Behavioral Impulse Responses

Bryan Kelly, Semyon Malamud, Emil Siriwardane, and Hongyu Wu*

This version: August 7, 2025

Abstract

We develop the concept of a Behavioral Impulse Response (BIR) to characterize how deviations from full-information rational expectations (FIRE) are corrected over time. BIRs generalize popular tests of FIRE by tracing out how forecast errors for a fixed target evolve following a revision. BIRs estimated on professional forecasts of macroeconomic outcomes and corporate earnings reveal that violations of FIRE occur far more frequently than suggested by existing tests. These violations tend to correct in a slow and gradual fashion, with sizable variation in correction speeds across forecast targets and forecasters. Our theoretical analysis shows that BIRs provide a simple yet powerful set of moments that help discipline models of belief formation.

JEL: C52, C53, D83, D84, E7, E17, G17, G14

*Bryan Kelly is at AQR Capital Management, Yale School of Management, and NBER. Semyon Malamud is at Swiss Finance Institute, EPFL, and CEPR, and is a consultant to AQR. Emil Siriwardane is at Harvard Business School and NBER. Hongyu Wu is at Yale School of Management. We are grateful to Nick Barberis, John Y. Campbell, Robert F. Engle, Leland Farmer (Discussant), Sam Hanson, Max Miller, Marco Sammon, Josh Schwartzstein, Andrei Shleifer, Jeremy Stein, Adi Sunderam, Luis Viceira, seminar participants at NBER 2024 Fall Behavioral Finance Working Group Meeting and Yale SOM for their helpful comments. AQR Capital Management is a global investment management firm that may or may not apply similar investment techniques or methods of analysis as described herein. The views expressed here are those of the authors and not necessarily those of AQR.

1 Introduction

A growing body of research in macroeconomics and finance documents systematic departures from full-information rational expectations (FIRE) in the forecasts of households, investors, and corporate managers.[1] Under this paradigm, individuals should form statistically efficient forecasts that optimally incorporate all available information. Yet, in reality, forecasts often deviate from FIRE due to behavioral biases (Bordalo et al., 2020b; Afrouzi et al., 2023), imperfect information (Coibion and Gorodnichenko, 2012, 2015), and model uncertainty (Timmermann, 1993; Schwartzstein and Sunderam, 2021; Farmer et al., 2023).

In this paper, we ask a simple question: how and when are these deviations corrected? Do forecasters "wake up" quickly after their predictions depart from FIRE or do corrections unfold more gradually? To appreciate the practical importance of this question, suppose that corporate managers are fully rational, but investors overreact to positive news (Stein, 1996). If investors quickly recognize and correct their error, then managers may ignore short-lived stock price inflation when making hiring, investment, and capital structure decisions. Conversely, managers may react more strongly when markets overreact if such errors are long-lived, thereby exerting tangible and distortive effects on the real economy.

To analyze the correction process rigorously, we develop a new tool called a behavioral impulse response (BIR). As its name suggests, BIRs are inspired by the long-standing use of impulse response functions in macroeconomics to study the dynamic effects of shocks (Sims, 1980). In a similar spirit, BIRs capture how forecast errors for a fixed target evolve following a forecast revision. BIRs can be easily estimated using OLS regressions with any dataset that contains a term structure of forecasts. More importantly, they provide a compact yet powerful set of moments that can be used to distinguish among competing models of belief formation.

[1] This research spans models of the macroeconomy (Sims, 2003; Woodford, 2003; Mankiw and Reis, 2002; Gabaix, 2020; Pflueger et al., 2020), corporate practice (Baker et al., 2007; Malmendier, 2018), and capital markets (Barberis et al., 2015a, 2021).

The idea behind BIRs is simple. Consider a sequence of forecasts $F_t, F_{t+1}, ..., F_{t+n-1}$ for a *fixed* target y_{t+n}, such as 2027Q4 GDP growth. Let $e_t = y_{t+n} - F_t$ denote the realized error associated with the forecast F_t and $R_t = F_t - F_{t-1}$ the one-period forecast revision made at time t. Focusing on a fixed target is crucial for isolating correction dynamics and distinguishes our approach from prior work that examines forecast errors for moving targets (Angeletos et al., 2021; Kučinskas and Peters, 2022). The BIR at horizon j is defined as:

$$B(j) = \mathbb{E}_t[e_{t+j}|R_t = 1] - \mathbb{E}_t[e_{t+j}|R_t = 0], \quad 0 \leq j < n,$$

which, in words, is the expected forecast error at $t + j$ given a unit revision at t. The first point on the BIR, $B(0)$, is the object of interest in the influential study of Coibion and Gorodnichenko (2015). When it is negative, the revision R_t is said to overreact, as it leads to a forecast F_t that reliably overshoots the target y_{t+n}.[2] The same logic applies to underreaction when $B(0) > 0$. Under FIRE, $B(j) = 0$ for all j, since forecast errors should be unpredictable by past revisions.

The remaining points on the BIR trace out how future errors evolve following a unit revision at time t. This means the shape of the BIR encodes the expected future path of revisions:

$$B(j-1) - B(j) = \mathbb{E}_t[R_{t+j}|R_t = 1] - \mathbb{E}_t[R_{t+j}|R_t = 0],$$

where $R_{t+j} = e_{t+j-1} - e_{t+j} = F_{t+j} - F_{t+j-1}$ is the one-period revision made at time $t + j$. To illustrate, suppose that forecasters overreact at t. If they are stubborn and never correct their mistake, all future revisions will be zero on average, resulting in a flat BIR (i.e., $B(0) = B(1) = ... = B(n-1) < 0$). In contrast, if forecasters immediately correct their mistake,

[2]There is some debate over how to interpret $B(0)$ (Afrouzi et al., 2023; Kučinskas and Peters, 2022). Following Coibion and Gorodnichenko (2015) and Bordalo et al. (2020b), we treat $B(0) < 0$ as overreaction. Regardless, $B(0) \neq 0$ constitutes a clear violation of FIRE.

next period's forecast will be revised downward, and all future forecast errors will be zero (i.e., $B(1) = ... = B(n-1) = 0$). This correction process is directly embedded in the shape of the BIR (see Figure 1), since $\mathbb{E}_t[R_{t+1}|R_t = 1] - \mathbb{E}_t[R_{t+1}|R_t = 0] = B(0) - B(1) = B(0) < 0$.

Every model of belief formation implies a corresponding BIR. Our main theoretical contribution is to derive BIRs for a general class of belief formation processes that encompasses most leading models in the behavioral economics literature (see Section 3). We develop intuition for the paper's main proposition by applying it to three prominent recent models, all of which generate overreaction in individual forecasts: diagnostic expectations (Bordalo et al., 2020a), costly information processing (Afrouzi et al., 2023), and extrapolative expectations (Hirshleifer et al., 2015; Nagel and Xu, 2022). We focus on these models because individual overreaction emerges as a pervasive feature in our empirical work.

BIRs exhibit stark differences across models, making them powerful tools for distinguishing among competing theories. For example, existing versions of diagnostic expectations feature abrupt corrections. While the precise timing depends on model calibration, standard parameterizations imply that forecasters fully reverse their initial overreaction after just one period, regardless of its magnitude. This leads to a sharp discontinuity in the BIR at $t+1$. In contrast, models of costly information processing yield smoother corrections. In these models, the speed of adjustment is closely tied to the degree of overreaction: less persistent processes generate larger overreactions but faster corrections. For both diagnostic expectations and costly information processing, deviations from FIRE at time t eventually fade as the realization date nears. Such convergence is absent in standard calibrations of extrapolative expectations, which feature an exceedingly slow, albeit smooth, correction process.

Empirically, BIRs are straightforward to estimate using local projection methods commonly employed in macroeconomics (see Jordà (2023) for a recent survey). We apply these methods to two standard datasets in the belief-formation literature: the Survey of

Professional Forecasters (SPF) and the Institutional Brokers' Estimate System (IBES). These datasets respectively contain individual forecasts of macroeconomic outcomes and corporate earnings for publicly traded U.S. stocks (e.g., Mankiw et al., 2003a; Coibion and Gorodnichenko, 2012; Bouchaud et al., 2019). One concern, particularly in IBES, is that forecasts may reflect strategic incentives rather than pure beliefs (Francis and Philbrick, 1993; Jackson, 2005). To address this concern, we highlight features of BIRs that are robust across both datasets. In addition, we report results using GDP forecasts from the Congressional Budget Office (CBO), where strategic motives are less likely. Our baseline analysis focuses on the longest available forecast horizons: $n = 3$ quarters in SPF and $n = 6$ in IBES.[3] As we discuss later, revisions at these horizons are more likely to reflect perceived news, making the resulting BIRs easier to interpret.

The estimated BIRs reveal five key facts about the dynamics of belief formation. First, most individual forecasts overreact in the Coibion and Gorodnichenko (2015) sense, at least for longer-horizon forecasts, and the degree of overreaction is decreasing in series persistence. These findings confirm the findings of Bordalo et al. (2020a) and Afrouzi et al. (2023). Our second, novel, fact is that corrections tend to be smooth and gradual. For example, in the average BIR from IBES, roughly 60% of the overreaction present at time t remains in the forecasts made one year later.

Our next set of facts sheds light on what triggers a correction. Correction speeds vary widely across forecasters and series, and are not tightly linked to the size of the initial error (Fact 3).[4] Moreover, forecasters are more likely to correct themselves when confronted with news that contradicts their initial revision (Fact 4), suggesting that the arrival and interpretation of new information also play a central role in shaping correction dynamics.

Lastly, BIRs evolve with experience and do not appear to be fixed individual traits. We

[3]Longer-horizon forecasts are available in IBES but are less populated. For completeness, these horizons are analyzed in Internet Appendix IA.3.1.

[4]Heterogeneity in BIRs across both forecasters and series exists within a subset of SPF forecasters who cover almost all series, suggesting that selection into particular forecast targets is not the primary driver.

document this fact by splitting each forecaster's history into early- and late-career periods. In both the SPF and IBES datasets, BIRs tend to move closer to the FIRE benchmark as forecasters gain experience.

What do these facts imply for behavioral models? As discussed earlier, the models we examine in our theoretical work were primarily designed to generate overreaction in individual forecasts (Fact 1), and our analysis confirms that they perform quantitatively well in this respect. In contrast, smooth and heterogeneous correction processes that are largely independent of the initial deviation from FIRE (Facts 2 and 3) pose greater challenges. Existing models of overreaction often imply correction dynamics that are too abrupt, too slow, or too tightly linked to the initial degree of overreaction. In our view, the next generation of behavioral models should therefore aim to accommodate more flexible correction behavior, likely through an expanded parameter set. The convergence of BIRs toward FIRE with experience (Fact 5) is also difficult to reconcile with models based on static behavioral traits and instead points to a potential role for learning.

To explore learning more directly, we also analyze the model of Farmer et al. (2023), in which Bayesian agents must form forecasts of a process that is the sum of a unit root and a transitory component, both unobserved. Farmer et al. (2023) show how model uncertainty of this kind severely slows Bayesian learning, leading to persistent deviations from FIRE, like those documented by Coibion and Gorodnichenko (2015). Building on their framework and simulation code, we ask whether model uncertainty can also produce realistic BIRs. Our results suggest that it can, depending on the initial priors used. Nevertheless, we are cautious about drawing strong conclusions from this exercise. The Bayesian framework's flexibility in specifying priors provides substantially more degrees of freedom than the parsimonious behavioral models we examine, making direct performance comparisons potentially misleading. Moreover, the choice of initial priors is difficult to discipline empirically, and a large body of observational and experimental evidence documents systematic departures

from perfect Bayesian updating in practice (e.g., Augenblick and Rabin, 2021; Gennaioli et al., 2025). The broader point of this exercise is that BIRs can be paired with simulations to evaluate a wide range of belief formation models.

Our empirical analysis reveals another important consideration for interpreting BIRs and related tests. Strictly speaking, impulse responses are defined as responses to structural shocks or other identified residuals in an economic model. BIRs, by contrast, capture responses to innovations in forecasts. A recurring theme in our analysis is that such innovations can stem from two sources: (i) perceived news and (ii) corrections of prior deviations from FIRE. The presence of this second component implies that BIRs are not structural in the traditional sense and complicates the standard interpretation of Coibion and Gorodnichenko (2015) (CG) regressions, which typically assume that revisions are responses to perceived news. This issue is particularly relevant for short-horizon forecasts, where a longer history of prior forecasts increases the likelihood that revisions reflect corrections.

To bring BIRs closer to a structural interpretation, the final part of the paper develops a generalized version that more credibly isolates the perceived news component of a revision R_t.[5] We argue that this can be achieved by conditioning on lagged revisions (i.e., $R_{t-1}, R_{t-2}, ...$). These generalized BIRs (GBIRs) reveal that short-horizon forecasts in IBES tend to underreact, while long-horizon forecasts overreact—patterns consistent with post-earnings announcement drift (Ball and Brown (1968), Bouchaud et al. (2019), and Augenblick et al. (2024)). We also show how BIRs and GBIRs can be used to implement more statistically powerful tests of FIRE. For example, at short horizons in IBES, less than 10% of forecasters reject the null of FIRE under the CG test, compared to 90% under the GBIR-based test.

Literature Review BIRs extend the influential tests of FIRE designed by Coibion and Gorodnichenko (2015). This extension yields two primary benefits. First, BIR-based tests of

[5]Much like vector autoregressions (VARs) in macroeconomics, BIRs still provide useful reduced-form moments for guiding theory even if revisions are not fully driven by perceived news.

FIRE have more statistical power because they leverage joint restrictions along the entire BIR curve, rather than a single restriction like that of Coibion and Gorodnichenko (2015). Second, and more importantly, BIRs shed light on the dynamics of belief correction by examining the temporal evolution of forecasts for a fixed target. Our theoretical analysis shows that these dynamics are useful for disciplining models of belief formation. Conceptually, our approach is closely related to recent work by Augenblick and Rabin (2021), who develop a Bayesian update test based on how the forecasts of a fixed target evolve over time.[6]

Our paper also complements previous research that has estimated the impulse response of forecast errors to identified macroeconomic shocks, mainly to characterize information frictions facing forecasters (Coibion and Gorodnichenko, 2012), explain discrepancies between individual and consensus forecasts (Angeletos et al., 2021), and examine forecasters' reactions to specific structural shocks (Kučinskas and Peters, 2022).[7] A key distinction is that these studies analyze forecast errors for a *moving* target—such as the time-series of one-period ahead forecast errors—while BIRs focus on the evolution of forecast errors for a *fixed* target. This subtle but important difference is precisely what allows BIRs to characterize the correction process. Moreover, BIRs do not rest on the assumption that forecasters can observe and identify structural shocks, a task that is particularly challenging in macroeconomic contexts (Bauer and Swanson, 2023). As with CG tests, BIRs rely on the much weaker assumption that forecasters are simply aware of their own forecast history.

The remainder of the paper is organized as follows. In Section 2, we formally define BIRs and describe a simple way to estimate them. Section 3 contains our theoretical analysis, which focuses primarily on how to use BIRs to distinguish between models of belief formation. Section 4 then estimates BIRs in both the SPF and IBES datasets. Section 5 highlights how BIRs can be paired with simulations to evaluate more complex models of belief formation,

[6] A large literature tests FIRE based on forecasts, including Mincer and Zarnowitz (1969), Lovell (1986), Thomas (1999), Mankiw et al. (2003b), Kohlhas and Walther (2021), Bianchi et al. (2022), and many others.

[7] In recent work, Bianchi et al. (2024b) embed forecasts from a machine learning model into a structural asset pricing model to study how the stock market reacts to different types of news.

like the one studied by Farmer et al. (2023). Section 6 presents a generalized version of BIRs and concludes. Additional details and results are also contained in a separate online appendix.

2 Behavioral Impulse Responses: Definition and Estimation

In this section, we formally define BIRs and highlight some of their basic properties. We then describe how to estimate BIRs on any dataset in which forecasts of a fixed target are tracked over time.

2.1 Definition and Basic Properties

The notation employed in this paper closely follows the conventions used in fixed-income asset pricing research (e.g., Cochrane and Piazzesi, 2005). Let y_t denote the realization of a stochastic time series, such as GDP or corporate earnings growth. $F_t^{(n)}$ is the forecast of y_{t+n} made at time t, or the n-period ahead forecast of y. For now, we consider only forecasts of a single time series made by a single individual. It is straightforward to extend the notation to accommodate multiple time series and forecasters, as we do in the empirical analysis contained in Section 4. $e_t^{(n)} = y_{t+n} - F_t^{(n)}$ is the realized forecast error associated with the forecast $F_t^{(n)}$. For a fixed target y_{t+n}, the sequence of forecasts from t onwards is therefore given by $F_t^{(n)}, F_{t+1}^{(n-1)}, ..., F_{t+n-1}^{(1)}$. Similarly, $e_t^{(n)}, e_{t+1}^{(n-1)}, ..., e_{t+n-1}^{(1)}$ is the sequence of forecast errors. Finally, $R_t^{(n)} = F_t^{(n)} - F_{t-1}^{(n+1)}$ is defined as the one-period forecast revision of y_{t+n} made at t.

Next, we formally define the behavioral impulse response function.

Definition 1 *The behavioral impulse response (BIR) function for forecast horizon n equals:*

$$B(n,j) = \mathbb{E}\left[e_{t+j}^{(n-j)} | R_t^{(n)} = 1\right] - \mathbb{E}\left[e_{t+j}^{(n-j)} | R_t^{(n)} = 0\right], \quad j = 0, 1, ..., n-1. \tag{1}$$

Given a unit revision made at t, $\{B(n,j)\}_{j=0}^{n-1}$ traces out the expected sequence of forecast errors, holding fixed the forecast target y_{t+n}. As their name suggests, BIRs are rooted in the vast macroeconomics literature on impulse response functions, dating back at least to Sims (1980). In the usual macroeconomic setting, an impulse response function traces how an outcome of interest, like inflation, responds dynamically to a shock, like a surprise change in the stance of monetary policy. Analogously, in our setting, a behavioral impulse response traces out how forecast errors evolve following a revision R_t at time t.

Under full-information rational expectations (FIRE), forecast errors are unpredictable by all information at t, implying that BIRs must be zero at all horizons and periods j after revision: $B(n,j) = 0 \quad \forall, n, j$. This simple property makes BIRs a useful tool for testing FIRE.

The initial point along the BIR curve, $B(n,0) = \mathbb{E}\left[e_t^{(n)} | R_t^{(n)} = 1\right] - \mathbb{E}\left[e_t^{(n)} | R_t^{(n)} = 0\right]$, is familiar to behavioral economics research on belief formation (Coibion and Gorodnichenko, 2015; Bordalo et al., 2020b; Farmer et al., 2023). It measures the expected forecast error associated with a revision made at time t. When $B(n,0) < 0$, an upward revision made today predicts a more negative future forecast error. As mentioned in the introduction, this is typically interpreted as an overreaction to the news because the information that prompted the upward revision results in a disappointing realization of y_{t+n}. By the same token, $B(n,0) > 0$ can be interpreted as underreaction to news relative to FIRE. The literature on belief formation has consistently found $B(n,0) < 0$ at the individual level and has built theories trying to explain this pattern (Bordalo et al., 2020b; Afrouzi et al., 2023).

While the extant literature focuses primarily on the link between *today's* error and today's revision, the main innovation in this paper is to also study the link between *future* errors and today's revision. For example, if a forecaster tends to overreact to the news today, how quickly, if ever, do they realize their mistake? Do they revise their forecast immediately, or does it take time? The remaining points ($j > 0$) along the BIR curve answer these questions.

To see this more formally, first note that the difference between the forecast error at times $t+1$ and t reveals the revision at $t+1$:

$$e_{t+1}^{(n-1)} - e_t^{(n)} = \left[y_{t+n} - F_{t+1}^{(n-1)}\right] - \left[y_{t+n} - F_t^{(n)}\right] = -R_{t+1}^{(n-1)},$$

Or, more generally, $e_{t+j}^{(n-j)} - e_{t+j-1}^{(n-j+1)} = -R_{t+j}^{(n-j)}$. From this identity, it follows that:

$$B(n, j-1) - B(n, j) = \mathbb{E}\left[R_{t+j}^{(n-j)} | R_t^{(n)} = 1\right] - \mathbb{E}\left[R_{t+j}^{(n-j)} | R_t^{(n)} = 0\right], \quad j > 0.$$

This equation says that the difference between two successive points on the BIR curve reveals the expected revision between those dates, following a unit revision made at time t. In other words, because they focus on a fixed forecast target, BIRs naturally embed the entire trajectory of subsequent revisions. Importantly, this feature that would be lost if we instead studied the time-series of one-period-ahead forecast errors (as in, for example, Kučinskas and Peters, 2022). Under FIRE, these revisions should also not be predictable by $R_t^{(n)}$, again implying a perfectly flat BIR.

Stylized Example Figure 1 shows hypothetical BIRs for three different forecasters, all of whom equally overreact at time t. For each forecaster, the initial point on the BIR curve is $B(n, 0) = -1$, indicating that an upward revision at time t reliably leads to a forecast that overshoots by one unit. What differs across the three forecasters is how they correct this overreaction.

The green BIR exemplifies a "stubborn" forecaster who never corrects their initial mistake, resulting in a flat BIR since all expected future revisions are zero. By contrast, the blue BIR reflects a forecaster who quickly undoes the overreaction at time t by revising their forecast down by one unit at time $t+1$, making all subsequent forecast errors zero. This immediate correction is formally encoded in the BIR via $B(n, 0) - B(n, 1) =$

$\mathbb{E}\left[R_{t+1}^{(n-1)}|R_t^{(n)}=1\right] - \mathbb{E}\left[R_{t+1}^{(n-1)}|R_t^{(n)}=0\right] = -1$. Finally, the orange BIR illustrates an intermediate case in which the forecaster gradually corrects their mistake over multiple periods.

As this example illustrates, BIRs capture both the initial degree of over- or underreaction at time t and how subsequent revisions unfold, thereby summarizing the dynamics of belief formation. In turn, the level and shape of a BIR offer a natural set of empirical moments against which to test and develop models of belief formation.

2.2 Estimation and Inference

Estimation BIRs can be easily estimated using any dataset that tracks forecasts of a fixed target through time. Given these datasets often contain multiple series and forecasters, we now extend the baseline BIR model as follows. Let $y_{s,t+n}$ denote the realization of series s at time $t+n$, where s could represent GDP growth or the earnings per share of a publicly listed firm. $F_{i,s,t}^{(n)}$ is the forecast made by individual i at time t of the realization of series s in n periods and $e_{i,s,t}^{(n)} = y_{s,t+n} - F_{i,s,t}^{(n)}$ is the associated forecast error. $R_{i,s,t}^{(n)} = F_{i,s,t}^{(n)} - F_{i,s,t-1}^{(n+1)}$ is the revision made by forecaster i at time t for the realization of series s at $t+n$.

The estimation of BIRs closely aligns with methodologies used in the macroeconomics literature. Specifically, assuming a linear conditional mean function, BIRs can be estimated by running n different regressions of forecast errors on past revisions:

$$e_{i,s,t+j}^{(n-j)} = c_j^{(n)} + \psi_j^{(n)} R_{i,s,t}^{(n)} + v_{i,s,t+j}^{(n-j)}, \quad \forall \quad 0 \leq j \leq n-1, \tag{2}$$

in which case the $B(n,j) = \psi_j^{(n)}$. This estimation strategy is akin to the local projection methods described originally by Jordà (2005) and summarized recently by Jordà (2023). Non-linear conditional mean functions are straightforward to accommodate as well. The alternative to estimating impulse responses with local projections is through a vector autoregression (VAR). As shown by Plagborg-Møller and Wolf (2021), the VAR-based and

local projection approaches deliver the same impulse response in population, though their finite-sample properties may differ. In general, VARs provide more efficient estimates but at the cost of higher bias, partly because they depend on appropriately defining the lag length of the model (Li et al., 2022). With these considerations in mind, we proceed with local projections mainly because they are easier to adapt to panel settings of the kind considered in our empirical work and because our applications use large panels in which precision is less of a concern.

The estimation strategy in regression (2) imposes homogeneity of BIRs across individuals, time, and targets (series). The resulting estimates should therefore be viewed as roughly the average BIR within a given dataset. The assumption of homogeneity can be easily relaxed by estimating BIRs on different subsamples of the data, like for each individual i or series s. We explore BIR heterogeneity later in our empirical work (Section 4).

Inference of Point Estimates Jordà (2023) discusses the main inference issues when estimating local projections using macroeconomic data. As he points out, the residuals of a local projection regression generally have a moving average structure that can affect the construction of standard errors. Consequently, corrections targeting serial correlation in errors, such as the popular Newey-West HAC estimators, are preferable. In our panel context, clustering standard errors within each forecaster-series cell, (i, s), is a simple way to account for these regression error structures, though more efficient solutions are possible. In our empirical work, we also cluster standard errors within each revision date and series cell, (t, s), to account for common information shocks that may induce revisions for a series to be cross-sectionally correlated at a given point in time.

Hypothesis Testing For many applications, it is desirable to test the coefficient restrictions of an entire BIR curve jointly. For instance, testing whether an estimated BIR can reject the null hypothesis of FIRE amounts to a joint test of whether $B(n, 0) = B(n, 1) =$

$... = B(n, n-1) = 0$. As we show in Section 6.2, an important advantage of this joint test is that it offers more power to reject FIRE than testing the single restriction of $B(n, 0) = 0$, as is commonly done in the literature (e.g., Coibion and Gorodnichenko, 2015).

This test can be implemented using a single stacked regression that estimates all $B(n, j)$ in one pass, allowing both intercepts and slopes to vary with j. Here, it is critical to cluster standard errors within each revision date and series cell, (t, s), since revisions will clearly be mechanically correlated along this dimension. Hypothesis testing on the vector $\boldsymbol{\psi} = (\psi_0^{(n)}, \psi_1^{(n)}, ..., \psi_{n-1}^{(n)})^\top$ of BIR coefficients then proceeds in the standard fashion. Testing whether two BIR curves from subsamples of the same overall dataset are equal can be achieved using a similar stacking procedure, then allowing the intercepts and slopes to vary with j and the subsamples of interest (e.g., a particular series).

3 BIRs in Leading Behavioral Models

Every model of belief formation will imply a BIR. In this section, we demonstrate this principle by deriving BIRs for a general class of behavioral models, one that encompasses many of the leading models in the extant literature (e.g., Hirshleifer et al., 2015; Nagel and Xu, 2022; Bordalo et al., 2020b; Afrouzi et al., 2023). Our theoretical results show that BIRs differ sharply across models, underscoring their usefulness for disciplining theories of belief formation.

3.1 General Setup

Throughout this section, we study the forecasting problem for a stochastic process $y_t,\ t \geq 0$. The information set of the forecaster is generated by a set of signals $S_t \in \mathbb{R}^P$, where some of the S_t signals can be lagged versions of y_t. Formally, this information set is described by the filtration $\mathcal{F}_t = \sigma(\{(y_\tau, S_\tau)\}_{\tau \leq t})$ generated by $(y_t, S_t) \in \mathbb{R}^{P+1}$. The stochastic process $Z_t,\ t \geq 0$, is adapted to $\mathcal{F} = \{\mathcal{F}_t\}$ if Z_t is $\mathcal{F}_t$-measurable for each $t \geq 0$. We use

$E_t[y_{t+n}] = E_t[y_{t+n}|\mathcal{F}_t]$ to denote the forecast of y_{t+n} under FIRE and $F_t[y_{t+n}]$ to denote the forecast based on potentially incomplete information or behavioral biases. When there is no ambiguity, we denote multi-period forecasts by $F_t^{(n)} = F_t[y_{t+n}]$ for $n \geq 1$.

To derive BIRs, we focus on a particular class of models in which forecasts depend linearly on the history of rational forecasts. This class, which we label as linear autonomous forecasts, is formally defined as follows.

Definition 2 *Forecasts F_t are defined as linear autonomous if they take the following form:*

$$F_t[y_{t+n}] = \sum_{\tau=0}^{\infty} \gamma(n,\tau) E_{t-\tau}[y_{t+n}] + \nu_t^{(n)} \tag{3}$$

for some coefficients $\gamma(n,\tau)$ and an independent noise process $\nu_t^{(n)}$ with $\text{Cov}(\nu_t^{(n)}, \nu_{t-1}^{(n)}) = 0$ and $\text{Var}[\nu_t^{(n)}] = \sigma_\varepsilon^2(n)$.

The coefficients $\gamma(n,\tau)$ define the loadings of the forecast on current and past rational forecasts. Though it is of course possible to define a broader class of processes in which $F_t[y_{t+n}]$ is an arbitrary function of rational forecasts, we think it is natural to focus on linear functions because they preserve the linearity of forecasts across targets, meaning $F_t[a_1 y_{t+n} + a_2 \tilde{y}_{t+n}] = a_1 F_t[y_{t+n}] + a_2 F_t[\tilde{y}_{t+n}]$.

A key property of linear autonomous forecasts is that they encompass many popular behavioral models of belief formation, including adaptive expectations (Cagan, 1956; Nerlove, 1958), extrapolative expectations (Barberis et al., 2015b; Hirshleifer et al., 2015), noisy information and sticky expectations (Giannoni and Woodford, 2003; Coibion and Gorodnichenko, 2012), diagnostic expectations (Bordalo et al., 2020a), smooth diagnostic expectations (Bianchi et al., 2024a), misestimation of persistence (Gabaix, 2019), and costly information processing (Afrouzi et al., 2023). Appendix A shows how to cast each of these models in the linear autonomous class. As one example, consider the model of diagnostic

expectations from Bordalo et al. (2020a). In this model, forecasts are given by:

$$F_t[y_{t+n}] = E_t[y_{t+n}] + \theta(E_t[y_{t+n}] - E_{t-J}[y_{t+n}]), \tag{4}$$

where the diagnosticity parameter θ modulates the size of the behavioral bias in the forecast, and J controls the number of revisions that drive the bias. $\theta = 0$ corresponds to FIRE and positive values of θ imply forecasts overreact in the CG sense. This model of beliefs clearly fits into the linear autonomous class in Definition 2. Specifically, for $J = 1$, $\gamma(n, 0) = 1 + \theta$, $\gamma(n, 1) = -\theta$, and $\gamma(n, \tau) = 0$ for $\tau > 1$.

3.2 BIRs for Linear Autonomous Forecasts

We now derive BIRs within the class of linear autonomous forecasts. To do so, we make the following assumptions about the data-generating process for y_t.

Assumption 1 *y_t evolves according to:*

$$y_{t+1} \;=\; \beta' S_t \;+\; \varepsilon_{t+1}\,, \tag{5}$$

for some vector $\beta \in \mathbb{R}^P$ and an i.i.d. stochastic process ε_t with zero mean and variance σ_ε^2.

Assumption 2 *The signals S_t evolve according to*

$$S_{t+1} \;= \rho S_t \;+\; \xi_{t+1}, \tag{6}$$

where $|\rho| < 1$ is a scalar that controls the (equal) persistence of signals, and ξ_{t+1} is an i.i.d. stochastic process with zero mean.

Assumption 3 *The process $q_t = \rho^{-1} E_t[y_{t+1}]$ is stationary with a long-run variance σ_q^2.*

Assumption 1 says that the true y_{t+1} is a linear function of lagged signals plus an error term. This assumption is without loss of generality because linearity can always be achieved by expanding the set of signals to include a non-linear transformation of the "true" set of signals (Kelly et al., forthcoming). Moreover, because S_t can include lags of y, Assumption 1 nests cases in which y_t is an arbitrary AR(q) process. Assumption 2 is stronger and puts a simple dynamic structure on the signals that inform each forecast. Combined with Assumption 1, it implies that the FIRE-forecast can be written as: $E_t[y_{t+n}] = \rho^{n-1} E_t[y_{t+1}] = \rho^n q_t$.[8]

Assumption 3 then imposes that this rational forecast is stationary. Finally, we make one additional assumption about the weight function $\gamma(n, \tau)$ to make the analysis of BIRs more tractable.

Assumption 4 (Separability) *The weight function $\gamma(n, \tau)$ that defines the belief formation process can be written as:*

$$\gamma(n, \tau) = k_0(n)\, \mathbf{1}_{\tau=0} + \delta(n)\kappa(\tau), \tag{7}$$

where, without loss of generality, we impose the normalization $\sum_{\tau=0}^{\infty} \kappa(\tau)\rho^{2\tau} = 1$.

The notion of separability in Assumption 4 is satisfied in all of the behavioral models discussed above. For example, continuing with the earlier example of diagnostic expectations where $J = 1$, we have $\kappa(0) = (1-\rho^2)^{-1}$, $\kappa(1) = -(1-\rho^2)^{-1}$, and $\kappa(\tau) = 0$ for $\tau > 1$. Moreover, $\delta(n) = \theta(1-\rho^2)$ and $k_0(n) = 1$. Under Assumptions 1-4, we can now derive the paper's main proposition regarding the shape of BIRs in the linear autonomous class of belief formation models.

Proposition 1 (BIRs in the Linear Autonomous Class) *Under Assumptions 1-4, BIRs*

[8] Under a more general structure in which $S_{t+1} = \Psi S_t + \xi_{t+1}$, it is straightforward to show that forecasts under FIRE are given by $E_t[y_{t+n}] = \beta'(\Psi)^{n-1} S_t$.

are given by:

$$B(n,j) \;=\; \frac{-\sigma_\varepsilon^2(n)\mathbf{1}_{j=0} + A(n,j) - \rho^2 A(n+1,j+1)}{D(n)}, \tag{8}$$

where

$$\begin{aligned}
A(n,j) &= (1-k_0(n-j))(k_0(n)+\delta(n)) \;-\; \delta(n-j)\Big[k_0(n)\Gamma(j)+\delta(n)\Xi(j)\Big] \\
D(n) &= D_{MZ}(n) + \rho^2 D_{MZ}(n+1) \\
&\quad - 2\rho^2(k_0(n+1)k_0(n) + k_0(n)\delta(n+1) + k_0(n+1)\delta(n)\Gamma(1) + \delta(n)\delta(n+1)\Xi(1)) \\
D_{MZ}(n) &= k_0(n)^2 + 2k_0(n)\delta(n) + \delta(n)^2\Xi(0) \\
\Xi(j) &= \sum_{\tau=0}^{j}\kappa(\tau) \;+\; \sum_{\tau_1=1}^{\infty}\sum_{\tau_2=0}^{\infty}\kappa(\tau_1+j)\kappa(\tau_2)\rho^{2\max(\tau_1,\tau_2)} \\
\Gamma(j) &= \sum_{\tau=0}^{j}\kappa(\tau) + \sum_{\tau=1}^{\infty}\kappa(\tau+j)\rho^{2\tau}
\end{aligned} \tag{9}$$

Proposition 1 shows that BIRs are a complex function of $\kappa(\tau)$ and $\delta(n)$, the two objects that define the linear autonomous class of forecasts. To develop more intuition for Proposition 1, we now use it to solve for the BIR in several leading behavioral models.

3.3 BIRs in Leading Behavioral Models

3.3.1 Diagnostic Expectations

To start, we continue with the example of diagnostic expectations from Bordalo et al. (2020a). In this case, Proposition 1 implies that the BIR is as follows.

Corollary 2 *In the model of diagnostic expectations (Bordalo et al., 2020a) given by Equation (4), the BIR for $J = 1$ and diagnosticity parameter θ equals:*

$$B(n,j) \;= -\frac{\theta(1+\theta)}{1+\theta(2+\theta(1+\rho^2))}\cdot\mathbf{1}_{j=0} \tag{10}$$

Figure 2a visualizes the implications of Corollary 2 by plotting BIRs for varying diagnosticity and signal persistence parameters. The diagnosticity parameters are set to $\theta \in \{0.5, 0.9\}$ and are based on previous estimates from Bordalo et al. (2019) and Bordalo et al. (2020a). For signal persistence, we explore two extremes, setting $\rho \in \{0.1, 0.9\}$.

By design, diagnostic expectations leads to overreaction at time t, as evidenced by the fact that $B(n, 0) < 0$ for all parameter values in Figure 2a. As we show later, this initial level of the diagnostic-BIR closely matches what we observe in the SPF and IBES data—an unsurprising result given that the model was originally designed with the SPF in mind. Moreover, for a fixed θ, a lower signal persistence ρ leads to a more pronounced overreaction, a fact that is directionally consistent with the empirical results in Bordalo et al. (2020a), and Afrouzi et al. (2023).

The most striking aspect of Figure 2a is the discontinuity in the diagnostic-BIR, which indicates that the diagnostic forecaster "wakes up" abruptly by $t+1$ and immediately corrects its time-t mistake. To understand the intuition for this result, notice that for $J = 1$, the n-period forecast at time t depends on the rational forecast and the rational revision: $F_t^{(n)} = \mathbb{E}_t[y_{t+n}] + \theta(\mathbb{E}_t[y_{t+n}] - \mathbb{E}_{t-1}[y_{t+n}])$. In turn, the error associated with the forecast made at $t+1$ depends on both the rational forecast error and revision:

$$\begin{aligned} e_{t+1}^{(n-1)} &= y_{t+n} - F_{t+1}^{(n-1)} = y_{t+n} - \Big\{\mathbb{E}_t[y_{t+n}] + \theta(\mathbb{E}_{t+1}[y_{t+n}] - \mathbb{E}_t[y_{t+n}])\Big\} \\ &= \underbrace{(y_{t+n} - \mathbb{E}_{t+1}[y_{t+n}])}_{\text{Rational Forecast Error}} - \theta \underbrace{(\mathbb{E}_{t+1}[y_{t+n}] - \mathbb{E}_t[y_{t+n}])}_{\text{Rational Forecast Revision}}. \end{aligned}$$

By definition, both components are unpredictable by *any* variable at time t, including the diagnostic time-t revision $R_t^{(n)}$, hence why the BIR goes to zero immediately. The abrupt discontinuity in the model-implied BIR turns out to be a more general feature of theories in which forecasts depend on finite lags of rational revisions, including the model of smooth

diagnostic expectations in Bianchi et al. (2024a). We demonstrate this point formally via the following proposition.

Proposition 3 *Suppose that $F_t[y_{t+n}]$ depends only on revisions in rational forecasts between t and $t-J$ (i.e., $\kappa(\tau) = 0 \quad \forall \tau > J$ and $\sum_{\tau=0}^{J} \kappa(\tau) = 0$). Then, for any $j \geq J$, $B(n, j) = 0$.*

Proposition 3 demonstrates that models like diagnostic and smooth diagnostic expectations rely on an abrupt correction mechanism. Specifically, these models imply that forecasters, given a lag structure J, fully rectify any overreaction at t by time $t+J$, resulting in a discontinuity in the model-implied BIR at $B(n, J)$. The purple line in Figure 2a illustrates this result by plotting the BIR for $J = 3$, $\theta = 0.9$, and $\rho = 0.9$. The flat and negative BIR for $j = 0, 1, 2$ indicates an initial overreaction to news followed by no correction for two periods. At $j = 3$, diagnostic forecasters then "wake up" and fully reverse their overreaction.

3.3.2 Costly Information Processing

Next, we consider the belief formation process proposed by Afrouzi et al. (2023), who develop a model to accommodate the fact that overreaction to contemporaneous news is larger for longer-horizon forecasts and series with less persistence, both in experimental and observational data. Afrouzi et al. (2023) focus on the special case in which y_t is an AR(1) process with mean μ and persistence parameter ρ. We show in Appendix A that this simple version of the model is a part of the linear autonomous class. Applying Proposition 1 to this model yields the following BIR.

Corollary 4 *Suppose $y_{t+1} = (1-\rho)\mu + \rho y_t + \varepsilon_t$ and that forecasts follow the model of costly information processing in Afrouzi et al. (2023), whereby:*

$$F_t[y_{t+n}] \;=\; const(n) \;+\; E_t[y_{t+n}] \;+\; \delta(n) E_t[y_{t+n}], \tag{11}$$

with

$$\delta(n) \;=\; (1-\rho^n)\rho^{-n}\min\left\{1,\;\left(\frac{\omega\underline{\tau}}{(1-\rho^n)^2}\right)^{1/(1+\gamma)}\right\} \tag{12}$$

for some $\omega, \underline{\tau}, \gamma \geq 0$. *Then, the BIR is given by:*

$$B(n,j) \;=\; A(n,j)\frac{1-\rho^2}{D(n)} \tag{13}$$

where

$$\begin{aligned} A(n,j) \;&=\; -\delta(n-j)\Big[1+\delta(n)\Big] \\ D_{MZ}(n) \;&=\; (1+\delta(n))^2 \\ D(n) \;&=\; D_{MZ}(n)+\rho^2 D_{MZ}(n+1) \;-\; 2\rho^2(1+\delta(n+1))(1+\delta(n))\,. \end{aligned} \tag{14}$$

Moreover, $B(n,j)$ *is monotone increasing and concave in* j.

Figure 2b uses Corollary 4 to plot BIRs under two parameterizations from Afrouzi et al. (2023): one with $\rho = 0.2$ and another with $\rho = 0.95$. These choices roughly mirror the average AR(1) coefficients we estimate for the low- and high-persistence SPF series in Internet Appendix IA.3.4. We fix the variance of ε_t at $\sigma_\varepsilon^2 = 0.2$ in both cases and set the scale and convexity of information-processing costs to $\omega = 1$ and $\gamma = 2$, respectively. Following Section 6.2 of Afrouzi et al. (2023), the baseline precision $\underline{\tau}$ that agents have about the long-run mean of y equals $(1-\rho^2)/\sigma_\varepsilon^2$.

In their empirical analysis, Afrouzi et al. (2023) document that overreaction is much larger for more persistent series, both in observational (i.e., the SPF) and experimental data. As expected, the blue and orange BIRs in Figure 2b show that their model of costly information processing can replicate this empirical pattern. Unlike in diagnostic expectations, however, the correction process here is smooth, not abrupt.

Figure 2b further demonstrates that, under costly information processing, corrections occur relatively quickly for low-persistence series (blue BIR) and more slowly for high-persistence series (orange BIR). In other words, the model's correction speed appears tightly linked to $B(0)$ through series persistence. Corollary 4 also highlights another important feature of costly information processing: the model-implied BIR is concave in j. After the initial overreaction, forecasters gradually adjust their forecasts at a decelerating rate, which produces the concavity evident in the figure.

3.3.3 Extrapolative Expectations

Lastly, we consider a model in which forecasters extrapolate from past realizations of y. The specific setup we consider largely draws on Hirshleifer et al. (2015) while also bearing similarities to Nagel and Xu (2022). Unlike Barberis et al. (2015b), who develop an asset pricing model centered on price extrapolation, both Hirshleifer et al. (2015) and Nagel and Xu (2022) focus on the extrapolation of past fundamentals. We follow both papers in assuming that the true data generating process has a constant mean, so that $y_{t+1} = \bar{y} + \varepsilon_{t+1}$, where ε_{t+1} is i.i.d.[9] In addition, forecasts are assumed to evolve according to:

$$\begin{aligned} F_t[y_{t+1}] &= \mu_t \\ \mu_{t+1} &= (1-\rho-\tilde{\rho})\bar{\mu} + \rho\mu_t + \tilde{\rho}y_{t+1}, \end{aligned} \tag{15}$$

where $\tilde{\rho}$ modulates the degree of extrapolation from the most recent realization of y, ρ controls the impact of extrapolation made from past realizations of y, and $\bar{\mu}$ is the forecasters perceived long-run mean. Hirshleifer et al. (2015) provide several possible microfoundations for this particular belief formation process, like, for instance, if the forecaster perceives a time-varying growth rate for y. Applying Proposition 1 to this model then yields the following BIR.

[9] In Appendix C, we allow y_t to follow an autoregressive process, so that $E_t[y_{t+1}] = \bar{y} + \rho_y(y_t - \bar{y})$. The expressions for BIRs are more involved in this case.

Corollary 5 (BIRs for extrapolative expectations) *In the model of extrapolative expectations given by* (15)*, the BIR is given by:*

$$B(n,j) \;=\; -\frac{\rho^j}{(\rho+\tilde{\rho})^j}\frac{1-\rho(\rho+\tilde{\rho})}{(1+(\rho+\tilde{\rho})^2-2\rho(\rho+\tilde{\rho}))} \;<\; 0\,. \tag{16}$$

In particular, $B(n,j)$ *is independent of* n*, and is increasing and concave in* j*.*

Consistent with Hirshleifer et al. (2015) and Nagel and Xu (2022), Corollary 5 shows that extrapolative expectations can indeed generate overreaction ($B(n,0) < 0$) at time t. However, a key feature of this framework is that the level of the BIR does not depend on the forecasting horizon n. As for the correction process, the corollary indicates that extrapolative expectations converge and correct smoothly to rational forecasts. As with costly information processing, this process is concave in j.

To better visualize the BIR implied by extrapolative expectations, Figure 2c plots the BIR using the following parameters: $n = 6$, $\bar{\mu} = 0$, $\tilde{\rho} = 0.018$, and $\rho = 1 - \tilde{\rho}$. Under this parameterization, the model of extrapolative expectations in (15) reduces to the one considered by Nagel and Xu (2022), who choose the value of $\tilde{\rho}$ based on the analysis of inflation expectations in Malmendier and Nagel (2015). The initial level of the BIR is negative and indicates overreaction at time t. Moreover, under this parameterization, the pace with which extrapolative agents correct their mistakes is extremely slow—essentially all of the initial overreaction remains even after five quarters.

4 BIRs in Observational Data

In this section, we estimate BIRs using forecasts from two widely used datasets in macroeconomics and finance: the SPF and IBES. Our analysis reveals five key facts about the dynamics of belief formation. First, consistent with prior research, individual forecasts tend to overreact in the sense of Coibion and Gorodnichenko (2015), especially for series with

low persistence (Bordalo et al., 2020b; Afrouzi et al., 2023). Second, these overreactions are corrected smoothly and gradually. Third, the speed of correction varies considerably across both forecasters and series, and is not tightly linked to the size of the initial over- or underreaction. Fourth, corrections occur more quickly when forecasters receive news that contradicts their initial revision. Fifth, BIRs tend to converge toward the FIRE benchmark as forecasters gain experience. We conclude by discussing the implications of these findings for leading behavioral models of belief formation.

4.1 Data

The SPF is a quarterly survey administered by the Federal Reserve Bank of Philadelphia that collects forecasts from professional forecasters on a range of macroeconomic and financial outcomes (e.g., nominal GDP). Our analysis includes the fifteen series listed in Table A1 of the Internet Appendix and largely follows Bordalo et al. (2020b). SPF forecasts extend up to four quarters ahead, which means that BIRs can be estimated for horizons up to $n = 3$.

IBES provides analyst-by-firm forecasts of quarterly, semi-annual, annual, and long-term EPS. We focus on quarterly EPS forecasts, as they are more densely populated at longer horizons and allow for a more direct comparison with the SPF. While IBES includes forecast horizons up to ten quarters ahead, the data become sparse at the longest horizons. As a result, we focus on BIRs with $n = 6$ in our baseline analysis.

Both datasets are cleaned and processed following standard practices from the literature. These steps are detailed in Internet Appendix IA.2, where we also define forecasted growth rates and forecast errors, describe how we handle outliers, and explain how we construct our main analysis samples.

Earnings forecasts may be influenced by strategic motives, such as currying favors with management or generating trading volume, rather than reflecting analysts' pure beliefs (e.g., Francis and Philbrick, 1993; Jackson, 2005). To limit the influence of such considerations on

our interpretation of BIRs, we focus on properties that are robust across both datasets, since strategic incentives are arguably less relevant for SPF forecasters. At the end of this section, we also provide robustness checks showing that our main conclusions hold in settings where strategic behavior is less likely— for example, in GDP forecasts from the Congressional Budget Office (CBO).

4.2 Aggregate BIRs

Figure 3 plots BIRs estimated from the SPF and IBES datasets, pooling forecasts across all series and forecasters. Loosely speaking, these BIRs reflect the behavior of the average forecaster in each dataset. Ninety-five percent confidence bands are based on standard errors clustered by forecaster-series, (i, s), and series-date, (s, t). Table 1 reports corresponding point estimates, standard errors, and sample sizes. Shorter-horizon BIRs are presented in Figure A1 of the Internet Appendix.

Individual forecasts in both datasets exhibit overreaction, meaning $B(n, 0) < 0$. This finding aligns with Bordalo et al. (2020a) and Afrouzi et al. (2023) for SPF.[10] At first glance, overreaction in IBES may seem at odds with Bouchaud et al. (2019), who document underreaction in consensus forecasts for 1-year-ahead forecasts of annual EPS. However, we show in Internet Appendix IA.3.7.1 that the difference is driven by their use of consensus forecasts, annual versus quarterly forecasts, and a shorter forecast horizon.

The novel content of BIRs lies in its shape, which reveals how overreaction corrects over time. To quantify correction speed, we define

$$W(n, j) = \frac{B(n, j)}{B(n, 0)}, \tag{17}$$

which equals 1 if no correction has occurred by $t + j$, and 0 if the initial overreaction has

[10] Our estimate of $B(3, 0)$ for the SPF is not comparable to any single Coibion and Gorodnichenko (2015) regression coefficient reported in Bordalo et al. (2020b) because we use slightly different definition of growth rates and pool forecasts across all series (see Internet Appendix IA.2.1).

fully reversed. In the SPF, $W(3,1) \approx 40\%$, indicating that 60% of the overreaction in the time-t forecast is corrected within one quarter, with little additional correction between $t+1$ and $t+2$. IBES corrections are much slower: $W(n,6) \approx 80\%$, with a half-life exceeding one year.

These slow corrections support Bordalo et al. (2024), who argue that stock return predictability stems from overreaction in forecasts of future cash flows, rather than time-varying discount rates (Campbell and Cochrane, 1999; Bansal and Yaron, 2004). In their framework, positive news generates overly optimistic earnings expectations and inflated stock prices. As actual earnings disappoint, expectations correct gradually, producing predictably low returns. Figure 3b offers direct evidence of this critical correction mechanism.[11]

Table 1 reports joint tests of FIRE. The null $B(n,j) = 0$ for all j is decisively rejected with p-values below 0.001 in both datasets. A weaker test of whether all BIR points are equal—equivalent to testing forecast revision predictability—is also rejected, indicating that forecasters gradually correct their deviations from FIRE rather than remaining "stubborn."

4.3 What Factors Influence Correction Speed?

We now investigate factors that influence how quickly forecasters correct their initial deviations from FIRE.

4.3.1 Forecaster Identity

We begin by studying the role of forecaster identity in shaping corrections. The basic approach divides forecasters in SPF into subgroups based on whether they overreact and how quickly they correct, measured by their individual $W(3,1)$. Overreacting forecasters

[11] Bordalo et al. (2024) focus primarily on forecasts of long-term growth (LTG) for S&P 500 firms. BIRs are harder to estimate for LTG forecasts because they have an ambiguous horizon, ranging from 3-5 years. Similar to BLGS, we find overreaction in longer-horizon individual forecasts, i.e., $B(6,0) < 0$. In Internet Appendix IA.3.7.3, we also confirm similar BIRs for large stocks, like those studied by BLGS.

are divided into terciles based on their $W(3,1)$, while underreactors are split into two groups to maintain balanced sample sizes.

For completeness, Table 2a presents summary statistics on how $B(0)$ varies across individual forecasters. The median SPF forecaster overreacts with a CG-coefficient of -0.27, and nearly three-quarters overreact. Table 2b then shows estimates of $W(3,j)$ based on pooled BIRs from each subgroup.

The variation in correction speeds is striking. The fast-correcting overreactors (first column in Table 2b) reverse nearly all overreaction within one quarter and fully eliminate it within two. In contrast, the slow-correcting overreactors (third column) retain 113% ($t = 9.40$) of their initial overreaction after one quarter and 90% ($t = 8.48$) after two quarters. Figure 4a visualizes these results by plotting BIRs for the slowest and fastest correcting forecasters.

Table 2c repeats this analysis for IBES analysts, sorting again by the sign of individual $B(6,0)$ and then dividing further into quartiles based on $W(6,1)$. The patterns are similar but even more pronounced: among analysts who overreact and correct slowly, an average of 117% ($t = 6.09$) of the overreaction at time t persists in forecasts made a year later. In contrast, for fast-correcting analysts, only 21% ($t = 2.90$) of the initial overreaction remains after one year. Figure 4b illustrates how this variation rise to both concave and convex BIRs across IBES analysts.

4.3.2 The Data-Generating Process

We next analyze whether the nature of the forecast target influences correction speed using the same methodology. Table 3a shows that twelve of the fifteen series in SPF overreact, consistent with the findings of Bordalo et al. (2020a).[12] Table 3b sorts series in SPF by the sign of their $B(3,0)$, then divides overreacting series into terciles based on $W(3,1)$.

[12]The three series that exhibit initial underreaction are the T-Bill yield, growth in GDP Price Index (PGDP), and growth in housing starts (HOUSING)

Each column reports estimates of $W(3, j)$ from a BIR that is pooled across series in a given subgroup.

Correction speeds vary meaningfully across series in the SPF. Fast-correcting series fully eliminate overreaction within one quarter, while slow-correcting series retain about half after two quarters, despite similar degress of initial overreaction. Figure 5a visualizes these differences between the two groups.

IBES exhibits even greater heterogeneity across stocks. Following the same approach, Table 3c sorts stocks based on the sign of $B(6, 0)$ and further divides based on $W(6, 1)$. Among overreacting stocks, the slowest quartile retains nearly all initial overreaction after one year, while the fastest quartile corrects almost fully within one quarter. A similarly wide range of correction dynamics is observed among underreacting series. Figure 5b illustrates this large variation for the slowest and fastest correcting stocks: nearly 20% of stocks of stocks in IBES overreact and correct so slowly that their BIRs are convex, while a similar fraction of series overreact but correct quickly enough to generate concave BIRs.

4.3.3 The Size of the Initial Reaction

Next, we examine the relationship between the size of the initial deviation from FIRE and the subsequent speed of correction. This link is important from the perspective of behavioral theories. If, for instance, forecaster correction speed were fully determined by initial deviation size—perhaps because large errors are more obvious and thus corrected faster—models of individual behavior could use a single parameter for both. If these vary independently, multiple behavioral parameters are needed. A similar logic applies when considering variation in correction speeds across series.

The results in Table 2 provide early evidence that the variation in correction speeds across forecasters is not fully explained by the magnitude of the initial deviation from FIRE. In Table 2b, overreacting SPF forecasters have similar $B(0)$ values but vastly different correction

speeds. Similarly, overreactive IBES analysts in the bottom and top quartiles of $W(6,1)$ have comparable levels of initial overreaction but strongly differ in correction speed. These two groups are not small, representing 32% of all analysts in IBES.

Series-level patterns tell a similar story. For series in the SPF, Table 3b shows virtually identical overreaction across the first three columns, yet correction speeds vary significantly. Formal tests reject the null hypothesis of equal correction speeds, as measured by $W(3,1)$, across these groups.[13] In Table 3c, stocks in IBES show some positive relationship between $B(6,0)$ and $W(6,1)$, but it is not monotonic.

Figure 5c provides the cleanest test using double sorts that hold $B(6,0)$ roughly constant across stocks while varying $W(6,1)$. We first divide overreacting stocks into two groups based on the magnitude of $B(6,0)$, then within each group, we further split stocks into fast and slow correctors based on $W(6,1)$. We repeat this process for underreacting stocks, yielding eight groups in total. Stocks with similar initial reactions are grouped by color, while those with slow (fast) corrections are depicted by a dashed (solid) line. Among the most overreacting stocks in blue, fast correctors retains only 42% of their initial overreaction versus 80% for the slow correctors ($p = 0.005$ when testing the difference), despite statistically indistinguishable initial reactions.

Similar patterns emerge across the other groups in the figure. One striking example of the weak link between $B(6,0)$ and $W(6,1)$ can be seen by comparing the dashed blue and dashed orange lines. Overreaction at time t is over ten times larger for the dashed blue group, yet the pace of correction is virtually identical across the two groups. We further confirm the relatively weak link between $B(n,0)$ and $W(n,1)$ at the series level for both SPF and IBES in Internet Appendix IA.3.4. There, we also explore additional factors that may influence correction speed, including the volatility and persistence of the underlying series.

[13] The null of equal $W(3,1)$ for the first and third groups is also rejected, as is the null of equal $W(3,2)$ across all three groups.

4.3.4 New Information: The Role of Reinforcing versus Contradicting Signals

Lastly, we ask how the nature of subsequent information shapes correction dynamics. To understand our approach, consider two forecasters who both overreact when revising their forecast of y_{t+6} upward. That is, both make positive revisions ($R_t^{(6)} > 0$) that generate negative forecast errors ($e_t^{(6)} < 0$). The forecasters differ only in their subsequent news: at $t+1$, forecaster A observes $e_t^{(1)} < 0$, which contradicts her initial optimism, while forecaster B observes $e_t^{(1)} > 0$, which reinforces his prior revision.

Figure 6 implements this comparison using the IBES data, which offers a sufficiently large sample size to support this type of conditioning. Figure 6a, labeled $\tau = 1$, splits observations based on the sign of the one-period-ahead forecast error at $t+1$. We restrict attention to cases where $R_t^{(6)}$ and $e_t^{(6)}$ have opposite signs (i.e., clear overreaction), and we retain only the middle three quintiles of both distributions to ensure comparable $B(0)$ across both groups. The results show a stark asymmetry: forecasters receiving contradictory signals begin correcting almost immediately, eliminating about 20% of the initial overreaction within one period. In contrast, those receiving reinforcing signals show virtually no correction through period three.

Figures 6b-6d extend this analysis to multiple periods of consistent signals. The asymmetry becomes even more pronounced. For example, in Figure 6b, forecasters receiving two consecutive contradicting signals eliminate roughly 40% of their initial error by period two, while those receiving reinforcing signals show minimal adjustment. In Figure 6d, after four consecutive reinforcing signals, later points on the BIR are more negative than the initial point, suggesting that overreaction worsens for these forecasters. Only at period five, when we no longer condition on news, is there a sign of any correction.

Taken together, these patterns provide strong evidence that corrections are triggered by the nature of incoming information: contradicting news prompts belief revision, while reinforcing news appears to entrench initial errors.

4.4 Experience

Given the large and heterogeneous deviations from FIRE documented in Figure 4, a natural question is whether BIRs are a static individual traits or evolve over time. BIRs offer a simple way to study this question because they can be estimated over different phases of an analyst's career. Figure 7 implements this idea for analysts with at least one decade for forecasts in either SPF (Figure 7a) or IBES (Figure 7b). The threshold of ten years is roughly the median experience in both datasets.

For a given forecaster i, we compute BIRs using the first five years of forecasts, requiring at least ten observations to be included. B_i^{early} is the average level of forecaster i's BIR during this "inexperienced" phase of their career. Forecasters are then classified into three terciles based on B_i^{early}. This means that the most overreacting forecasters are in the bottom tercile and the underreacting are in the top. Within each tercile, we then estimate two pooled BIRs: one using data during the first five years of each forecaster's career and another based on forecasts for all subsequent years when they are more experienced. The figure shows the resulting BIRs from both career phases for the group of forecasters who overreacted the most early in their careers. In the Internet Appendix Section IA.3.7.4, we document similar results for the smaller set of underreactors.

Figure 7 reveals a robust relationship between BIRs and experience in both the SPF and IBES datasets: forecasters appear to become more rational as their careers progress. In SPF, overreaction is fairly large and relatively slow correcting for inexperienced forecasters, with nearly two-thirds of their overreaction at time t still present in forecasts made one quarter later. For experienced forecasters in SPF, overreaction in time-t forecasts is economically small and statistically insignificant, and any overreaction is fully corrected within one period. Figure 7b paints a similar picture for equity analysts in IBES, confirming that the findings

from the SPF are not driven by any issues with how individual contributors are tracked in that dataset.[14]

4.5 Robustness

Strategic Considerations A potential concern with our empirical analysis thus far is that forecasts may reflect strategic rather than belief-based motivations, particularly in IBES where analysts may face incentives to curry favors with management or generate trading volume for affiliated investment banks (e.g., Francis and Philbrick, 1993; Jackson, 2005).[15] We address this in three ways. First, our main analysis largely emphasizes patterns that common acorss SPF and IBES, since strategic incentives are likely less relevant for SPF. Second, Figure 8 shows that GDP forecasts from the Congressional Budget Office—which faces minimal strategic pressure—exhibit similar overreaction and correction dynamics (see Internet Appendix IA.2.3 for details). Third, IBES results are robust to restricting the sample to post-Regulation FD data after October 2000, when analysts incentives to distort forecasts were arguably reduced (Internet Appendix Figure A11).

Selection Heterogeneity in correction speeds could reflect forecaster specialization rather than genuine individual differences. If forecasters cover different series and correction speeds depend only on the properties of the DGP, observed individual variation might be spurious. However, restricting to SPF forecasters who cover virtually all series yields similar heterogeneity patterns, both across forecasters and series. This suggests the variation documented in Table 2 and Figure 4 likely reflects genuine differences across forecasters, while the variation in Table 3 and Figure 5 reflects underlying differences in the DGPs.

[14] For a complete discussion of these issues, see the SPF documentation.

[15] Of course, forecasters also presumably face strong career incentives to produce accurate forecasts, which may counterbalance these pressures.

Heterogeneity vs New Information The heterogeneity in correction speeds documented in Sections 4.3.1 and 4.3.2 could also be confounded if certain forecasters or series systematically receive different types of news following initial revisions. To address this concern, we replicate our forecaster- and series-level analyses on a restricted sample where subsequent information shocks are minimal. We proxy for minimal news using small one-period-ahead forecast errors following the initial revision. Internet Appendix Figure A5 indicates that correction speeds vary meaningfully across both forecasters and series in this minimal-news sample, and remain largely untethered to the initial reaction. These robust patterns suggest suggest variation in correction speeds across forecasters, series, and initial reactions do not reflect systematic differences in new information.

4.6 Implications for Behavioral Theories

To summarize, our empirical analysis in this section highlights the following five facts about belief formation:

1. Most individual forecasts overreact in the CG sense, especially for series with low persistence.

2. On average, deviations from FIRE are corrected smoothly and gradually over time.

3. Correction speeds vary widely across forecasters and series, and are not tightly linked to the size of the initial over- or underreaction.

4. Corrections occur more quickly when forecasters receive news that contradicts their initial revision, and more slowly when the news reinforces it.

5. BIRs tend to converge to FIRE as forecasters gain experience.

What do these facts imply for behavioral models of belief formation? As noted earlier, Fact 1 is not new to our paper and has been well-established in the literature (Bordalo et

al., 2020a; Afrouzi et al., 2023). Accordingly, most leading behavioral models were designed with the goal of matching this feature of the data. Indeed, our theoretical analysis in Section 3.3 confirms that many of these models do so successfully in a quantitative sense.

Facts 2 and 3, however, present greater challenges. Many leading models predict correction speeds that are either too abrupt (Bordalo et al., 2019; Bianchi et al., 2024a), too slow (Hirshleifer et al., 2015; Nagel and Xu, 2022), or too tightly linked to the size of the initial deviation from FIRE (Afrouzi et al., 2023). These findings suggest that the next generation of behavioral models may benefit from additional flexibility to generate smooth correction dynamics where initial reaction size and correction speed can vary independently. This need not be viewed as a limitation of existing work, as these models were intentionally designed with parsimony in mind. Fact 4 is broadly consistent with a range of belief formation models, both rational and behavioral, though it rules out extreme forms of inattention or rigidity. Finally, Fact 5 is difficult to reconcile with purely static behavioral biases and suggests a potential role for learning. We explore this possibility in the next section.

5 BIRs and Model Uncertainty

In this section, we demonstrate how BIRs can be paired with simulations to explore models of belief formation that lie outside of the linear autonomous class studied in Section 3. We specifically consider belief formation for Bayesian forecasters in the presence of model uncertainty (e.g., Timmermann (1993)). Our specific setup mirrors that of Farmer et al. (2023), henceforth FNS. Using their formulation of model uncertainty, which operates through an unobserved components model, we characterize the conditions under which a Bayesian forecaster can generate BIRs that are observed in the SPF and IBES datasets.

5.1 Setup

Farmer et al. (2023) study belief formation about the following data generating process (DGP) for y_t:

$$\begin{aligned} y_t &= \mu_t + x_t, \\ \mu_t &= \mu_{t-1} + \sqrt{\gamma}\sigma\eta_t, \quad \eta_t \sim N(0,1) \\ x_t &= \rho x_{t-1} + \sqrt{1-\gamma}\sigma\omega_t, \quad \omega_t \sim N(0,1) \end{aligned} \tag{17}$$

In this model, the target series y_t is the sum of a random walk, μ_t, and an AR(1) process, x_t, both of which are unobserved. The parameters governing y_t are ρ, σ, and γ. They respectively control the persistence of x_t, the conditional variance of y_t, and the share of this variance attributable to μ_t shocks. Like the states, x_t and μ_t, these parameters are also unknown to forecasters, who only observe quarterly realizations of y_t.

FNS endow forecasters with beliefs about the initial states, μ_1 and x_1, which we specify below. In addition, forecasters have the following initial priors about the parameters governing y_t:

$$\rho \sim N(\mu_p, \sigma_p^2), \quad \gamma \sim \mathcal{B}(\alpha_\gamma, \beta_\gamma), \quad \sigma^2 \sim \mathcal{IG}(1.25, 0.5625), \tag{18}$$

where $\mathcal{B}$ is the beta distribution and $\mathcal{IG}$ is the inverse-gamma distribution. We define $\Theta = (\mu_p, \sigma_p^2, \alpha_\gamma, \beta_\gamma)$ as the parameter vector that defines the forecasters' initial priors about ρ and γ. As in FNS, the priors for σ^2 are held fixed throughout our analysis.

In the latter part of their paper, FNS simulate y_t for fixed values of ρ, γ, and σ and then study how the resulting time-series of forecasts depends on initial priors Θ. Using their replication kit, we conduct a similar exercise to explore the conditions under which Bayesian forecasters in this setting will generate BIRs that resemble those observed in the

SPF and IBES. Our simulation proceeds as follows. First, we simulate five different paths of y_t according to the DGP in (17), each of length T. For the SPF, we set $T = 190$ and for IBES, $T = 260$, corresponding to the length of the respective samples in the data. For SPF forecasts, we set the true values $\rho = 0.02$ and $\gamma = 0.02$; for IBES, we set $\rho = 0.5$ and $\gamma = 0.25$. In both cases, we fix $\sigma = 0.5$. As we discuss below, these parameters were chosen so that the simulated y_t approximates the average realized behavior of series in the SPF and IBES datasets, at least in terms of persistence.

Next, we set forecasters' beliefs about the initial states to be $\mu_1 \sim N(y_1, 1)$ and $x_1 \sim N(0, 1)$, such that the two are perfectly negatively correlated given the DGP in Equation (17). Armed with their initial beliefs about the states (μ_1, x_1) and the parameters of the DGP contained in Θ, the forecasters observe simulated realizations of y_t and update their beliefs according to Bayes law. After ten years of realizations, they then form forecasts of y_{t+n} in each quarter t based on their beliefs at that time. This update sequence exactly follows FNS and generates a panel of forecasts for each simulated sample of y_t. For each sample of simulated forecasts, we estimate a BIR, denoted by $\tilde{B}(n, j; \Theta)$. We use this notation to reiterate the dependence of the forecasts, and hence the BIR, on the initial beliefs Θ.

We then search for the initial belief vector Θ to best match the average individual BIR observed in the SPF or IBES datasets, as plotted in Figures 3a and 3b, respectively. That is, we select Θ to minimize $\sum_j [B(n, j) - \tilde{B}(n, j; \Theta)]^2$. As in the data, we use $n = 3$ for the SPF and $n = 6$ for the IBES BIRs. For each optimal Θ^*, we then simulate $n = 100$ different samples of y and estimate the BIRs as described above for the given initial beliefs Θ^*. Finally, we compute the average BIR across the n samples, along with confidence bands based on the standard deviation of each point on the BIR divided by $\sqrt{n}$.

5.2 Main Findings

Figure 9a plots the resulting simulated BIR along with the true one from the SPF. The figure shows that the model fits the data rather well. The initial levels of the BIRs are nearly identical, as is the speed of correction. Figure 9c shows the distribution of initial beliefs, governed by Θ^*, and the true parameter values (vertical dotted line). As in FNS, overreaction occurs when beliefs about ρ and γ are biased upward. The true ρ and γ equal 0.02, whereas agents in the model have priors for ρ and γ that are centered around 0.45 and 0.14, respectively.

Figure 9b presents results for IBES. The model captures the initial level of overreaction and generates slow corrections: 80% of the overreaction from time t persists after one quarter, closely matching the data. However, the model corrects too quickly thereafter, generating an additional 25 percentage points of correction from $t+1$ to $t+2$, whereas the data show virtually none. Unlike SPF, overreaction here stems from downward-biased beliefs about ρ (0.15 vs. true 0.5) combined with upward-biased beliefs about γ (0.99 vs. true 0.25).

To develop some intuition for how beliefs about how γ and ρ modulate the speed of correction, the green line in Figure 9b shows a simulated BIR for initial beliefs that are centered around $\rho = 0.95$ and $\gamma = 0.90$. These beliefs were chosen to match the initial level of overreaction in the data while also ensuring a rapid correction process. The results of this exercise suggest upward-biased beliefs about γ may be the driving force behind overreaction at time t, whereas beliefs about ρ may play a bigger role in the correction process.

It may be tempting to view Figure 9 as evidence that model uncertainty can generate realistic BIRs at the individual level, even when corrections to overreactions at time t are slow. However, this interpretation should be viewed cautiously given that the Bayesian framework allows considerable flexibility in fitting the data through prior specification, whereas the behavioral models examined earlier were intentionally designed with parsimony

in mind. The comparison is thus not entirely fair given the much higher effective degrees of freedom in the model uncertainty approach.

In addition, the apparent success of the Bayesian framework hinges on two key questions. First, does our simulated DGP reasonably approximate the forecasted series? While a decisive answer to this question is beyond the scope of our paper, Figure 10 offers a preliminary answer by comparing simulated and actual data paths for both datasets. Visually, the series appear similar in terms of persistence, with average and simulated AR(1) coefficients of 0.33 vs. 0.37 for SPF, and 0.95 vs 0.88 in IBES, respectively.[16]

Second, are forecasters' initial priors Θ^* reasonable? This question is challenging to answer definitively and presents a valuable direction for future research. For the time being, we take Figure 9 as evidence of the potential for model uncertainty to account for deviations from FIRE observed in survey data. As FNS argue, model uncertainty is especially appealing for understanding deviations from FIRE in professional forecasts given strong career incentives for accuracy. More broadly, our analysis in this section demonstrates how BIRs can evaluate belief formation models through simulation when closed-form solutions are unavailable.

6 Extensions and Conclusion

6.1 Generalized BIRs

Our empirical analysis consistently shows that deviations from FIRE at time t are corrected only gradually. As discussed in Section 2, this fact implies that revisions reflect two distinct components: (i) a forecaster's reaction to newly perceived information; and (ii) corrections of deviations from FIRE made in previous forecasts. The presence of this second component complicates the standard interpretation of Coibion and Gorodnichenko (2015) regressions,

[16]For both the SPF and IBES simulations, the AR(1) coefficient is based on the average AR coefficient across the $n = 100$ simulations.

which typically assumes that revisions reflect pure news. This concern is arguably more acute at shorter horizons because there is a longer history of prior forecasts that could induce corrections. In contrast, revisions of longer-horizon forecasts are more likely to reflect newly perceived information, as there are fewer prior forecasts to revise. For this reason, our baseline BIR analysis has focused on longer horizons in both SPF and IBES.

It is worth emphasizing that BIRs remain valuable even when revisions do not reflect pure news. Much like reduced-form VARs in macroeconomics, they still provide a minimum set of internally consistent moments that models must match to account for the empirical dynamics of belief formation. With that said, isolating the news component of revisions may be desirable in many applications, as it enables a cleaner structural interpretation of BIRs as responses to newly perceived information. To that end, we now introduce a generalized version of a BIR:

Definition 3 *Let* $\mathbf{H}_{t-1}^{(n)} = [R_{t-1}^{(n+1)}, R_{t-2}^{(n+2)}, ..]^\top$ *collect the history of revisions up to time* t *for a forecast target* y_{t+n}*. The generalized behavioral impulse response function (GBIR) for forecast horizon* n *is given by*

$$GB(n,j) = \mathbb{E}\left[e_{t+j}^{(n-j)} | R_t^{(n)} = 1; \mathbf{H}_{t-1}^{(n)}\right] - \mathbb{E}\left[e_{t+j}^{(n-j)} | R_t^{(n)} = 0; \mathbf{H}_{t-1}^{(n)}\right], \quad j = 0, 1, ..., n-1$$

Why condition on past revisions $\mathbf{H}_{t-1}^{(n)}$? If deviations from FIRE are corrected gradually, revisions will exhibit autocorrelation through time (Section 2). Thus, conditioning on $\mathbf{H}_{t-1}^{(n)}$ helps strip out this predictable component of $R_t^{(n)}$, thereby isolating the forecaster's response to newly perceived information. In this sense, $GB(n,0)$ provides a more precise measure of overreaction or underreaction than its unconditioned counterpart, $B(n,0)$.

Practically speaking, Generalized BIRs can be estimated simply by adding lagged revisions $\mathbf{H}_{t-1}^{(n)}$ to our baseline regression:

$$e_{t+j}^{(n-j)} = c_j^{(n)} + \psi_j^{(n)} R_t^{(n)} + \mathbf{\Gamma}_j^{(n)} \mathbf{H}_{t-1}^{(n)} + v_{t+j}^{(n-j)}, \quad \forall j = 0, , , , .n-1, \tag{19}$$

where $GB(n,j) = \psi_j^{(n)}$. When implementing regression (19) in the SPF data, we control for one lagged revision when $n = 2$ and two lagged revisions when $n = 1$. In the IBES data, $\mathbf{H}_{t-1}^{(n)}$ includes only one lag when $n = 6$. Although additional lags in IBES are technically available, they are increasingly sparse, so we limit to one lag to preserve sample size. For shorter horizons, we include more lags: two lags for $n = 5$, three for $n = 4$, and so on.[17]

Figure 11a plots the average GBIR in the SPF, estimated by pooling over all series and forecasts. For comparison, we also include the GBIR for $n = 3$, which is identical to the BIR because there are no lags on which to condition. In the SPF, there is overreaction at all horizons, with $GB(n,0) < 0$ for $n = 1,2,3$. The magnitude of overreaction decreases slightly with the forecast horizon, though the pattern is not strictly monotonic. Regarding correction speeds, for $n = 2$, more than half of the initial overreaction is reversed within one quarter.

Figure 11b contains estimates of the average GBIR in IBES, again pooling over all series and forecasts. We begin by plotting $GBIR(n = 6, j)$ to ensure comparability with our baseline BIR. This GBIR conditions on a single lagged revision. While the overall shape resembles the BIR shown in Figure 3b, it features less overreaction and slower correction. Figure A12b shows that this difference is primarily driven by sample selection. Estimating the GBIR for $n = 6$ requires data on 7- and 8-quarter-ahead forecasts in order to form lagged revisions, and such long-horizon forecasts are more sparse in the IBES. By contrast, the baseline BIR in Figure 3b is estimated on a much larger sample because it does not rely on these additional horizons. When re-estimated on the restricted GBIR sample (Figure A12b), the baseline BIR yields a similar degree of overreaction and correction speed.

Figure 11b also includes GBIRs for shorter horizons ($n = 3,2,1$) to explore how forecast behavior varies with horizon length. Interestingly, we find underreaction at shorter horizons, followed by gradual correction. This pattern is consistent with the findings of Bouchaud et al.

[17] Figure A12a in the Internet Appendix estimates that control for only one lagged revision, which results in a much larger sample size. The resulting GBIRs are similar to the ones in Figure 11.

(2019) and with the large literature on the post-earnings announcement drift (e.g. Ball and Brown, 1968; Bernard and Thomas, 1989; Sadka, 2006). The combination of underreaction at short horizons and overreaction at longer horizons also aligns with recent work by Augenblick et al. (2024), who propose a model in which individuals underreact to strong signals and overreact to weak ones. Understanding why this pattern appears in IBES but not in the SPF is an interesting direction for future research.

6.2 How pervasive are deviations from FIRE?

Relative to existing methods, statistical tests of FIRE based on BIRs are more powerful because they are based on joint restrictions on the conditional dynamics of current and future forecast errors. GBIRs take this logic one step further by imposing a restriction that current *and* prior revisions should have no forecasting power for current and future forecast errors. In this subsection, we leverage this feature of BIRs and GBIRs to test the null hypothesis of FIRE in the cross-section of series and forecasters. We then benchmark the resulting rejection rates against those from the popular test proposed by Coibion and Gorodnichenko (2015), henceforth the CG-test.

Forecast Targets (Series) Table 4a summarizes the fraction of series in the SPF and IBES that reject one of three tests of FIRE. The first is CG-test of whether the initial point $B(n, 0)$ on the BIR equals zero. The second is based on whether the entire BIR curve equals zero. As discussed in Section 2, this is equivalent to jointly testing whether forecast revisions at time t predict contemporaneous forecast errors and subsequent revisions. The third test asks whether the entire GBIR curve is zero and whether lagged revisions help predict forecast errors. Formally, this involves jointly testing if $\psi_j^{(n)} = 0$ and $\mathbf{\Gamma}_j^{(n)} = \mathbf{0}$ for all j in Equation (19). The logic of this final test is that nothing, including current and past revisions, should predict forecast errors under FIRE.

The sample for all of our tests is constructed as follows. For a given forecast target

y_{t+n} and forecaster, we include only observations for which a complete set of forecasts are available from time $t-1$ through $t+n$, including the nowcast. This means we can construct a continuous set of revisions and forecast errors beginning at time t. For the SPF data, we set $n = 3$; for IBES, we set $n = 6$. An important feature of this sample construction is that it ensures all three tests are estimated on the same set of observations. This consistency is crucial for meaningfully comparing rejection rates across tests. Note that this means the sample for IBES differs from the one used to construct the GBIR in Figure 11b, which required forecasts from $t-2$ to $t+6$ to condition on a single lagged revision. We also require a series to have at least ten observations to avoid small sample issues.

Using this sample, we cycle through all available series and estimate CG-style coefficients, BIRs, and GBIRs for each value of n. For $n = 6$, the GBIR and BIR coincide by construction, as there are no available lagged revisions on which to condition. For shorter horizons, the GBIR includes an increasing number of lagged revisions: one lag for $n = 5$, two lags for $n = 4$, and so on. For each horizon n, the table reports the fraction of series for which the null of FIRE is rejected under each test at a 5% confidence level. To account for the multiple testing problem, we adjust individual p-values using the procedure in Benjamini and Hochberg (1995).

Consistent with Bordalo et al. (2020b), the first row of Table 4a shows that the majority of series in the SPF reject the CG-test of FIRE, regardless of the horizon. The second row reports the fraction of series that reject FIRE based on the full BIR curve. Rejection rates are moderately higher across all horizons under this more stringent test. Results are nearly identical when using GBIRs, as shown in the third row.

The third through fifth rows of Table 4a present the results for series in IBES. Rejection rates from the CG-test are low across horizons, typically around 5%. In contrast, BIR-based tests yield substantially higher rejection rates, particularly at longer horizons. At short horizons, about 30% of series reject FIRE using the BIR, roughly four times the CG

rate. For horizons beyond four quarters, BIR-based rejections rise to nearly 50%, or ten to twelve times higher than CG-tests. GBIRs yield even stronger results: rejection rates are stable around 60% across all horizons—nine times higher than CG-tests at short horizons and double those from BIRs.

Forecasters Table 4b summarizes tests of FIRE across the cross-section of forecasters. Unlike Table 4a, which reports BIRs estimated for each series, this table is based on BIRs estimated separately for each forecaster i. Under the CG-test, fewer than 20% of SPF forecasters reject the null of FIRE at both $n = 2$ and $n = 3$. Rejection rates increase threefold when using the full BIR, and nearly fivefold when using GBIRs. The contrast is even starker for IBES forecasters: across all horizons, only about 7% reject FIRE under the CG-test. The fifth row of Table 4b shows that rejection rates rise to roughly 35% under BIR-based tests for shorter horizons ($n \leq 3$), a fourfold increase. More strikingly, nearly 90% of forecasters reject FIRE at $n = 2$ when using GBIRs. At longer horizons, GBIR-based rejection rates exceed 70%, nearly ten times higher than those from the CG-test.

To summarize, the results in Table 4 show that deviations from FIRE are exceedingly common in both the SPF and IBES datasets, occurring much more frequently than suggested by existing tests in the literature.[18] Rejection rates are substantially higher when using BIRs and GBIRs, which offer greater statistical power by imposing joint restrictions on the dynamics of forecast errors and revisions. This observation parallels Jordà (2023), who argues that hypothesis testing using traditional impulse responses offer more statistical power than tests based on individual point estimates.

[18] Table A5 in the Internet Appendix shows that correcting for multiple hypothesis testing is important. Without the Benjamini and Hochberg (1995) adjustment, nearly 30% of SPF forecasters reject FIRE under CG-tests, compared to about 20% with the correction. The impact is even larger in IBES, where the larger number of series and forecasters leads to rejection rates that drop by roughly a factor of four after adjustment.

6.3 Concluding Thoughts

In this paper, we develop the concept of a behavioral impulse response function, which reveals both initial deviations from FIRE and subsequent correction dynamics. BIRs generalize the test of FIRE proposed by Coibion and Gorodnichenko (2015) and offer greater statistical power through joint restrictions on future forecast errors and revisions. They can be easily estimated using OLS regressions on any dataset with a term structure of forecasts.

We estimate BIRs using individual forecasts of macroeconomic and financial variables from the SPF and earnings per share from the IBES. Our analysis reveals that BIRs from both datasets strongly reject the null hypothesis of FIRE, showing a clear tendency for individual forecasters to overreact when forecasting distant targets. On average, deviations from FIRE are corrected gradually in both datasets, though correction speeds vary meaningfully across forecast targets and forecasters. Although popular behavioral models of belief formation can match the level of BIRs, they often have a harder time matching their shape. The slow correction that is observed for many forecasters (and series) poses a particular challenge for existing models, as does the fact that correction speed varies meaningfully even holding fixed the size of the initial deviation from FIRE.

We also highlight how BIRs can be paired with simulations to explore belief formation in more complicated settings, such as when forecasters are uncertain about the true model of the world. Here, we uncover suggestive evidence that model uncertainty may explain the dynamic deviations from FIRE observed in BIRs. However, further research is needed to assess the plausibility of the initial beliefs underpinning these sorts of models. Moving forward, we recommend that researchers studying belief formation incorporate BIRs to refine and validate their models, comparing derived or simulated BIRs against empirical estimates like those presented in this study.

References

Afrouzi, Hassan, Spencer Y Kwon, Augustin Landier, Yueran Ma, and David Thesmar, "Overreaction in expectations: Evidence and theory," *The Quarterly Journal of Economics*, 2023, p. qjad009.

Angeletos, George-Marios, Zhen Huo, and Karthik A Sastry, "Imperfect macroeconomic expectations: Evidence and theory," *NBER Macroeconomics Annual*, 2021, *35* (1), 1–86.

Augenblick, Ned and Matthew Rabin, "Belief Movement, Uncertainty Reduction, and Rational Updating," *The Quarterly Journal of Economics*, 02 2021, *136* (2), 933–985.

___, Eben Lazarus, and Michael Thaler, "Overinference from Weak Signals and Underinference from Strong Signals*," *The Quarterly Journal of Economics*, 10 2024, *140* (1), 335–401.

Baker, Malcolm, Richard S Ruback, and Jeffrey Wurgler, "Behavioral corporate finance," in "Handbook of empirical corporate finance," Elsevier, 2007, pp. 145–186.

Ball, Ray and Philip Brown, "An Empirical Evaluation of Accounting Income Numbers," *Journal of Accounting Research*, 1968, *6* (2), 159–178.

Bansal, Ravi and Amir Yaron, "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *The Journal of Finance*, 2004, *59* (4), 1481–1509.

Barberis, Nicholas, Lawrence J. Jin, and Baolian Wang, "Prospect Theory and Stock Market Anomalies," *The Journal of Finance*, 2021, *76* (5), 2639–2687.

___, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, "X-CAPM: An extrapolative capital asset pricing model," *Journal of Financial Economics*, 2015, *115* (1), 1–24.

___, ___, ___, and ___, "X-CAPM: An extrapolative capital asset pricing model," *Journal of financial economics*, 2015, *115* (1), 1–24.

Bauer, Michael D. and Eric T. Swanson, "An Alternative Explanation for the 'Fed Information Effect'," *American Economic Review*, March 2023, *113* (3), 664–700.

Benjamini, Yoav and Yosef Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1995, *57* (1), 289–300.

Bernard, Victor L. and Jacob K. Thomas, "Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium?," *Journal of Accounting Research*, 1989, *27*, 1–36.

Bianchi, Francesco, Cosmin L Ilut, and Hikaru Saijo, "Smooth Diagnostic Expectations," Technical Report, National Bureau of Economic Research 2024.

___, Sydney C. Ludvigson, and Sai Ma, "Belief Distortions and Macroeconomic Fluctuations," *American Economic Review*, July 2022, *112* (7), 2269–2315.

___, ___, and ___, "What Hundreds of Economic News Events Say About Belief Overreaction in the Stock Market," Working paper 2024.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, "Diagnostic expectations and stock returns," *The Journal of Finance*, 2019, *74* (6), 2839–2874.

___, ___, ___, and ___, "Belief overreaction and stock market puzzles," 2020.

___, ___, Rafael LaPorta, and Andrei Shleifer, "Belief Overreaction and Stock Market Puzzles," *Journal of Political Economy*, 2024.

___, ___, Yueran Ma, and Andrei Shleifer, "Overreaction in macroeconomic expectations," *American Economic Review*, 2020, *110* (9), 2748–2782.

Bouchaud, Jean-Philippe, Philipp Krueger, Augustin Landier, and David Thesmar, "Sticky expectations and the profitability anomaly," *The Journal of Finance*, 2019, *74* (2), 639–674.

Cagan, Phillip, "The monetary dynamics of hyperinflation," *Studies in the Quantity Theory if Money*, 1956.

Campbell, John Y. and John H. Cochrane, *Journal of Political Economy*, 1999, *107* (2), 205–251.

Cochrane, John H. and Monika Piazzesi, "Bond Risk Premia," *American Economic Review*, March 2005, *95* (1), 138–160.

Coibion, Olivier and Yuriy Gorodnichenko, "What Can Survey Forecasts Tell Us about Information Rigidities?," *Journal of Political Economy*, 2012, *120* (1), 116–159.

___ **and** ___ , "Information rigidity and the expectations formation process: A simple framework and new facts," *American Economic Review*, 2015, *105* (8), 2644–2678.

Farmer, Leland E., Emi Nakamura, and Jón Steinsson, "Learning about the Long Run," Working paper 2023.

Francis, Jennifer and Donna Philbrick, "Analysts' Decisions As Products of a Multi-Task Environment," *Journal of Accounting Research*, 1993, *31* (2), 216–230.

Gabaix, Xavier, "Behavioral inattention," in "Handbook of behavioral economics: Applications and foundations 1," Vol. 2, Elsevier, 2019, pp. 261–343.

___ , "A Behavioral New Keynesian Model," *American Economic Review*, August 2020, *110* (8), 2271–2327.

Gennaioli, Nicola, Marta Leva, Raphael Schoenle, and Andrei Shleifer, "How Inflation Expectations De-Anchor: The Role of Selective Memory Cues," 2025. Working paper.

Giannoni, Marc P and Michael Woodford, "How forward-looking is optimal monetary policy?," *Journal of Money, Credit and Banking*, 2003, pp. 1425–1469.

Hirshleifer, David, Jun Li, and Jianfeng Yu, "Asset pricing in production economies with extrapolative expectations," *Journal of Monetary Economics*, 2015, *76*, 87–106.

Jackson, Andrew R., "Trade Generation, Reputation, and Sell-Side Analysts," *The Journal of Finance*, 2005, *60* (2), 673–717.

Kelly, Bryan T, Semyon Malamud, and Kangying Zhou, "The virtue of complexity in return prediction," *Journal of Finance*, forthcoming.

Kohlhas, Alexandre N. and Ansgar Walther, "Asymmetric Attention," *American Economic Review*, September 2021, *111* (9), 2879–2925.

Kučinskas, Simas and Florian S. Peters, "Measuring Under- and Overreaction in Expectation Formation," *The Review of Economics and Statistics*, 11 2022, pp. 1–45.

Lazarus, Eben, Daniel J Lewis, James H Stock, and Mark W Watson, "HAR inference: Recommendations for practice," *Journal of Business & Economic Statistics*, 2018, *36* (4), 541–559.

Li, Dake, Mikkel Plagborg-Møller, and Christian K Wolf, "Local Projections vs. VARs: Lessons From Thousands of DGPs," Working Paper 30207, National Bureau of Economic Research July 2022.

Lovell, Michael C., "Tests of the Rational Expectations Hypothesis," *The American Economic Review*, 1986, *76* (1), 110–124.

Malmendier, Ulrike, "Behavioral Corporate Finance," in B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds., *Handbook of Behavioral Economics - Foundations and Applications 1*, Vol. 1 of Handbook of Behavioral Economics: Applications and Foundations 1, North-Holland, 2018, pp. 277–379.

___ **and Stefan Nagel**, " Learning from Inflation Experiences *," *The Quarterly Journal of Economics*, 10 2015, *131* (1), 53–87.

Mankiw, N. Gregory and Ricardo Reis, "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve*," *The Quarterly Journal of Economics*, 11 2002, *117* (4), 1295–1328.

___ **,** ___ **, and Justin Wolfers**, "Disagreement about Inflation Expectations," *NBER Macroeconomics Annual*, 2003, *18*, 209–248.

___ **,** ___ **, and** ___ , "Disagreement about Inflation Expectations," *NBER Macroeconomics Annual*, 2003, *18*, 209–248.

Mincer, Jacob A and Victor Zarnowitz, "The evaluation of economic forecasts," in "Economic forecasts and expectations: Analysis of forecasting behavior and performance," NBER, 1969, pp. 3–46.

Nagel, Stefan and Zhengyang Xu, "Asset pricing with fading memory," *The Review of Financial Studies*, 2022, *35* (5), 2190–2245.

Nerlove, Marc, "Adaptive expectations and cobweb phenomena," *The Quarterly Journal of Economics*, 1958, *72* (2), 227–240.

Òscar Jordà, "Local Projections for Applied Economics," *Annual Review of Economics*, 2023, *15* (1), 607–631.

Pflueger, Carolin, Emil Siriwardane, and Adi Sunderam, "Financial Market Risk Perceptions and the Macroeconomy*," *The Quarterly Journal of Economics*, 03 2020, *135* (3), 1443–1491.

Plagborg-Møller, Mikkel and Christian K. Wolf, "Local Projections and VARs Estimate the Same Impulse Responses," *Econometrica*, 2021, *89* (2), 955–980.

Sadka, Ronnie, "Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk," *Journal of Financial Economics*, 2006, *80* (2), 309–349.

Schwartzstein, Joshua and Adi Sunderam, "Using Models to Persuade," *American Economic Review*, January 2021, *111* (1), 276–323.

Sims, Christopher A., "Macroeconomics and Reality," *Econometrica*, 1980, *48* (1), 1–48.

___ , "Implications of rational inattention," *Journal of Monetary Economics*, 2003, *50* (3), 665–690. Swiss National Bank/Study Center Gerzensee Conference on Monetary Policy under Incomplete Information.

Stein, Jeremy C., "Rational Capital Budgeting in an Irrational World," *Journal of Business*, 1996, *69* (4), 429–455.

Thomas, Lloyd B., "Survey Measures of Expected U.S. Inflation," *Journal of Economic Perspectives*, December 1999, *13* (4), 125–144.

Timmermann, Allan G., "How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices," *The Quarterly Journal of Economics*, 1993, *108* (4), 1135–1145.

Woodford, Michael, *Imperfect Common Knowledge and the Effects of Monetary Policy*, Princeton University Press,

Òscar Jordà, "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, March 2005, *95* (1), 161–182.

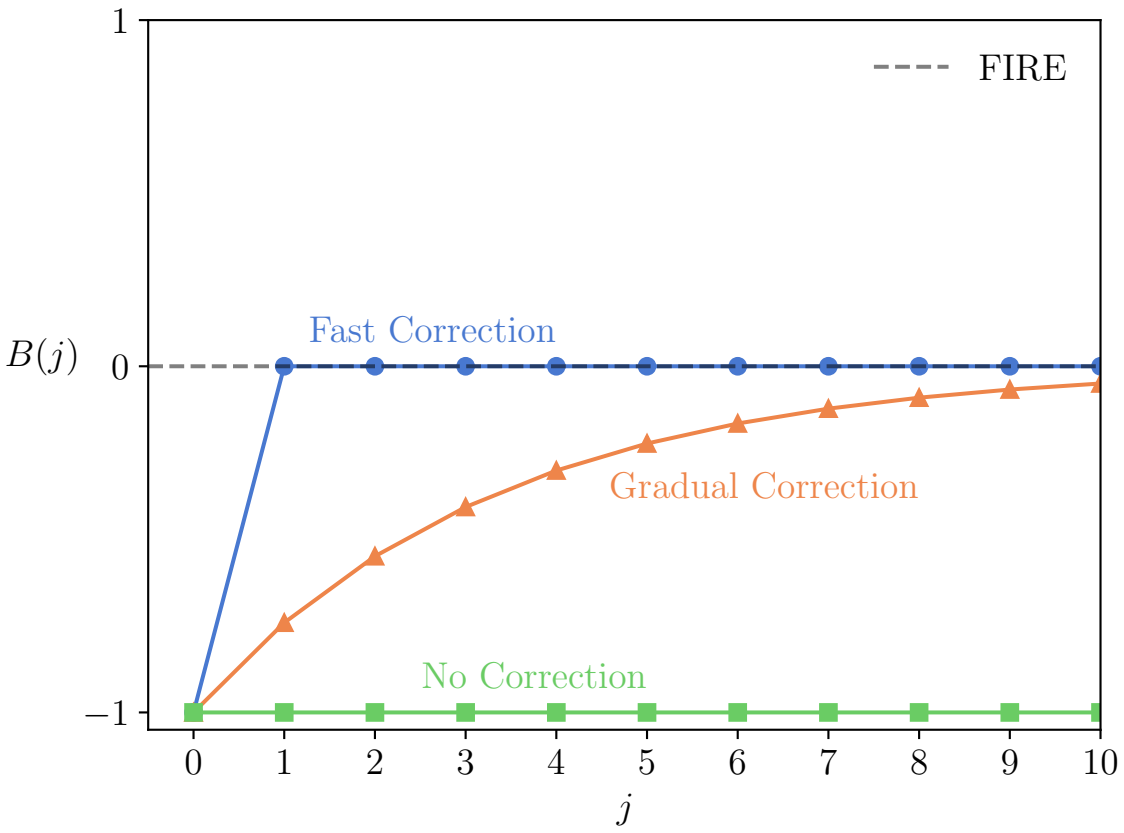


Figure 1: Stylized Example of a BIR. This figure shows stylized BIRs for three different forecasters, all of whom overreact by the same amount at time t. The three forecasters differ in how quickly they correct their overreaction.

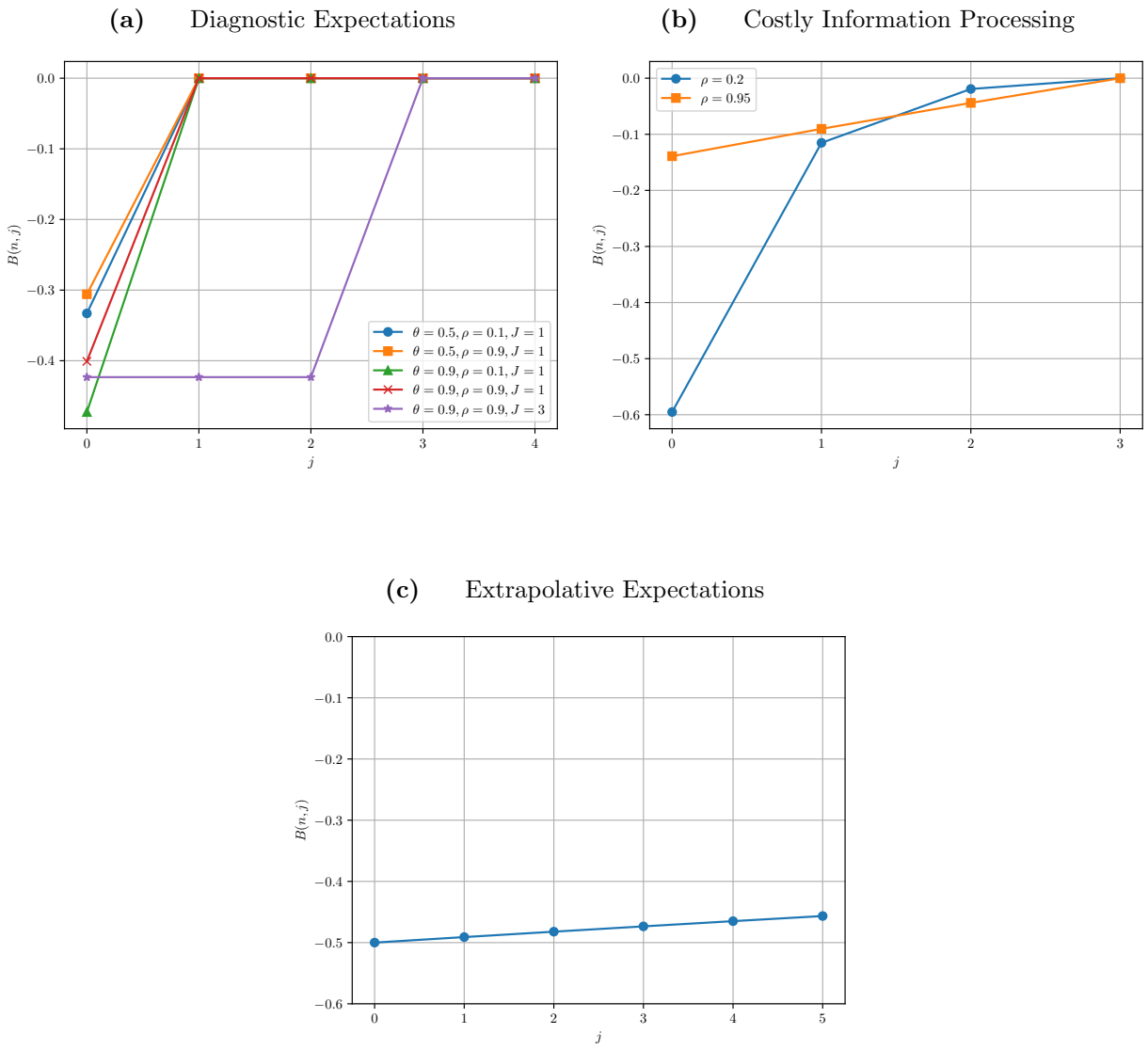
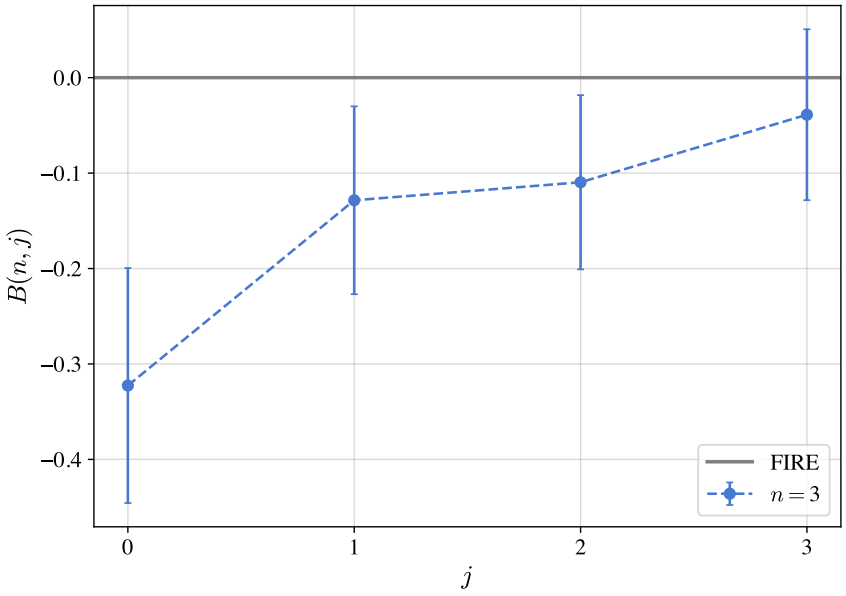


Figure 2: BIRs in Leading Behavioral Theories. Figure 2a shows BIRs for different levels of diagnosticity (θ) and signal persistence (ρ) under the model of diagnostic expectations from Bordalo et al. (2020a). Figure 2b shows the BIR from Afrouzi et al. (2023), with the exact calibration detailed in Section 3.3.2. Figure 2c shows the BIR implied by the model of Nagel and Xu (2022), using the constant-gain learning parameter of 0.018 estimated in Malmendier and Nagel (2015).

(a) Average BIR in SPF



(b) Average BIR in IBES

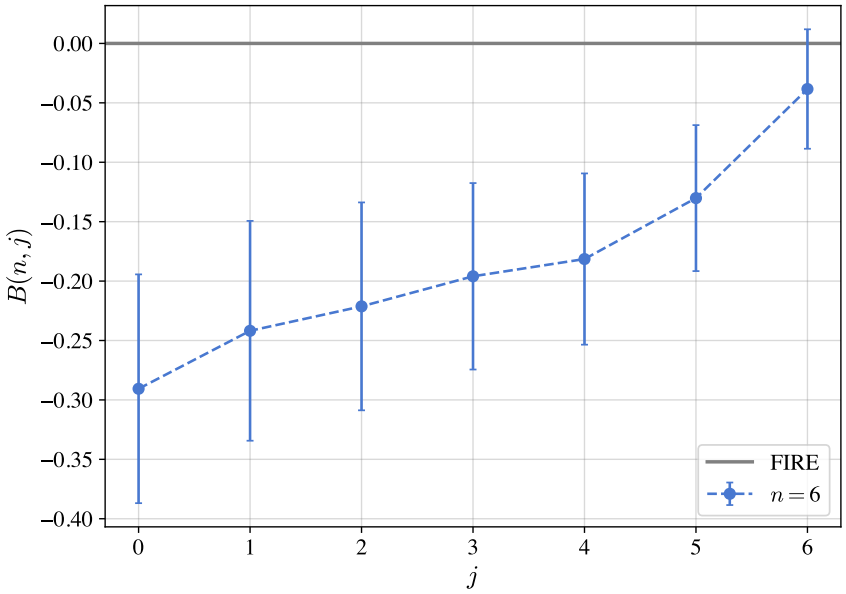
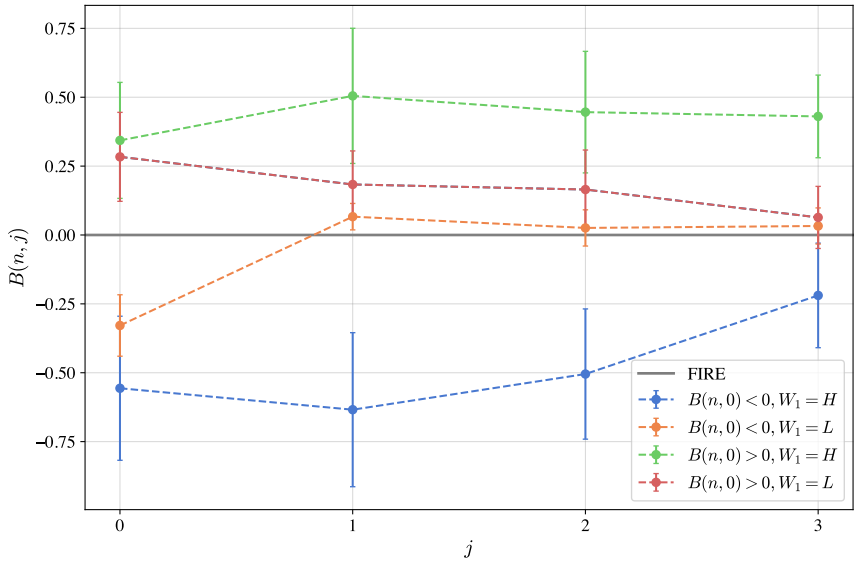


Figure 3: Average BIRs from SPF and IBES. Figures 3a and 3b show BIR estimates when pooling across all forecasts in SPF ($n = 3$) and IBES ($n = 6$), respectively. 95% confidence intervals in both panels are based on standard errors that are clustered by forecaster-series (i, s) and series-date (s, t), where t is the date of the first observed revision for a given forecast horizon n. The black line in both plots shows the BIR under FIRE. See Section 2.2 for complete estimation details.

(a) SPF



(b) IBES

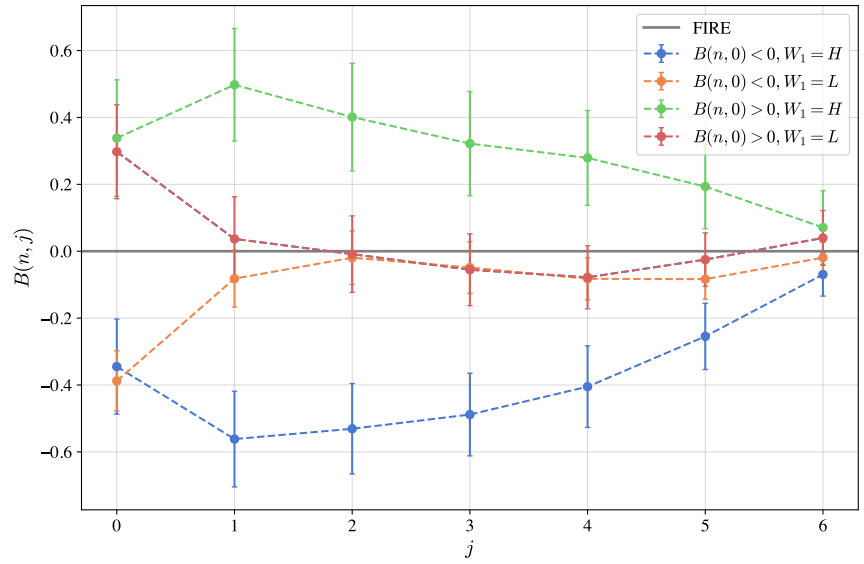


Figure 4: Variation in BIRs Across Forecasters. Figures 4a and 4b visualize variation in BIRs across forecasters in SPF and IBES, respectively. For both datasets, BIRs are first estimated for each forecaster, then forecasters are split based on whether they initially over or underreact. Within the set of overreactors, forecasters are further split in to k bins based on $W_1 = B(n,1)/B(n,0)$. The same process is repeated for underreactors. A pooled BIR is then estimated for each bin and plotted in the figure. For SPF, $k = 3$ for forecasters that overreact, with BIRs shown only for the top and bottom terciles, and $k = 2$ for those that underreact. For IBES, $k = 4$ for both types of forecasters, and BIRs are shown only for the top and bottom quartiles of W_1. 95% confidence intervals are based on standard errors that are clustered by forecaster-series, (i, s), and series-date, (s, t). The black line in each plot shows the BIR under FIRE. See Section 2.2 for complete estimation details.

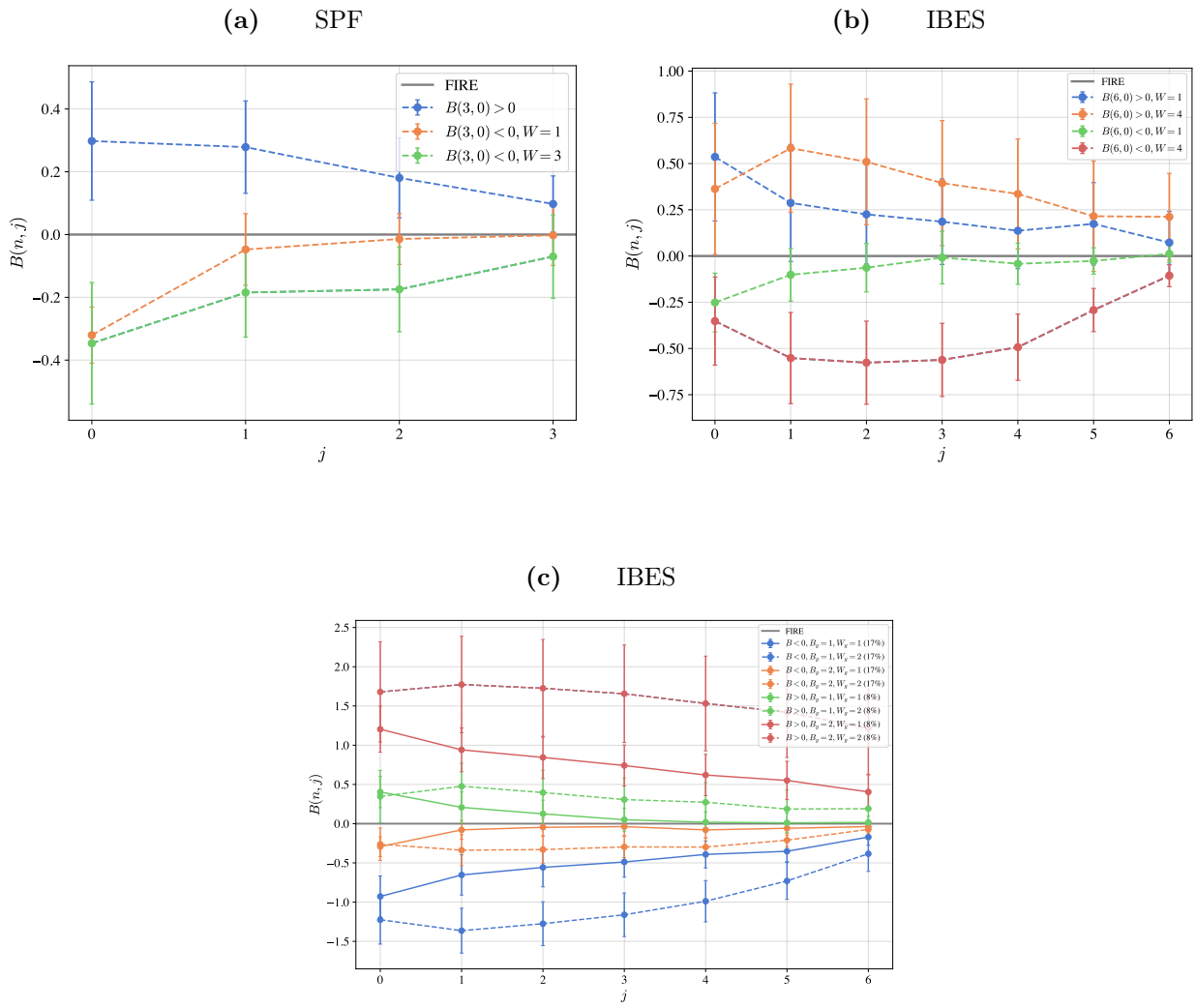


Figure 5: Variation in BIRs Across Series. Figure 5a visualizes variation in BIRs across series in SPF. To construct the plot, we estimate series-by-series BIRs and divide them based on the sign of $B(n,0)$. Within the set of series that exhibit underreaction ($B(n,0) > 0$), we estimate a pooled BIR and plot it in blue. Within the set of series that exhibit overreaction, we further split them into three groups based on $W(n,1) = B(n,1)/B(n,0)$, and plot the top and bottom terciles. Figure 5b shows BIRs resulting from an analogous process in IBES, splitting both underreacting and overreacting series into four groups and plotting the top and bottom quartiles. Figure 5c applies a double-sorting methodology, first dividing overreacting and underreacting series separately into two groups based on the magnitude of $B(n,0)$, then further splitting each group based on $W(n,1)$, creating eight groups total. The idea is to isolate variation in $W(n,1)$ while holding $B(n,0)$ roughly fixed. The legend in Figure 5c shows the percentage of series assigned to each group. 95% confidence intervals are based on standard errors that are clustered by forecaster-series, (i,s), and series-date, (s,t). The black line in each plot shows the BIR under FIRE. See Section 2.2 for complete estimation details.

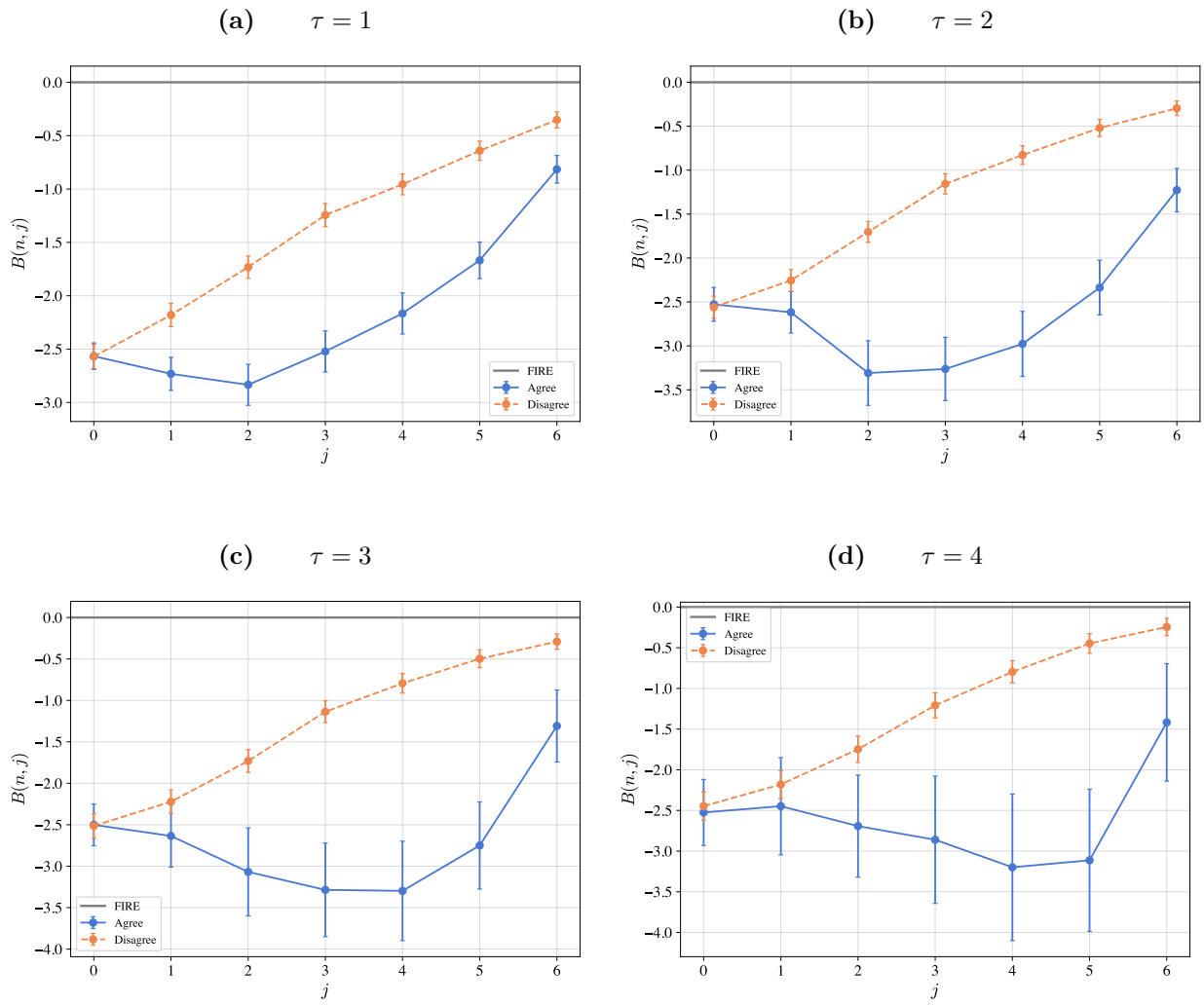


Figure 6: How New Information Shapes Corrections. This figure shows the estimates for BIRs of forecasts with different subsequent signals in IBES. The sample for all plots is based on forecasts where the revision $R_t^{(6)}$ and associated forecast error $e_t^{(6)}$ have opposite signs (overreaction). We further retain the middle three terciles of both distributions to ensure comparable levels of overreaction across subsamples. Each panel then conditions on whether the subsequent sequence of one-period forecast errors have the same sign as the initial revision, $R_t^{(6)}$. When they do, the forecaster receives reinforcing signals; when they don't, the forecaster receives contradictory signals. Figure 6a conditions on $e_t^{(1)}$. Figure 6b conditions on both $e_t^{(1)}$ and $e_{t+1}^{(1)}$ having the same (or opposite) sign as $R_t^{(6)}$. Subsequent panels extend this logic to longer sequences of consistent signals. 95% confidence intervals in the plot are based on standard errors clustered by analyst-stock, (i, s), and stock-date, (s, t). The black line shows the full information rational expectations benchmark.

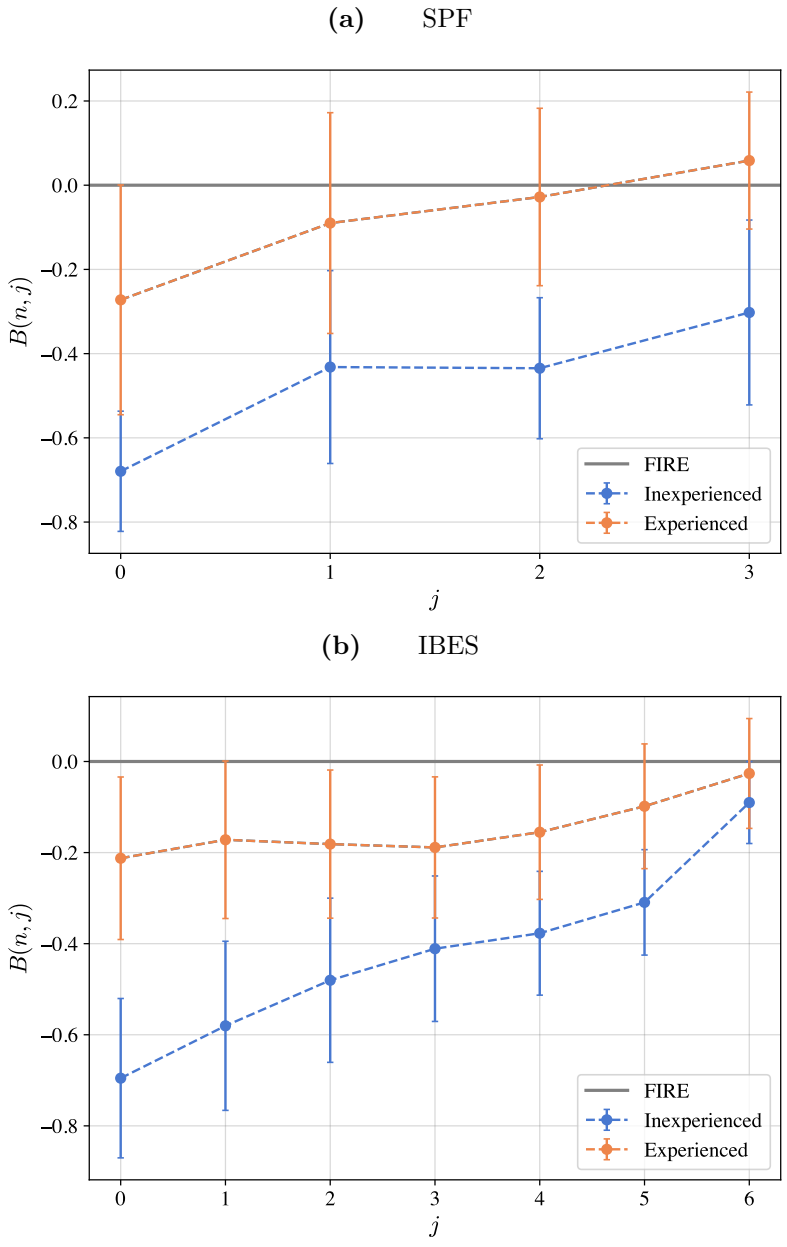


Figure 7: BIRs and Experience. This figure shows how BIRs vary with forecaster experience. We start by estimating BIRs for each forecaster using their first five years of data, requiring at least ten forecasts. We sort forecasters i by the average level of their BIR, denoted as B_i^{early}, and assign them to terciles. Within each tercile, we then estimate pooled BIRs using: (i) the first five-years of forecasts made by the forecaster (labeled as the "inexperienced" BIRs in blue); and (ii) all forecasts made after year five (labeled as the "experienced" BIR in orange). Figures 7a and 7b show the estimates for forecasters in SPF and IBES, respectively. For both plots, estimates are shown for forecasters that overreact early in their careers (i.e., the bottom tercile of B_i^{early}). Results for underreactors can be found in Internet Appendix Figure A9. 95% confidence intervals in the plot are based on standard errors clustered by analyst-stock, (i, s), and stock-date, (s, t). The black line shows the full information rational expectations benchmark.

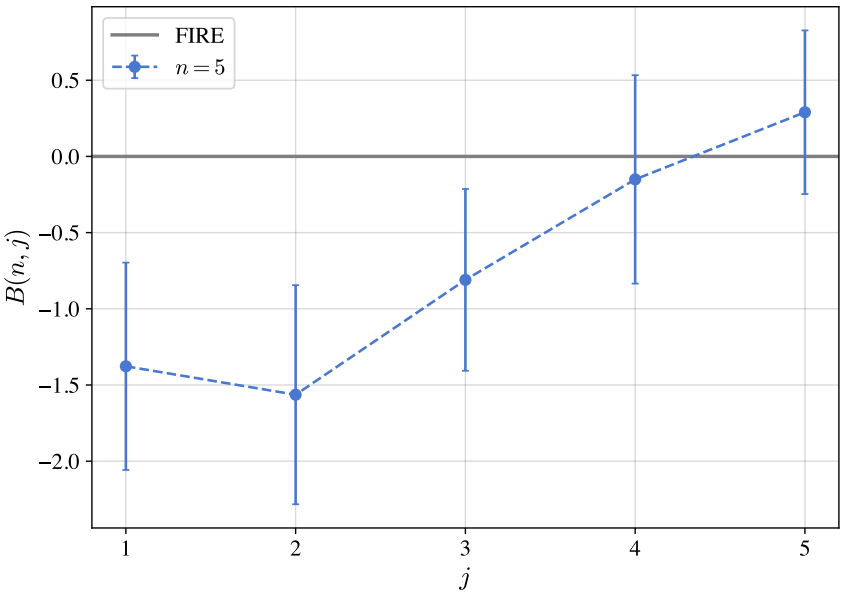


Figure 8: BIRs from CBO. This figure shows the estimated BIR ($n = 5$ years) for real GDP growth forecasts from the CBO. 95% confidence intervals are based on Newey-West standard errors with lags specified as $L = \lceil 1.3 \times T^{\frac{1}{2}} \rceil$, as proposed in Lazarus et al. (2018). The black line shows the BIR under FIRE. See Section 2.2 for complete estimation details.

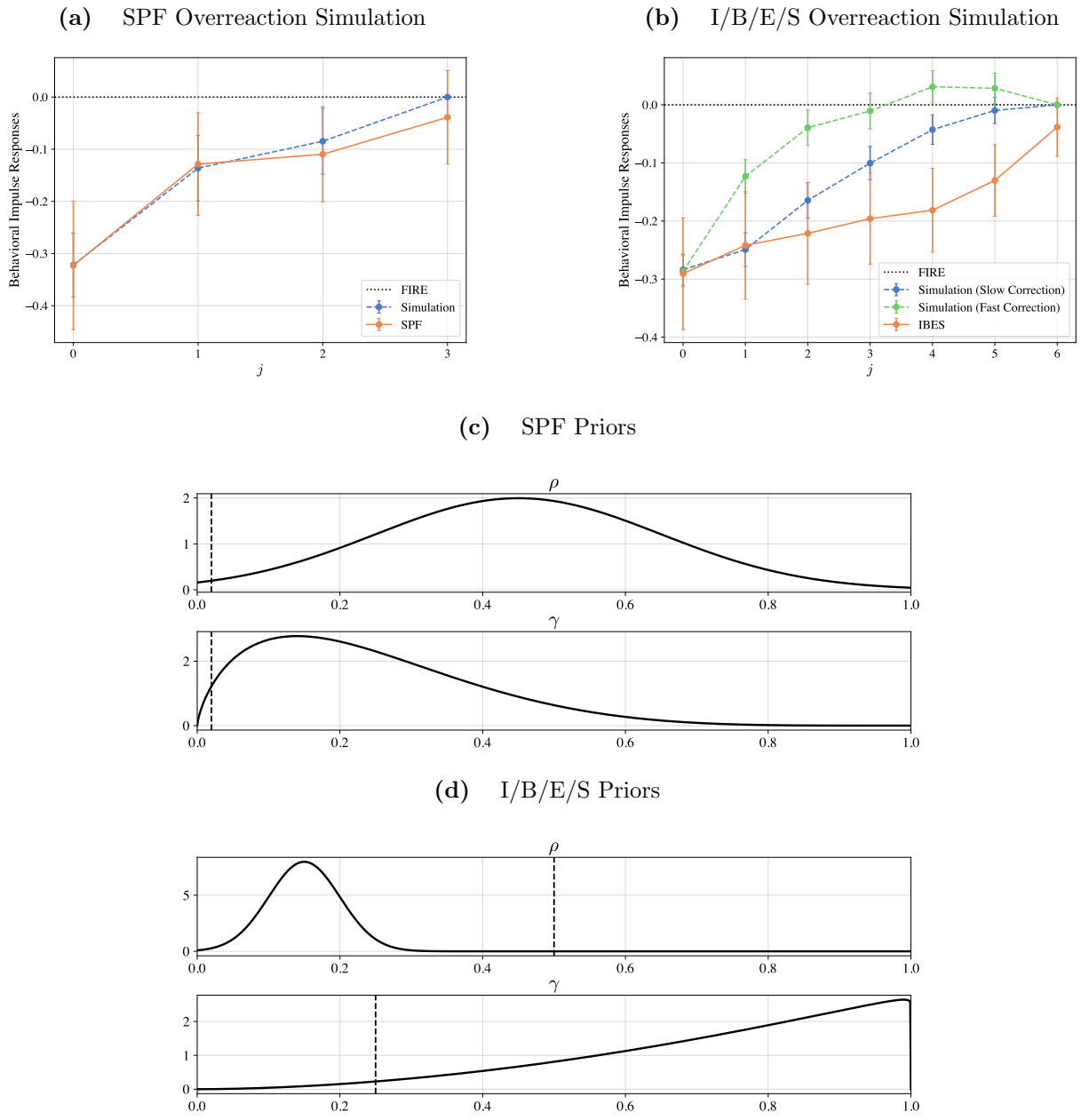
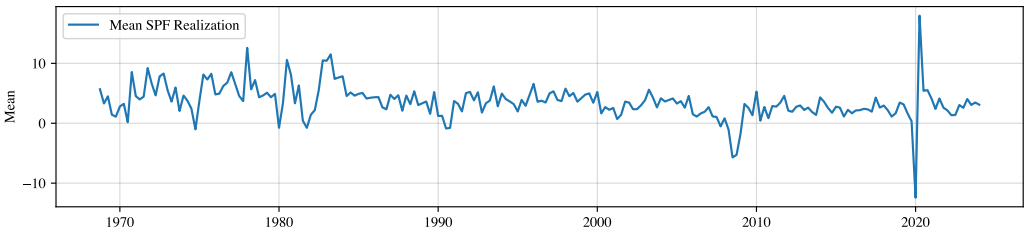
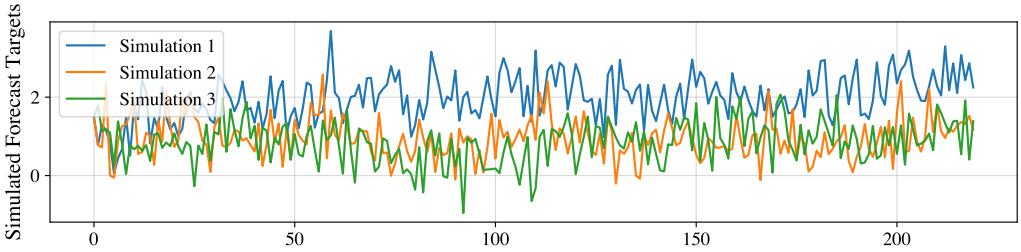


Figure 9: BIRs with Model Uncertainty. Figure 9a shows BIRs estimated on the pooled SPF forecasts with $n = 3$. The plot also shows BIRs from the simulated model of Farmer et al. (2023) with the best-matched priors. Figure 9b shows the BIR for the IBES pooled forecasts with $n = 6$ and the simulated Bayesian forecasters in Farmer et al. (2023) with the best-matched priors. Figures 9c and 9d show the best-matched priors with respect to the SPF and IBES BIRs, respectively. See Section 5 for complete details.

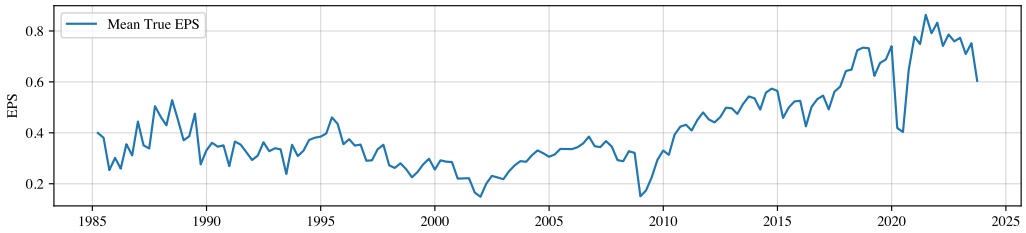
(a) Quarterly Mean SPF Realization



(b) SPF Sample Simulated Series



(c) Quarterly Mean EPS



(d) EPS Sample Simulated Series

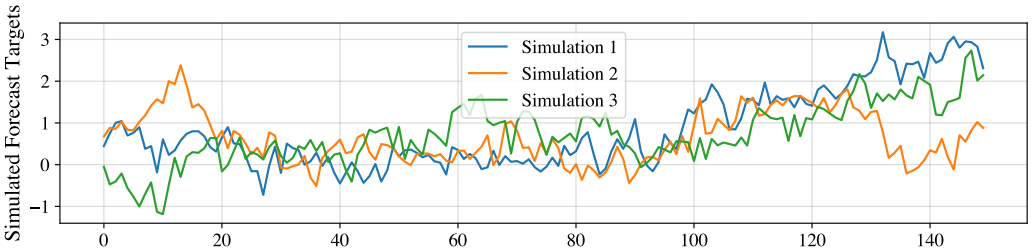


Figure 10: BIRs with Model Uncertainty. Figure 10a shows the quarterly cross-sectional average realization of the SPF forecast targets. Figure 10b shows three randomly simulated paths according to Farmer et al. (2023) with $\rho = 0.02, \gamma = 0.02$. Figure 10c shows the quarterly cross-sectional average EPS in the IBES sample. Figure 10d shows three randomly simulated paths according to Farmer et al. (2023) with $\rho = 0.50, \gamma = 0.25$.

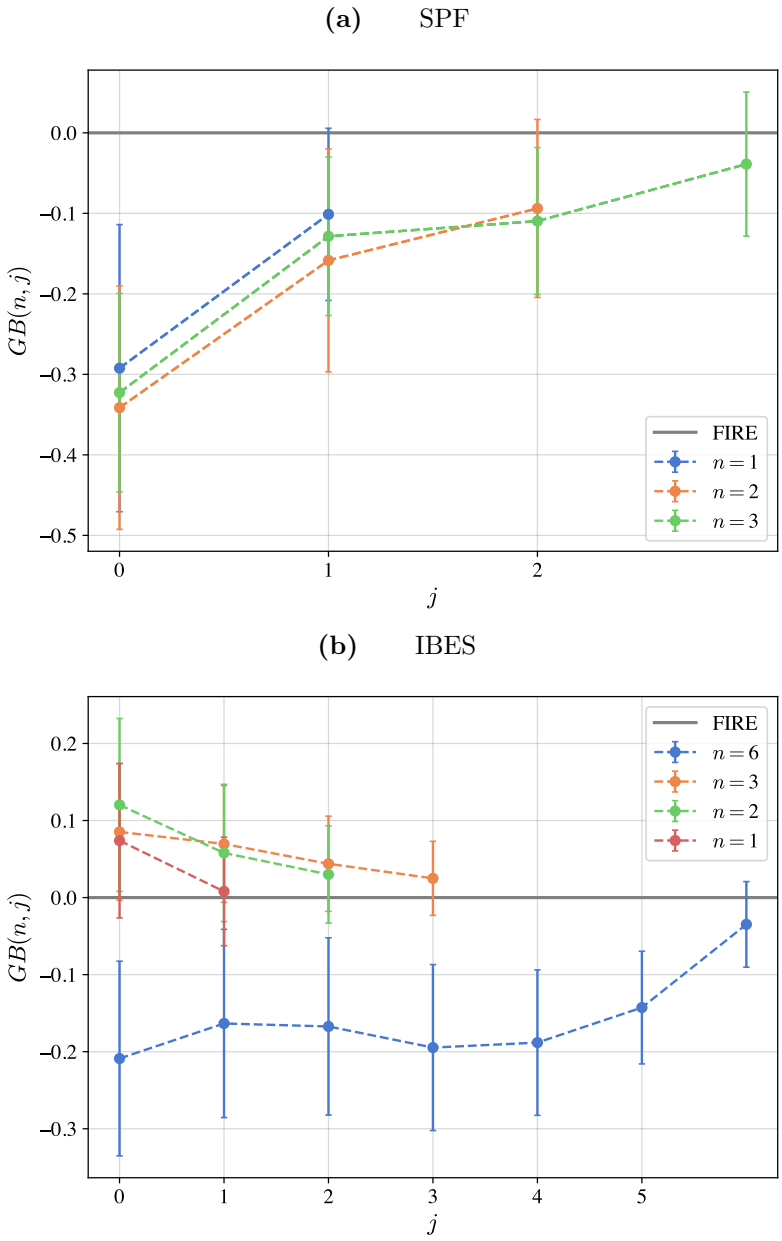


Figure 11: Generalized BIRs. This figure shows the estimates for generalized BIRs. The generalized BIRs projects forecast errors on forecast revisions while controlling for past revisions of the same target. Figures 11a and 11b show the estimates for forecasters in SPF and IBES, respectively. 95% confidence intervals in the plot are based on standard errors clustered by analyst-stock, (i, s), and stock-date, (s, t). The black line shows the full information rational expectations benchmark. See Section 6.1 for complete details on estimation and sample selection.

	I/B/E/S	SPF
$B(n,0)$	-0.291^{***}	-0.323^{***}
	(0.049)	(0.063)
$B(n,1)$	-0.242^{***}	-0.128^{**}
	(0.047)	(0.050)
$B(n,2)$	-0.221^{***}	-0.110^{**}
	(0.045)	(0.047)
$B(n,3)$	-0.196^{***}	-0.039
	(0.040)	(0.046)
$B(n,4)$	-0.181^{***}	
	(0.037)	
$B(n,5)$	-0.130^{***}	
	(0.031)	
$B(n,6)$	-0.038	
	(0.026)	
n	6	3
Adj. R^2 (%)	0.30	0.36
Observations	49,081	42,150
p(FIRE)	<0.001	<0.001
p(Equal B)	<0.001	<0.001

Table 1: BIRs Estimated from Data. This table shows point estimates for BIRs based on all forecasts of quarterly EPS and macroeconomic outcomes from IBES and SPF. Columns (1) and (2) show IBES and SPF BIRs estimates for $n = 6$ and $n = 3$, respectively. Each row in the table corresponds to a different j. Standard errors are listed in parentheses below each point estimate and are clustered by forecaster-series, (i, s), and series-date, (s, t). The row with $p(FIRE)$ reports the p-value from testing the null hypothesis of FIRE, i.e. $B(n, j) = 0, \forall j$. The row with p(Equal B) reports the p-value from testing the null hypothesis that all points along the BIR curve are equal, which is a weaker test of FIRE. The adjusted-R^2 is based on averaging over the individual regressions used to estimate each point on the BIR curve (e.g., $B(n,0), ..., B(n,n)$). See Section 2.2 for complete estimation details. *, **, and *** denote p-values of less than 0.10, 0.05, and 0.01, respectively.

Data	N	mean	$p10$	$p25$	$p50$	$p75$	$p90$	$\mathbb{P}(B(n,0)<0)$
SPF	190	-0.28	-1.14	-0.67	-0.36	0.05	0.47	72
IBES	970	-0.25	-1.58	-0.79	-0.27	0.27	1.01	63

(a) Distribution of Initial Reaction in SPF and IBES

	W_1 bins for Overreactors			W_1 bins for Underreactors	
	1	2	3	1	2
W_1	-0.20**	0.57***	1.14***	0.65***	1.47***
W_2	-0.08	0.48***	0.91***	0.58***	1.30***
W_3	-0.10	0.28**	0.39***	0.24	1.25***
$B(3,0)$	-0.33***	-0.61***	-0.56***	0.28***	0.34***
% Forecasters	24	24	24	14	14

(b) SPF: Correction Speed

	W_1 bins for Overreactors				W_1 bins for Underreactors			
	1	2	3	4	1	2	3	4
W_1	0.21**	0.81***	1.01***	1.63***	0.12	0.77***	0.98***	1.47***
W_2	0.05	0.67***	0.99***	1.54***	-0.03	0.66***	0.94***	1.19***
W_3	0.13	0.59***	0.86***	1.42***	-0.19	0.60***	0.95***	0.95***
W_4	0.21***	0.49***	0.77***	1.17***	-0.26	0.44***	0.89***	0.82***
W_5	0.22***	0.37***	0.66***	0.74***	-0.08	0.38***	0.89***	0.57***
W_6	0.05	0.21***	0.35***	0.20	0.13	0.29***	0.67***	0.21
$B(6,0)$	-0.39***	-0.78***	-0.99***	-0.34***	0.30***	1.03***	1.04***	0.34***
% Forecasters	16	16	16	16	9	9	9	9

(c) IBES: Correction Speed

Table 2: BIR Heterogeneity Across Forecasters. This table summarizes the size and correction of deviations from FIRE across forecasters in the SPF and IBES. In Table 2a, we estimate forecaster-level BIRs, requiring at least ten observations. We then report the summary statistics on the resulting $B(n,0)$. In Table 2b, we first sort forecasters in SPF based on whether their $B(n,0)$ is negative (overreactors) or positive (underreactors). Within the set of overreactors (underreactors), we further sort forecasters into terciles (halves) on $W(n,1) = B(n,1)/B(n,0)$. Within each group, we then estimate a pooled BIR and report the resulting $W(n,j) = B(n,j)/B(n,0)$, a measure of correction speed. Table 2c repeats this exercise for forecasters in IBES, separately dividing forecasters that overreact or underreact into quartiles based on $W(n,1)$. In all tables $n = 3$ for SPF and $n = 6$ for IBES. *, **, and *** denote p-values of less than 0.10, 0.05, and 0.01, respectively.

Data	N	mean	$p10$	$p25$	$p50$	$p75$	$p90$	$\mathbb{P}(B(n,0)<0)$
SPF	15	-0.22	-0.46	-0.40	-0.33	-0.14	0.18	80
IBES	1299	-0.32	-1.46	-0.88	-0.37	0.13	0.76	69

(a) Distribution of Initial Reaction in SPF and IBES

	W_1 bins for Overreactors			Underreactors
	1	2	3	
W_1	0.15	0.40***	0.53***	0.93***
W_2	0.04	0.16	0.50***	0.61***
W_3	0.01	-0.11	0.20	0.33**
$B(3,0)$	-0.32***	-0.35***	-0.35***	0.30***
# Series	4	4	4	3

(b) SPF: Correction Speed

	W_1 bins for Overreactors				W_1 bins for Underreactors			
	1	2	3	4	1	2	3	4
W_1	0.20	0.74***	1.00***	1.57***	0.39***	0.86***	1.03***	1.61***
W_2	0.01	0.67***	0.89***	1.64***	0.24	0.79***	0.96***	1.40***
W_3	0.08	0.57***	0.75***	1.60***	0.14	0.70***	0.92***	1.09***
W_4	0.18*	0.49***	0.68***	1.40***	0.07	0.62***	0.83***	0.92***
W_5	0.10	0.45***	0.56***	0.83***	0.04	0.57***	0.78***	0.59**
W_6	0.03	0.23***	0.30***	0.30**	0.08	0.47***	0.64***	0.58**
$B(6,0)$	-0.37***	-0.65***	-0.90***	-0.35***	0.49***	1.25***	1.16***	0.36**
% Forecasters	15	16	16	15	8	11	9	9

(c) IBES: Correction Speed

Table 3: BIR Heterogeneity Across Data-Generating Processes. This table summarizes the size and correction of deviations from FIRE across series in the SPF and IBES. In Table 3a, we estimate series-level BIRs, requiring at least ten observations. We then report the summary statistics on the resulting $B(n,0)$. In Table 3b, we first sort series in SPF based on whether their $B(n,0)$ is negative (overreactors) or positive (underreactors). Within the set of overreactors, we further sort series into terciles on $W(n,1) = B(n,1)/B(n,0)$. Within each group, we then estimate a pooled BIR and report the resulting $W(n,j) = B(n,j)/B(n,0)$, a measure of correction speed. Because there is a limited number of series that exhibit underreaction in SPF ($B(n,0) > 0$), we do not sort further and instead report the estimated $W(n,j)$ when pooling across all series that underreact. Table 3c repeats this exercise for series in IBES, separately dividing series that overreact or underreact into quartiles based on $W(n,1)$. In all tables $n = 3$ for SPF and $n = 6$ for IBES. *, **, and *** denote p-values of less than 0.10, 0.05, and 0.01, respectively.

Data	Test	n				
		2	3	4	5	6
SPF	$B(n,0)=0$	0.73	0.80			
SPF	$B(n,j)=0,\forall j$	0.80	0.80			
SPF	Generalized BIR	0.80	0.80			
IBES	$B(n,0)=0$	0.07	0.04	0.03	0.04	0.06
IBES	$B(n,j)=0,\forall j$	0.28	0.31	0.35	0.43	0.48
IBES	Generalized BIR	0.60	0.60	0.57	0.57	0.48
# Series SPF		15	15			
# Series IBES		1299	1299	1299	1299	1299

(a) By Series

Data	Test	n				
		2	3	4	5	6
SPF	$B(n,0)=0$	0.14	0.17			
SPF	$B(n,j)=0,\forall j$	0.44	0.62			
SPF	Generalized BIR	0.65	0.62			
IBES	$B(n,0)=0$	0.09	0.07	0.07	0.03	0.07
IBES	$B(n,j)=0,\forall j$	0.32	0.40	0.53	0.62	0.72
IBES	Generalized BIR	0.90	0.91	0.91	0.90	0.72
# Forecasters SPF		190	190			
# Forecasters IBES		970	970	970	970	970

(b) By Forecaster

Table 4: How Pervasive are Deviations from FIRE? This table summarizes tests of FIRE across series and forecasters. Table 4a is based on BIRs estimated for each series s and forecast horizon n, whereas Table 4b is based on BIRs estimated for each forecaster i and forecast horizon n. For each BIR, we conduct three tests for FIRE based on whether: (i) $B(n,0)=0$, following Coibion and Gorodnichenko (2015); (ii) $B(n,j)=0$ for all j; and (iii) the regression coefficients on all past revisions are jointly zero in the generalized BIR framework (see Equation (19)). We include only observations with complete forecasts from time $t-1$ to $t+n$ (with $n=3$ for SPF and $n=6$ for IBES), and focus on forecasters/series with at least ten observations. For each test, we report the fraction of BIRs for which the null of FIRE is rejected with 95% confidence, adjusting p-values for multiple testing following Benjamini and Hochberg (1995). The bottom rows of each table show the unique number of series and forecasters, respectively, in the cross-section for every n.

A Linear Autonomous Forecasts and Leading Behavioral Models

We now show that many popular models of belief formation produce forecasts that belong to the linear autonomous class.

- **Adaptive Expectations** of Cagan (1956) and Nerlove (1958) model forecasts as:

$$F_t[y_{t+1}] = \delta y_t + (1-\delta) F_{t-1}[y_{t+1}]\,.$$

If y_t is an AR(1) process with the autocorrelation coefficient ψ, then $y_t = \psi^{-1} E_t[y_{t+1}]$. In this case, it is straightforward to show that:

$$F_t[y_{t+1}] = \sum_{\tau=0}^{\infty} (1-\delta)^\tau \delta \psi^{-1} E_{t-\tau}[y_{t+1}]\,,$$

which clearly satisfies the conditions of Definition 2.

- **Extrapolative Expectations** (Barberis et al., 2015b; Hirshleifer et al., 2015) take the form:

$$F_t[y_{t+1}] \;=\; y_t + \phi(y_t - y_{t-1}) \tag{20}$$

Assuming y_t follows an AR(1) process with coefficient ψ as above, we can rewrite extrapolative expectations as

$$F_t[y_{t+1}] \;=\; \psi^{-1} E_t[y_{t+1}] + \phi(\psi^{-1} E_t[y_{t+1}] - \psi^{-2} E_{t-1}[y_{t+1}]), \tag{21}$$

which again satisfies the conditions of Definition 2.

- **Noisy Information/Sticky Expectations** (e.g., Giannoni and Woodford (2003); Coibion and Gorodnichenko (2012)) postulate

$$F_t[y_{t+n}] \;=\; (1-\lambda)E_t[y_{t+n}] \;+\lambda F_{t-1}[y_{t+n}] + \;\varepsilon_t(n).$$

  Iterating this backward yields:

$$F_t[y_{t+n}] \;=\; \sum_{\tau=0}^{\infty}\lambda^{\tau}(1-\lambda)E_{t-\tau}[y_{t+n}] \;+\; noise_t\,. \tag{22}$$

- **Diagnostic Expectations** of Bordalo et al. (2020b) postulate that

$$F_t[y_{t+n}] = E_t[y_{t+n}] \;+\; \theta(E_t[y_{t+n}] - E_{t-J}[y_{t+n}]). \tag{23}$$

  for some lag $J \geq 1$, which is clearly a special case of Definition 2.

- **Smooth Diagnostic Expectations** introduced in Bianchi et al. (2024a) postulate that:

$$F_t[y_{t+n}] = E_t[y_{t+n}] \;+\; \delta(n)(E_t[y_{t+n}] - E_{t-J}[y_{t+n}]). \tag{24}$$

  for some lag $J \geq 1$, which is clearly a special case of Definition 2. Bianchi et al. (2024a) provide a micro-foundation for the dependence of $\delta(n)$ on n, linking it to the volatility of the underlying stochastic process.

- **Mis-estimation of persistence.** As described in Gabaix (2019), the agent anchors the persistence parameter ψ to a default level of persistence, ψ_*:

$$\hat{\psi} = m\psi + (1-m)\psi_*, \tag{25}$$

where $m \in (0,1)$ controls the degree of irrationality. Assuming y follows an AR(1) process

$$y_{t+1} \;=\; \psi y_t \;+\; \varepsilon_{t+1}, \tag{26}$$

means that:

$$F_t[y_{t+n}] \;=\; (\hat{\psi}/\psi)^{n-1} E_t[y_{t+n}], \tag{27}$$

which falls under Definition 2.

- **Limited Information Processing Capacity.** Afrouzi et al. (2023) develop a model of limited information capacity that generates deviations from FIRE, with the strength of the deviations depending on the persistence ψ of the process as well as the forecasting horizon. They consider a simple AR(1) model,

$$y_t \;=\; (1-\rho)\mu \;+\; \rho y_{t-1} \;+\; \varepsilon_t, \tag{28}$$

so that

$$E_t[y_{t+n}] \;=\; \rho^n y_t \;+\; (1-\rho^n)\mu. \tag{29}$$

With limited information capacity, forecasts are given by:

$$F_t[y_{t+n}] \;=\; \rho^j(\delta_1(n) y_t \;+\; \delta_2(n)\mu), \tag{30}$$

where the dependence of δ_1, δ_2 on n is subtle and nonlinear. We discuss this model in Section 3.3.2. Here, we only note that it obviously satisfies Definition 2:

$$F_t[y_{t+n}] \;=\; \rho^n(\delta_1(n)y_t \;+\; \delta_2(n)\mu) \;=\; \delta_1(n)E_t[y_{t+n}] \;+\; \hat{\mu}(n) \tag{31}$$

for some constants $\hat{\mu}(n)$.

B Proofs for BIR Theory

Proposition 6 *Under the hypotheses of Proposition 1,*

$$\begin{aligned}
&\mathrm{Cov}(y_{t+n},\, F_t[y_{t+n}]) \;=\; \rho^{2n}\sigma_q^2(k_0(n)+\delta(n)\Gamma) \;>\; 0 \\
&\mathrm{Cov}(F_{t+j}[y_{t+n}],\, F_t[y_{t+n}]) \;=\; \sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \;+\; \rho^{2n}\sigma_q^2(k_0(n)k_0(n-j)+k_0(n-j)\delta(n)\Gamma+k_0(n)\delta(n-j)\Gamma(j)) \\
&+\; \rho^{2n}\delta(n-j)\delta(n)\sigma_q^2\Xi(j),
\end{aligned} \tag{32}$$

where

$$\Xi(j) \;=\; \sum_{\tau=0}^{j}\kappa(\tau) \;+\; \sum_{\tau_1=1}^{\infty}\sum_{\tau_2=0}^{\infty}\kappa(\tau_1+j)\kappa(\tau_2)\rho^{2\max(\tau_1,\tau_2)} \tag{33}$$

and

$$\Gamma(j) \;=\; \sum_{\tau=0}^{j}\kappa(\tau)+\sum_{\tau=1}^{\infty}\kappa(\tau+j)\rho^{2\tau} \tag{34}$$

In particular,

$$\begin{aligned}
&\mathrm{Var}[F_t[y_{t+n}]] \;=\; \sigma_\varepsilon^2(n) \;+\; \rho^{2n}\sigma_q^2(k_0(n)^2+k_0(n)\delta(n)\Gamma+k_0(n)\delta(n)\Gamma(0)) \\
&+\; \rho^{2n}\delta(n)^2\sigma_q^2\Xi(0)
\end{aligned} \tag{35}$$

and

$$
\begin{aligned}
&\mathrm{Var}[F_t[y_{t+n}] - F_{t-1}[y_{t+n}]] \\
&= \mathrm{Var}[F_t[y_{t+n}] \ + \ \mathrm{Var}[F_t[y_{t+n+1}] \ + \ 2\mathrm{Cov}(F_t[y_{t+n}], F_{t-1}[y_{t+n}]) \\
&= \sigma_\varepsilon^2(n) \ + \ \rho^{2n}\sigma_q^2(k_0(n)^2 + k_0(n)\delta(n)\Gamma + k_0(n)\delta(n)\Gamma(0)) \\
&+ \ \rho^{2n}\delta(n)^2\sigma_q^2\Xi(0) \\
&+ \ \sigma_\varepsilon^2(n+1) \ + \ \rho^{2(n+1)}\sigma_q^2(k_0(n+1)^2 + k_0(n+1)\delta(n+1)\Gamma + k_0(n+1)\delta(n+1)\Gamma(0)) \\
&+ \ \rho^{2(n+1)}\delta(n+1)^2\sigma_q^2\Xi(0) \\
&- 2\Big(\rho^{2(n+1)}\sigma_q^2(k_0(n+1)k_0(n) + k_0(n)\delta(n+1)\Gamma + k_0(n+1)\delta(n)\Gamma(1)) \\
&+ \ \rho^{2(n+1)}\delta(n)\delta(n+1)\sigma_q^2\Xi(1)\Big)
\end{aligned}
\tag{36}
$$

Proof of Propositions 1 and 6. We have

$$
\begin{aligned}
E_t[F_{t+j}[y_{t+n}]] \ &= \ E_t[\sum_{\tau=0}^{\infty}\gamma(n-j,\tau)\rho^{n+\tau-j}q_{t+j-\tau}] \\
&= \ \sum_{\tau=j+1}^{\infty}\gamma(n-j,\tau)\rho^{n+\tau-j}q_{t+j-\tau} \\
&+ \ q_t\sum_{\tau=0}^{j}\gamma(n-j,\tau)\rho^n \\
&= \ q_t\sum_{\tau=0}^{j}\gamma(n-j,\tau)\rho^n \ + \ \sum_{\tau=1}^{\infty}\gamma(n-j,\tau+j)q_{t-\tau}\rho^{n+\tau}
\end{aligned}
\tag{37}
$$

and

$$
\begin{aligned}
E_t[y_{t+n} - F_{t+j}[y_{t+n}]] \ &= \ E_t[y_{t+n}] - E_t[F_{t+j}[y_{t+n}]] \\
&= \ \rho^n q_t \ - \ q_t\sum_{\tau=0}^{j}\gamma(n-j,\tau)\rho^n \ - \ \sum_{\tau=1}^{\infty}\gamma(n-j,\tau+j)q_{t-\tau}\rho^{n+\tau} \\
&= \ \rho^n q_t\left(1 - \sum_{\tau=0}^{j}\gamma(n-j,\tau)\right) \ - \ \sum_{\tau=1}^{\infty}\gamma(n-j,\tau+j)q_{t-\tau}\rho^{n+\tau}
\end{aligned}
\tag{38}
$$

and hence

$$\begin{aligned}
&\mathrm{Cov}(E_t[y_{t+n} - F_{t+j}[y_{t+n}]], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \\
&= \mathrm{Cov}(E_t[y_{t+n}], F_t[y_{t+n}]) - \mathrm{Cov}(E_t[y_{t+n}], F_{t-1}[y_{t+n}]) \\
&- \mathrm{Cov}(F_{t+j}[y_{t+n}], F_t[y_{t+n}]) + \mathrm{Cov}(F_{t+j}[y_{t+n}], F_{t-1}[y_{t+n}])\,.
\end{aligned} \tag{39}$$

We have

$$\begin{aligned}
&\mathrm{Cov}(y_{t+n}, F_t[y_{t+n}]) = \mathrm{Cov}(E_t[y_{t+n}], F_t[y_{t+n}]) = \rho^{2n}\mathrm{Cov}(q_t, k_0(n)q_t + \delta(n)\sum_{\tau=0}^{\infty}\kappa(\tau)\rho^{\tau}q_{t-\tau}) \\
&= \rho^{2n}\sigma_q^2(k_0(n) + \delta(n)\Gamma)
\end{aligned} \tag{40}$$

and, similarly, by stationarity,

$$\mathrm{Cov}(y_{t+n}, F_{t-1}[y_{t+n}]) = \mathrm{Cov}(y_{t+n+1}, F_t[y_{t+n+1}]) = \rho^{2(n+1)}(k_0(n+1) + \delta(n+1)\Gamma)\,. \tag{41}$$

Then,

$$
\begin{aligned}
\mathrm{Cov}(F_{t+j}[y_{t+n}], F_t[y_{t+n}]) \;&=\; \sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \\
&+\; \mathrm{Cov}(E_t[F_{t+j}[y_{t+n}]], F_t[y_{t+n}]) \\
&\underbrace{=}_{(37)}\; \sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \\
&+\; \mathrm{Cov}(q_t \sum_{\tau=0}^{j} \gamma(n-j,\tau)\rho^n \\
&+\; \sum_{\tau=1}^{\infty} \gamma(n-j,\tau+j) q_{t-\tau}\rho^{n+\tau}, \sum_{\tau=0}^{\infty} \gamma(n,\tau) q_{t-\tau}\rho^{n+\tau}) \\
&=\; \sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \\
&+\; \mathrm{Cov}(q_t k_0(n-j)\rho^n +\; q_t \delta(n-j) \sum_{\tau=0}^{j} \kappa(\tau)\rho^n \\
&+\; \delta(n-j) \sum_{\tau=1}^{\infty} \kappa(\tau+j) q_{t-\tau}\rho^{n+\tau}, \rho^n k_0(n) q_t + \delta(n) \sum_{\tau=0}^{\infty} \kappa(\tau) q_{t-\tau}\rho^{n+\tau}) \\
&=\; \sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \;+\; \rho^{2n}\sigma_q^2(k_0(n)k_0(n-j) + k_0(n-j)\delta(n)\Gamma + k_0(n)\delta(n-j)\Gamma(j)) \\
&+\; \rho^{2n}\delta(n-j)\delta(n)\sigma_q^2 \sum_{\tau=0}^{j} \kappa(\tau)\Gamma \\
&+\; \rho^{2n}\delta(n-j)\delta(n)\sigma_q^2 \sum_{\tau_1=1}^{\infty}\sum_{\tau_2=0}^{\infty} \kappa(\tau_1+j)\kappa(\tau_2)\rho^{2\max(\tau_1,\tau_2)} \\
&=\; \sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \;+\; \rho^{2n}\sigma_q^2(k_0(n)k_0(n-j) + k_0(n-j)\delta(n)\Gamma + k_0(n)\delta(n-j)\Gamma(j)) \\
&+\; \rho^{2n}\delta(n-j)\delta(n)\sigma_q^2\Xi(j)\,,
\end{aligned}
\tag{42}
$$

where we have defined

$$
\Gamma(j) \;=\; \sum_{\tau=0}^{j} \kappa(\tau) + \sum_{\tau=1}^{\infty} \kappa(\tau+j)\rho^{2\tau},\; \Gamma = \Gamma(0). \tag{43}
$$

Thus, by stationarity,

$$\begin{aligned}
&\mathrm{Cov}(F_{t+j}[y_{t+n}], F_{t-1}[y_{t+n}]) \;=\; \mathrm{Cov}(F_{t+j+1}[y_{t+n+1}], F_t[y_{t+n+1}]) \\
&= \;\rho^{2(n+1)}\sigma_q^2(k_0(n+1)k_0(n-j) + k_0(n-j)\delta(n+1)\Gamma + k_0(n+1)\delta(n-j)\Gamma(j+1)) \\
&+ \;\rho^{2(n+1)}\delta(n-j)\delta(n+1)\sigma_q^2\Xi(j+1)
\end{aligned} \tag{44}$$

Thus,

$$\begin{aligned}
&\mathrm{Cov}(y_{t+n} - F_{t+j}[y_{t+n}], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \\
&= \;\mathrm{Cov}(E_t[y_{t+n}], F_t[y_{t+n}]) \;-\; \mathrm{Cov}(E_t[y_{t+n}], F_{t-1}[y_{t+n}]) \\
&- \mathrm{Cov}(F_{t+j}[y_{t+n}], F_t[y_{t+n}]) \;+\; \mathrm{Cov}(F_{t+j}[y_{t+n}], F_{t-1}[y_{t+n}]) \\
&= \;\rho^{2n}\sigma_q^2(k_0(n) + \delta(n)\Gamma) - \rho^{2(n+1)}(k_0(n+1) + \delta(n+1)\Gamma) \\
&- \sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \;-\; \rho^{2n}\sigma_q^2(k_0(n)k_0(n-j) + k_0(n-j)\delta(n)\Gamma + k_0(n)\delta(n-j)\Gamma(j)) \\
&- \;\rho^{2n}\delta(n-j)\delta(n)\sigma_q^2\Xi(j) \\
&+ \;\rho^{2(n+1)}\sigma_q^2(k_0(n+1)k_0(n-j) + k_0(n-j)\delta(n+1)\Gamma + k_0(n+1)\delta(n-j)\Gamma(j+1)) \\
&+ \;\rho^{2(n+1)}\delta(n-j)\delta(n+1)\sigma_q^2\Xi(j+1) \\
&= \;-\sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \\
&+ \;\rho^{2n}\sigma_q^2(1 - k_0(n-j))((k_0(n) - \rho^2 k_0(n+1)) + \Gamma(\delta(n) - \rho^2\delta(n+1))) \\
&- \rho^{2n}\,\sigma_q^2((k_0(n)\delta(n-j)\Gamma(j) - \rho^2 k_0(n+1)\delta(n-j)\Gamma(j+1)))) \\
&- \;\rho^{2n}\,\sigma_q^2\,\delta(n-j)(\delta(n)\Xi(j) \;- \rho^2\delta(n+1)\Xi(j+1)) \\
&= \;-\sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \\
&+ \;\rho^{2n}\sigma_q^2(1 - k_0(n-j))((k_0(n) - \rho^2 k_0(n+1)) + \Gamma(\delta(n) - \rho^2\delta(n+1))) \\
&- \rho^{2n}\,\sigma_q^2\delta(n-j)\Big[(k_0(n)\Gamma(j) - \rho^2 k_0(n+1)\Gamma(j+1)) \\
&+ (\delta(n)\Xi(j) \;- \rho^2\delta(n+1)\Xi(j+1))\Big]
\end{aligned} \tag{45}$$

Then, the regression coefficient is given by

$$\frac{\text{Cov}(y_{t+n} - F_{t+j}[y_{t+n}], F_t[y_{t+n}] - F_{t-1}[y_{t+n}])}{\text{Var}[F_t[y_{t+n}] - F_{t-1}[y_{t+n}]]} \tag{46}$$

and the claim follows once we define

$$\begin{aligned}
D(n) &= D_{MZ}(n) + \rho^2 D_{MZ}(n+1) - 2\rho^2(k_0(n+1)k_0(n) + k_0(n)\delta(n+1) + k_0(n+1)\delta(n)\Gamma(1) \\
&\quad + \delta(n)\delta(n+1)\Xi(1)) \\
D_{MZ}(n) &= k_0(n)^2 + 2k_0(n)\delta(n) + \delta(n)^2\Xi(0) \\
\Xi(j) &= \sum_{\tau=0}^{j} \kappa(\tau) \; + \; \sum_{\tau_1=1}^{\infty}\sum_{\tau_2=0}^{\infty} \kappa(\tau_1 + j)\kappa(\tau_2)\rho^{2\max(\tau_1,\tau_2)} \\
\Gamma(j) &= \sum_{\tau=0}^{j} \kappa(\tau) + \sum_{\tau=1}^{\infty} \kappa(\tau + j)\rho^{2\tau}\,.
\end{aligned} \tag{47}$$

□

Proof of Proposition 3. The claim follows directly from Proposition 1. □

Corollary 7 (BIRs in (Smooth) Diagnostic Expectations) *Suppose that*

$$F_t[y_{t+n}] \;=\; E_t[y_{t+n}] \;+\; \theta(n)(E_t[y_{t+n}] - E_{t-J}[y_{t+n}]). \tag{48}$$

Then,

$$\begin{aligned}
\Xi(j) &= (1-\rho^{2J})^{-1}(1 - (1-\rho^{2J})^{-1}(\rho^{2(J-j)} - \rho^{2J}))\mathbf{1}_{j<J} \\
\Gamma(j) &= (1-\rho^{2J})^{-1}(1 - \rho^{2(J-j)})\mathbf{1}_{j<J}
\end{aligned} \tag{49}$$

and

$$\kappa(\tau) = (1-\rho^{2J})^{-1} \begin{cases} 1, & \textit{if } \tau = 0 \\ -1, & \textit{if } \tau = J \\ 0, & \textit{if } \tau \notin \{0, J\} \end{cases}$$

and $\delta(n) = \theta(n)(1-\rho^{2J})$.

Proof of Corollaries 2 and 7. We perform the proof directly for smooth diagnostic expectations.

We have that $\kappa(\tau)$ is given by:

$$\kappa(\tau) = (1-\rho^{2J})^{-1} \begin{cases} 1, & \text{if } \tau = 0 \\ -1, & \text{if } \tau = J \\ 0, & \text{if } \tau \notin \{0, J\} \end{cases}$$

and $\delta(n) = \theta(1-\rho^{2J})$. Then,

$$\gamma(n,\tau) \;=\; \mathbf{1}_{\tau=0} \;+\; \delta(n)\kappa(\tau)\,. \tag{50}$$

Then,

$$
\begin{aligned}
\Xi(j) &= \sum_{\tau=0}^{j} \kappa(\tau) + \sum_{\tau_1=1}^{\infty}\sum_{\tau_2=0}^{\infty} \kappa(\tau_1+j)\kappa(\tau_2)\rho^{2\max(\tau_1,\tau_2)} \\
&= (1-\rho^{2J})^{-1}\mathbf{1}_{j<J} - (1-\rho^{2J})^{-2}\sum_{\tau_1=1}^{\infty}\mathbf{1}_{\tau_1\geq 1}\mathbf{1}_{\tau_1+j=J}(\rho^{2\max(\tau_1,0)}-\rho^{2\max(\tau_1,J)}) \\
&= (1-\rho^{2J})^{-1}\mathbf{1}_{j<J} - (1-\rho^{2J})^{-2}(\rho^{2(J-j)}-\rho^{2J})\mathbf{1}_{j<J} \\
&= (1-\rho^{2J})^{-1}(1-(1-\rho^{2J})^{-1}(\rho^{2(J-j)}-\rho^{2J}))\mathbf{1}_{j<J} \\
\Gamma(j) &= \sum_{\tau=0}^{j}\kappa(\tau) + \sum_{\tau=1}^{\infty}\kappa(\tau+j)\rho^{2\tau} \\
&= (1-\rho^{2J})^{-1}(1-\rho^{2(J-j)})\mathbf{1}_{j<J}
\end{aligned}
\tag{51}
$$

□

Proof of Corollary 4. Recall that

$$
B(n,j) = \frac{A(n,j)-\rho^2 A(n+1,j+1)}{D(n)}, \tag{52}
$$

where

$$
\begin{aligned}
A(n,j) &= -\delta(n-j)\Big[\Gamma(j)+\delta(n)\Xi(j)\Big] \\
D(n) &= D_{MZ}(n)+\rho^2 D_{MZ}(n+1) - 2\rho^2(1+\delta(n+1) \\
&\quad + \delta(n)\Gamma(1)+\delta(n)\delta(n+1)\Xi(1)) \\
D_{MZ}(n) &= 1+2\delta(n)+\delta(n)^2\Xi(0)
\end{aligned}
\tag{53}
$$

Thus, for $J > 1$, we have

$$\begin{aligned} A(n,j) &= -\theta(1+\theta)[1-\rho^{2(J-j)}]\mathbf{1}_{j<J} \\ D_{MZ} &= 1 + 2\theta(1-\rho^{2J}) + \theta^2(1-\rho^{2J}) \\ D(n) &= (1+\rho^2)D_{MZ} - 2\rho^2(1+\theta(1-\rho^{2J}) + \theta(1+\theta)(1-\rho^{2(J-1)})) \end{aligned} \tag{54}$$

where

$$\begin{aligned} \Xi(j) &= \sum_{\tau=0}^{j} \kappa(\tau) \; + \; \sum_{\tau_1=1}^{\infty}\sum_{\tau_2=0}^{\infty} \kappa(\tau_1+j)\kappa(\tau_2)\rho^{2\max(\tau_1,\tau_2)} \\ \Gamma(j) &= \sum_{\tau=0}^{j} \kappa(\tau) + \sum_{\tau=1}^{\infty} \kappa(\tau+j)\rho^{2\tau} \end{aligned} \tag{55}$$

where $\kappa(\tau) = \delta_{0,\tau}$, so that

$$\begin{aligned} \Xi(j) &= 1 \\ \Gamma(j) &= 1\,. \end{aligned} \tag{56}$$

The claim follows. □

Corollary 8 *For $J > 1$, we have*

$$B(n,j) = const \begin{cases} 1-\rho^2, & j \le J-1 \\ 0, & j > J-1 \end{cases} \tag{57}$$

where

$$const = -\theta(1+\theta)((1+\rho^2)D_{MZ} - 2\rho^2(1+\theta(1-\rho^{2J}) + \theta(1+\theta)(1-\rho^{2(J-1)})))^{-1} \tag{58}$$

with

$$D_{MZ} = 1 + 2\theta(1-\rho^{2J}) + \theta^2(1-\rho^{2J}) \tag{59}$$

Proof of Corollary 8. Recall that

$$B(n,j) \;=\; \frac{A(n,j) - \rho^2 A(n+1,j+1)}{D(n)}, \tag{60}$$

where

$$\begin{aligned} A(n,j) &= -\delta(n-j)\Big[\Gamma(j) + \delta(n)\Xi(j)\Big] \\ D(n) &= D_{MZ}(n) + \rho^2 D_{MZ}(n+1) - 2\rho^2(1 + \delta(n+1) \\ &\quad + \delta(n)\Gamma(1) + \delta(n)\delta(n+1)\Xi(1)) \\ D_{MZ}(n) &= 1 + 2\delta(n) + \delta(n)^2\Xi(0) \end{aligned} \tag{61}$$

Thus, for $J > 1$, we have

$$\begin{aligned} A(n,j) \;&=\; -\theta(1+\theta)[1-\rho^{2(J-j)}]\mathbf{1}_{j<J} \\ D_{MZ} &= 1 + 2\theta(1-\rho^{2J}) + \theta^2(1-\rho^{2J}) \\ D(n) \;&=\; (1+\rho^2)D_{MZ} - 2\rho^2(1 + \theta(1-\rho^{2J}) + \theta(1+\theta)(1-\rho^{2(J-1)})) \end{aligned} \tag{62}$$

□

C Extrapolative Expectations

Recall that, in the main text (Corollary 5), we study simple extrapolative expectations in the following setting: The true data generating process has a constant mean so that

$y_{t+1} = \bar{y} + \varepsilon_{t+1}$, where ε_{t+1} is i.i.d., while forecasts are assumed to evolve according to:

$$\begin{aligned} F_t[y_{t+1}] &= \mu_t \\ \mu_{t+1} &= (1-\rho-\tilde{\rho})\bar{\mu} + \rho\mu_t + \tilde{\rho}y_{t+1}, \end{aligned} \tag{63}$$

where $\tilde{\rho}$ modulates the degree of extrapolation from the most recent realization of y, ρ controls the impact of extrapolation made from the past.

In this appendix, we introduce an extension of the model of extrapolative expectations studied in the main text, allowing for persistence in the underlying process.

Corollary 9 *Suppose that* $y_{t+1} = \bar{y} + \rho_y(y_t - \bar{y}) + \varepsilon_{t+1}$ *with* $|\rho_y| < 1$ *and* ε_{t+1} *are i.i.d., whereas the agent believes*

$$\begin{aligned} F_t[y_{t+1}] &= \mu_t \\ \mu_{t+1} &= (1-\rho-\tilde{\rho})\bar{\mu} + \rho\mu_t + \tilde{\rho}y_{t+1} \\ F_t[\mu_{t+1}] &= (1-\rho-\tilde{\rho})\bar{\mu} + (\rho+\tilde{\rho})\mu_t\,, \end{aligned} \tag{64}$$

with $|\rho| < 1$. *The model is linear, autonomous, and separable, with*

$$\delta(n) = \frac{\tilde{\rho}}{\rho+\tilde{\rho}}\frac{1}{1-\rho\rho_y}((\rho+\tilde{\rho})/\rho_y)^n,\ \kappa(\tau) = (1-\rho\rho_y)(\rho/\rho_y)^\tau \tag{65}$$

We have

$$B(n,j) = \frac{A(n,j)}{D(n)}, \tag{66}$$

where

$$\begin{aligned} A(n,j) &= \mathrm{Cov}(y_{t+n} - F_{t+j}[y_{t+n}], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \\ &= \hat{\rho}^{n-1}\left(\rho_y^n - \hat{\rho}^{n-j-1}\tilde{\rho}\frac{\rho^{j+1} - \rho_y^{j+1}}{\rho - \rho_y}\right)\left(\tilde{\rho} - \tilde{\rho}^2\rho_y\frac{1}{1-\rho_y\rho}\right) \\ &- \hat{\rho}^{n-1}\hat{\rho}^{n-j-1}\tilde{\rho}\left(\tilde{\rho}\rho^j\frac{\rho\rho_y}{1-\rho\rho_y} - \rho^j\tilde{\rho}^2\frac{\rho(1+\rho\rho_y)}{(1-\rho^2)(1-\rho\rho_y)}\right) \end{aligned} \tag{67}$$

and

$$\begin{aligned} D(n) &= \mathrm{Var}[F_t[y_{t+n}] - F_{t-1}[y_{t+n}]] \\ &= \hat{\rho}^{2(n-1)}\tilde{\rho}^2\left(\frac{1-\rho\hat{\rho} - \hat{\rho}(\rho_y - \tilde{\rho} - \rho\rho_y\hat{\rho})}{(1-\rho^2)(1-\rho\rho_y)}\right). \end{aligned} \tag{68}$$

The rational expectations case corresponds to $\rho = 0, \tilde{\rho} = \rho_y$.

Proof of Corollaries 5 and 9. We have

$$F_t[y_{t+n}] = F_t[\mu_{t+n-1}] = const + (\rho + \tilde{\rho})^{n-1}\mu_t \tag{69}$$

Without loss of generality, we normalize $\mathrm{Var}[y_t] = 1$. The AR(1) structure of y_t implies that

$$\mathrm{Cov}(y_t, y_{t+h}) = \rho_y^{|h|} \tag{70}$$

We have

$$\mu_t = \sum_{\tau=0}^{\infty}\rho^\tau\tilde{\rho}y_{t-\tau} + const\,. \tag{71}$$

and

$$E_t[y_{t+n}] \;=\; const \;+\; \rho_y^n y_t \tag{72}$$

and, hence,

$$E_{t-\tau}[y_{t+n}] \;=\; const \;+\; \rho_y^{n+\tau} y_{t-\tau} \tag{73}$$

Let

$$\hat{\rho} = (\rho + \tilde{\rho}) \tag{74}$$

Thus,

$$\begin{aligned}
\mathrm{Cov}(y_{t+n}, F_t[y_{t+n}]) \;&=\; (\rho+\tilde{\rho})^{n-1}\mathrm{Cov}(y_{t+n}, \mu_t) \\
&\underbrace{=}_{\mathrm{Cov}(y_{t+n},y_{t-\tau})=\rho_y^{n+\tau}\ by\ (70)} (\rho+\tilde{\rho})^{n-1}\mathrm{Cov}(y_{t+n}, \sum_{\tau=0}^{\infty}\rho^\tau\tilde{\rho}y_{t-\tau}) \\
&=\; (\rho+\tilde{\rho})^{n-1}\tilde{\rho}\sum_{\tau=0}^{\infty}\rho^\tau\rho_y^{\tau+n}\sigma_y^2 \\
&=\; (\rho+\tilde{\rho})^{n-1}\tilde{\rho}\rho_y^n\frac{1}{1-\rho\rho_y}\sigma_y^2,
\end{aligned} \tag{75}$$

while, by (69), with $b = \rho/\rho_y$, we get

$$
\begin{aligned}
E_t[F_{t+j}[y_{t+n}]] \;&=\; E_t[(\rho+\tilde\rho)^{n-j-1}\mu_{t+j}] \;=\; E_t[(\rho+\tilde\rho)^{n-j-1}\mu_{t+j}] \\
&=\; E_t[(\rho+\tilde\rho)^{n-j-1}\mu_{t+j}] \\
&\underbrace{=}_{(71)}\; const + E_t[(\rho+\tilde\rho)^{n-j-1}\sum_{\tau=0}^{\infty}\rho^\tau\tilde\rho y_{t-\tau+j}] \\
&=\; const \;+\; (\rho+\tilde\rho)^{n-j-1}\sum_{\tau=0}^{j}\rho^\tau\tilde\rho\rho_y^{j-\tau}y_t \;+\; (\rho+\tilde\rho)^{n-j-1}\tilde\rho\sum_{\tau=j+1}^{\infty}\rho^\tau y_{t-\tau+j} \\
&=\; const \;+\; \hat\rho^{n-j-1}\tilde\rho\rho_y^j\frac{b^{j+1}-1}{b-1}y_t \;+\; \hat\rho^{n-j-1}\tilde\rho\sum_{\tau=j+1}^{\infty}\rho^\tau y_{t-\tau+j}\,,
\end{aligned}
\tag{76}
$$

and, hence,

$$
\begin{aligned}
&\mathrm{Cov}(y_{t+n}-F_{t+j}[y_{t+n}], F_t[y_{t+n}]-F_{t-1}[y_{t+n}]) \\
&=\; \mathrm{Cov}(E_t[y_{t+n}-F_{t+j}[y_{t+n}]], F_t[y_{t+n}]-F_{t-1}[y_{t+n}]) \\
&=\; \mathrm{Cov}\left(\rho_y^n y_t - \hat\rho^{n-j-1}\tilde\rho\Big(\rho_y^j\frac{b^{j+1}-1}{b-1}y_t \;+\; \sum_{\tau=j+1}^{\infty}\rho^\tau y_{t-\tau+j}\Big), \hat\rho^{n-1}\mu_t - \hat\rho^n\mu_{t-1}\right).
\end{aligned}
\tag{77}
$$

Now, by (71),

$$
\begin{aligned}
\mu_t - \hat\rho\mu_{t-1} \;&=\; const \;+\; \sum_{\tau=0}^{\infty}\rho^\tau\tilde\rho y_{t-\tau} \\
&- \hat\rho\sum_{\tau=0}^{\infty}\rho^\tau\tilde\rho y_{t-1-\tau} \\
&=\; const \;+\; \tilde\rho y_t \;+\sum_{\tau=1}^{\infty}\rho^{\tau-1}(\rho-\hat\rho)\tilde\rho y_{t-\tau} \\
&=\; const + \tilde\rho y_t - \sum_{\tau=1}^{\infty}\rho^{\tau-1}\tilde\rho^2 y_{t-\tau}
\end{aligned}
\tag{78}
$$

because

$$\rho - \hat{\rho} \;=\; -\tilde{\rho}. \tag{79}$$

Without loss of generality, we normalize $\mathrm{Var}[y_t] = \sigma_y^2 = 1$. Then,

$$\begin{aligned}
&\mathrm{Cov}(y_{t+n} - F_{t+j}[y_{t+n}], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \\
&= \mathrm{Cov}(E_t[y_{t+n} - F_{t+j}[y_{t+n}]], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \\
&= \hat{\rho}^{n-1}\mathrm{Cov}\left(\rho_y^n y_t - \hat{\rho}^{n-j-1}\tilde{\rho}\Big(\rho_y^j \frac{b^{j+1}-1}{b-1} y_t \;+\; \sum_{\tau=j+1}^{\infty} \rho^\tau y_{t-\tau+j}\Big), \mu_t - \hat{\rho}\mu_{t-1}\right) \\
&= \hat{\rho}^{n-1}\left(\rho_y^n - \hat{\rho}^{n-j-1}\tilde{\rho}\rho_y^j \frac{b^{j+1}-1}{b-1}\right)\mathrm{Cov}\left(y_t, \tilde{\rho} y_t - \sum_{\tau=1}^{\infty} \rho^{\tau-1}\tilde{\rho}^2 y_{t-\tau}\right) \\
&- \hat{\rho}^{n-1}\hat{\rho}^{n-j-1}\tilde{\rho}\mathrm{Cov}\left(\sum_{\tau=j+1}^{\infty} \rho^\tau y_{t-\tau+j}, \tilde{\rho} y_t - \sum_{\tau=1}^{\infty} \rho^{\tau-1}\tilde{\rho}^2 y_{t-\tau}\right).
\end{aligned} \tag{80}$$

Now,

$$\begin{aligned}
&\mathrm{Cov}\left(y_t, \tilde{\rho} y_t - \sum_{\tau=1}^{\infty} \rho^{\tau-1}\tilde{\rho}^2 y_{t-\tau}\right) \\
&= \tilde{\rho} \;-\; \sum_{\tau=1}^{\infty} \rho_y^\tau \rho^{\tau-1}\tilde{\rho}^2 \\
&= \tilde{\rho} - \tilde{\rho}^2 \rho_y \frac{1}{1-\rho_y\rho}\,.
\end{aligned} \tag{81}$$

Similarly,

$$
\begin{aligned}
&\mathrm{Cov}\left(\sum_{\tau_1=j+1}^{\infty}\rho^{\tau_1}y_{t-\tau_1+j},\tilde{\rho}y_t-\sum_{\tau=1}^{\infty}\rho^{\tau-1}\tilde{\rho}^2y_{t-\tau}\right)\\
&=\sum_{\tau_1=j+1}^{\infty}\rho^{\tau_1}\mathrm{Cov}\left(y_{t-\tau_1+j},\tilde{\rho}y_t-\sum_{\tau=1}^{\infty}\rho^{\tau-1}\tilde{\rho}^2y_{t-\tau}\right)\\
&=\tilde{\rho}\sum_{\tau_1=j+1}^{\infty}\rho^{\tau_1}\rho_y^{|-\tau_1+j|}\\
&-\sum_{\tau_1=j+1}^{\infty}\rho^{\tau_1}\sum_{\tau=1}^{\infty}\rho^{\tau-1}\tilde{\rho}^2\rho_y^{|\tau-\tau_1+j|}\\
&=\tilde{\rho}\sum_{\tau=1}^{\infty}\rho^{\tau+j}\rho_y^{\tau}\;-\;\sum_{\tau_1=1}^{\infty}\rho^{\tau_1+j}\sum_{\tau=1}^{\infty}\rho^{\tau-1}\tilde{\rho}^2\rho_y^{|\tau-\tau_1|}\\
&=\tilde{\rho}\rho^j\frac{\rho\rho_y}{1-\rho\rho_y}\\
&-\rho^j\tilde{\rho}^2\sum_{\tau_1=1}^{\infty}\rho^{\tau_1}\left(\sum_{\tau=1}^{\tau_1}\rho^{\tau-1}\rho_y^{\tau_1-\tau}\;+\;\sum_{\tau=\tau_1+1}^{\infty}\rho^{\tau-1}\rho_y^{\tau-\tau_1}\right) \qquad (82)\\
&=\tilde{\rho}\rho^j\frac{\rho\rho_y}{1-\rho\rho_y}-\rho^j\tilde{\rho}^2\sum_{\tau_1=1}^{\infty}\rho^{\tau_1}\left(\rho_y^{\tau_1}(1/\rho_y)\frac{(\rho/\rho_y)^{\tau_1}-1}{(\rho/\rho_y)-1}\;+\;\sum_{\tau=1}^{\infty}\rho^{\tau_1+\tau-1}\rho_y^{\tau}\right)\\
&=\tilde{\rho}\rho^j\frac{\rho\rho_y}{1-\rho\rho_y}-\rho^j\tilde{\rho}^2\sum_{\tau_1=1}^{\infty}\rho^{\tau_1}\left(\frac{\rho^{\tau_1}-\rho_y^{\tau_1}}{\rho-\rho_y}\;+\;\rho^{\tau_1}\rho_y\frac{1}{1-\rho\rho_y}\right)\\
&=\tilde{\rho}\rho^j\frac{\rho\rho_y}{1-\rho\rho_y}-\rho^j\tilde{\rho}^2\sum_{\tau_1=1}^{\infty}\frac{\rho^{2\tau_1}-(\rho\rho_y)^{\tau_1}}{\rho-\rho_y}\\
&-\rho^j\tilde{\rho}^2\sum_{\tau_1=1}^{\infty}\rho^{2\tau_1}\rho_y\frac{1}{1-\rho\rho_y}\\
&=\tilde{\rho}\rho^j\frac{\rho\rho_y}{1-\rho\rho_y}-\rho^j\tilde{\rho}^2\frac{\frac{\rho^2}{1-\rho^2}-\frac{\rho\rho_y}{1-\rho\rho_y}}{\rho-\rho_y}\\
&-\rho^j\tilde{\rho}^2\frac{\rho^2}{1-\rho^2}\rho_y\frac{1}{1-\rho\rho_y}
\end{aligned}
$$

A final transformation gives

$$\frac{\frac{\rho^2}{1-\rho^2} - \frac{\rho\rho_y}{1-\rho\rho_y}}{\rho - \rho_y} \;=\; \frac{\rho^2(1-\rho\rho_y) - \rho\rho_y(1-\rho^2)}{(1-\rho^2)(1-\rho\rho_y)(\rho-\rho_y)} \;=\; \frac{\rho}{(1-\rho^2)(1-\rho\rho_y)} \tag{83}$$

and hence,

$$\begin{aligned} &\mathrm{Cov}\left(\sum_{\tau=j+1}^{\infty} \rho^{\tau} y_{t-\tau+j}, \tilde{\rho} y_t - \sum_{\tau=1}^{\infty} \rho^{\tau-1}\tilde{\rho}^2 y_{t-\tau}\right) \\ &= \tilde{\rho}\rho^j \frac{\rho\rho_y}{1-\rho\rho_y} - \rho^j\tilde{\rho}^2 \frac{\rho(1+\rho\rho_y)}{(1-\rho^2)(1-\rho\rho_y)}. \end{aligned} \tag{84}$$

Note also that

$$\rho_y^j \frac{b^{j+1}-1}{b-1} \;=\; \frac{\rho^{j+1} - \rho_y^{j+1}}{\rho - \rho_y} \tag{85}$$

so that, summarizing, we get

$$\begin{aligned} &\mathrm{Cov}(y_{t+n} - F_{t+j}[y_{t+n}], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \\ &= \mathrm{Cov}(E_t[y_{t+n} - F_{t+j}[y_{t+n}]], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \\ &= \mathrm{Cov}\left(\rho_y^n y_t - \hat{\rho}^{n-j-1}\tilde{\rho}\Big(\rho_y^j \frac{b^{j+1}-1}{b-1} y_t \;+\; \sum_{\tau=j+1}^{\infty} \rho^{\tau} y_{t-\tau+j}\Big), \hat{\rho}^{n-1}\mu_t - \hat{\rho}^n \mu_{t-1}\right) \\ &= \hat{\rho}^{n-1}\left(\rho_y^n - \hat{\rho}^{n-j-1}\tilde{\rho}\frac{\rho^{j+1}-\rho_y^{j+1}}{\rho-\rho_y}\right)\left(\tilde{\rho} - \tilde{\rho}^2\rho_y\frac{1}{1-\rho_y\rho}\right) \\ &- \hat{\rho}^{n-1}\hat{\rho}^{n-j-1}\tilde{\rho}\left(\tilde{\rho}\rho^j\frac{\rho\rho_y}{1-\rho\rho_y} - \rho^j\tilde{\rho}^2\frac{\rho(1+\rho\rho_y)}{(1-\rho^2)(1-\rho\rho_y)}\right) \end{aligned} \tag{86}$$

Similarly,

$$\begin{aligned} \mathrm{Var}[F_t[y_{t+n}] - F_{t-1}[y_{t+n}]] \;&=\; \mathrm{Cov}(F_t[y_{t+n}] - F_{t-1}[y_{t+n}], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \\ &= \mathrm{Cov}(F_t[y_{t+n}] - F_{t-1}[y_{t+n}], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \end{aligned} \tag{87}$$

Above, we have computed

$$\begin{aligned}
&\mathrm{Cov}(F_t[y_{t+n}], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \\
&= \mathrm{Cov}\left(\hat{\rho}^{n-1}\tilde{\rho}\Big(y_t + \sum_{\tau=1}^{\infty}\rho^{\tau}y_{t-\tau}\Big), \hat{\rho}^{n-1}\mu_t - \hat{\rho}^{n}\mu_{t-1}\right) \\
&= \hat{\rho}^{2(n-1)}\tilde{\rho}^2\mathrm{Cov}\left(\Big(y_t + \sum_{\tau=1}^{\infty}\rho^{\tau}y_{t-\tau}\Big), y_t - \sum_{\tau=1}^{\infty}\rho^{\tau-1}\tilde{\rho}y_{t-\tau}\right)
\end{aligned} \tag{88}$$

From (84), we know that

$$\begin{aligned}
&\mathrm{Cov}\left(\sum_{\tau=1}^{\infty}\rho^{\tau}y_{t-\tau}, \tilde{\rho}y_t - \sum_{\tau=1}^{\infty}\rho^{\tau-1}\tilde{\rho}^2y_{t-\tau}\right) \\
&= \tilde{\rho}\frac{\rho\rho_y}{1-\rho\rho_y} - \tilde{\rho}^2\frac{\rho(1+\rho\rho_y)}{(1-\rho^2)(1-\rho\rho_y)},
\end{aligned} \tag{89}$$

while, by (81),

$$\begin{aligned}
&\mathrm{Cov}\left(y_t, \tilde{\rho}y_t - \sum_{\tau=1}^{\infty}\rho^{\tau-1}\tilde{\rho}^2y_{t-\tau}\right) \\
&= \tilde{\rho} - \tilde{\rho}^2\rho_y\frac{1}{1-\rho_y\rho},
\end{aligned} \tag{90}$$

implying that

$$
\begin{aligned}
&\mathrm{Cov}(F_t[y_{t+n}], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \\
&= \hat{\rho}^{2(n-1)}\tilde{\rho}^2\mathrm{Cov}\left(\left(y_t \ + \ \sum_{\tau=1}^{\infty}\rho^{\tau}y_{t-\tau}\right), y_t - \sum_{\tau=1}^{\infty}\rho^{\tau-1}\tilde{\rho}y_{t-\tau}\right) \\
&= \hat{\rho}^{2(n-1)}\tilde{\rho}^2\left(\frac{\rho\rho_y}{1-\rho\rho_y} - \tilde{\rho}\frac{\rho(1+\rho\rho_y)}{(1-\rho^2)(1-\rho\rho_y)}\right. \\
&\left. + 1 \ - \ \tilde{\rho}\rho_y\frac{1}{1-\rho_y\rho}\right) \\
&= \hat{\rho}^{2(n-1)}\tilde{\rho}^2\left(\frac{\rho\rho_y}{1-\rho\rho_y} - \tilde{\rho}\frac{\rho(1+\rho\rho_y)}{(1-\rho^2)(1-\rho\rho_y)}\right. \\
&\left. + 1 \ - \ \tilde{\rho}\rho_y\frac{1}{1-\rho_y\rho}\right) \\
&= \hat{\rho}^{2(n-1)}\tilde{\rho}^2\left(\frac{1}{1-\rho\rho_y} - \tilde{\rho}\frac{\rho}{(1-\rho^2)(1-\rho\rho_y)}\right) \\
&= \hat{\rho}^{2(n-1)}\tilde{\rho}^2\left(\frac{1-\rho\tilde{\rho}-\rho^2}{(1-\rho^2)(1-\rho\rho_y)}\right)
\end{aligned}
\tag{91}
$$

Similarly, by (89),

$$
\begin{aligned}
&\mathrm{Cov}(F_{t-1}[y_{t+n}], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \\
&= \hat{\rho}^{2n-1}\tilde{\rho}^2\mathrm{Cov}\left(\left(y_{t-1} \ + \ \sum_{\tau=1}^{\infty}\rho^{\tau}y_{t-1-\tau}\right), y_t - \sum_{\tau=1}^{\infty}\rho^{\tau-1}\tilde{\rho}y_{t-\tau}\right) \\
&= \hat{\rho}^{2n-1}\tilde{\rho}^2\rho^{-1}\mathrm{Cov}\left(\left(\sum_{\tau=1}^{\infty}\rho^{\tau}y_{t-\tau}\right), y_t - \sum_{\tau=1}^{\infty}\rho^{\tau-1}\tilde{\rho}y_{t-\tau}\right) \\
&= \hat{\rho}^{2n-1}\tilde{\rho}^2\rho^{-1}\left(\frac{\rho\rho_y}{1-\rho\rho_y} - \tilde{\rho}\frac{\rho(1+\rho\rho_y)}{(1-\rho^2)(1-\rho\rho_y)}\right) \\
&= \hat{\rho}^{2n-1}\tilde{\rho}^2\rho^{-1}\left(\frac{\rho\rho_y}{1-\rho\rho_y} - \tilde{\rho}\frac{\rho(1+\rho\rho_y)}{(1-\rho^2)(1-\rho\rho_y)}\right) \\
&= \hat{\rho}^{2n-1}\tilde{\rho}^2\frac{\rho_y - \tilde{\rho} - \rho\rho_y(\rho+\tilde{\rho})}{(1-\rho^2)(1-\rho\rho_y)}
\end{aligned}
\tag{92}
$$

Thus,

$$\begin{aligned}
&\mathrm{Var}[F_t[y_{t+n}] - F_{t-1}[y_{t+n}]] \\
&= \hat{\rho}^{2(n-1)}\tilde{\rho}^2\left(\frac{1-\rho\tilde{\rho}-\rho^2}{(1-\rho^2)(1-\rho\rho_y)}\right) - \hat{\rho}^{2n-1}\tilde{\rho}^2\frac{\rho_y-\tilde{\rho}-\rho\rho_y(\rho+\tilde{\rho})}{(1-\rho^2)(1-\rho\rho_y)} \\
&= \hat{\rho}^{2(n-1)}\tilde{\rho}^2\left(\frac{1-\rho\tilde{\rho}-\rho^2}{(1-\rho^2)(1-\rho\rho_y)} - \hat{\rho}\frac{\rho_y-\tilde{\rho}-\rho\rho_y(\rho+\tilde{\rho})}{(1-\rho^2)(1-\rho\rho_y)}\right) \\
&= \hat{\rho}^{2(n-1)}\tilde{\rho}^2\left(\frac{1-\rho\hat{\rho}-\hat{\rho}(\rho_y-\tilde{\rho}-\rho\rho_y\hat{\rho})}{(1-\rho^2)(1-\rho\rho_y)}\right).
\end{aligned} \tag{93}$$

In the simpler case of $\rho_y = 0$, we have

$$F_t[y_{t+n}] \;=\; (\rho+\tilde{\rho})^{n-1}\mu_t,\;\; \mu_t \;=\; const \;+\; \sum_{\tau=0}^{\infty}\rho^{\tau}\tilde{\rho}y_{t-\tau} \tag{94}$$

while

$$E_t[\mu_{t+1}] \;=\; const \;+\; \rho\mu_t \tag{95}$$

$$\begin{aligned}
&E_t[F_{t+j}[y_{t+n}]] \;=\; const \;+\; (\rho+\tilde{\rho})^{n-1-j}\rho^j\mu_t \\
&E_t[y_{t+n} - F_{t+j}[y_{t+n}]] \;=\; \bar{\mu} \;-\; const \;-\; (\rho+\tilde{\rho})^{n-1-j}\rho^j\mu_t \\
&\mathrm{Cov}(E_t[y_{t+n} - F_{t+j}[y_{t+n}]], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \\
&= \mathrm{Cov}(-(\rho+\tilde{\rho})^{n-1-j}\rho^j\mu_t, (\rho+\tilde{\rho})^{n-1}\mu_t - (\rho+\tilde{\rho})^n\mu_{t-1}) \\
&= -(\rho+\tilde{\rho})^{2n-2-j}\rho^j\mathrm{Var}[\mu] + (\rho+\tilde{\rho})^{2n-1-j}\rho^{j+1}\mathrm{Var}[\mu] \\
&\mathrm{Var}[F_t[y_{t+n}] - F_{t-1}[y_{t+n}]] \;=\; \mathrm{Var}[\mu]((\rho+\tilde{\rho})^{2(n-1)} + (\rho+\tilde{\rho})^{2n} - 2(\rho+\tilde{\rho})^{2n-1}\rho)
\end{aligned} \tag{96}$$

so that

$$B(n,j) \;=\; \frac{-(\rho+\tilde{\rho})^{2n-2-j}\rho^{j}+(\rho+\tilde{\rho})^{2n-1-j}\rho^{j+1}}{(\rho+\tilde{\rho})^{2(n-1)}+(\rho+\tilde{\rho})^{2n}-2(\rho+\tilde{\rho})^{2n-1}\rho} \;=\; \rho^{j}\frac{\rho(\rho+\tilde{\rho})-1}{(\rho+\tilde{\rho})^{j}(1+(\rho+\tilde{\rho})^{2}-2\rho(\rho+\tilde{\rho}))} \tag{97}$$

□