first day of the corresponding month before the publication of the effective US CPI inflation. 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 point of time, allowing forecasters to react to each other. Therefore, any deviation from the *herd* can be considered as deliberate. This allows us to capture both the quality (Fig. 2) and boldness (deviating from the herd) of forecasts.

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

$$L_{i,t} = \left(\pi_t - \mathbb{E}^i_{t-\tau}\left[\pi_t\right]\right)^2, \tag{8}$$

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.

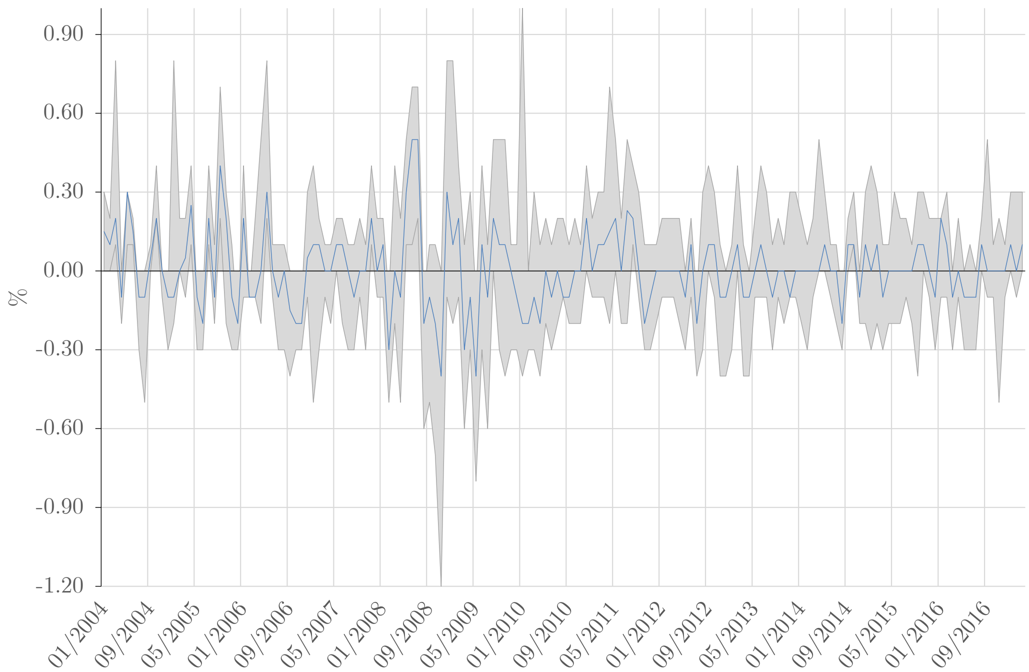
Second, we assess *boldness*, which we define as deviating from the "herd" of other forecasters. In other words,

$$B_{i,t} = \left(\mathbb{E}^i_{t-\tau}\left[\pi_t\right] - M_{i,t}\left[\mathbb{E}^i_{t-\tau}\left[\pi_t\right]\right]\right)^2, \tag{9}$$

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

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

Figure 2: Expert Forecast Performance



Note: 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.

Recall that an expert forecast can be updated until the last minute, and is public after submission. Since forecasts are typically not submitted in the last second, but we take 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 are overconfident in their forecasts since they disregard information from other forecasts.

There is, however, one caveat. Since we use nowcasts, where an excellent 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 mostly 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.

4.2 Forecaster Characteristics

Our databases on expert characteristics are based on an automated collection, which is manually augmented where needed. The primary source is LinkedIn.[11] 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, such as private and institutional websites. Partial database entries from LinkedIn were completed in the same fashion.

In essence, the data we collected about experts encompasses 151 experts and 164 institutions that published at least one US CPI inflation forecast. Since our panel specification includes forecaster fixed-effects, we drop all forecasters providing only a single forecast, leaving us with 112 experts.

For each forecaster, we collect data on past job experience (type and duration[12]), separated in to (a) academic experience, (b) central bank experience, and (c) experience in financial institutions. 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). 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 the highest degree, our database also includes the Shanghai Ranking classification for the respective university following both the economics and finance rankings.[13]

Last but not least, we collect additional information, such as origin and age. However, since we want to be consistent across all database entries, and the date of birth is typically not reported in US-style CVs, age is proxied using total experience plus 18 years plus the typical time necessary to obtain the final degree (three years for a Bachelor's degree, five for a Master's, and 11 for doctoral degrees).

4.3 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

[11] LinkedIn is a US employment-oriented Internet service, founded in 2002. It is mainly used for professional networking, including employers posting jobs, job seekers posting their CVs, and people who want to broaden their network.

[12] The expert's experience corresponds to the full period in which the forecaster published forecasts.

[13] We use the Shanghai Ranking's Global Ranking, Academic Subjects, 2017. We decompose each ranking (global, economics, and finance) to four levels: first tier, second tier, third tier, and not ranked.

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.

5 Results and Interpretation

This section uses the data described in Section 4 to detect the characteristics leading to forecasting performance and boldness, and experts' sentiments. We apply a proportional hazards model to examine expert's survival (Section 5.1), a cross-sectional analysis to identify the role of inherent traits (Section 5.2), a panel estimation to explore the role of experience (Section 5.3), a probit model to assess expert's pessimistic and optimistic behaviors (Section 5.4), and forecasting ability tests to identify the forecasting performance of characteristics-based groups of experts.

5.1 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 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:

$$h_{it} = h_0\left(t\right) \times e^{\lambda_0 f(Y_{it-1}) + X_i \Lambda}, \tag{10}$$

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.[14] 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

[14] Raw, MA5, and wMA5. Raw stands for the last period dependent variable, MA5 for five-period moving average (backward-looking) and wMA5 for weighted moving average where the most recent periods have a higher weight.

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 1 to capture our research question's intuition. Tables 1 to 3 present three transformations for each dependant variable. The Raw column stands for the last period performance or boldness[15] (raw data), MA5 for five-period moving average (backward-looking), and wMA5 for weighted moving average where the most recent periods have a higher weight. In these tables, positive significant coefficients show that a low forecasting performance or boldness increases the probability of removal (i.e., not being a forecaster in the next period).

Table 1: Survival Estimates: Intuition.

	Performance		
	Raw	MA5	wMA5
Dependent variable	1.801	5.119**	3.59**
Local forecaster (LF)	-0.425**	-0.417**	-0.419**
Education: Economics (EC)	-0.586**	-0.577**	-0.568**

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

Table 1 presents the survival estimates for US CPI experts, without interactions, related only to the expert's location and education field. 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 field and 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 financial analysts literature (Malloy, 2005). We also find that an expert with

[15] See Eq. 8 and Eq. 9, respectively, for the definitions of performance and boldness.

a degree in economics may survive longer than an expert without an economics degree, a result interpreted below (Table 2).

Table 2 presents the survival estimates for US CPI experts without interactions with all the variables at our disposal.[16]

Table 2: Survival Estimates Without Interactions.

	Performance			Boldness		
	Raw	MA5	wMA5	Raw	MA5	wMA5
Dependent variable	2.657	6.87**	4.655*	4.139	8.577	7.395
Age > 50 years old	0.561	0.555	0.554	0.586*	0.573*	0.583*
Local forecaster (LF)	-0.929**	-0.939**	-0.927**	-0.935**	-0.954**	-0.943**
Local institution (LI)	-0.063	-0.026	-0.046	-0.048	-0.038	-0.041
Financial institution (FI)	-0.5	-0.459	-0.482	-0.465	-0.444	-0.452
Experience: Central bank (CB)	-0.318	-0.304	-0.312	-0.298	-0.277	-0.279
Experience: Academia (AC)	0.691*	0.667*	0.692*	0.69*	0.657*	0.687*
Gender (G)	0.546	0.614	0.57	0.521	0.544	0.525
Education: MA (MA)	0.167	0.172	0.181	0.157	0.182	0.171
Education: PhD (PhD)	-0.795	-0.828	-0.795	-0.784	-0.766	-0.773
Education: Economics (EC)	-1.548***	-1.528***	-1.521***	-1.562***	-1.551***	-1.544***
Education: Finance (EF)	-1.288*	-1.332*	-1.273*	-1.358**	-1.366**	-1.34**
Ranking: Economics (RE)	0.287*	0.288*	0.283*	0.263*	0.261*	0.257
Ranking: Finance (RF)	-0.188	-0.183	-0.19	-0.177	-0.175	-0.178

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

Table 2 shows that a low past forecasting performance (MA5 and wMA5) increases the probability of being removed. However, providing bolder forecasts does not seem to decrease expert survival significantly.[17] According to these results, the expert's future seems to be determined partly by current and past performance, but not necessarily by herding behavior. Herding does not seem to significantly influence the experts' survival rate.[18]

Table 2 also shows that having a degree in economics immunizes the expert from being removed as an economist. Interestingly, a degree in finance does not offer the same protection as having graduated in economics. Experts with a degree in

[16] See Section 4 for more details.

[17] 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.

[18] 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.

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,[19] increasing their survival rate.[20] Having graduated from a top-ranked university in economics[21] or having experience in academia seems to deteriorate the expert's survival rate.

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 from top-ranked universities or academia to leave the profession more frequently than other experts. This might also mean that these experts are fired more often because, for instance, they are hired based on higher expectations than others, which they may not meet.

The expert's age appears to be a crucial factor for survival when dealing with boldness. Bolder experts will survive less the farther they are above fifty. While this result is not necessarily related to accumulated experience, it shows that an old expert forecasting outside the consensus has a higher chance of not being a forecaster in the next period than one who is in the consensus or younger.

Table 2 also shows that being a foreign expert, i.e., not living in the US, affects the survival of the expert. This bias in favor of local experts reflects the average inability of experts not living in the US to survive as a forecaster.

Table 3 presents the survival estimates for US CPI experts with interactions.

Table 3 provides additional insights about expert survivals. First, it shows that having graduated with a Master's degree from a top finance institution plays a positive role in the expert's survival. This result sharply contrasts with Master's degrees from top institutions in economics. Second, this result is mitigated when it interacts with a higher education level. PhD graduation from a top institution in economics increases the survival probability of the expert. However, this increase in the expert's survival rate is less significant for a PhD from a top institution in economics than for a Master's degree from a top institution in finance.

One might think the expert's institution is more an intermediary than a factor. An expert with a PhD might increase his chances of working in a better institution, with better information-gathering function, than an expert without a PhD.

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

[20] 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 with an education in both economics and finance.

[21] See Section 4.2.

Table 3: Survival Estimates with Interactions.

	Performance			Boldness		
	Raw	MA5	wMA5	Raw	MA5	wMA5
Dependent variable	2.661	6.569**	4.584	4.083	6.785	6.803
Age > 50 years old	0.969***	0.915***	0.938***	0.958***	0.914***	0.93***
Local forecaster (LF)	-1.48***	-1.481***	-1.472***	-1.501***	-1.502***	-1.499***
Local institution (LI)	0.387	0.413	0.394	0.385	0.379	0.377
Financial institution (FI)	-0.63	-0.592	-0.61	-0.716*	-0.7*	-0.699*
Experience: Central bank (CB)	-0.231	-0.183	-0.21	-0.219	-0.176	-0.19
Experience: Academia (AC)	1.228***	1.183***	1.217***	1.15***	1.105***	1.134***
Gender (G)	0.345	0.46	0.391	0.335	0.397	0.355
Education: MA (MA)	-0.251	-0.225	-0.237	-0.25	-0.223	-0.241
Education: PhD (PhD)	-1.574**	-1.579**	-1.565**	-1.512**	-1.488**	-1.497**
Education: Economics (EC)	-1.409***	-1.342**	-1.361**	-1.424***	-1.41***	-1.393***
Education: Finance (EF)	-1.448*	-1.491**	-1.429*	-1.521**	-1.548**	-1.493**
Ranking: Economics (RE)	0.188	0.172	0.178	0.196	0.19	0.186
Ranking: Finance (RF)	-0.043	-0.019	-0.035	-0.043	-0.035	-0.04
Interactions						
MA × RE	0.384	0.421	0.39	0.339	0.341	0.329
MA × RF	-0.982***	-0.997***	-0.992***	-0.968***	-0.959***	-0.97***
PhD × RE	-0.927**	-0.941**	-0.937**	-0.9**	-0.92**	-0.918**
FI × RF	0.383	0.424	0.398	0.378	0.394	0.385
FI × LI	0.635	0.41	0.523	0.503	0.382	0.409
G × LI	0.25	0.328	0.324	0.219	0.347	0.258

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

Financial institutions may have a 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 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.

Interestingly, Table 3 shows that the expert's age significantly affects the expert's survival rate for both low performances and bolder forecasts. As in Table 2, it indicates that low average performance (MA5) significantly decreases the expert's survival rate.

Tables 2 and 3 mitigate 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, older and probably more experienced experts are more likely to lose their jobs than younger and perhaps less experienced experts. Among other things, 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.

5.2 Cross-Section Estimation

Because most of the forecaster traits we collect in our data do not vary over their forecast history, we opt for a simple cross-sectional approach to assess their relevance in this second set of analyses. However, we have to account for the fact that a large part of forecast errors comes from unpredictable shocks to inflation, affecting all forecasts simultaneously. This would not be an issue if we had a balanced panel (i.e., all forecasters were affected by this equally), but we have some forecasters only active during the so-called Great Moderation, while others provided forecasts during the turbulent times of the GFC.

Essentially mirroring the idea of time-fixed-effects in panel analysis, we control for the average performance of other forecasters during the period in which the forecaster under consideration was active. For simplicity, we refer to this control as the *active period-specific effect*[22] for the remainder of the paper. This yields the regression equation:

$$\bar{y}_i = X_i\Gamma + \gamma_0 \frac{\sum_{t\in\mathbb{T}_i}\sum_{j\in\mathbb{F}_t} y_{jt}}{\sum_{t\in\mathbb{T}_i}(|\mathbb{F}_t|-1)} + \epsilon_i, \tag{11}$$

where X_i is the set of forecaster traits, $\mathbb{T}_i$ is the set of periods when forecaster i was active, $\mathbb{F}_t$ is the set of forecasters active at time t, and $|\mathbb{F}_t|$ is the cardinality of that set. The bar operator indicates the arithmetic mean. γ_0 is a regression coefficient and ϵ is the error term.

By controlling for institutional characteristics, we also guarantee that the effect of forecaster characteristics is not merely driven by some institution's easier access to forecasters with specific traits (e.g., forecasters with a PhD or from the best institutions). We decompose our sample (1997:Q1-2017:Q4) into two subsamples, the pre-GFC (1997:Q1-2008:Q1) and post-GFC (2008:Q1-2017:Q4) subsamples. Table 4 presents the cross-section estimations for the post-GFC US CPI inflation expert forecasts, with the correction for time-effects outlined in Eq. 11, for the variables presented in Section 4. The results presented in this section are based

[22] Namely, Dependent variable (mean).

on demeaned variables so that the point estimate of the non-interacted variables corresponds to the marginal effect at the mean.

Table 4: Cross-Section Estimates: Intuition

	Performance
Dependent variable (mean)	0.45
Local forecaster (LF)	-0.006*
Local institution (LI)	0.003
Financial institution (FI)	-0.007*
Experience: Central bank (CB)	-0.001
Experience: Academia (AC)	0.001
Gender (G)	0.003
Education: MA (MA)	-0.011**
Education: PhD (PhD)	-0.015***
Education: Economics (EC)	-0.006
Education: Finance (EF)	-0.008
Ranking: Economics (RE)	-0.002*
Ranking: Finance (RF)	0.001
Observations	58
Adjusted R^2	0.068

Note: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 4 shows that education level and field, and institution type and location, could have a role in the expert's forecasting performance. An expert with a Master's or PhD degree in Economics, local or affiliated with a financial institution, may provide better inflation forecasts. These results, which provide a clue about our research question, are developed and discussed below.

Table 5 presents the cross-section estimations for the US CPI inflation expert forecasts with the correction for time-effects.

The cross-section estimates presented in Table 5 show that performance and boldness are influenced differently.

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

Table 5: Cross-Section Estimates

	Performance			Boldness		
	Full sample	Pre-GFC	Post-GFC	Full sample	Pre-GFC	Post-GFC
Dependent variable (mean)	1.23***	0.471	1.303***	0.969***	-1.011	0.978***
Local forecaster (LF)	-0.012	0.006	-0.013	-0.002	0.011	-0.001
Local institution (LI)	-0.012	-0.105***	-0.005	-0.007	-0.077***	-0.004
Financial institution (FI)	-0.015	-0.101**	-0.007	-0.003	-0.055*	0
Experience: Central bank (CB)	0	-0.005	0.002	-0.003	0.009	-0.002
Experience: Academia (AC)	-0.001	0.003	0	-0.003	-0.001	-0.003
Gender (G)	-0.003	-0.015	0.001	-0.002	-0.027	0
Education: MA (MA)	0.007	-0.013	0.008	-0.001	-0.026	-0.002
Education: PhD (PhD)	-0.003	-0.018	-0.005	-0.005	-0.012	-0.006
Education: Economics (EC)	-0.014	0.004	-0.022	-0.008	0.003	-0.016*
Education: Finance (EF)	-0.016	0.025	-0.027	-0.013	0.05*	-0.022**
Ranking: Economics (RE)	-0.001	-0.006	-0.002	-0.003	-0.022	-0.003
Ranking: Finance (RF)	-0.008	-0.089**	-0.009	0.001	-0.054*	0.002
Interactions						
MA × RE	-0.002	-0.007	-0.002	0.003	0.031	0.004
MA × RF	-0.007	0.025	-0.011	-0.006	-0.004	-0.008*
PhD × RE	-0.002	0.015	0.001	0.001	0.02	0.003
FI × RF	0.011	0.081**	0.013	0	0.06***	-0.001
FI × LI	0.022	0.101*	0.014	0.01	0.046	0.006
G × LI	0.023	0.143**	0.003	0.021**	0.138***	0.007
Observations	67	30	63	67	30	63
Adjusted R^2	0.35	0.53	0.38	0.11	0.53	0.08

Note: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

at a financial institution and 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 of all, working at a local or financial institution decreases the expert's boldness only before the GFC. However, having graduated in economics or finance decreases bold expert forecasts after the GFC, while having graduated in finance before the GFC increased the expert's boldness. Interestingly, the boldness of female experts from local institutions was greater before the GFC, while the robustness of this result is still questioned due to the paucity of female experts before the GFC. Over the full sample, female experts from local institutions herd more than male experts. After the GFC, having graduated with a Master's degree from a top university in finance slightly but significantly reduced the expert's boldness.

In line with Clarke and Subramanian (2006), performance and boldness results in Table 5 are often similar in terms of significance and sign,[23] 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 sometimes 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 5.3).

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 5.3.

5.3 Panel Estimation

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:

$$y_{it} = \beta_0 \ln x_{it} + u_i + v_t + \varepsilon_{it}, \tag{12}$$

where y_{it} is one of our two loss functions discussed in Section 4.1 (performance and boldness) and x_{it} is one of our experience measures discussed in Section 4. β_0 is the regression coefficient. 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 Section 5.1. While the evidence for the existence of such an effect is mixed, and if it exists 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[24] (Heckman,

[23] This is also the case with our next results (Table 7).

[24] 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.

1979). The resulting equation is given by:

$$y_{it} = \beta_0 \ln x_{it} + \beta_1 \mathbb{1}_{t \in last_5(i)} + u_i + v_t + \varepsilon_{it}, \tag{13}$$

where $last_5(i)$ is the set of the last five periods in which forecaster i submits a forecast to our dataset. β_0 and β_1 are regression coefficients.

The model controls for individual specific and time effects. Table 6 presents simplified panel estimations for the pre-GFC US CPI inflation expert forecasts with time and experience fixed-effects. The results presented in this section are also based on demeaned variables to identify the partial effects.

Table 6: Panel Estimates: Intuition

	Performance
Experience (E)	-0.031***
Age > 50 years old	0.17***
Two-way interactions with experience	
Local institution (LI)	-0.019**
Financial institution (FI)	0.026***
Gender (G)	-0.031***
Education: MA (MA)	-0.041***
Education: Economics (EC)	-0.031*
Ranking: Economics (RE)	0.015***
Ranking: Finance (RF)	-0.011***
Observations	339
Adjusted R^2	0.01

Note: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

The simplified results presented in Table 6 show that experience generally improves the expert's forecasting performance. It also provides an intuition about the results discussed below, namely the positive effects of a top-ranked educational institution in Finance or the affiliated institution location on the expert's performance. These initial results confirm the relevance of our research question and results.

Table 7 presents our full results, the panel estimations for the US CPI inflation expert forecasts with time and experience fixed-effects.

The panel estimation presented in Table 7 shows that being a local forecaster, working in a financial institution, or having experience at a central bank or in

Table 7: Panel Estimates

	Performance			Boldness		
	Full sample	Pre-GFC	Post-GFC	Full sample	Pre-GFC	Post-GFC
Experience (E)	-0.017***	-0.008	0.007	-0.009***	0.009	-0.001
Age > 50 years old	0.004	0.195**	0.008	-0.003	0.119**	-0.002
Two-way interactions with experience						
Local forecaster (LF)	0.033**	0.014	0.199*	0.002	0	0.091
Local institution (LI)	-0.088***	-0.066	-0.292***	-0.036***	-0.042	-0.131***
Financial institution (FI)	0.031***	0.061	0.077***	0.016***	0.015	0.041***
Experience: Central bank (CB)	0.018***	0.036	0.019***	0.007***	0.028	0.006
Experience: Academia (AC)	0.074***	0.011	0.306*	0.029***	0.009	0.17**
Gender (G)	-0.045***	-0.053*	-0.032	-0.029***	-0.035*	-0.002
Education: MA (MA)	-0.012	-0.005	-0.01	-0.007	-0.021	0.013
Education: PhD (PhD)	-0.025***	0.02	-0.037***	-0.012***	-0.006	-0.015***
Education: Economics (EC)	0.056*	0.039	0.128	0.016	0.034	0.094
Education: Finance (EF)	-0.081***	0.087	-0.288***	-0.036***	0.142***	-0.124***
Ranking: Economics (RE)	0.027***	-0.007	0.059***	0.014***	-0.001	0.025***
Ranking: Finance (RF)	-0.01***	-0.002	-0.006	-0.004***	-0.01	0.002
Three-way interactions with experience						
FI × RE	0.005	0.056	-0.018	0.004	0.027	-0.007
MA × RE	-0.128***	-0.046	-0.163***	-0.07***	-0.048	-0.05***
PhD × RE	-0.132***	-0.022	-0.19***	-0.073***	-0.027	-0.061***
EC × RE	-0.074***	-0.043*	-0.272***	-0.031***	-0.031**	-0.128***
FI × RF	-0.006	-0.093**	0.023	-0.002	-0.07***	0.004
G × RF	-0.065***	0.001	-0.029	-0.047***	-0.005	0.003
MA × RF	-0.003	-0.017	-0.007	-0.002	-0.005	-0.007**
EF × RF	0.004	-0.027	0.008	0.009***	-0.087***	0.002
FI × LI	-0.029**	-0.062	-0.064***	-0.014*	0.028	-0.001
AC × LI	-0.362***	0.059	-1.434**	-0.156***	0.027	-0.8**
MA × LI	0.246***	-0.099*	0.35***	0.145***	-0.052	0.111***
PhD × LI	0.245***	0	0.28***	0.165***	0	0.139***
FI × LF	-0.106***	0	-0.147***	-0.057***	0	-0.057***
G × LF	-0.085**	0	0.093	-0.084***	0	0.06
AC × LF	0.357***	0	1.458**	0.155***	0	0.805**
MA × LF	0.152***	0	0.191***	0.081***	0	0.072***
PhD × LF	0.167***	0	0.263***	0.067***	0	0.049**
EC × LF	0.182***	0	0.598***	0.082***	0	0.273***
G × FI	-0.349***	0	-0.413***	-0.178***	0	-0.144***
AC × FI	-0.018	0	0.106***	0.002	0	0.017
EC × FI	-0.014	0	-0.415***	0.001	0	-0.22***
G × CB	0.137***	0	0.211***	0.072***	0	0.066***
EC × AC	0.073***	0.04	0.295***	0.034***	0.016	0.135***
EC × MA	-0.13**	0	-0.206	-0.023	0	-0.152
Observations	3073	350	2723	3077	350	2727
Adjusted R^2	0.04	0.08	0.07	0.02	0.02	0.03

Note: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

academia decrease the expert's forecasting performance after the GFC or over the full sample. It also shows that experience, working in a local institution, or being educated in finance increase the expert's forecasting performance after the GFC or over the full sample. Experts may underreact less to prior CPI information as their experience increases, suggesting one reason why experts, like analysts, become more accurate with experience (Mikhail et al., 2003). Similar to analysts' firm experience, which is strongly and positively associated with analysts forecast boldness (Clarke and Subramanian, 2006; Huang et al., 2017), our results complement these findings for CPI experts by differentiating the experience type, size, and somehow the reputation accumulated through their education.

The pre-GFC period significantly differs from the post-GFC one. Over the pre-GFC period, only being a local forecaster appeared to explain the expert's forecasting performance significantly. Being a female expert over the full or pre-GFC sample decreases the forecasting performance, but like all these results, they are mitigated by interactions. However, our results show that the GFC contributed significantly to change the distribution of the effects of characteristics on forecasting performance and boldness, a loosely documented fact.[25]

Interactions show that having a Master's degree or a PhD in economics, education in a top university (economics ranking), working in a financial and local institution, having experience in academia and working in a local institution, or being a local forecaster working in a local institution, generally improve the expert's forecasting performance over the full sample but also during the post-GFC period. Before the GFC, having a degree in economics from a top university (economics ranking), a degree in finance from a top university (finance ranking), or holding a Master's degree and working in a local institution improved expert's forecasting performances.

Putting aside interactions, the results reported in Table 7 are close to the herding results except for one interesting instance: Having a degree in finance led to greater boldness before the GFC but to more herding behaviors after the GFC and over the full sample. 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). However, unlike forecasting performance, being a local forecaster does not seem to influence experts' herding behaviors, which slightly mitigates Clarke and Subramanian (2006).

Interactions reveal a more detailed picture. Better education quality in economics[26] improves experts' forecasting performance among those with a Master's

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

[26] Measured with the ShanghaiRanking's Global Ranking of Academic Subjects in Economics,

or a PhD degree. 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 and finance generally accentuates experts' herding behaviors, except for experts who graduated in economics from a top-ranked university in finance. The latter characteristics tend to increase experts' boldness.

Working in a financial and local institution improves experts' forecasting performances while it weakly influences herding behaviors. However, working in a local institution while having experience in academia increases both forecasting performance and herding behaviors. This is mitigated if local experts working in a local institution are considered. Working in a local institution and having a Master's degree or a PhD decreases both forecasting performance and bolder behaviors. The previous findings are confirmed for local experts having a Master's degree or a PhD. Local forecasters with a degree in economics tend to provide low-quality forecasts that are far from the consensus. Interestingly, having experience in academia and working in a financial institution does not significantly influence experts' herding behavior, but does affect their forecasting performance. Similarly, having a degree in economics and having a Master's degree affect forecasting performance and herding behaviors differently. The findings that mix geography and education relate to several strands of the literature. Our results demonstrate that the likelihood of boldness increases with the expert's forecasting performance and 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 having central bank experience outperform men having central bank experience (E $\times$ G $\times$ CB), which may confirm a labor market entry selection bias. Female experts in a market segment in which their concentration is lower (central banking) 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 strongest under recessions, our results bring another layer to this conclusion. The GFC not only changed differences in opinion but also modified the influence of expert's characteristics on their forecasting performance and boldness.

---

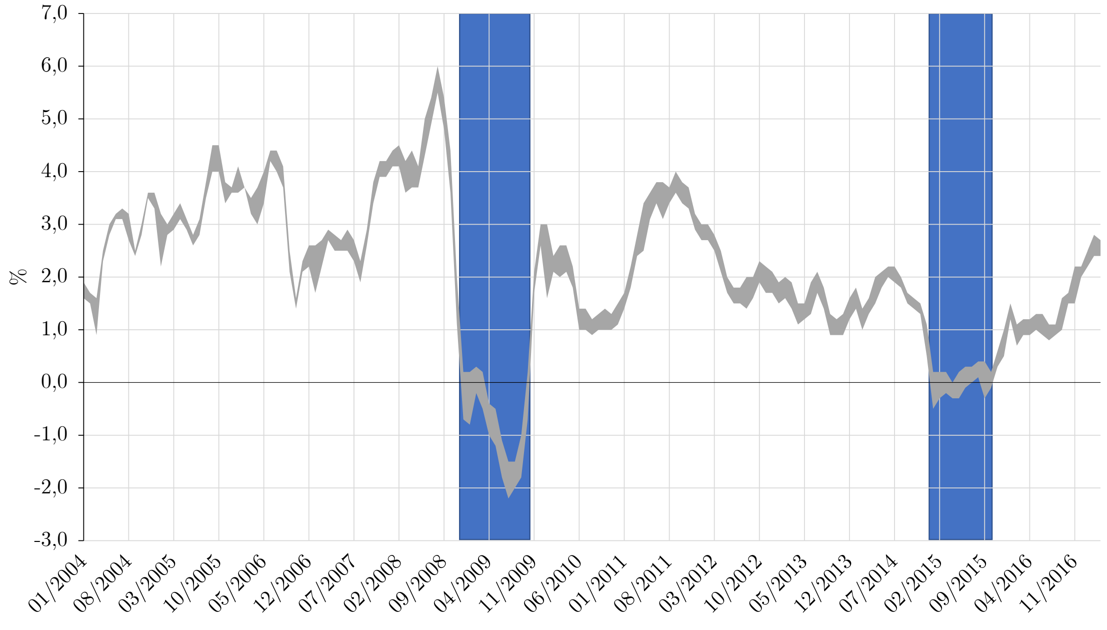
see Section 4.

5.4 Pessimism

In our last analysis, we estimate "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 2.0%.

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 3 shows how in the early period of the GFC (late 2008 to early 2009), and then again over most of 2015, forecasters disagreed on whether or not there would be deflation.

Figure 3: Forecast spread and periods of mixed inflation/deflation expectations.



Note: 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.

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:

$$p\left(\mathbb{E}_{it}\left(\pi_i t\right) < 0\right) = \Phi\left(\psi_0 \ln x_{it} + X_i \Psi\right), \tag{14}$$

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

We then split this sample even further 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) vs. periods where no deflation was observed, i.e., when expecting inflation can genuinely be seen as overly pessimistic.

Table 8 presents the probit panel regressions for a general situation (i.e., during both inflationary and deflationary periods) and during only deflationary periods.

Table 8: Probit Estimates for Pessimism

	Performance	
	All	Deflation
Experience	-0.628***	-0.557***
Local forecaster (LF)	-0.165	-0.088
Local institution (LI)	-0.228	-0.272
Financial institution (FI)	-0.172	-0.067
Experience: Central bank (CB)	-0.965***	-0.576
Experience: Academia (AC)	0.598	0.616
Gender (G)	0.116	-0.166
Education: MA (MA)	0.398	0.138
Education: PhD (PhD)	0.375	0.294
Education: Economics (EC)	14.538	15.208
Education: Finance (EF)	14.668	15.542
Ranking: Economics (RE)	0.129	0.184
Ranking: Finance (RF)	-0.095	-0.197
Observations	336	212

Note: ***, **, 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.

Table 8 shows that experts with more experience or with central bank experience are less likely to predict deflation. However, this propensity becomes insignificant under deflation for experts with central bank experience, which strongly influences private information. These experts are less pessimistic, but this is mitigated when pessimism turns out to be justified.

In line with our model (Section 3), 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, exposed in Section 3, 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).

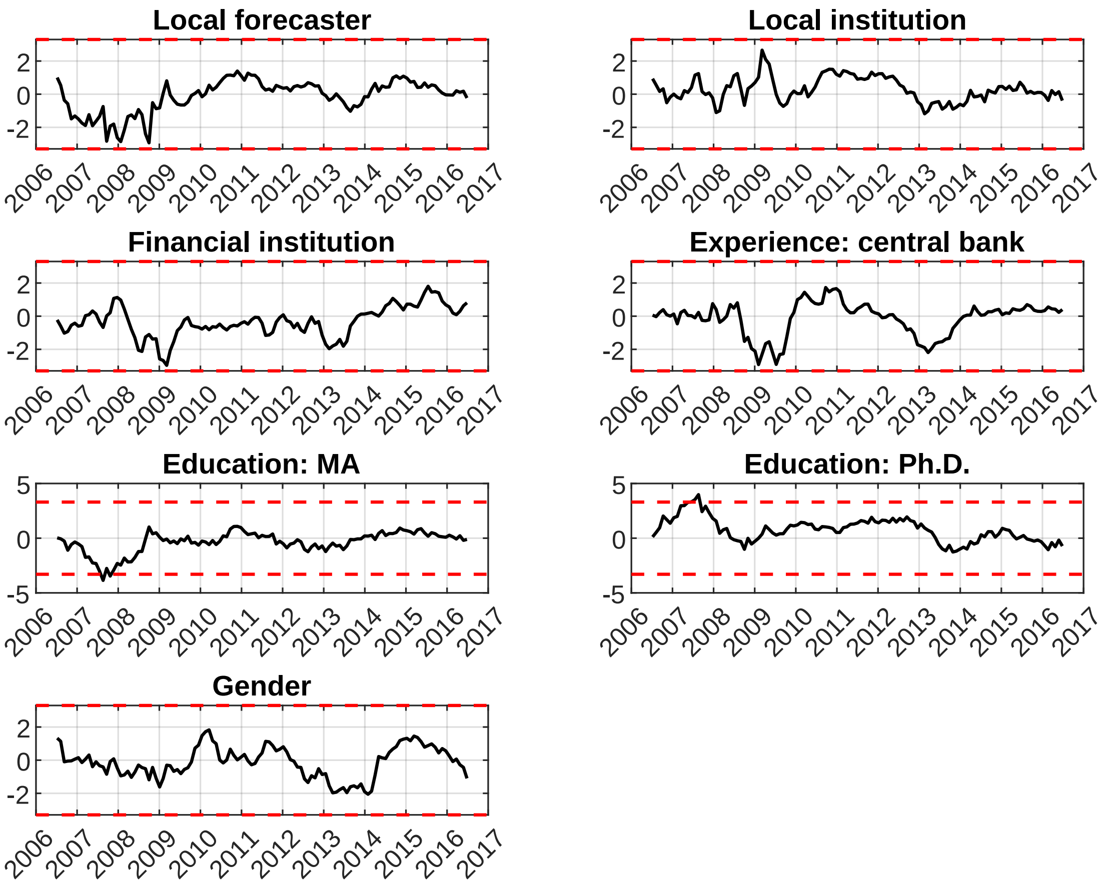
5.5 Forecasting Ability

The characteristics identified in Section 5 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 characteritics-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 merely that there is a meaningful difference in predictive ability for a subsample. Fig. 4 presents these fluctuation tests for forecasts grouped by characteristics.

The test rejects only for education level characteristics (MA and PhD, see Fig. 4) 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 for the remaining characteristics, confirming that experts' characteristics may have a significant role in their out-of-sample forecasting outcomes even in unstable environments. Our findings may have implications for policymakers. They may decide to select their inputs (multiple expert

Figure 4: Forecasting Ability Tests



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

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.

6 Policy Implications and conclusion

In line with our model, the characteristics we identify influence the 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) 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 decisions. The more accurate the forecast, the less a surprise could occur, minimizing the required adjustment costs of the investment portfolio and the corresponding market volatility when the data will be made public (Laster et al., 1999).

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 5.5) 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 5.4).

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 channels 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. Related to career concerns and institutional labor market expectations, degrees in finance or economics do not offer the same protection, while having graduated from a top university decreases the expert's survival rate.[27] Our survival analysis also sheds light on the expert's age, unrelated to accumulated experience, which appears to decrease expert survival. The older the expert (above 50 years old), the less a bold expert will survive.

The comparison of 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, or education field or quality change their forecasting performance. Our panel estimations also show that before the GFC, only the expert's age, institution type, education field, and quality matter for forecasting performance and boldness.

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

We interpret our results as evidence of a characteristics effect in inflation expert 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.

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.

[27] See MA × RF and PhD × RE in Table 7.

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 (U.S.).

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 Characteristics and Herding Behavior in Forecasting. Journal of Finance 60 (1), 307–341.

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.

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.

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.

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.

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.