Behavioral Impulse Responses

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

This version: August 28, 2024

Abstract

We develop the concept of a Behavioral Impulse Response (BIR), which tracks how future forecast errors and revisions of a fixed target dynamically evolve in reaction to current news. BIRs offer more powerful tests of Full-Information Rational Expectations (FIRE) compared to existing methods and quantify how quickly deviations from FIRE are corrected over time. Utilizing BIRs estimated from the Survey of Professional Forecasters and equity analyst earnings forecasts, we establish several new facts about belief formation, focusing primarily on correction speeds. Our theoretical analysis uses estimated BIRs to assess the strengths and weaknesses of leading models of belief formation.

Keywords:

JEL: C3, C58, C61, G11, G12, 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 University and NBER. Hongyu Wu is at Yale School of Management. We are grateful to seminar participants at 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

How do people form beliefs? For nearly half a century, the dominant model of human behavior in economics—particularly in macroeconomics and finance—has rested on the assumption of full information rational expectations (FIRE). Under this paradigm, individuals perfectly observe the structural shocks hitting the economy and fully understand their subsequent equilibrium effects. This knowledge leads them to form efficient forecasts of future economic outcomes. In reality, though, forecasting can be clouded by behavioral biases (Bordalo et al., 2020b), model uncertainty (Schwartzstein and Sunderam, 2021; Farmer et al., 2023), and imperfect information (Coibion and Gorodnichenko, 2012). As a result, over the past two decades, there has been a steady push to integrate more realistic features of belief formation into models of the macroeconomy (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).

Understanding the speed and process by which deviations from FIRE are corrected is critical to these efforts. Consider, for instance, a scenario where stockholders are prone to overreact to positive news, but corporate managers are fully rational. If investors correct themselves quickly after an overreaction, 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 effects on the economy.

In this paper, we introduce a new tool, called a behavioral impulse response (BIR), to characterize the process by which deviations from FIRE are corrected over time. As their name suggests, BIRs draw inspiration from a long tradition in macroeconomics that uses vector autoregressions (VARs) to study how shocks propagate dynamically through a system (Sims, 1980). In a similar manner, BIRs capture how an individual's beliefs evolve in response to news. BIRs are extremely portable and easy to estimate, relying only on OLS

regressions and an observable term structure of forecasts. Furthermore, as we show more precisely below, they deliver a simple yet powerful set of moments that can help differentiate between models of belief formation and thus guide the development of theory.

To understand how BIRs work, consider a sequence of forecasts $F_t, F_{t+1}, ..., F_{t+n-1}$ for a fixed target y_{t+n}, such as 2025Q4 GDP growth, and define $e_t = y_{t+n} - F_t$ as the realized error associated with the time-t forecast, F_t. In an influential study, Coibion and Gorodnichenko (2015) develop a test for whether F_t adheres to FIRE by studying the conditional expectation of e_t following a forecast revision, $R_t = F_t - F_{t-1}$, at time t. BIRs build on this test by jointly studying the conditional mean of *all* future forecast errors e_t, e_{t+1}, ..., e_{t+n-1} following a unit revision R_t. More formally, we define the BIR, $B(j)$, at horizon j as follows:

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

The first point on the BIR, $B(0)$, is the object of interest in Coibion and Gorodnichenko (2015). Under FIRE, it — and the rest of the BIR curve—should equal zero because forecast errors are not predictable. If, however, $B(0) < 0$, F_t can be said to overreact or overshoot because it is reliably followed by a disappointing realization of y_{t+n}. Conversely, $B(0) > 0$ indicates that upward revisions made at time t are consistently associated with positive forecast errors e_t, meaning the forecast F_t underreacts to the news.

Each subsequent point along the BIR curve then measures how future forecast errors are expected to evolve following a unit revision at time t. Equivalently, because the difference between two consecutive forecast errors reveals the revision between the two periods, the shape of the BIR captures how long an over- or under-reaction at time t remains in future forecasts. To see this more explicitly, note that:

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

where $R_{t+j} = F_{t+j} - F_{t+j-1}$ is the one-period revision made at $t+j$. In the extreme, if over- or under-reaction at time t is permanent, then all future revisions will be zero on average so that $B(0) = ... = B(n-1) \neq 0$. This type of behavior will manifest as a perfectly flat BIR, whose level is determined by the initial degree of over- or under-reaction.

The BIR will take on a different shape when over- or under-reaction at time t is short-lived. For illustration, assume that forecasts overreact at time t, meaning $B(0) < 0$. If forecasters correct this error immediately, all future forecast errors should average to zero, resulting in $B(1) = ... = B(n-1) = 0$. Additionally, the fact that forecasters correct their time-t overreaction after one period implies that their revision at $t+1$ is predictably negative. This correction process is directly embedded in the shape of the BIR (see Figure 1), since $\mathbb{E}_t[R_{t+1}|R_t = 1] - \mathbb{E}_t[R_{t+1}|R_t = 0] = B(0) - B(1) = B(0) < 0$.

The estimation of BIRs is straightforward and parallels the large literature on local projection methods in macroeconomics (see Jordà (2023) for a recent survey). In our empirical work, we estimate them using two datasets that are commonly used in the belief-formation literature. The first is the Survey of Professional Forecasters (SPF), a survey that contains forecasts by multiple individuals of various macroeconomic and financial series (e.g., Mankiw et al. (2003), Coibion and Gorodnichenko (2012)). The second is the Institutional Brokers' Estimate System (IBES), a panel of individual equity analyst forecasts of earnings for publicly traded U.S. stocks (e.g., Bouchaud et al. (2019), Van Binsbergen et al. (2023)). Both datasets are sampled quarterly. We use the estimated BIRs from both datasets to establish five new facts about belief formation, focusing primarily on how deviations from FIRE are corrected over time.

The first fact is that correction speeds depend on the data generating process for y. Consistent with prior work, individual SPF and IBES forecasters both tend to overreact to contemporaneous news (Bordalo et al., 2020b, 2024), especially at longer horizons (Giglio and Kelly, 2017) and for less persistent series (Afrouzi et al., 2023). What differs across

the two datasets is the speed of correction. BIRs from the SPF, which is dominated by macroeconomic time series, exhibit a relatively fast pace of correction. On average, less than 20% of the overreaction in the time-t forecast remains in the forecast made at $t+1$. IBES analysts, by contrast, correct their overreaction at a relatively slow pace, at least for longer-horizon forecasts. For example, when equity analysts forecast earnings that will be realized in six quarters ($n = 6$), half of the overreaction embedded in their time-t forecast is still contained in their forecasts made four quarters later. We also document substantial variation in correction speeds across series within each dataset.[1]

The second fact is that correction speeds are influenced by factors beyond the magnitude of the initial reaction. To highlight this point, we non-parametrically search for subsets of stocks in the IBES dataset that exhibit similar degrees of overreaction, yet differ in their speed of correction. For the fast-correcting stocks, roughly half of the initial overreaction at time t is undone within two quarters, whereas for the slow-correcting stocks virtually none is undone. Similar patterns emerge when studying stocks that exhibit underreaction. Correction speeds also vary with stock characteristics that are commonly studied in the finance literature (e.g., Fama and French (1993)). For example, while earnings forecasts for value and growth stocks show similar levels of overreaction at longer horizons, corrections in value stocks occur at a significantly slower pace. By the logic of Bordalo et al. (2024), these differences in correction speeds are likely to influence the extent and timing of return predictability (Shiller, 1981; Campbell and Shiller, 1988; Cochrane, 2011), a topic we explore further in a companion paper.

Our third fact is that forecaster behavior, as measured by the level and shape of BIRs, strongly depends on the nature of news. Strikingly, at horizons over a year, equity analysts appear to underreact to news that they perceive as negative but overreact to news that is perceived as positive. Here, we define the sentiment of the news received by each analyst

[1]While series peristence affects the degree of overreaction (Afrouzi et al., 2023), it has less of an impact on the speed of correction (see Internet Appendix IA.3.1).

according to the sign of their forecast revision at time t. Moreover, over and underreaction are more pronounced for smaller news events, and correction speeds vary meaningfully with the sentiment of news. We also document that IBES analysts are more overreactive and slower to correct deviations from FIRE during recessions. Interestingly, SPF forecasters exhibit far less state and sentiment dependence than IBES analysts.

A fourth aspect of BIRs that we explore empirically is heterogeneity across forecasters. Consistent with the findings of Bouchaud et al. (2019), some equity analysts consistently overreact to news, while others tend to underreact. A novel result from our study is that the speed of correction varies widely, even among analysts with similar initial tendencies to over or underreact to news. For instance, among overreactive analysts, some correct the bulk of their errors within one or two quarters, whereas others may take twice as long. This fact suggests that the initial reaction to news and the subsequent process of correction are governed by distinct cognitive mechanisms.

The fifth and final fact that we document characterizes how forecaster behavior changes over time. To study this phenomenon, we estimate BIRs for each equity analyst in IBES using the first five years of their careers and compare them to BIRs estimated using data later in their careers. Our analysis reveals that over time, the BIRs of all types of analysts tend to converge towards the FIRE benchmark, weighing against the notion that BIRs are a static personal trait. Instead, this convergence suggest that experience enhances analysts' abilities to avoid forecast errors and to correct them more quickly when they occur.

In the latter part of the paper, we use these facts to assess the strengths and weaknesses of different models of belief formation. Our analysis is organized around a general class of models—what we call the "linear autonomous class"—that encompasses a wide range of belief formation processes that have been studied in prior work, including adaptive expectations (Cagan, 1956; Nerlove, 1958), diagnostic expectations (Bordalo et al., 2020a; Bianchi et al., 2024), noisy information and sticky information (Giannoni and Woodford, 2003; Coibion and

Gorodnichenko, 2012), and many others. Our primary theoretical contribution is to derive a closed-form expression for the BIR of the entire linear autonomous class.

To illustrate the paper's main proposition, we focus on three popular models of belief formation, starting with the model of diagnostic expectations used by Bordalo et al. (2020a). A calibrated version of the model using estimates from the literature shows that diagnostic expectations succeed in matching the level of BIRs observed empirically, both in the SPF and IBES datasets. Our theoretical analysis further reveals that models of diagnostic expectations imply an abrupt correction process: in standard applications, any overreaction at time t is fully corrected within one period. While this appears to fit well for SPF forecasts, it is less successful at matching IBES forecasts, which exhibit a smooth and slow correction process at longer forecast horizons.

Next, we analyze the model from Afrouzi et al. (2023), who study belief formation when forecasters face information processing costs. Similar to diagnostic expectations, costly information processing can successfully match the level and shape of BIRs from SPF but has a harder time with the IBES. The rate of correction under costly information processing is smooth and decelerating, with most of the gains occurring in the first few periods. However, we find several instances in the IBES data in which the correction process is extremely slow at first, only to accelerate after many periods.

The third model we study features extrapolative expectations, as in Hirshleifer et al. (2015). Under common calibrations used in the literature (e.g., Nagel and Xu (2022)), forecasters with extrapolative expectations overreact to a degree that resembles what is observed in SPF and IBES. Nevertheless, these models imply a correction process that is much too slow relative to the data.

In our last theoretical exercise, we move outside the class of linear autonomous models and demonstrate how BIRs can be paired with simulations to study virtually any model of belief formation. We specifically analyze the model from Farmer et al. (2023), in which Bayesian

agents must form forecasts of a process consisting of a unit root and a transitory component, both of which are 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). Using their simulation code, we ask whether model uncertainty can also generate realistic BIRs. We find that model uncertainty shows some promise in matching empirical BIRs, even those with a slow correction process like in IBES, but the success depends critically on the initial priors of the forecasters. Whether the initial priors we consider are plausible is an open and interesting question for future research.

Literature Review As mentioned above, BIRs extend the influential tests of FIRE designed by Coibion and Gorodnichenko (2015) to study belief dynamics. This extension yields two primary benefits. First, tests of FIRE via BIRs offer more statistical power because they are based on joint restrictions across the entire BIR curve, as opposed to tests of a single restriction like the one proposed by Coibion and Gorodnichenko (2015). Second, and more importantly, BIRs shed light on the dynamics of belief formation 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. In this respect, our approach is closely related to recent work by Augenblick and Rabin (2021), who develop a test of Bayesian updating based on how forecasts of a fixed target evolve through time.

Impulse responses of forecasts to identified macroeconomic shocks have been used before in research on belief formation, mainly to characterize information frictions facing forecasters (Coibion and Gorodnichenko, 2012), explain discrepancies between individual and consensus forecasts (Angeletos et al., 2021), and study forecasters' reactions to specific structural shocks (Kučinskas and Peters, 2022). A large benefit of BIRs—and one that makes them more portable across different datasets—is that they do not require knowledge or identification of structural shocks, which is especially challenging in macroeconomic settings (Bauer and

Swanson, 2023). The cost is that BIRs capture how forecasts respond to all information driving a revision, whereas traditional shock-based impulse responses isolate reactions to a specific piece of information. In many applications, this feature of BIRs may actually be desirable since economic decisions and market prices are ultimately determined by how beliefs respond to all forms of news. Additionally, BIRs unveil shortcomings in leading models of belief formation that went previously unnoticed by shock-based analyses, thus underscoring their value for refining existing models.

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 then estimates BIRs in both the SPF and IBES datasets. Section 4 contains our theoretical analysis, which focuses primarily on how to use BIRs to assess models of belief formation. Section 5 highlights how BIRs can be paired with simulations to evaluate more complex models of belief formation, like the one studied by Farmer et al. (2023). Section 6 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 discuss why BIRs are useful for testing theories of belief formation. Finally, we describe how to estimate BIRs on any dataset in which forecasts of a fixed target (e.g., 2024Q4 GDP) 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 3. $e_t^{(n)} = y_{t+n} - F_t^{(n)}$ 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 function (BIR) for forecast horizon n is given by*

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

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 react dynamically to news received at time t, as measured by the revision $R_t^{(n)}$.

Under full-information rational expectations (FIRE), BIRs take a very simple form that is useful for hypothesis testing: they are zero for all horizons, n, and periods since forecast, j. This follows from the fact that forecast errors are unpredictable under FIRE.

Remark 1 Under FIRE, $B(n,j) = 0, \quad \forall n, j$

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 (Bordalo et al. (2020b); Coibion and Gorodnichenko (2015); 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 negative future forecast error. This is naturally interpreted as an overreaction to the news because the information that prompted the upward revision results in a forecast that overshoots the target. By the same logic, $B(n,0) > 0$ corresponds to under-reaction 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 phenomenon (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 $t+1$:

$$\begin{aligned} e_{t+1}^{(n-1)} - e_t^{(n)} &= \left[y_{t+n} - F_{t+1}^{(n-1)}\right] - \left[y_{t+n} - F_t^{(n)}\right] \\ &= -\left[F_{t+1}^{(n-1)} - F_t^{(n)}\right] \\ &= -R_{t+1}^{(n-1)}. \end{aligned}$$

Or, more generally, $e_{t+1}^{(n-j)} - e_t^{(n-j+1)} = -R_{t+1}^{(n-j)}$. 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, conditional upon a unit revision made at time

t. Thus, BIRs embed the trajectory of future revisions following a revision made at time t. Under FIRE, these revisions should also not be predictable by $R_t^{(n)}$, implying a perfectly flat BIR.

Stylized Example. Returning to the question posed above, consider a data-generating process where the rational forecast of y_{t+n} is always zero, like, for example, an i.i.d process with zero means. Next, suppose that at time t the forecaster irrationally revises its forecast up from $F_{t-1}^{(n+1)} = 0$ to $F_t^{(n)} = 1$, implying $R_t^{(n)} = 1$. Further, suppose that the forecaster comes to its senses in the next period, revising the forecast back down and keeping it at the rational forecast after that.

Figure 1 plots the implied BIR in this stylized example for $n = 10$. The first point on the BIR curve is $B(n, 0) = -1$ because the revision made at time t leads to a forecast that overshoots by one unit. The remaining points along the BIR curve are all zero because the forecaster realizes the mistake and makes no more errors. The mistake correction also implies that the revision at that time $t + 1$ is $R_{t+1}^{(n-1)} = -1$. The magnitude of this correction is naturally embedded in the BIR because $B(n, 0) - B(n, 1) = \mathbb{E}\left[R_{t+1}^{(n-1)} | R_t^{(n)} = 1\right] - \mathbb{E}\left[R_{t+1}^{(n-1)} | R_t^{(n)} = 0\right] = -1$

As this example highlights, BIRs measure the initial degree of over- or under-reaction to news at time t and how subsequent revisions then respond. This is the sense in which BIRs summarize the dynamics of belief formation. Consequently, any theory of belief formation must match the level and shape of BIRs that are estimated empirically. BIRs are also useful because they offer a natural way to study how belief formation varies across forecasting environments. For instance, they can be used to characterize how belief formation varies across individuals, stocks, or the nature of news. We highlight these applications later in Section 3, after describing how to estimate BIRs.

2.2 Estimation and Inference

Estimation BIRs can be easily estimated using any data 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 be GDP growth or the earnings per share of a publicly listed firm. $F^{(n)}_{i,s,t}$ is the forecast made by individual i at time t of the realization of series s in n periods and $e^{(n)}_{i,s,t} = y_{s,t+n} - F^{(n)}_{i,s,t}$ is the associated forecast error. $R^{(n)}_{i,s,t} = F^{(n)}_{i,s,t} - F^{(n+1)}_{i,s,t-1}$ is the revision made by forecaster i at time t for series s's realization at $t+n$.

The estimation process for BIRs closely aligns with the 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:

$$e^{(n-j)}_{i,s,t+j} = c^{(n)}_j + \psi^{(n)}_j R^{(n)}_{i,s,t} + v^{(n-j)}_{i,s,t+j}, \quad \forall j = 0,,,,.n-1, \tag{2}$$

in which case the $B(n,j) = \psi^{(n)}_j$. 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 the 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.

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 time.

Joint Inference 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 $B(n, 0) = B(n, 1) = ... = B(n, n-1) = 0$. 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)).

To implement such a test, notice that equation (2) can be estimated for all j in a single panel regression that appropriately stacks the data. Specifically, suppose there are $i = 1, ..., I$ forecasters, $s = 1, ..., S$ series, and $t = 1, ..., T$ periods. This means the last n-period forecast is made at $T - n$. The entire BIR curve can then be estimated in one pass by running the

following stacked regression:

$$
\begin{aligned}
e_{i,s,t}^{(n)} &= c_0^{(n)} + \psi_0^{(n)} R_{i,s,t}^{(n)} + v_{i,s,t}^{(n)} \\
e_{i,s,t+1}^{(n-1)} &= c_1^{(n)} + \psi_1^{(n)} R_{i,s,t}^{(n)} + v_{i,s,t+1}^{(n-1)} \\
e_{i,s,t+2}^{(n-2)} &= c_2^{(n)} + \psi_2^{(n)} R_{i,s,t}^{(n)} + v_{i,s,t+1}^{(n-2)} \\
&\vdots \\
e_{i,s,t+n-2}^{(2)} &= c_{n-2}^{(n)} + \psi_{n-2}^{(n)} R_{i,s,t}^{(n)} + v_{i,s,t+n-2}^{(2)} \\
e_{i,s,t+n-1}^{(1)} &= c_{n-1}^{(n)} + \psi_{n-1}^{(n)} R_{i,s,t}^{(n)} + v_{i,s,t+n-1}^{(1)},
\end{aligned}
$$

for all i, s, t. This one-shot regression approach simply requires stacking the data used to run each individual regression in (2) and then allowing the intercept and slope 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 $\boldsymbol{\psi} = (\psi_0^{(n)}, \psi_1^{(n)}, ..., \psi_{n-1}^{(n)})'$ 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 the same stacking procedure, then allowing the intercepts and slopes to vary with j and the subsamples of interest (e.g., a particular series).

3 New Facts About Belief Formation

In this section, we estimate BIRs using forecasts from the SPF and IBES. The SPF is a quarterly survey, administered by the Federal Reserve Bank of Philadelphia, that contains predictions from professional forecasters on several macroeconomic and financial outcomes (e.g., nominal GDP). IBES contains analyst-by-firm level forecasts of quarterly EPS, semi-annual EPS, annual EPS, and long-term EPS growth over a number of horizons. We focus on quarterly EPS forecasts because they are the most populated for longer forecast horizons. We largely follow standard practice from the literature when cleaning and processing both

datasets. These steps, along with how we define forecast errors and other variables, are detailed in Internet Appendix IA.1.

3.1 Fact 1: Correction Speeds Depend on the Data Generating Process

We now estimate from the SPF and IBES datasets based on the regressions described in Equation (2). Our baseline estimates pool data across all forecasters and series, so they should be interpreted as the BIR for the average forecaster in each respective dataset.

3.1.1 Baseline BIRs from the SPF

Figure 2a contains estimated BIR curves from the SPF dataset for $n = 3, 2, 1$, along with 95% confidence bands based on standard errors that are clustered by forecaster-series, (i, s), and series-date, (s, t). Table 1 presents the full set of point estimates contained in the plot, their standard errors, and the average number of observations used to estimate each BIR curve.

The green line in Figure 2a shows the BIR for $n = 3$, the longest available horizon in the SPF. The first point on the curve, $B(n, 0)$, is negative and statistically different from zero. This finding reflects the well-known fact that individual forecasters in the SPF typically overreact to new information (Bordalo et al., 2020b). It is important to note, however, that our baseline estimate of $B(3, 0)$ is not directly 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.1.1).

The novelty of BIRs lies in the subsequent parts of the curve, starting with $B(3, 1)$. The estimated coefficient implies that following a revision made at t for a realization at $t + 3$, the expected forecast error at $t + 1$ is roughly one-fifth the size of the error at t. In other words, on average, 80% of the initial overreaction in the time-t forecast is undone within one

period. The next point on the BIR curve, $B(3,2)$, is roughly equal to $B(3,1)$, suggesting no further correction from $t+1$ and $t+2$. By $t+3$, forecasters have more or less corrected their initial overreaction. This is not surprising given that $B(3,3)$ is based on the nowcast of y_{t+3}, which is made during the quarter in which y_{t+3} is realized.

The remaining BIR curves in Figure 2a correspond to shorter forecast horizons, with $n=2$ in orange and $n=1$ in blue. The ordering of the curves indicates that overreaction decreases with the horizon. For instance, the magnitudes of $B(1,0)$ and $B(3,0)$ suggest that overreaction to a contemporaneous revision is about 25% less for short-horizon forecasts ($n=1$) compared to long-horizon ones ($n=3$). Additionally, a formal test of whether $B(3,0) < B(2,0) < B(1,0)$ yields a p-value of 0.998. This finding accords with past research arguing that overreaction grows with the forecasting horizon, such as Giglio and Kelly (2017), Bouchaud et al. (2019), Bordalo et al. (2019a), and Afrouzi et al. (2023).

The bottom two rows of Table 1 show p-values for two different joint tests of each BIR curve. The first, labeled $p(FIRE)$, tests the null hypothesis that the estimated BIR is generated under FIRE. As mentioned above, this is a joint test of whether $B(n,j)=0$ for all j. At all horizons, we reject the null of FIRE with p-values below 0.001. The p-values that arise from this joint test highlight why BIRs offer stronger tests of FIRE compared to those proposed by Coibion and Gorodnichenko (2015). To see why, notice that the Coibion and Gorodnichenko (2015) tests are based on $B(n,0)=0$, which for $n=3$ yields a p-value of 9.0e-39. While this is small, it is nonetheless orders of magnitude larger than the p-value of 4.6e-51 that obtains when testing whether $B(3,j)=0$ for all j. In the traditional macroeconomic setting, the observation that joint tests of impulse responses are more powerful than tests of individual point estimates has also been made by Jordà (2023).

The last row of Table 1 presents p-values from a weaker test of FIRE, namely that all points along the BIR curve are equal, even if they differ from zero. Recall from the discussion in Section 2.1 that this is equivalent to testing whether future revisions can be predicted by

current ones, which should not be the case under FIRE. One might also label this as a test of forecaster "stubbornness" since it uncovers whether forecasters tend to stick with their forecast after a revision. Unsurprisingly, for all n, we easily reject this null as well.

3.1.2 Baseline BIRs from IBES

Figure 2b shows BIRs for different n in the IBES data. The confidence bands in the plot are based on standard errors that are clustered by analyst-stock and stock-date, and the full set of point estimates are also reported in Table 2. The blue line shows the BIR for $n = 6$. Its first point, $B(6, 0)$, indicates that the average IBES forecaster overreacts to news received at t, much like the average SPF forecaster.[2] The remaining points of the curve show that that IBES forecasters gradually realize their initial mistake. Unlike SPF forecasters, however, this correction occurs at a relatively slow pace. For instance, by $t + 1$, only about 20% of the overreaction in the time-t forecasts has been corrected, and the initial error's half-life extends beyond a year. Remarkably, even the forecasts made in the quarter of the realization still contain about 20% of the initial overreaction. Unsurprisingly, the null of FIRE is easily rejected for IBES forecasts (see the bottom two rows of Table 2).

The BIR for $n = 6$ supports a crucial aspect of the hypothesis proposed by Bordalo et al. (2024), henceforth BGLS, which centers on the established observation that high stock valuation ratios at time t are predictive of lower future returns (Campbell and Shiller, 1988). Contrary to theories positing time-varying discount rates as the explanation, BLGS argue that the predictability of returns stems from overreaction in the market's forecasts of future cash flows. The basic idea is that the market irrationally exaggerates its response to the news at time t; positive news leads to overly optimistic expectations of future cash flows and unwarranted inflation of stock prices. This overvaluation is corrected over time as earnings growth fails to meet expectations, resulting in a reversion in prices and the predictability of

[2] Bouchaud et al. (2019) document underreaction in consensus forecasts for 1-year forecasts of annual, not quarterly, EPS. We show in Internet Appendix IA.2.1 that their focus on consensus forecasts and annual EPS drives the differences between our results.

returns. At the heart of this process is the market's eventual acknowledgment of its initial error. The BIR for $n = 6$ offers direct evidence of this correction, illustrating that IBES forecasters do adjust their views, albeit slowly, after an initial overreaction to the news.[3]

The remaining BIR curves in Figure 2b further show that overreaction strongly decreases with forecast horizon, aligning with our early results from the SPF data. The speed of this decay is quite rapid, and we can easily reject the null that the BIRs are identical for all n. While the point estimate for $B(5, 0)$ indicates overreaction, it is half the size of $B(6, 0)$. Similarly, $B(4, 0)$, while negative, is less than 20% of the magnitude of $B(6, 0)$. For $n \leq 3$, deviations from FIRE are generally small and have an unstable sign.

The fast decay of overreaction as n decreases potentially explains BGLS's finding that forecasts of one-year ahead earnings have weaker predictive power for future stock market returns compared to long-term growth forecasts. The simple reason is that shorter-term earnings are subject to less overreaction and are a relatively small portion of the future cash flows that drive stock market valuations. In turn, they serve as less reliable indicators of impending price reversals.[4]

3.1.3 Takeaways

The baseline BIRs in Figure 2 highlight several interesting properties of belief formation. For one, Figures 2a and 2b confirms prior findings that forecasts in SPF and IBES both overreact to contemporaneous news, especially for longer horizons. The main difference between SPF and IBES is correction speed. In SPF, which contains mostly macroeconomic time series, overreaction is rapidly correcting, disappearing after essentially one quarter. In contrast,

[3] 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, however, we find overreaction in individual forecasts at this horizon, i.e., $B(6, 0) < 0$. In Internet Appendix IA.2.3, we also confirm similar BIRs for large stocks, like those studied by BLGS.

[4] Table A6 of the Internet Appendix shows estimate for $n < 6$ using only the set of forecasters and stocks that are used for estimation when $n = 6$. The BIRs in this restricted sample are similar to those found in Figure 2b, indicating that the composition of forecasters is not driving our results.

longer-horizon BIRs in IBES reveal a much slower average correction process for equity analysts. Our preferred interpretation of these results is that the speed with deviations from FIRE are corrected depends on the underlying data generating process.

We reinforce this conclusion by studying variation in correction speeds within each dataset, which also mitigates concerns that the variation in correction speeds between Figures 2a and 2b arises because SPF and IBES forecasteres are inherently different types of people or face different incentives. In each dataset, we estimate series-by-series BIRs and, for illustration purposes, retain only those that overreact initially. Within this set, we then sort series in K bins by the convexity of their BIRs, a natural measure of correction speed. Within each convexity bin, we estimate a pooled BIR to maximize precision. $K = 5$ for SPF and $K = 10$ for IBES, reflecting the greater number of stocks in IBES compared to macroeconomic series in SPF. In addition, we set the horizon of the BIR for IBES equal $n = 5$ to focus on a relatively long forecast horizon while still maintaing reasonable analyst coverage.

Figure 2c indicate that correction speeds vary across series in the SPF. Series with fast correction speeds (blue) overreact slightly more than those with slow correction speeds (orange), though less than 15% of this overreaction remains in the forecast one period later and all of it is gone within two periods. For the slower correcting series, almost 40% of the initial overreaction is still embedded in forecasts after two periods.

Figure 2d reveals even more variability within the IBES dataset. Here, almost all of the initial overreaction in the slow-correcting series (blue) remains in the forecast made four periods later. This slow correction is followed by an abrupt correction in the nowcast made during the quarter of the target's realization. In contrast, the initial overreaction in the fast-correcting series (orange) is essentially undone within a single period.

Together, the evidence in Figure 2 supports the notion that the speed with which forecasters correct deviations from FIRE depends on the underlying data generating process.[5]

[5]One possible explanation for this finding is that correction speed depends on the persistence of the

3.2 Fact 2: Correction speed depends on more than just the size of the reaction

Figure 2d illustrates that deviations from FIRE are corrected at varying speeds within the IBES dataset. A straightforward explanation for this pattern is that smaller initial overreactions (orange) are corrected more quickly than larger ones. If this were consistently true, structural models could potentially employ a single parameter to encapsulate the dynamics of belief formation as reflected in BIRs. However, if the speed of correction significantly varies even when the degree of initial overreaction or underreaction is held constant, models would need to distinguish between these two aspects of belief formation separately.

To better understand this issue, we proceed in three steps. Using the IBES dataset, we first sort stocks with a negative $B(5,0)$—equivalently, those that overreact—into terciles based on the level of $B(5,0)$. Second, within the most overreactive stocks, we further divide stocks into terciles based on their speed of correction, as measured by $W = B(5,1)/B(5,0)$. By this measure, higher values of W indicate slower correction. Within each tercile of W, we then estimate BIRs by pooling across all stock-analyst forecasts. Finally, we repeat this process for stocks that demonstrate initial underreaction. Stocks must include at least twenty corresponding forecasts with $n = 5$ to be included in the analysis. We focus on the IBES dataset because its large cross-section of stocks and analysts affords more statistical precision compared to the SPF.

Figure 3a presents the results. The first thing to notice is that within the set of stocks for which forecasters tend to initially overreact the most (red and green lines), the degree of overreaction is virtually identical. Nevertheless, the speed of correction between the two groups is materially different. For the fast-correcting stocks, roughly half of the initial overreaction is undone within two quarters, whereas for the slow-correcting stocks virtually none is undone. Forecasts of slow-correcting stocks converge to FIRE more rapidly in the

underlying series Afrouzi et al. (2023), though we show in Internet Appendix IA.3.1 that this is unlikely to be the full story.

last two periods, leading to the two BIR curves to converge by $t+5$. The blue and orange lines in Figure 3a reveal a similar pattern for stocks with underreactive forecasts at time t.

The remaining panels in Figure 3 study how correction speeds vary with specific stock characteristics, focusong on a few that are popular in the asset pricing literature (Fama and French, 2015). In what follows, we assign each stock s to a tercile based on a given characteristic $C_{s,t-1}$ at time $t-1$. Using revisions made at time t, we then estimate pooled BIRs for all forecasts made on all stocks in each characteristic-tercile. We use lagged characteristic values when assigning stocks to terciles to ensure that they are observable by the forecaster at the time of the revision. Data on stock characteristics is taken directly from Jensen et al. (2023).

Figure 3b displays BIRs for stocks that differ in their book-to-market ratios. The orange line in the plot shows the estimated BIR for so-called growth stocks—those with low book-to-market ratios–and the blue line shows it for value stocks—those with high book-to-market ratios. The plot shows that, following a revision at t, contemporaneous forecasts for both the average value and growth stocks exhibit overreaction, with no significant difference in the initial magnitude of their responses. However, the distinction between the two manifests in the pace at which forecasters adjust this initial overreaction. Corrections for growth stocks occur relatively quickly, in contrast to value stocks, where the adjustment process is markedly slower. Notably, forecasts for value stocks maintain a considerable portion of their initial overreaction throughout the quarter leading up to the realization (i.e., the nowcast).

Figure 3c compares BIRs between high and low-volatility stocks, with volatility determined by trailing 60-day returns. It reveals that high-volatility stocks exhibit a significantly larger initial overreaction to news—nearly double that observed in low-volatility stocks. Furthermore, this overreaction in high-volatility stocks persists, showing negligible correction over three quarters. This pattern could arise if it is more difficult for analysts to learn about stocks with high volatility.

Finally, Figure 3d compares BIRs for stocks with high and low profitability, as defined by the ratio of operating profits to book equity (Fama and French (2015)). The BIR of stocks with low profitability lines up closely with the FIRE benchmark. In contrast, forecasts for high-profitability stocks exhibit an initial overreaction (the first point on the BIR) and then a very slow correction — the half-life of the initial overreaction is over four quarters.[6]

3.3 Fact 3: Forecaster behavior depends on the nature of news

We now explore whether the level and shape of BIRs varies with the nature of news received by forecasters. A simple way to measure the sentiment of news received by each forecaster is based on the sign and size of their revision. For example, a large and positive revision is an indication that a forecaster received very positive news. Figure 4a implements this idea in IBES by estimating BIRs for $n = 5$ based on: (i) whether $R_{i,s,t}$ is positive or negative; (ii) its absolute magnitude, $|R_{i,s,t}|$. Specifically, large positive revisions are defined based on whether $|R_{i,s,t}|$ is in the top tercile of all positive revisions. All remaining positive revisions are considered small, and small and large negative revisions are defined analogously. For reasons that we explain below, we run this analysis using forecasts made outside of the Covid-19 pandemic (i.e., excluding forecasts made in 2020Q1 and 2020Q2).

The most striking thing from Figure 4a is that equity analysts *underreact* to negative news but *overreact* to positive news. Moreover, the size of the deviation from FIRE is decreasing in the size of the news. That is, over and underreaction are larger—at least in the Coibion and Gorodnichenko (2015) sense—when the news is smaller. In Internet Appendix IA.3.2, we show this result holds for different forecast horizons and is not driven by outliers. One possible explanation is that negative news is less pleasant and thus easier to ignore. Pinning down the microfoundations of this result is an interesting avenue for future theory.[7]

[6] Bouchaud et al. (2019) develop a model attributing the profitability premium to under-reaction. However, as we demonstrate in Internet Appendix IA.2.1, overreaction in their model also yields a profitability premium, given a plausible degree of overreaction. Thus, the critical distinction between over- and under-reaction hinges on the direction of return predictability. See the appendix for additional discussion.

[7] This type of pattern is harder to rationalize under the view that deviations from FIRE arise due to

Figure 4a also shows that correction speeds vary with the sentiment of perceived news. For example, the speed of correction is slowest when forecasters receive small positive news (orange). On average, after two periods, most of the relatively large overreaction at time t remains in forecasts. On the other hand, overreaction is both smaller and more quickly corrected for large negative news. While the magnitude of underreaction to small negative news is larger than the overreaction to small positive news, this deviation from FIRE also corrects much faster. Roughly sixty percent of the initial overreaction to small negative news is corrected within two periods.

Figure 4b probes whether the sentiment of news has a similar impact on the BIRs of SPF forecasters. Unlike with equity analysts, SPF forecasters appear to overreact to both positive and negative news, though their response to both types of news still differs in meaningful ways. In particular, they initially overreact more to positive news, but the speed with which they correct this overreaction is also faster compared to negative news. A formal statistical test of whether the two BIRs are equal is rejected at conventional confidence levels. [To do: break out by size of revision]

Figure 4c asks a related question: do equity analysts react differently to news during recessions? Recession dates are taken directly from the Business Cycle Dating committee of the National Bureau of Economic Research (NBER). The figure shows that BIRs for equity analysts depend heavily on the state of the economy. Outside of recessions (orange line), equity analysts exhibit a modest amount of overreaction that is then corrected almost entirely in one quarter. In contrast, are almost seven times more overreactive during recessions (blue line). The speed of correction during recessions is somewhat linear. For $n = 6$, roughly half of the overreaction at time t is still embedded.

A significant portion of the variation in BIRs across recessions versus non-recession periods can be attributed to the singular impact of the COVID-19 pandemic. This becomes

information frictions (Mankiw and Reis, 2002; Coibion and Gorodnichenko, 2012), given the ability of professional forecasters to observe the realization of y should presumably not depend on the sentiment of the news.

evident when examining the green line in Figure 4c, which represents the estimated BIR for recessions, excluding the pandemic-induced recession. Although the initial point on the BIR curve confirms that overreaction is more pronounced during recessions, the magnitude of this effect is halved once the pandemic is excluded. Moreover, the pace of correction when excluding the pandemic is somewhat slower, with only about a third of the initial overreaction corrected after three quarters. In other words, overreaction during the pandemic was much larger than normal, but this overreaction was corrected relatively quickly. This finding is perhaps not surprising given the rapid decline and rebound of the aggregate stock market that occurred during this time. For this reason, we excluded the pandemic when estimating Figure 4a. In Internet Appendix IA.3.2, we present results including the pandemic. Unsurprisingly, the only material thing that changes is forecasts made following large negative shocks are on average slightly overreactive.

Unlike equity analysts, Figure 4d shows that the belief formation process does not vary meaningfully with the state of the economy for SPF forecasters.[8] The BIRs for recession and non-recession periods are closely aligned, and we cannot reject the null hypothesis that they are different ($p = 0.302$). Both BIRs exhibit the same pattern of relatively fast correction seen in our baseline estimates.

3.4 Fact 4: BIRs vary widely across forecasters

Next, we study heterogeneity across forecasters. We again focus our attention on the IBES dataset, due to its relative size and ability to consistently identify individual forecasters.[9] Bouchaud et al. (2019) document heterogeneity in Coibion and Gorodnichenko (2015) style regression coefficients (i.e., $B(n, 0)$) across equity analysts. Building on their work, the main

[8]For the SPF sample, the recession periods are 1969Q4-1970Q4, 1973Q4-1975Q1, 1980Q1-1980Q3, 1981Q3-1982Q4, 1990Q3-1991Q1, 2001Q1-2001Q4, 2007Q4-2009Q2, and 2020Q1-2020Q2.

[9]Earlier versions of the SPF do not necessarily have unique identifiers for each individual forecaster. See Section 4 of the SPF documentation for more details.

question we seek to answer in this subsection is whether individual forecasters differ in their correction speeds.

We proceed in a similar manner to our analysis of stock-level heterogeneity. We start by estimating BIRs for each analyst, requiring at least ten forecasts per analyst, and separating them further based on whether they initially over or underreact on average ($B(5,0)$). Within each group, we further assign analysts to terciles based on $W_i = B_i(5,1) - B(5,0)$, a measure of how quickly they correct deviations from FIRE. Lastly, we estimate a pooled BIR for analysts in each W-tercile.

For illustration, Figure 5 plots BIRs for slow and fast-correcting analysts from the median B_i group, both for over and underreactors. The analysts who tend to overreact are depicted with blue and orange BIRs. By design, the average initial overreaction of both groups is nearly identical. The primary distinction lies in their correction speeds. Fast-correcting analysts reverse the majority of their initial overreaction within a single quarter, though a residual overreaction persists until the quarter of realization. The convex shape of the slow-correcting analysts' BIR stands in stark contract to the concave shape of the fast-correcting analysts. The initial overreaction by slow-correcting analysts remains entirely uncorrected in their forecasts for three quarters, with a significant portion still evident one year later. Both BIRs only converge by the quarter of realization. A formal statistical test easily rejects the null that the BIRs are identical.

In a similar manner, the red and green lines in Figure 5 show BIRs for underreactive analysts. The magnitude of underreaction for these analysts is comparable to that of the overreactive analysts. In addition, much like overreactive analysts, underreactive analysts vary strongly in their speed of correction. The fast-correcting underreactors (red) correct over half of their initial deviation from FIRE within one period and fully correct it within a year. The slow-correcting analysts (green) correct much slower, with fifty percent of their initial reaction still present in their forecasts made a year later. Overall, Figure 5

suggests analysts differ meaningfully in how quickly they correct their deviations from FIRE. Moreover, the figure reinforces the idea that a sufficient model of belief formation must allow for correction effects to be at least partly decoupled from the mechanism that generates over or underreaction in the first place.

[To-Do: Harmonize the way we measure correction speed across analyses]

3.5 Fact Five: BIRs converge to FIRE over time

Given the large and heterogenous deviations from FIRE revealed by Figure 5, a natural question is whether BIRs are a static feature of IBES analysts or whether behavioral responses to information change over time. BIRs offer a natural way to study this question because they can be estimated over different phases of an analyst's career. Figure 6 implements this idea for analysts with at least one decade for forecasts in IBES.[10] For each, we compute BIRs based on their first five years of forecasts. Let $B_i^{early}(n, j)$ be the BIR for analyst i during this "inexperienced" period, with B_i^{early} denoting the average BIR. Analysts are then sorted into three terciles based on B_i^{early}. Within each, we estimate pooled BIRs first using data from the initial five years and then from forecasts across all subsequent years. The BIRs for the low-B_i^{early} group (overreactors) are displayed in Figure 6a, with BIRs from the early career phase labeled "Inexperienced" and those from the later labeled "Experienced." A similar analysis for the high-B_i^{early} group (underreactors) is presented in 6b.

Consistent with Bouchaud et al. (2019), Figures 6a and 6b both suggest that IBES analysts become more rational as their careers progress. In Figure 6a, analysts are overreactive and slow to correct their mistakes when they are early in their careers. For example, it takes nearly four quarters for them to correct the overreaction embedded in the revision made at time t. However, later in their careers, we cannot reject the null that the BIRs for these analysts conform to FIRE at a 95% confidence level ($p = 0.163$). Similarly, Figure 6b

[10] For this particular analysis, we exclude forecasts made during the Covid-19 pandemic because of the singularly large deviations from FIRE during this period (see Section 3.3).

reveals that analysts with underreactive BIRs early in their careers also approach the FIRE benchmark as their careers advance.

4 Using BIRs to Assess Behavioral Models of Belief Formation

In this section, we derive BIRs for a general class of behavioral models that allows for fairly flexible departures from rationality. This class encompasses a wide range of the most popular models in the literature, including extrapolative expectations (Hirshleifer et al., 2015; Nagel and Xu, 2022), diagnostic expectations (Bordalo et al., 2020b), and costly information acquisition (Afrouzi et al., 2023). We then assess the strengths and weaknesses of different models by comparing their implied BIRs with those estimated from the SPF and IBES datasets.

4.1 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 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 $\nu_t^{(n)}$.

The coefficients $\gamma(n,\tau)$ define the loading 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., 2024), 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. This model of beliefs clearly fits into the linear autonomous class in Definition 2. Specifically, for $J = 1$, the $\gamma(n,\tau)$ coefficients are given

by:

$$\gamma(n,\tau) = \begin{cases} 1+\theta, & \text{if } \tau = 0 \\ -\theta, & \text{if } \tau = 1 \\ 0, & \text{if } \tau > 1 \end{cases}$$

4.2 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^2_ε.

Assumption 2 *The signals S_t evolve according to*

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

where I is the $P \times P$ identity matrix, $|\rho| < 1$ 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^2_q.*

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:[11]

$$E_t[y_{t+n}] \;=\; \rho^{n-1} E_t[y_{t+1}] \;=\; \rho^n q_t. \tag{7}$$

Assumption 3 then imposes that this rational forecast is stationary. Finally, we make one more additional assumption on 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:*

$$\gamma(n,\tau) = \mathbf{1}_{\tau=0} \;+\; \delta(n)\kappa(\tau), \tag{8}$$

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 example of diagnostic expectations from above and setting $J = 1$, $\kappa(\tau)$ is given by:

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

and $\delta(n) = \theta(1-\rho^2)$. Given these assumptions, we can derive the paper's main proposition regarding the shape of BIRs in the linear autonomous class of belief formation models.

[11] 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 $E_t[y_{t+n}] = \beta'(\Psi)^{n-1} S_t$.

Proposition 2 (BIRs in the Linear Autonomous Class) *Under Assumptions 1-4, BIRs are given by:*

$$B(n,j) = \frac{A(n,j) - \rho^2 A(n+1,j+1)}{D(n)}, \tag{9}$$

where

$$\begin{aligned}
A(n,j) &= -\delta(n-j)\Big[\Gamma(j) + \delta(n)\Xi(j)\Big] \\
D(n) &= D_{MZ}(n) + \rho^2 D_{MZ}(n+1) - 2\rho^2(1 + \delta(n+1) \\
&\quad + \delta(n)\Gamma(1) + \delta(n)\delta(n+1)\Xi(1)) \\
D_{MZ}(n) &= 1 + 2\delta(n) + \delta(n)^2\Xi(0) \\
\Xi(j) &= \sum_{\tau=0}^{j} \kappa(\tau) \; + \; \sum_{\tau_1=1}^{\infty}\sum_{\tau_2=0}^{\infty} \kappa(\tau_1+j)\kappa(\tau_2)\rho^{2\max(\tau_1,\tau_2)} \\
\Gamma(j) &= \sum_{\tau=0}^{j} \kappa(\tau) + \sum_{\tau=1}^{\infty} \kappa(\tau+j)\rho^{2\tau}
\end{aligned} \tag{10}$$

Proposition 2 shows that BIRs are a complex function of $\kappa(\tau)$ and $\delta(n)$, the two objects that define the linear autonomous class of forecasts.

4.3 BIRs in Leading Behavioral Models

We now apply Proposition 2 to solve for the BIR in several leading behavioral models.

4.3.1 Diagnostic Expectations

To start, we continue with the example of diagnostic expectations from Bordalo et al. (2020a). In this case, Proposition 2 implies that the BIR is as follows.

Corollary 3 *In the model of diagnostic expectations (Bordalo et al., 2020a) given by Equa-*

tion (4), the BIR for $J = 1$ and diagnosticity parameter θ equals:

$$B(n, j) = \frac{A}{D}\mathbf{1}_{j=0} \tag{11}$$

where $\delta = \theta(1 - \rho^2)$ and:

$$\begin{aligned} A &= -\delta[1 + \delta\Xi] \\ D &= D_{MZ} + \rho^2 D_{MZ} - 2\rho^2(1 + \delta) \\ D_{MZ} &= 1 + 2\delta + \delta^2\Xi \\ \Xi &= (1 - \rho^2)^{-1} \end{aligned} \tag{12}$$

To visually illustrate the implications of Corollary 3, Figure 7a displays BIRs for varying diagnosticity and signal persistence parameters. Using diagnosticity estimates from Bordalo et al. (2020a) and Bordalo et al. (2019b), we set $\theta \in \{0.5, 0.9\}$. For signal persistence, we explore two extremes, setting $\rho \in \{0.1, 0.9\}$.

Figure 7a reveals several strengths of the diagnostic expectations framework. First and foremost, the initial level $B(n, 0)$ of the model-implied BIR lines up well with those found in the SPF and IBES datasets (Figures 2a and 2b). Moreover, holding θ fixed, a lower signal persistence ρ results in a more pronounced overreaction, which is directionally consistent with Bordalo et al. (2020a), Afrouzi et al. (2023), and the analysis of persistence that we present in Internet Appendix IA.3.1.

While Figure 7a demonstrates how diagnostic expectations can match the level of observed BIRs, it also highlights a potential shortcoming of the model. To understand the core issue, note that for $J = 1$, the belief distortion introduced into the time-t diagnostic forecast $F_t[y_{t+n}]$ depends on the one-period rational revision, $\mathbb{E}_t[y_{t+n}] - \mathbb{E}_{t-1}[y_{t+n}]$. Consequently, as shown in Figure 7a and Corollary 3, diagnostic forecasters should, on average, fully correct their overreaction to news within a single period. 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. (2024). We demonstrate this point formally via the following proposition.

Proposition 4 *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 4 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 7a illustrates this result by plotting the BIR for $J=3$, $\theta=0.9$, and $\rho=0.9$. The formula for diagnostic BIRs with $J>1$ is derived formally in Appendix [INSERT]. For $J=3$, 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.

For SPF forecasters, this pattern of abrupt correction appears to be a reasonable approximation of reality. For instance, Figure 2a shows that SPF forecasters correct nearly 80% of their overreaction at time t within one period, consistent with $J=1$ in the diagnostic expectations model. However, the same is not true for IBES forecasters, who exhibit a much more gradual and smooth correction process (Figure 2b): half of the overreaction in the average IBES time-t forecast for $n=5$ remains in forecasts made one year later, at $t+4$. Moreover, there is no visible discontinuity in the BIR for equity analysts.

Heterogeneous speeds of correction, like those seen in Figure 5, also pose a challenge for models featuring diagnostic expectations. The reason why is that the level and shape of the diagnostic-BIR are largely pinned down by the diagnosticity and persistence parameters, θ and ρ. Consequently, it is difficult for such models to produce two BIRs that have identical initial levels of overreaction but differ in their speeds of correction.[12]

[12]One possibility is that the BIR with fast correction (blue) in Figure 5 is generated from a diagnostic

In summary, standard parameterizations of diagnostic expectations are effective at explaining the BIRs derived from individual forecasts in the SPF. However, these models have a harder time matching the smoother correction processes observed in the IBES forecasts. Furthermore, they face challenges in accounting for the heterogeneity in correction speeds observed across both the SPF and IBES datasets, particularly the pronounced variability seen in the latter.

4.3.2 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 model is a part of the linear autonomous class. Applying Proposition 2 to this model yields the following BIR.

Corollary 5 *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:*

$$F_t[y_{t+n}] \;=\; const(n) \;+\; E_t[y_{t+n}] \;+\; \delta(n)E_t[y_{t+n}], \tag{13}$$

with

$$\delta(n) \;=\; (1-\rho^n)\rho^{-n} \min\left\{1,\; \left(\frac{\omega\underline{\tau}}{(1-\rho^n)^2}\right)^{1/(1+\gamma)}\right\} \tag{14}$$

model with $J = 1$, whereas the BIR with slow correction (orange) has a larger J. However, under Proposition 4, the slow-correcting BIR should also have a discontinuity at $B(n, J)$, yet it is smooth in the data.

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

$$B(n,j) \;=\; A(n,j)\frac{1-\rho^2}{D(n)} \tag{15}$$

where

$$\begin{aligned} A(n,j) \;&=\; -\delta(n-j)\Big[1+\delta(n)\Big] \\ D_{MZ}(n) \;&=\; (1+\delta(n))^2 \\ D(n) \;&=\; D_{MZ}(n)+\rho^2 D_{MZ}(n+1) \;-\; 2\rho^2(1+\delta(n+1))(1+\delta(n))\,. \end{aligned} \tag{16}$$

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

Figure 7b uses Corollary 5 to plot BIRs for two different parameterizations from Afrouzi et al. (2023), one for when $\rho = 0.2$ and another for $\rho = 0.95$. These values roughly correspond to the average AR(1) coefficients we estimate for the low- and high-persistence series, respectively, in Internet Appendix IA.3.1. In both setups, we fix the variance σ_ε^2 of ε_t at 0.2. Additionally, we set $\omega = 1$ and $\gamma = 2$, which define the scale and convexity of information processing costs, respectively. Following Section 6.2 of Afrouzi et al. (2023), we set the baseline precision $\underline{\tau}$ that agents have about the long-run mean of y equal to $(1-\rho^2)/\sigma_\varepsilon^2$.

The orange line in Figure 7b depicts the model-implied BIR when y exhibits high persistence ($\rho = 0.95$). This model's initial level of overreaction is comparable to what is observed in Figure A6 for SPF forecasts of series with high persistence. However, the correction process in the model is slow, with the majority of the time-t overreaction still present in the $t+1$ forecast. In contrast, SPF forecasts of highly persistent series demonstrate a faster correction, although the standard errors are wide enough that the model's $B(n,1)$ falls within a 2-standard error range of the empirical $B(n,1)$ in the data.

The blue line in Figure 7b shows the model-implied BIR when y exhibits low persistence

($\rho = 0.2$). As demonstrated by Afrouzi et al. (2023), the initial degree of overreaction is significantly higher in this scenario, aligning closely but slightly more negatively with its empirical counterpart in Figure A6. Furthermore, the speed of correction for the high-ρ series is similar to that observed empirically, with the vast majority of the overreaction at time-t being corrected by $t+1$.

Overall, Figure 7b shows that the model of costly information processing in Afrouzi et al. (2023) is fairly successful at matching both the level and shape of BIRs from the SPF data, as well as their dependence on the persistence of the process being forecast. Nevertheless, Corollary 5 also highlights a weakness of the model, namely that the model-implied BIR is concave in j. In other words, following an initial overreaction, forecasters in the model gradually correct their mistakes but at a decelerating rate. This means that the model cannot match many of the BIRs observed in the IBES data. For instance, in Figure 5, we find that for some forecasters and series in IBES, the BIR is convex in j and thus implies an extremely slow correction process.

4.3.3 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. As in Hirshleifer et al. (2015), we assume 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.[13] In addition, forecasts are assumed to evolve according to:

$$\begin{aligned} F_t[y_{t+1}] &= \mu_t \\ \mu_{t+1} &= (1-\rho-\tilde{\rho})\bar{\mu} + \rho\mu_t + \tilde{\rho}y_{t+1}, \end{aligned} \tag{17}$$

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 2 to this model then yields the following BIR.

Corollary 6 (BIRs for extrapolative expectations) *In the model of extrapolative expectations given by* (17)*, the BIR is given by:*

$$B(n,j) = -\frac{\rho^j}{(\rho+\tilde{\rho})^j}\frac{1-\rho(\rho+\tilde{\rho})}{(1+(\rho+\tilde{\rho})^2-2\rho(\rho+\tilde{\rho}))} < 0\,. \tag{18}$$

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 6 shows that extrapolative expectations can generate both overreaction ($B(n,0)<0$)) and the pattern of smooth decay observed in empirically estimated BIRs. However, the corollary also reveals that the extrapolative expectations model in (17) fails to accurately reflect how the level of BIRs varies with the forecasting horizon n (see Figure 2b). This follows from the fact that n appears nowhere in the expression for the BIR.

To visualize the BIR implied by extrapolative expectations, Figure 7c plots the BIR using the following parameters: $n=6$, $\bar{\mu}=0$, $\tilde{\rho}=0.018$, and $\rho=1-\tilde{\rho}$. Under this parameteriza-

[13] In Appendix D, we allow for an autoregressive structure in y_t, 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.

tion, the model of extrapolative expectations in (17) 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, in this case, equals -0.5 and is comparable to what is observed empirically. However, in addition to its lack of dependence on forecast horizon n, the BIR is essentially flat under this parameterization of extrapolative expectations. The implication is that following overreaction to the news at time t, extrapolative agents counterfactually correct their mistakes at a very slow pace — essentially, all of their initial overreaction remains after five quarters.

4.4 Discussion

The analysis in this section leverages BIRs to assess the strengths and weaknesses of prominent behavioral models of belief formation. A key finding is that models of diagnostic expectations Bordalo et al. (2020b) and costly information processing Afrouzi et al. (2023) produce BIRs that are relatively realistic when compared to those observed in the SPF, though the latter model most accurately captures how the level of BIRs depends on the persistence of the forecast target. However, both models have a harder time matching the BIRs from the IBES data, which suggests a much slower correction process. Moreover, BIRs from IBES often display heterogeneity that is incompatible with these models, such as their dependence on the state of the economy or the sentiment of news.

At a broader level, the results in this section also highlight the benefits that BIRs offer over traditional impulse responses of forecast errors to identified structural shocks. For one, the discontinuous correction process that is inherent to models of diagnostic expectations is not readily apparent in the impulse responses studied by Angeletos et al. (2021) and Kučinskas and Peters (2022). In addition, our analysis of how BIRs vary across forecasters and stocks would be considerably more difficult to implement with traditional impulse response methods because of the latter's dependence on well-identified shocks. For instance,

with traditional impulse responses, analyzing whether equity analysts overreact or underreact to specific stocks would require identifying exogenous shocks to those stocks. Furthermore, as mentioned in the introduction, BIRs capture how forecast errors dynamically respond to all news received at time t, not just a particular shock, which is arguably what matters for economic outcomes (e.g., prices).

5 BIRs and Model Uncertainty

In this section, we show how BIRs can be paired with simulations to explore models of belief formation that lie outside of the linear autonomous class studied in Section 4. We specifically consider belief formation for Bayesian forecasters in the presence of model uncertainty, mirroring the setup 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:

$$\begin{aligned} y_t &= \mu_t + x_t, \\ \mu_t &= \mu_{t-1} + \sqrt{\gamma}\sigma\eta_t, \quad \eta_t \sim N(0,1) \\ x_t &= \rho x_{t-1} + \sqrt{1-\gamma}\sigma\omega_t, \quad \omega_t \sim N(0,1) \end{aligned} \tag{18}$$

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:

$$\rho \sim N(\mu_p, \sigma_p^2), \quad \gamma \sim \mathcal{B}(\alpha_\gamma, \beta_\gamma), \quad \sigma^2 \sim \mathcal{IG}(1.25, 0.5625), \tag{19}$$

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 by assuming parameter values for ρ, γ, and σ and then study the time-series of forecasts for different 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 the ones found in the SPF and IBES datasets. Our simulation exercise specifically proceeds as follows. First, we simulate 5 different paths of y_t according to the DGP in (18), each of length $T = 260$ (for the IBES, we use $T = 190$ to match the time frame of the real data). When matching SPF forecasts, we set the true $\rho = 0.02$ and $\gamma = 0.02$. For IBES forecasts, we set $\rho = 0.5$ and $\gamma = 0.25$. For both SPF and IBES analysis, we set $\sigma = 0.5$. As we discuss below, these parameters were chosen so that the simulated y_t roughly resembles the average realized series in the SPF and IBES datasets, at least in terms of their observed persistence. Next, we set forecasters' beliefs about the initial states to be $\mu_1 \sim N(y_1, 1)$ and $x_1 \sim N(0, 1)$, with the two perfectly negatively correlated given the DGP in Equation (18). Armed with their initial beliefs about the states (μ_1, x_1) and parameters of the DGP contained in Θ, 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 updating 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.

Next, we search over the initial belief vector Θ to best match the average individual BIR observed in the SPF or IBES datasets, as plotted in Figures 2a and 2b, 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 8a plots the resulting simulated BIR along with the true one from the SPF. The figure shows that the model fits the data fairly well. The initial levels of the BIRs are nearly identical, indicating a similar level of initial overreaction at time t. The speed of correction in the simulated data is also similar, albeit a bit slower relative to the data. In the model, roughly half of the overreaction in the time-t forecast is eliminated by $t + 1$, whereas in the data, it is closer to 75%. Figure 8c 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 0.2 and 0.3, respectively.

Figure 8b presents the simulated BIR (in blue) when targeting the average individual BIR from IBES (in orange). Once again, the model is able to generate a relatively realistic BIR. The initial level of overreaction in the model, as captured by $B(n, 0)$, is very close to the data. In addition, the slow speed of correction embedded in the true BIR is also present

in the model. In the data, following an overreaction at time t, roughly eighty percent of the overreaction is still present in the $t+2$ forecast, compared to about 65% in the model. This is an interesting result given that, as we showed in Section 4, standard behavioral models have a hard time matching the pace of correction in IBES forecasts. With that said, the speed of correction does accelerate too quickly in the model for $j > 2$ relative to the data.

Figure 8d displays the initial belief distributions that underlie Figure 8b. Unlike the SPF, overreaction, in this case, occurs through a mixture of downward biased beliefs about ρ and upward biased beliefs about γ: beliefs about ρ are centered 0.05, well below the true $\rho = 0.5$, and beliefs about γ are centered around 0.99, well above the true $\gamma = 0.25$. To develop some intuition for how beliefs about how γ and ρ modulate the speed of correction, the green line in Figure 8b 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 exercise suggests that upward-biased beliefs about γ are the driving force behind the initial overreaction, whereas beliefs about ρ are a key determinant of the correction process.

It may be tempting to view Figure 8 as evidence that model uncertainty can generate realistic BIRs at the individual level, even when corrections to overreactions at time t are slow. This interpretation hinges on two key questions. First, does our simulated DGP realistically approximate the series forecasted by SPF forecasters and IBES analysts? While a definitive answer to this question is outside of the scope of this paper, Figure 9 offers a preliminary answer by comparing sample paths of the simulated y with the average realized y in both the SPF and IBES datasets. For instance, Figures 9a and 9b show the cross-sectional average realized SPF series alongside three sample paths of the simulated y, and Figure 9c and 9d do the same for the IBES series. Visually, the true and simulated series appear similar, at least in terms of their persistence. The AR(1) coefficients for the average

and simulated realized SPF series equal 0.33 and 0.38, respectively. Similarly, the AR(1) coefficients for the average and simulated IBES series are 0.95 and 0.87, respectively.[14]

The second question is whether forecasters' initial beliefs Θ^* about the parameters governing y in our simulation are reasonable, as they are crucial for driving both the initial degree of overreaction and the speed of correction. This question is challenging to answer definitively and presents a valuable direction for future research. For the time being, we take Figure 8 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 explaining deviations from FIRE in professional forecasts, as there are strong career incentives to generate forecasts that are free of behavioral biases. More broadly, the results in this section highlight how BIRs can be paired with simulations to explore belief formation in situations where a closed-form BIR is not readily available.

6 Conclusion

In this paper, we develop the concept of a behavioral impulse response function, which captures how forecast errors and revisions of a fixed target dynamically respond following a forecast revision made at time t. Intuitively, BIRs capture how much a forecast overreacts or underreacts at time t, as well as how quickly they then correct their initial mistake. BIRs offer more powerful tests of FIRE relative to existing methods like the one proposed by Coibion and Gorodnichenko (2015) because they impose joint restrictions on how all future forecast errors and revisions respond to a revision at time t. They can be easily estimated using OLS regressions on any dataset that features 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

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

individual forecasters to overreact when forecasting distant targets. While overreactions are corrected relatively quickly in the SPF data, corrections in the IBES data are notably slower. Additionally, BIRs demonstrate significant variation across different economic states, forecasters, and stocks within the IBES dataset. Although popular behavioral models of belief formation align well with the BIRs from the SPF, they are less successful at matching those from the IBES. We also highlight how BIRs can be used 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.

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

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.

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

Binsbergen, Jules H Van, Xiao Han, and Alejandro Lopez-Lira, "Man versus machine learning: The term structure of earnings expectations and conditional biases," *The Review of Financial Studies*, 2023, *36* (6), 2361–2396.

Bordalo, Pedro, Gennailoli Nicola, Rafael La Porta, and Andrei Shleifer, "Diagnostic Expectations and Stock Returns," *The Journal of Finance*, 2019, *74* (6), 2839–2874.

___ **, 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 Robert J. Shiller, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *The Review of Financial Studies*, 1988, *1* (3), 195–228.

Cochrane, John H., "Presidential Address: Discount Rates," *The Journal of Finance*, 2011, *66* (4), 1047–1108.

___ **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.

Fama, Eugene F and Kenneth R French, "Common risk factors in the returns on stocks and bonds," *Journal of financial economics*, 1993, *33* (1), 3–56.

___ **and** ___ , "A five-factor asset pricing model," *Journal of financial economics*, 2015, *116* (1), 1–22.

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

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.

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

Giglio, Stefano and Bryan Kelly, "Excess Volatility: Beyond Discount Rates*," *The Quarterly Journal of Economics*, 08 2017, *133* (1), 71–127.

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

Jensen, Theis Ingerslev, Bryan Kelly, and Lasse Heje Pedersen, "Is there a replication crisis in finance?," *The Journal of Finance*, 2023, *78* (5), 2465–2518.

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

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

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.

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.

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.

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

Shiller, Robert J., "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?," *The American Economic Review*, 1981, *71* (3), 421–436.

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

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

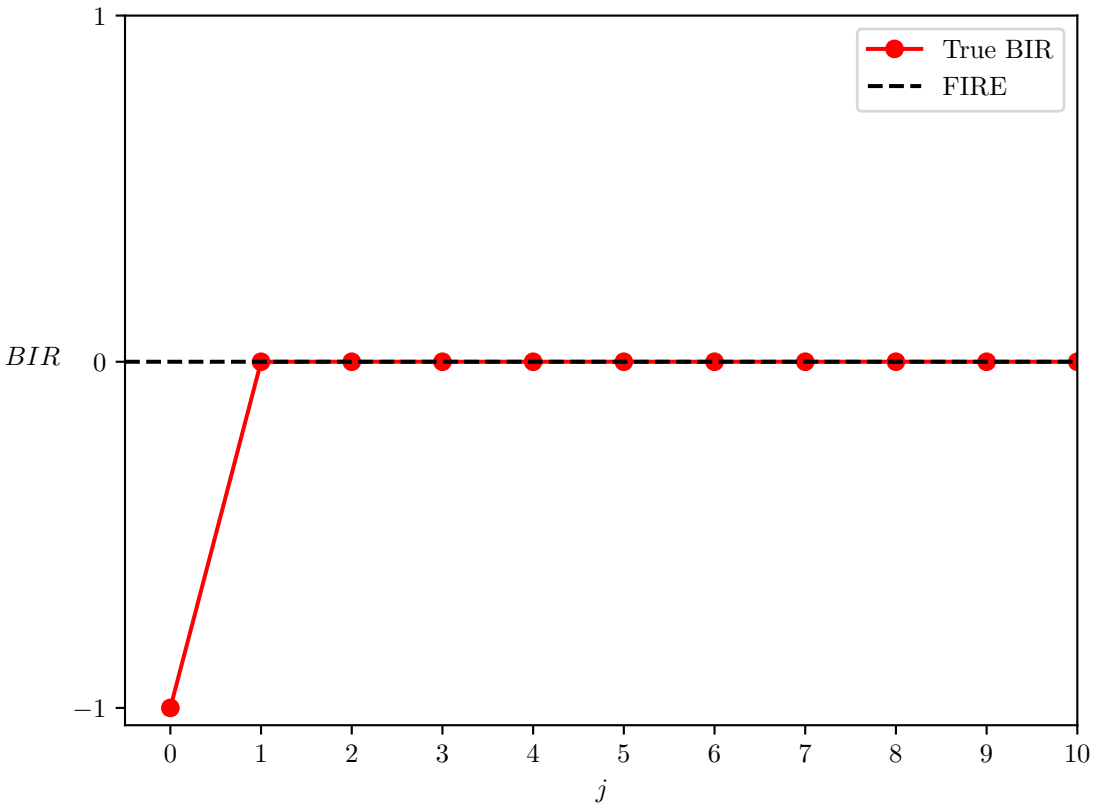


Figure 1: Stylized Example of a BIR. This figure shows a stylized BIR for an iid, mean zero forecast target. The forecaster generating the BIR is assumed to have a rational forecast up to time t, after which they irrationally revise their forecast up by one unit. They then realize their mistake at time $t+1$ and revise their forecast accordingly.

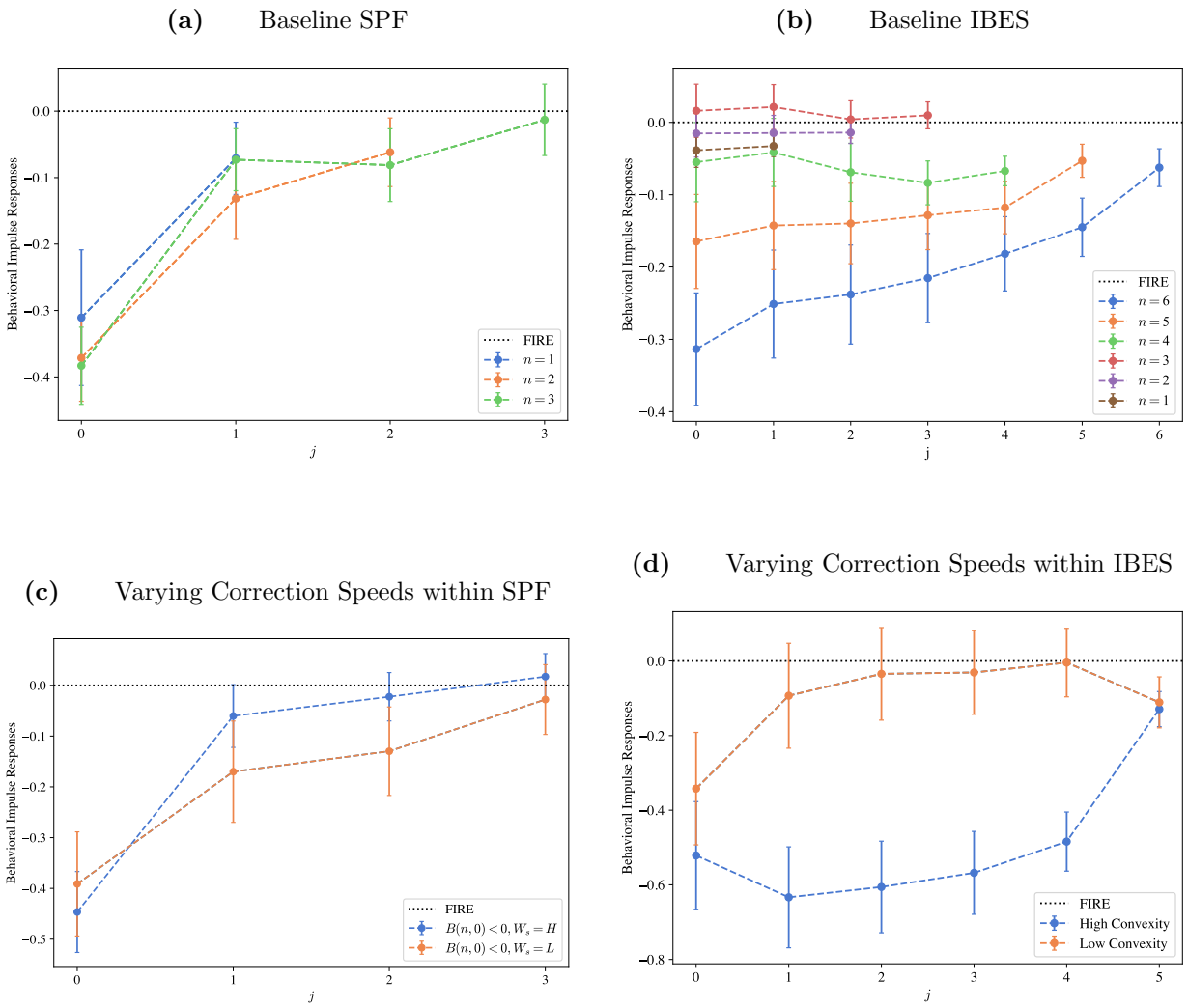


Figure 2: BIRs from SPF and IBES. Panels (a) and (b) show baseline BIR estimates for different in n in SPF and IBES, respectively. Panel (c) contains BIR estimates for overreacting series in SPF that exhibit high and low individual BIR-convexity. Panel (d) repeats the exercise from panel (c) for IBES stocks. In panels (a) and (c), 95% confidence intervals are based on standard errors that are clustered by forecaster-series, (i, s), and series-date, (s, t). In panels (b) and (d), they are based on standard errors clustered by analyst-stock, (i, s), and stock-date, (s, t). In all plots, the black dashed line shows the BIR under FIRE. See Section 2.2 for complete estimation details.

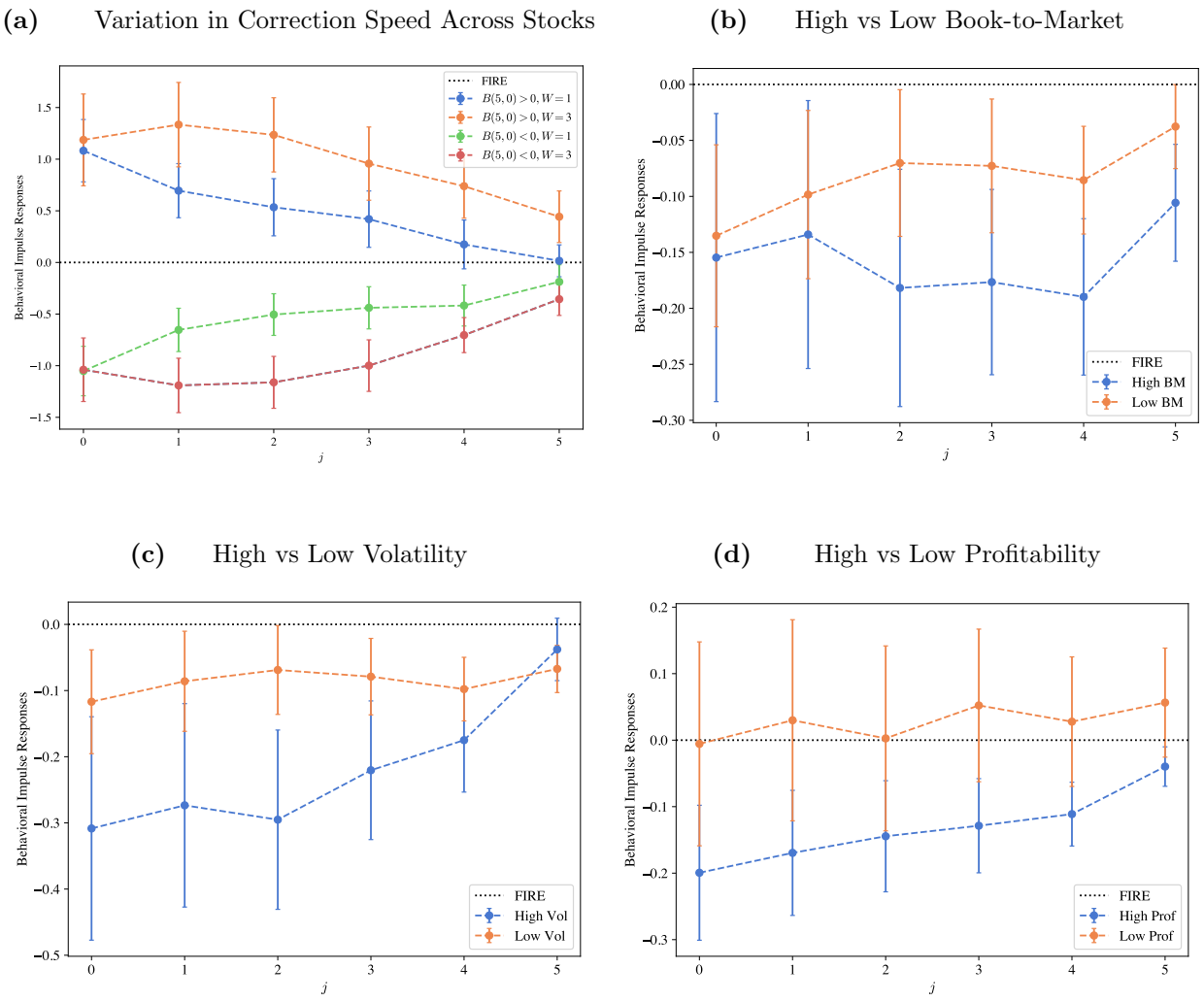


Figure 3: Heterogeneity in Correction Speeds This figure shows how BIRs vary within subsamples of the I/B/E/S quarterly earnings forecast data. Panel (a) sorts stocks based on whether their forecasts initially over or underreact, then further sorts them into terciles based on how quickly deviations from FIRE correct, as measured by $W = B(5,1)/B(5,0)$. BIRs are then estimated by pooling all stock within a given tercile of W. In panels (b)-(d), we first assign a firm to a group based on a monthly cross-sectional factor sort using the firm-level characteristic data from Jensen et al. (2023); the top (bottom) one-thrid goes to the high (low) category. Then, we separate the full sample by the stock category at the end of the quarter prior to the quarter when the forecast on quarter $t+n$ earnings is submitted. Panels A to C thus show the BIRs on the factor-based subsamples for book-to-market (BM), monthly realized volatility (Vol), and profitability (Prof). The black dashed line shows the full information rational expectations benchmark.95% confidence intervals in the plot are based on standard errors clustered by analyst-stock, (i,s), and stock-date, (s,t)

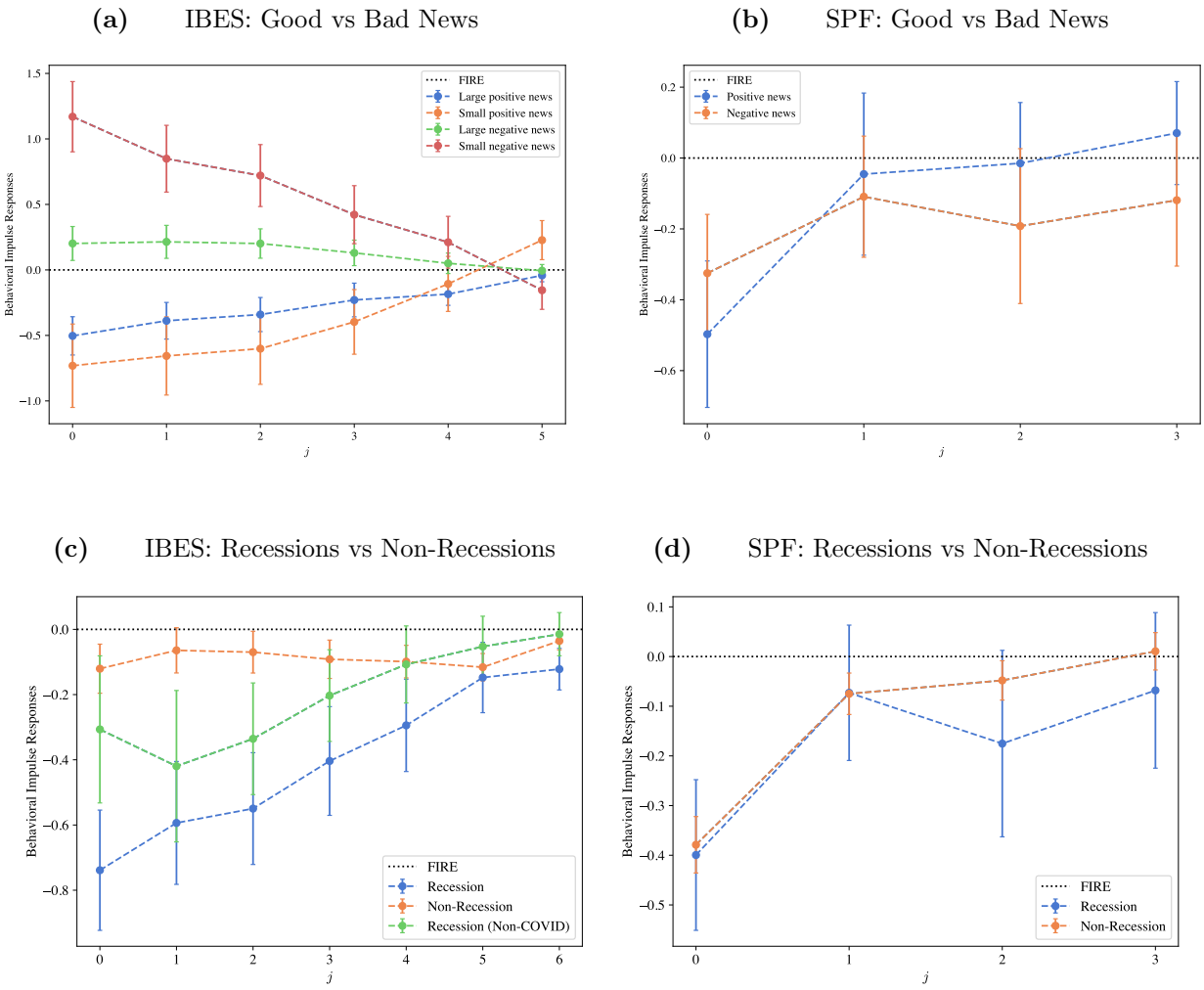


Figure 4: BIRs and the Nature of Shocks. This figure shows how BIRs vary with the nature of news received by forecasters. Panel (a) shows BIR that condition on the sign and size of the revision made by each forecaster. Large positive (negative) revisions are defined as the top tercile of positive (negative) revisions and small positive (negative) revisions are all others. Panel (b) shows a similar result for the SPF, but only conditions on the sign of revisions. Panel (c) conditions on whether revisions in IBES are made during NBER recessions. Panel (d) repeats panel (c) for SPF.

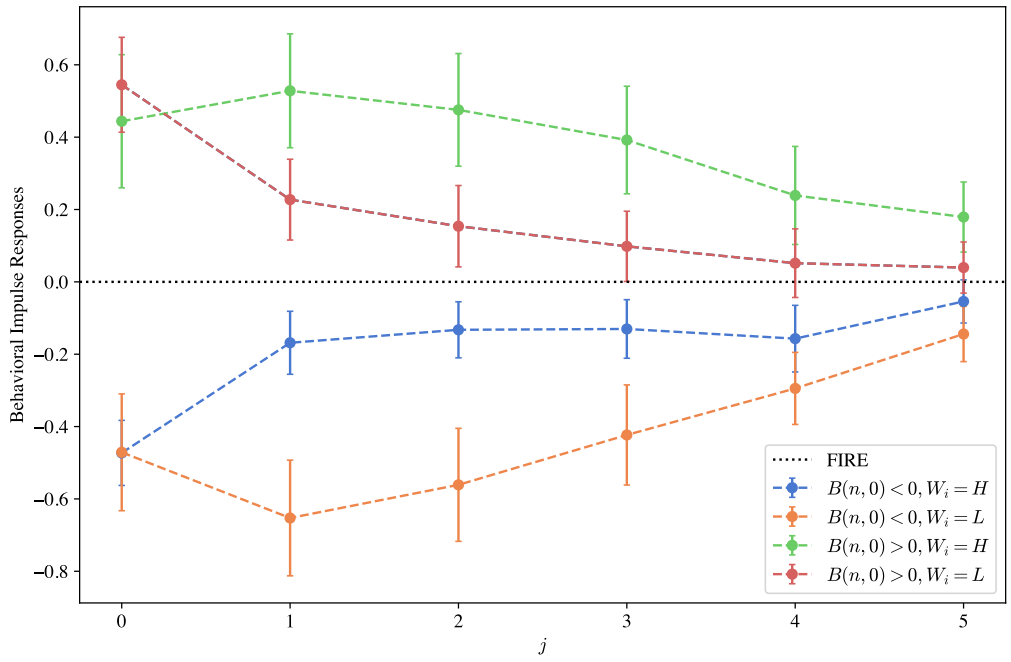


Figure 5: Heterogeneity Across Equity Analysts This figure shows how BIRs vary across I/B/E/S forecasters. Each curve shows an estimated BIR($n = 5$) for a subset of analysts, along with 95% confidence intervals with standard errors clustered by analyst-stock, (i, s), and stock-date, (s, t). The blue and orange lines arise from the following sorting procedure. First, we estimate BIRs for each analyst i, retaining only those where $B_i(5, 0) < 0$, and sort analysts into terciles based on $B_i(5, 0)$. Within the middle $B_i(5, 0)$-tercile of analysts, we then further sort analysts into terciles based on $W_i = B_i(5, 1) - B_i(5, 0)$, a measure of how quickly analysts correctly adjust future forecasts. The orange line shows the BIR estimated using data for all analysts in the high W_i group, and the blue line shows the BIR estimated for the low W_i group. The red and green lines repeat the same process but for analysts who initially underreact ($B_i(5, 0) > 0$). Analysts must include at least ten forecasts and forecasts made during the Covid-19 pandemic are excluded. The black dashed line shows the full information rational expectations benchmark.

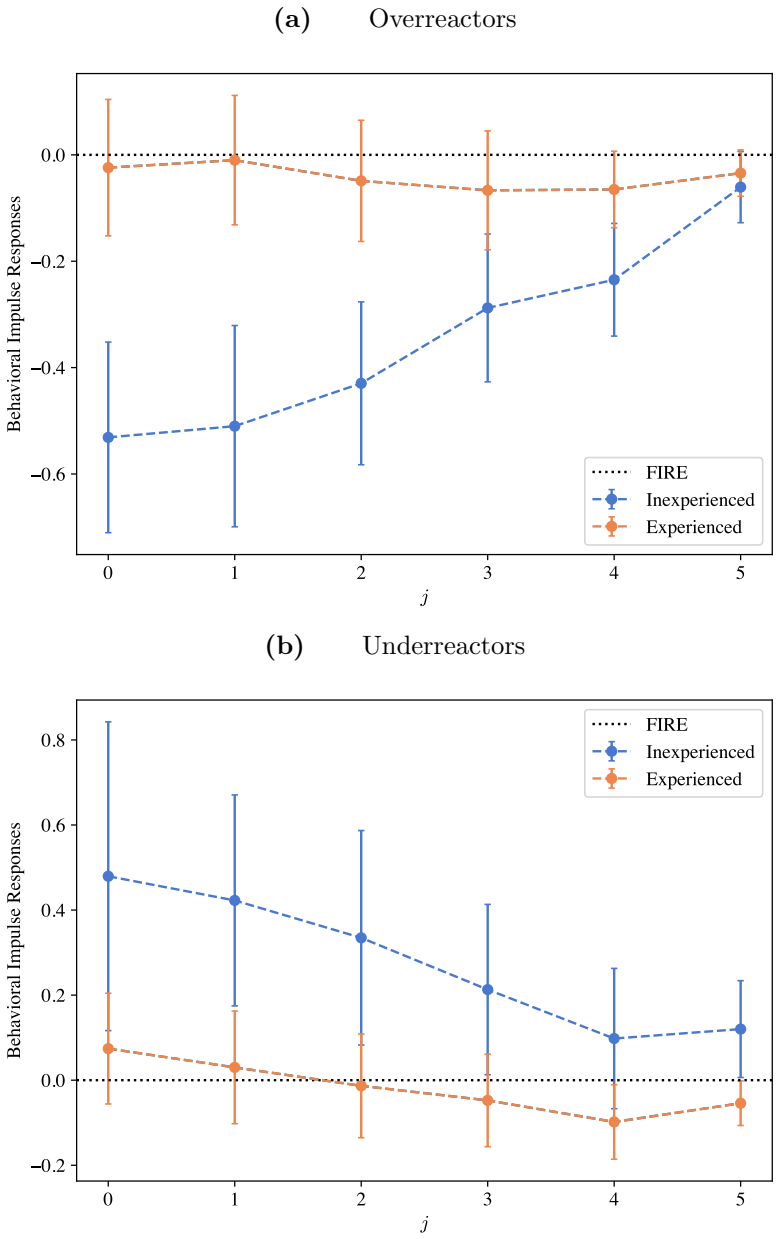


Figure 6: BIRs and Learning. This figure shows how BIRs vary across phases of an analyst's career in the I/B/E/S data. Each curve shows an estimated BIR using data pooled across several forecasts and analysts, all for $n = 5$. The bars show the corresponding 95% confidence intervals with standard errors clustered by analyst-stock, (i, s), and stock-date, (s, t). We focus this analysis on analysts who have at least ten years of submissions. We use the first five-year of data of each individual analyst i to estimate their BIR, $B_i^{early}(n, j)$, and take the average for each individual to get B_i^{early}; for each $B_i^{early}(n, j)$ estimation, we require the analyst to have at least ten contributions. We sort analysts by their B_i^{early} and assign them to terciles. Within each tercile, we then estimate BIRs using: (i) the first five-years of forecasts made by the analysts (labeled as the "inexperienced" BIRs in blue); and (ii) all forecasts made after year five (labeled as the "experienced" BIR in orange). Panel (a) panel shows the estimates for analysts in the bottom B_i^{early}-tercile, and panel (b) shows them for the top B_i^{early}-tercile. Forecasts made during the COVID-19 pandemic are excluded because all forecasts deviated wildly from FIRE during this time (see Section **??**). The black dashed line shows the full information rational expectations benchmark.

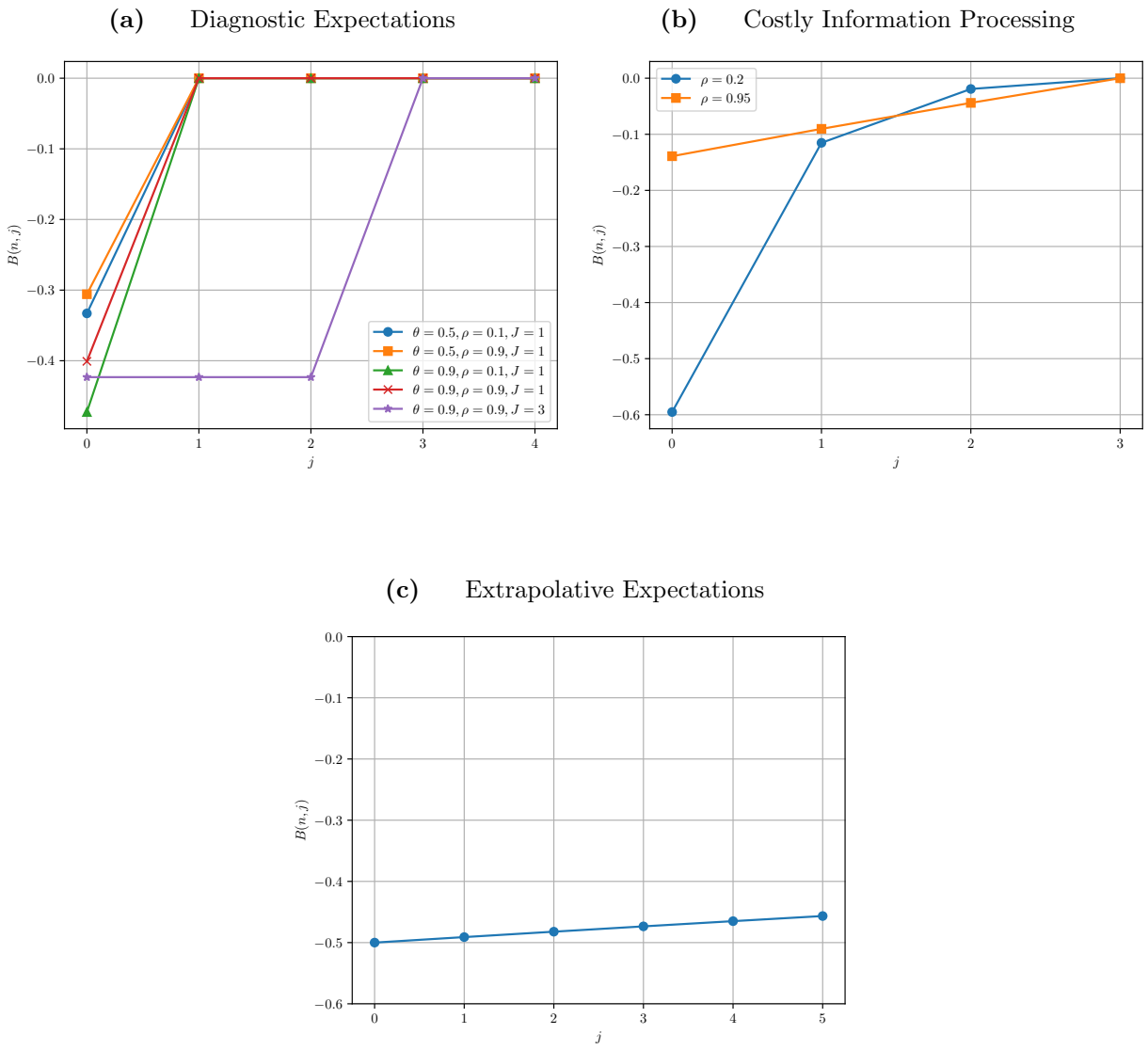


Figure 7: BIRs under Diagnostic and Extrapolative Expectations Panel (a) of this figure shows BIRs for different levels of diagnosticity (θ) and signal persistence (ρ) under the model of diagnostic expectations from Bordalo et al. (2020a). Panel (d) shows the BIR from Afrouzi et al. (2023), with the exact calibration detailed in Section 4.3.2. Panel (c) 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).

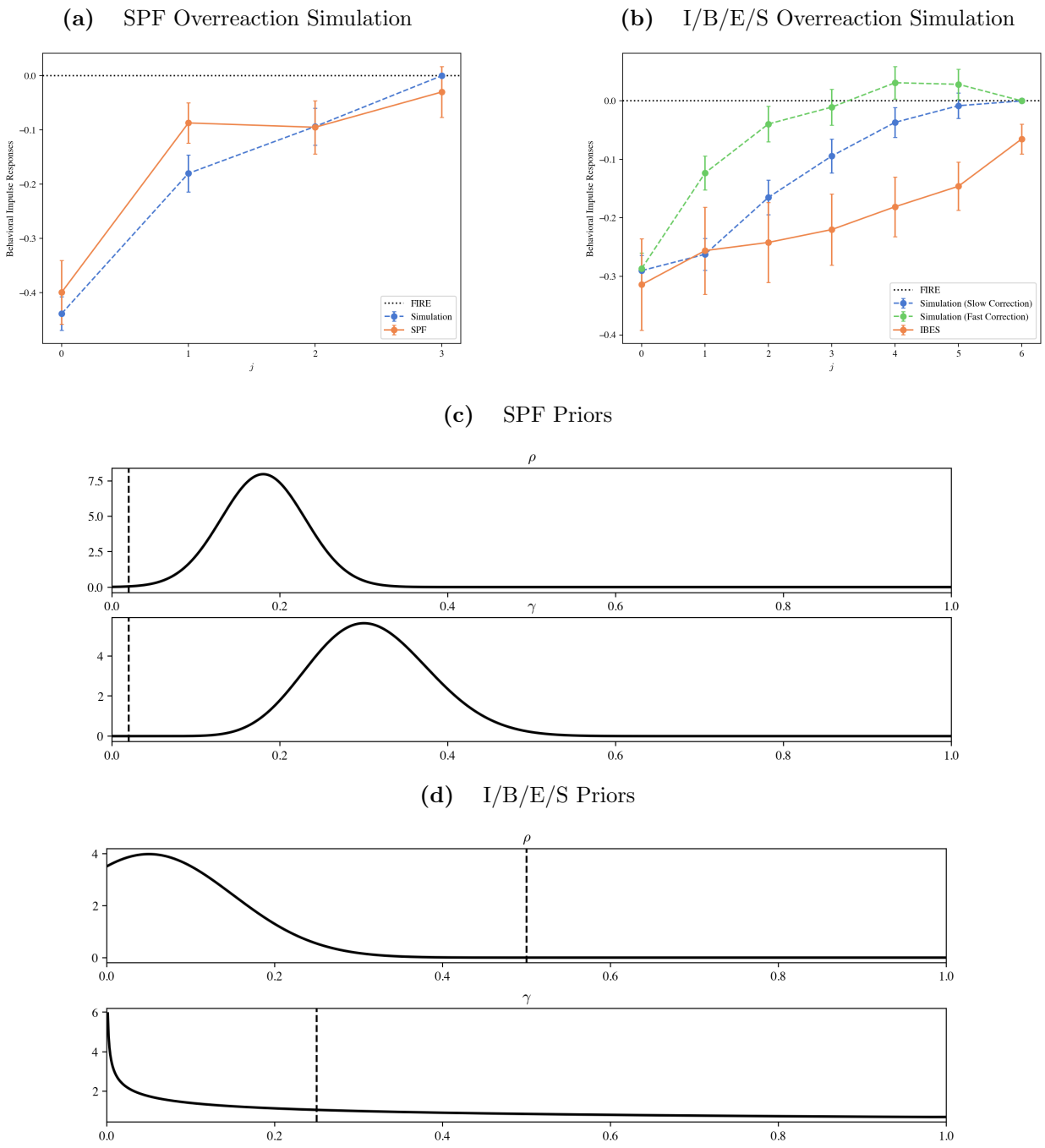
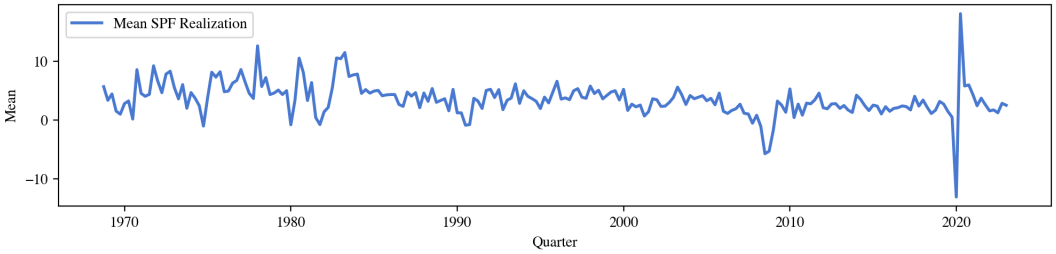
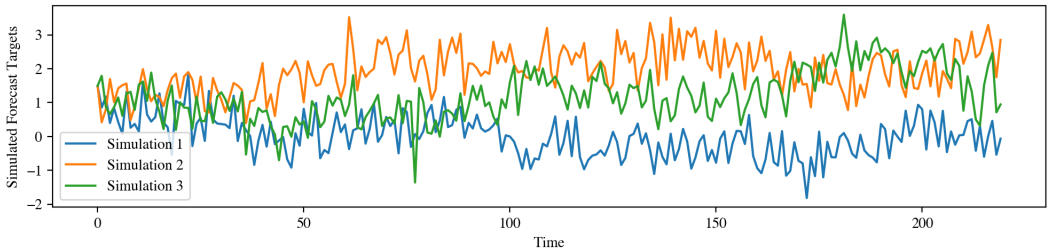


Figure 8: BIRs with Model Uncertainty Panel (a) of this figure shows BIRs estimated on the pooled SPF forecasts with $n = 3$. The plot also shows BIRs from the simulated model of Farmer et al. (2023) with the best-matched priors. Panel (b) shows the BIR for the I/B/E/S pooled forecasts with $n = 6$ and the simulated Bayesian forecasters in Farmer et al. (2023) with the best-matched priors. Panel (c) and (d) show the best-matched priors with respect to the SPF and I/B/E/S 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

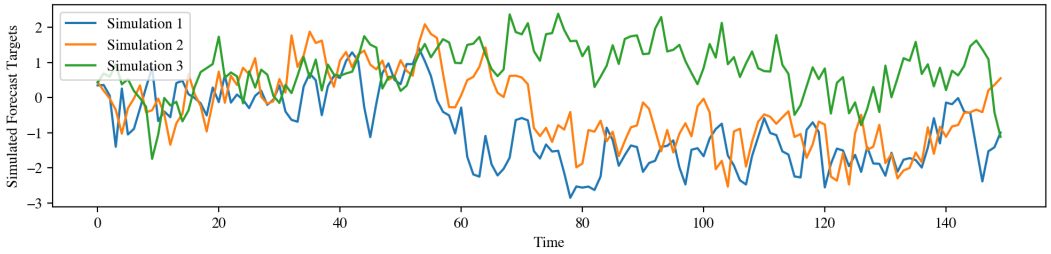


Figure 9: BIRs with Model Uncertainty Panel (a) of this figure shows the quarterly cross-sectional average realization of the SPF forecast targets and counts of stocks in the I/B/E/S sample from 1985Q1 to 2022Q3. Panel (b) shows three randomly simulated paths according to Farmer et al. (2023) with $\rho = 0.02, \gamma = 0.02$. Panel (c) shows the quarterly cross-sectional average EPS in the I/B/E/S sample from 1968Q4 to 2023Q1. Panel (d) shows three randomly simulated paths according to Farmer et al. (2023) with $\rho = 0.50, \gamma = 0.30$.

	(1)	(2)	(3)
$B(n,0)$	-0.383^{***}	-0.371^{***}	-0.311^{***}
	(0.030)	(0.033)	(0.052)
$B(n,1)$	-0.073^{***}	-0.132^{***}	-0.071^{**}
	(0.024)	(0.031)	(0.028)
$B(n,2)$	-0.081^{***}	-0.062^{**}	
	(0.028)	(0.026)	
$B(n,3)$	-0.013		
	(0.027)		
n	3	2	1
Adj. R^2 (%)	0.73	1.05	1.39
Observations	59,881	63,276	66,192
p(FIRE)	<0.001	<0.001	<0.001
p(Equal B)	<0.001	<0.001	<0.001

Table 1: BIRs Estimated from SPF Data. This table shows point estimates for BIRs based on the SPF forecast data. Columns (1), (2), and (3) show BIR estimates for $n = 3, 2, 1$, 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, namely that all points along the BIR curve equal zero, 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 number of observations and adjusted-R^2 are 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.

	(1)	(2)	(3)	(4)	(5)	(6)
$B(n,0)$	-0.314^{***}	-0.162^{***}	-0.052^{*}	0.017	-0.014	-0.036^{***}
	(0.040)	(0.033)	(0.028)	(0.019)	(0.017)	(0.012)
$B(n,1)$	-0.256^{***}	-0.140^{***}	-0.039	0.022	-0.014	-0.033^{***}
	(0.038)	(0.031)	(0.024)	(0.016)	(0.013)	(0.007)
$B(n,2)$	-0.242^{***}	-0.137^{***}	-0.067^{***}	0.005	-0.014^{*}	
	(0.035)	(0.029)	(0.021)	(0.013)	(0.008)	
$B(n,3)$	-0.220^{***}	-0.125^{***}	-0.082^{***}	0.009		
	(0.031)	(0.024)	(0.016)	(0.009)		
$B(n,4)$	-0.181^{***}	-0.115^{***}	-0.067^{***}			
	(0.026)	(0.019)	(0.010)			
$B(n,5)$	-0.146^{***}	-0.051^{***}				
	(0.021)	(0.012)				
$B(n,6)$	-0.065^{***}					
	(0.013)					
n	6	5	4	3	2	1
Adj. R^2 (%)	0.66	0.26	0.11	0.00	0.01	0.06
Observations	39,166	77,492	139,012	234,203	412,021	635,570
p(FIRE)	<0.001	<0.001	<0.001	0.005	0.019	<0.001
p(Equal B)	<0.001	<0.001	<0.001	0.007	0.992	0.420

Table 2: BIRs from IBES Data. This table shows point estimates for BIRs based on the I/B/E/S quarterly earnings forecast data. 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, namely that all points along the BIR curve equal zero, 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 number of observations and adjusted-R^2 are 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.

A Linear Autonomous Forecasts and Leading Behavioral Models

We now show that many popular models of belief formation produce forecasts that a part of 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., 2024) 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., 2024) 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 that

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

holds, we get

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

which is clearly a special case of 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

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

and show that, with limited information capacity,

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

where the dependence of δ_1, δ_2 on n is subtle and non-linear. We discuss this model below. Here, we only note that it obviously fits 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

Proposition 7 *Assuming that*

(i) The market discount rate is constant at r,

(ii) The stock analysts' expectations represent the market's, i.e., the price of the dividend strip at $t+n$ is

$$P_t(n) \;=\; e^{-rn}F_t(n) \;\equiv\; e^{-rn}F_t[D_{t+n}].$$

Then the expected price of $P_{t+j}(n-j)$ conditional on a unit revision made today is

$$\mathbb{E}[P_{t+j}(n-j)|R_t(n)=1] \;=\; e^{-r(n-j)}\left(\theta(n) \;-\; B(n,j)\right), \quad \forall\ 0 \le j \le n, \tag{32}$$

where

$$\theta(n) \;=\; E[D_{t+n}|R_t(n)=1]. \tag{33}$$

Proof of Proposition 7. By definition, we have

$$F_{t+j}(n-j) \;=\; D_{t+n} - e_{t+j}(n-j). \tag{34}$$

Therefore, by the assumption (i) of Proposition 7, there is

$$\begin{aligned}
\mathbb{E}[P_{t+j}(n-j)|R_t(n)=1] \;&=\; \mathbb{E}[e^{-r(n-j)}F_{t+j}(n-j)|R_t(n)=1] \\
&=\; \mathbb{E}[e^{-r(n-j)}(D_{t+n} - e_{t+j}(n-j))|R_t(n)=1] \\
&=\; \mathbb{E}[e^{-r(n-j)}D_{t+n}|R_t(n)=1] - \mathbb{E}[e^{-r(n-j)}e_{t+j}(n-j)|R_t(n)=1]
\end{aligned} \tag{35}$$

Hence, Equation (32) holds by Equation (33) and by the definition of $B(n,j)$. □

C Proofs for BIR Theory

Proposition 8 *Under the hypotheses of Proposition 2,*

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

where

$$\Xi(j) \;=\; \sum_{\tau=0}^{j}\kappa(\tau) \;+\; \sum_{\tau_1=1}^{\infty}\sum_{\tau_2=0}^{\infty}\kappa(\tau_1+j)\kappa(\tau_2)\rho^{2\max(\tau_1,\tau_2)} \tag{37}$$

and

$$\Gamma(j) \;=\; \sum_{\tau=0}^{j}\kappa(\tau)+\sum_{\tau=1}^{\infty}\kappa(\tau+j)\rho^{2\tau} \tag{38}$$

Proof of Propositions 8 and 2. We have

$$\begin{aligned}
E_t[F_{t+j}[y_{t+n}]] \;&=\; E_t[\sum_{\tau=0}^{\infty}\gamma(n-j,\tau)\rho^{n+\tau-j}q_{t+j-\tau}] \\
&=\; \sum_{\tau=j+1}^{\infty}\gamma(n-j,\tau)\rho^{n+\tau-j}q_{t+j-\tau} \\
&+\; q_t\sum_{\tau=0}^{j}\gamma(n-j,\tau)\rho^{n} \\
&=\; q_t\sum_{\tau=0}^{j}\gamma(n-j,\tau)\rho^{n} \;+\; \sum_{\tau=1}^{\infty}\gamma(n-j,\tau+j)q_{t-\tau}\rho^{n+\tau}
\end{aligned} \tag{39}$$

and

$$
\begin{aligned}
E_t[y_{t+n} - F_{t+j}[y_{t+n}]] \;&=\; E_t[y_{t+n}] - E_t[F_{t+j}[y_{t+n}]] \\
&=\; \rho^n q_t \;-\; q_t \sum_{\tau=0}^{j} \gamma(n-j,\tau)\rho^n \;-\; \sum_{\tau=1}^{\infty} \gamma(n-j,\tau+j) q_{t-\tau}\rho^{n+\tau} \\
&=\; \rho^n q_t \left(1 - \sum_{\tau=0}^{j} \gamma(n-j,\tau)\right) \;-\; \sum_{\tau=1}^{\infty} \gamma(n-j,\tau+j) q_{t-\tau}\rho^{n+\tau}
\end{aligned}
\tag{40}
$$

and hence

$$
\begin{aligned}
&\mathrm{Cov}(E_t[y_{t+n} - F_{t+j}[y_{t+n}]], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \\
&=\; \mathrm{Cov}(E_t[y_{t+n}], F_t[y_{t+n}]) \;-\; \mathrm{Cov}(E_t[y_{t+n}], F_{t-1}[y_{t+n}]) \\
&-\mathrm{Cov}(F_{t+j}[y_{t+n}], F_t[y_{t+n}]) \;+\; \mathrm{Cov}(F_{t+j}[y_{t+n}], F_{t-1}[y_{t+n}])\,.
\end{aligned}
\tag{41}
$$

We have

$$
\begin{aligned}
&\mathrm{Cov}(y_{t+n}, F_t[y_{t+n}]) \;=\; \mathrm{Cov}(E_t[y_{t+n}], F_t[y_{t+n}]) \;=\; \rho^{2n}\mathrm{Cov}(q_t, q_t + \delta(n)\sum_{\tau=0}^{\infty} \kappa(\tau)\rho^{\tau} q_{t-\tau}) \\
&=\; \rho^{2n}\sigma_q^2(1 + \delta(n)\Gamma)
\end{aligned}
\tag{42}
$$

and, similarly, by stationarity,

$$
\mathrm{Cov}(y_{t+n}, F_{t-1}[y_{t+n}]) \;=\; \mathrm{Cov}(y_{t+n+1}, F_t[y_{t+n+1}]) = \;\rho^{2(n+1)}(1 + \delta(n+1)\Gamma)\,. \tag{43}
$$

Then,

$$
\begin{aligned}
\mathrm{Cov}(F_{t+j}[y_{t+n}], F_t[y_{t+n}]) \;&=\; \sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \\
&+\; \mathrm{Cov}(q_t \sum_{\tau=0}^{j} \gamma(n-j,\tau)\rho^n \\
&+\; \sum_{\tau=1}^{\infty} \gamma(n-j,\tau+j) q_{t-\tau}\rho^{n+\tau}, \sum_{\tau=0}^{\infty} \gamma(n,\tau) q_{t-\tau}\rho^{n+\tau}) \\
&=\; \sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \\
&+\; \mathrm{Cov}(q_t\rho^n +\; q_t\delta(n-j)\sum_{\tau=0}^{j}\kappa(\tau)\rho^n \\
&+\; \delta(n-j)\sum_{\tau=1}^{\infty}\kappa(\tau+j)q_{t-\tau}\rho^{n+\tau}, \rho^n q_t + \delta(n)\sum_{\tau=0}^{\infty}\kappa(\tau)q_{t-\tau}\rho^{n+\tau}) \\
&=\; \sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \;+\; \rho^{2n}\sigma_q^2(1+\delta(n)\Gamma + \delta(n-j)\Gamma(j)) \\
&+\; \rho^{2n}\delta(n-j)\delta(n)\sigma_q^2\sum_{\tau=0}^{j}\kappa(\tau)\Gamma \\
&+\; \rho^{2n}\delta(n-j)\delta(n)\sigma_q^2\sum_{\tau_1=1}^{\infty}\sum_{\tau_2=0}^{\infty}\kappa(\tau_1+j)\kappa(\tau_2)\rho^{2\max(\tau_1,\tau_2)} \\
&=\; \sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \;+\; \rho^{2n}\sigma_q^2(1+\delta(n)\Gamma + \delta(n-j)\Gamma(j)) \;+\; \rho^{2n}\delta(n-j)\delta(n)\sigma_q^2\Xi(j)
\end{aligned}
\tag{44}
$$

where we have defined

$$
\Gamma(j) \;=\; \sum_{\tau=0}^{j}\kappa(\tau) + \sum_{\tau=1}^{\infty}\kappa(\tau+j)\rho^{2\tau} \tag{45}
$$

Thus, by stationarity,

$$
\begin{aligned}
&\mathrm{Cov}(F_{t+j}[y_{t+n}], F_{t-1}[y_{t+n}]) \;=\; \mathrm{Cov}(F_{t+j+1}[y_{t+n+1}], F_t[y_{t+n+1}]) \\
&=\; \rho^{2(n+1)}\sigma_q^2(1+\delta(n+1)\Gamma + \delta(n-j)\Gamma(j+1)) \;+\; \rho^{2(n+1)}\delta(n-j)\delta(n+1)\sigma_q^2\Xi(j+1)
\end{aligned}
\tag{46}
$$

Thus,

$$
\begin{aligned}
&\mathrm{Cov}(y_{t+n} - F_{t+j}[y_{t+n}], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \;=\; -\sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \\
&+\; \rho^{2n}\sigma_q^2((1-\rho^2) + \Gamma(\delta(n) - \rho^2\delta(n+1))) \\
&-\rho^{2n}\,\sigma_q^2((1-\rho^2) + (\delta(n)\Gamma + \delta(n-j)\Gamma(j) - \rho^2(\delta(n+1)\Gamma + \delta(n-j)\Gamma(j+1)))) \\
&-\;\rho^{2n}\,\sigma_q^2\,\delta(n-j)(\delta(n)\Xi(j) \;-\rho^2\delta(n+1)\Xi(j+1)) \qquad (47) \\
&=\; -\sigma_\varepsilon^2(n)\mathbf{1}_{j=0} \\
&-\rho^{2n}\,\sigma_q^2\delta(n-j)\Big[(\Gamma(j) - \rho^2\Gamma(j+1)) \\
&+(\delta(n)\Xi(j) \;-\rho^2\delta(n+1)\Xi(j+1))\Big]
\end{aligned}
$$

□

Proof of Proposition 4. The claim follows directly from Proposition 2. □

Corollary 9 (BIRs in (Smooth) Diagnostic Expectations) *Suppose that*

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

Then,

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

and

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

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

Proof of Corollaries 3 and 9. 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,

$$\gamma(n,\tau) \;=\; \begin{cases} 1+\theta, & \tau = 0 \\ -\theta, & \tau = J \end{cases} . \tag{50}$$

Then,

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

□

Corollary 10 *For $J > 1$, we have*

$$B(n,j) \;=\; const \begin{cases} 1-\rho^2, & j \;\leq\; J-1 \\ 0, & j \;>\; J-1 \end{cases} \tag{52}$$

where

$$const \;=\; -\theta(1+\theta)((1+\rho^2)D_{MZ} - 2\rho^2(1+\theta(1-\rho^{2J}) + \theta(1+\theta)(1-\rho^{2(J-1)})))^{-1} \tag{53}$$

with

$$D_{MZ} = 1 + 2\theta(1-\rho^{2J}) + \theta^2(1-\rho^{2J}) \tag{54}$$

Proof of Corollary 10. Recall that

$$B(n,j) \;=\; \frac{A(n,j) - \rho^2 A(n+1,j+1)}{D(n)}, \tag{55}$$

where

$$\begin{aligned} A(n,j) &= -\delta(n-j)\Big[\Gamma(j) + \delta(n)\Xi(j)\Big] \\ D(n) &= D_{MZ}(n) + \rho^2 D_{MZ}(n+1) - 2\rho^2(1 + \delta(n+1) \\ &\quad + \delta(n)\Gamma(1) + \delta(n)\delta(n+1)\Xi(1)) \\ D_{MZ}(n) &= 1 + 2\delta(n) + \delta(n)^2\Xi(0) \end{aligned} \tag{56}$$

Thus, for $J > 1$, we have

$$\begin{aligned} A(n,j) &= -\theta(1+\theta)[1-\rho^{2(J-j)}]\mathbf{1}_{j<J} \\ D_{MZ} &= 1 + 2\theta(1-\rho^{2J}) + \theta^2(1-\rho^{2J}) \\ D(n) &= (1+\rho^2)D_{MZ} - 2\rho^2(1+\theta(1-\rho^{2J}) + \theta(1+\theta)(1-\rho^{2(J-1)})) \end{aligned} \tag{57}$$

□

D Extrapolative Expectations

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 11 *Suppose that* $E_t[y_{t+1}] = \bar{y} + \rho_y(y_t - \bar{y})$*, whereas the agent believes*

$$\begin{aligned} F_t[y_{t+1}] &= \mu_t \\ \mu_{t+1} &= (1-\rho-\tilde{\rho})\bar{\mu} + \rho\mu_t + \tilde{\rho}y_{t+1} \\ F_t[\mu_{t+1}] &= (1-\rho-\tilde{\rho})\bar{\mu} + (\rho+\tilde{\rho})\mu_t\,, \end{aligned} \tag{58}$$

The model is linear, autonomous, and separable, with

$$\delta(n) = \frac{\tilde{\rho}}{\rho+\tilde{\rho}}\frac{1}{1-\rho\rho_y}((\rho+\tilde{\rho})/\rho_y)^n,\ \kappa(\tau) = (1-\rho\rho_y)(\rho/\rho_y)^\tau \tag{59}$$

where const makes sure that $\Gamma(0) = 1$*. We have*

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

where

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

$$\begin{aligned} A(n,j) &= \delta(n-j)[\Gamma(j) + \delta(n)\Xi(j)] \\ \Xi(j) &= \frac{1-\rho\rho_y}{1-(\rho/\rho_y)} + (\rho/\rho_y)^{j+1}(1-\rho\rho_y)\frac{\rho_y^2 - 1 + \rho^3\rho_y}{(1-\rho^2)(1-(\rho/\rho_y))} \\ \Gamma(j) &= \frac{1-\rho\rho_y}{1-\rho/\rho_y} + (\rho/\rho_y)^{j+1}\frac{\rho_y^2-1}{(1-\rho/\rho_y)} \end{aligned} \tag{62}$$

Proof of Corollaries 11 and 6. We have

$$F_t[y_{t+n}] = F_t[\mu_{t+n-1}] = cost + (\rho+\tilde{\rho})^{n-1}\mu_t \tag{63}$$

while

$$\begin{aligned} \mu_t &= (1-\rho-\tilde{\rho})\bar{\mu} + \rho\mu_{t-1} + \tilde{\rho}\rho_y^{-1}(E_t[y_{t+1}] - \bar{y}) = const + \sum_{\tau=0}^{\infty}\rho^\tau\tilde{\rho}\rho_y^{-1}E_{t-\tau}[y_{t-\tau+1}] \\ &= \tilde{const} + \sum_{\tau=0}^{\infty}\rho^\tau\rho_y^{-(\tau+n-1)}\tilde{\rho}\rho_y^{-1}E_{t-\tau}[y_{t+n}] \end{aligned} \tag{64}$$

Let $b = (\rho/\rho_y)$. Then, we have

$$
\begin{aligned}
\Xi(j) &= \sum_{\tau=0}^{j} \kappa(\tau) \ + \ \sum_{\tau_1=1}^{\infty}\sum_{\tau_2=0}^{\infty} \kappa(\tau_1+j)\kappa(\tau_2)\rho_y^{2\max(\tau_1,\tau_2)} \\
&= \ (1-\rho\rho_y)\sum_{\tau=0}^{j} b^{\tau} \ + \ (1-\rho\rho_y)^2 \sum_{\tau_1=1}^{\infty}\sum_{\tau_2=0}^{\infty} b^{\tau_1+j+\tau_2}\rho_y^{2\max(\tau_1,\tau_2)} \\
&= \ (1-\rho\rho_y)\frac{1-b^{j+1}}{1-b} \\
&+ \ (1-\rho\rho_y)^2 \sum_{\tau_1=1}^{\infty}\sum_{\tau_2=0}^{\tau_1} b^{\tau_1+j+\tau_2}\rho_y^{2\tau_1} \ + \ (1-\rho\rho_y)^2 \sum_{\tau_1=1}^{\infty}\sum_{\tau_2=\tau_1+1}^{\infty} b^{\tau_1+j+\tau_2}\rho_y^{2\tau_2} \\
&= \ (1-\rho\rho_y)\frac{1-b^{j+1}}{1-b} \\
&+ \ (1-\rho\rho_y)^2 \sum_{\tau_1=1}^{\infty} b^{j+\tau_1}\rho_y^{2\tau_1}\frac{1-b^{\tau_1+1}}{1-b} \ + \ (1-\rho\rho_y)^2 \sum_{\tau_1=1}^{\infty} b^{j+\tau_1}\frac{(b\rho_y^2)^{\tau_1+1}}{1-b\rho_y^2} \\
&= \ (1-\rho\rho_y)\frac{1-b^{j+1}}{1-b} \\
&+ \ (1-\rho\rho_y)^2 \frac{b^{j+1}\rho_y^2}{1-b}\left(\frac{1}{1-b\rho_y^2} - \frac{b^2}{1-b^2\rho_y^2}\right) \\
&+ \ (1-\rho\rho_y)^2 \frac{b^{j+3}\rho_y^4}{(1-b^2\rho_y^2)(1-b\rho_y^2)} \\
&= \ \frac{1-\rho\rho_y}{1-b} \ + \ b^{j+1}\Big(- \frac{1-\rho\rho_y}{1-b} \ + \ (1-\rho\rho_y)^2 \frac{\rho_y^2}{1-b}\left(\frac{1}{1-b\rho_y^2} - \frac{b^2}{1-b^2\rho_y^2}\right) \\
&+ \ (1-\rho\rho_y)^2 \frac{b^2\rho_y^4}{(1-b^2\rho_y^2)(1-b\rho_y^2)}\Big) \\
&= \ \frac{1-\rho\rho_y}{1-(\rho/\rho_y)} \\
&+ \ (\rho/\rho_y)^{j+1}\Big(- \frac{1-\rho\rho_y}{1-(\rho/\rho_y)} \ + \ (1-\rho\rho_y)^2 \frac{\rho_y^2}{1-(\rho/\rho_y)}\left(\frac{1-\rho^2-(1-\rho\rho_y)(\rho/\rho_y)^2}{(1-\rho\rho_y)(1-\rho^2)}\right) \\
&+ \ (1-\rho\rho_y)^2 \frac{\rho^2\rho_y^2}{(1-\rho^2)(1-\rho\rho_y)}\Big) \\
&= \ \frac{1-\rho\rho_y}{1-(\rho/\rho_y)} \\
&+ \ (\rho/\rho_y)^{j+1}(1-\rho\rho_y)\frac{\rho_y^2-1+\rho^3\rho_y}{(1-\rho^2)(1-(\rho/\rho_y))}
\end{aligned}
\tag{65}
$$

Similarly,

$$\begin{aligned}
\Gamma(j) &= \sum_{\tau=0}^{j} \kappa(\tau) + \sum_{\tau=1}^{\infty} \kappa(\tau+j)\rho_y^{2\tau} \\
&= (1-\rho\rho_y)\frac{1-b^{j+1}}{1-b} + (1-\rho\rho_y)\sum_{\tau=1}^{\infty} b^{\tau+j}\rho_y^{2\tau} \\
&= (1-\rho\rho_y)\frac{1-(\rho/\rho_y)^{j+1}}{1-\rho/\rho_y} + (1-\rho\rho_y)(\rho/\rho_y)^{j+1}\rho_y^2\frac{1}{1-\rho\rho_y} \\
&= \frac{1-\rho\rho_y}{1-\rho/\rho_y} + (\rho/\rho_y)^{j+1}\frac{\rho_y^2-1}{(1-\rho/\rho_y)}
\end{aligned} \tag{66}$$

Thus,

$$\begin{aligned}
A(n,j) &= \delta(n-j)[\Gamma(j)+\delta(n)\Xi(j)] \\
&= \frac{\tilde{\rho}}{\rho+\tilde{\rho}}\frac{1}{1-\rho\rho_y}((\rho+\tilde{\rho})/\rho_y)^{n-j}\left[\frac{1-\rho\rho_y}{1-\rho/\rho_y}\right. \\
&+ (\rho/\rho_y)^{j+1}\frac{\rho_y^2-1}{(1-\rho/\rho_y)} + \frac{\tilde{\rho}}{\rho+\tilde{\rho}}\frac{1}{1-\rho\rho_y}((\rho+\tilde{\rho})/\rho_y)^{n}\Big(\frac{1-\rho\rho_y}{1-(\rho/\rho_y)} \\
&\left. + (\rho/\rho_y)^{j+1}(1-\rho\rho_y)\frac{\rho_y^2-1+\rho^3\rho_y}{(1-\rho^2)(1-(\rho/\rho_y))}\Big)\right]
\end{aligned} \tag{67}$$

$$B(n,j) = \frac{A(n,j)-\rho^2 A(n+1,j+1)}{D(n)}, \tag{68}$$

where

$$\begin{aligned}
D(n) &= D_{MZ}(n) + \rho^2 D_{MZ}(n+1) - 2\rho^2(1+\delta(n+1) \\
&\quad + \delta(n)\Gamma(1) + \delta(n)\delta(n+1)\Xi(1)) \\
D_{MZ}(n) &= 1 + 2\delta(n) + \delta(n)^2\Xi(0)
\end{aligned} \tag{69}$$

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

$$F_t[y_{t+n}] \;=\; (\rho+\tilde{\rho})^{n-1}\mu_t,\;\; \mu_t \;=\; const \;+\; \sum_{\tau=0}^{\infty}\rho^{\tau}\tilde{\rho}y_{t-\tau} \tag{70}$$

while

$$E_t[\mu_{t+1}] \;=\; const \;+\; \rho\mu_t \tag{71}$$

$$\begin{aligned}
&E_t[F_{t+j}[y_{t+n}]] \;=\; const \;+\; (\rho+\tilde{\rho})^{n-1-j}\rho^j\mu_t \\
&E_t[y_{t+n} - F_{t+j}[y_{t+n}]] \;=\; \bar{\mu} \;-\; const \;-\; (\rho+\tilde{\rho})^{n-1-j}\rho^j\mu_t \\
&\mathrm{Cov}(E_t[y_{t+n} - F_{t+j}[y_{t+n}]], F_t[y_{t+n}] - F_{t-1}[y_{t+n}]) \\
&=\; \mathrm{Cov}(-(\rho+\tilde{\rho})^{n-1-j}\rho^j\mu_t, (\rho+\tilde{\rho})^{n-1}\mu_t - (\rho+\tilde{\rho})^{n}\mu_{t-1}) \\
&=\; -(\rho+\tilde{\rho})^{2n-2-j}\rho^j\mathrm{Var}[\mu] + (\rho+\tilde{\rho})^{2n-1-j}\rho^{j+1}\mathrm{Var}[\mu] \\
&\mathrm{Var}[F_t[y_{t+n}] - F_{t-1}[y_{t+n}]] \;=\; \mathrm{Var}[\mu]((\rho+\tilde{\rho})^{2(n-1)} + (\rho+\tilde{\rho})^{2n} - 2(\rho+\tilde{\rho})^{2n-1}\rho)
\end{aligned} \tag{72}$$

so that

$$B(n,j) \;=\; \frac{-(\rho+\tilde{\rho})^{2n-2-j}\rho^j + (\rho+\tilde{\rho})^{2n-1-j}\rho^{j+1}}{(\rho+\tilde{\rho})^{2(n-1)} + (\rho+\tilde{\rho})^{2n} - 2(\rho+\tilde{\rho})^{2n-1}\rho} \;=\; \rho^j\frac{\rho(\rho+\tilde{\rho})-1}{(\rho+\tilde{\rho})^j(1+(\rho+\tilde{\rho})^2 - 2\rho(\rho+\tilde{\rho}))} \tag{73}$$

□

Proof of Corollary 5. Recall that Recall that

$$B(n,j) \;=\; \frac{A(n,j) - \rho^2 A(n+1,j+1)}{D(n)}, \tag{74}$$

where

$$\begin{aligned} A(n,j) &= -\delta(n-j)\Big[\Gamma(j) + \delta(n)\Xi(j)\Big] \\ D(n) &= D_{MZ}(n) + \rho^2 D_{MZ}(n+1) - 2\rho^2(1 + \delta(n+1) \\ &\quad + \delta(n)\Gamma(1) + \delta(n)\delta(n+1)\Xi(1)) \\ D_{MZ}(n) &= 1 + 2\delta(n) + \delta(n)^2\Xi(0) \end{aligned} \tag{75}$$

Thus, for $J > 1$, we have

$$\begin{aligned} A(n,j) &= -\theta(1+\theta)[1-\rho^{2(J-j)}]\mathbf{1}_{j<J} \\ D_{MZ} &= 1 + 2\theta(1-\rho^{2J}) + \theta^2(1-\rho^{2J}) \\ D(n) &= (1+\rho^2)D_{MZ} - 2\rho^2(1 + \theta(1-\rho^{2J}) + \theta(1+\theta)(1-\rho^{2(J-1)})) \end{aligned} \tag{76}$$

where

$$\begin{aligned} \Xi(j) &= \sum_{\tau=0}^{j} \kappa(\tau) \; + \; \sum_{\tau_1=1}^{\infty}\sum_{\tau_2=0}^{\infty} \kappa(\tau_1 + j)\kappa(\tau_2)\rho^{2\max(\tau_1,\tau_2)} \\ \Gamma(j) &= \sum_{\tau=0}^{j} \kappa(\tau) + \sum_{\tau=1}^{\infty} \kappa(\tau + j)\rho^{2\tau} \end{aligned} \tag{77}$$

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

$$\begin{aligned} \Xi(j) &= 1 \\ \Gamma(j) &= 1\,. \end{aligned} \tag{78}$$

The claim follows. □