A Hansen-Jagannathan Bound for Beliefs

Jiacheng Li and Sean Myers*

August 2026

ABSTRACT

What can forecast errors tell us about deviations from full-information rational expectations (FIRE)? For a large class of belief mechanisms, we show that commonly studied moments of forecast errors have a "problem of arbitrariness": any belief mechanism that deviates from FIRE can match these moments if we apply the mechanism to some latent component of the forecasted variable (e.g., extrapolation of a latent trend in yields rather than extrapolation of yields). Thus, even a belief mechanism with an arbitrarily small deviation from FIRE can match any values of these moments. To address this problem, we propose a new moment, the forecast error Sharpe ratio (FESR). We derive a Hansen-Jagannathan bound for beliefs which shows that the empirical FESR provides a lower bound on the magnitude of the required deviation from FIRE and that this bound can be used to evaluate a belief mechanism across all possible latent variables. Applying our framework to professional forecasts of the 10-year yield, we find that the minimum required deviation from FIRE is substantially larger than the deviations generated by models of learning and behavioral expectations for standard parameter values, with some notable exceptions.

*Jiacheng Li and Sean Myers are with the Wharton School at the University of Pennsylvania. We thank Gideon Bornstein, Leland Bybee, Laura Castillo-Martinez, Aditya Chaudhry, Thummim Cho, Carter Davis, Ricardo Delao, Andrei Goncalves, Theis Jensen, Paymon Khorrami, Lukas Kremens, Xinwei Li, Ernest Liu, Stefan Nagel, Alessandra Peter, Nikolai Roussanov, and Tom Sargent for valuable comments. Email addresses: jl10727@wharton.upenn.edu, semyers@wharton.upenn.edu.

What do data on forecast errors tell us about deviations from Full-information Rational Expectations (FIRE)? As a concrete example, consider the secular decline in long-term yields over the last 30-40 years. The blue line of Figure 1 shows the realized ten-year maturity Treasury yield and, at any given point in time, the dashed black lines show the forecasted path for the yield over the next four quarters from the consensus Blue Chip professional forecasters. There are multiple moments that we can use to summarize the resulting forecast errors, such as the Mincer-Zarnowitz coefficient, the serial autocorrelation of forecast errors, and the Coibion-Gorodnichenko coefficient. These three moments measure how forecast errors relate to forecasts, past forecast errors, and forecast revisions, respectively.

The problem, which we formalize in this paper, is that for a large class of belief mechanisms these moments can always be matched if we allow the mechanism to be applied to a latent component of the yield. We call this the "problem of arbitrariness" in reference to Sargent (1993). Once a researcher departs from FIRE, there are infinitely many degrees of freedom in how a belief mechanism is related to a forecasted variable (e.g., the agent could be extrapolating the level of yields, a latent trend in yields, a latent transitory component, a latent persistent component, etc.). Even belief mechanisms that are arbitrarily close to FIRE cannot be rejected by these moments, such as an agent who believes a latent component has persistence $\rho + \epsilon$ rather than ρ for some minuscule $\epsilon \neq 0$.

To address this problem, we propose a Hansen-Jagannathan bound for beliefs and a new moment for evaluating forecast errors: the forecast error Sharpe ratio (FESR), which is simply the average forecast error divided by its standard deviation. The original Hansen and Jagannathan (1991) bound allows researchers to use financial Sharpe ratios for observed assets to reject potential stochastic discount factors without needing to know how the stochastic discount factor relates to the observed asset. In a similar way, our belief-based Hansen-Jagannathan bound allows us to use FESR for observed forecasts to reject belief mechanisms without needing to know how the mechanism relates to the forecasted variable. Further, just like financial Sharpe ratios provide a lower bound for the variance of

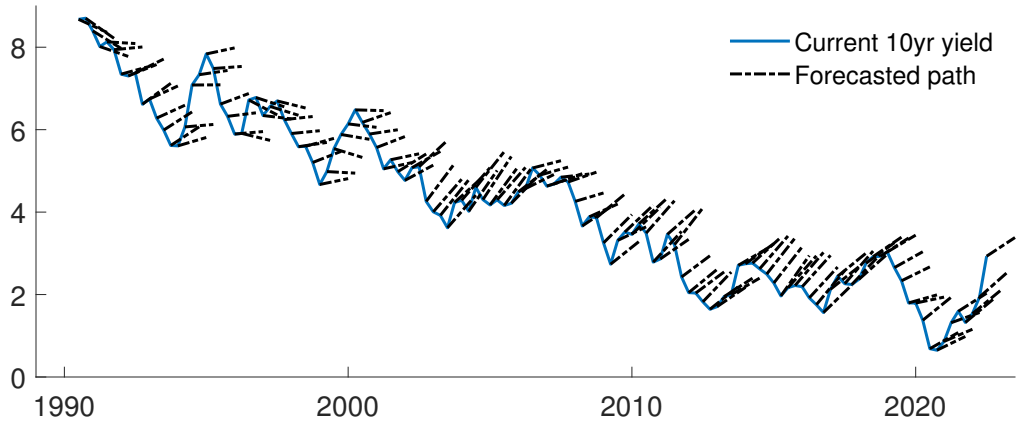


Figure 1. Realized and forecasted yield. This figure shows the realized 10-year maturity Treasury yield (blue) in percentage points. In black, the figure shows the forecasted path for the next four quarters from the consensus expectations for the Blue Chip professional forecasters.

the stochastic discount factor, we show that the FESR provides a lower bound for how much a belief mechanism must deviate from FIRE.

Applying our approach to the empirical yield forecasts, we show that we can reject Bayesian learning mechanisms in which the agent believes in real time that her forecast errors are almost entirely noise (i.e., mechanisms in which patterns in forecast errors are only clear ex post). For behavioral mechanisms, the broad takeaway is that the parameters controlling the degree of bias must be large. For adaptive expectations, over/underreaction, and extrapolation, the empirical FESR places a lower bound on the adaptive gain, the absolute value of the over/underreaction parameter, and the difference between the subjective persistence and the FIRE persistence, respectively.

Our central framework is the following. We characterize a belief mechanism as a believed probability distribution $\pi_t^*(\ell_{t+1})$ that may differ from the FIRE probability distribution $\pi_t(\ell_{t+1})$ for some potentially latent variable ℓ_{t+1}. Note that ℓ_{t+1} can be a vector that captures non-FIRE beliefs about many variables. We then assume that the agent's forecast for all other variables y_{t+1} is

$$E_t^*[y_{t+1}] = \int_{\ell_{t+1}} E_t[y_{t+1}|\ell_{t+1}]\,\pi_t^*(\ell_{t+1})\,d\ell_{t+1}$$

where $E_t[y_{t+1}|\ell_{t+1}]$ is the FIRE conditional expectation. In words, we assume that the agent

understands everything about the model except for the distribution of future ℓ_{t+1} (i.e., all of the non-FIRE aspects of the model are contained in ℓ_{t+1}).

This framework gives us the structure that we need to show the "problem of arbitrariness" for any persistently biased belief mechanism.[1] For clarity, we will continue to use the concrete example of yield forecasts but the logic applies for any forecasted variable. Given a belief mechanism $(\pi_t^*(\ell_{t+1}), \pi_t(\ell_{t+1}))$ and empirical values for the Mincer-Zarnowitz coefficient, the serial autocorrelation of forecast errors, and the Coibion-Gorodnichenko coefficient (MZ, AC, CG) for the yield, we show how to construct a model yield $y_{t+1} = a_t + b_t \ell_{t+1}$ such that (i) y_{t+1} will match the empirical dynamics of the yield and (ii) forecast errors for y_{t+1} will match the empirical MZ, AC, and CG coefficients. In other words, we can always propose a relationship between yields and ℓ_{t+1} which is empirically plausible (i.e., it matches the empirical dynamics of the yield) and perfectly replicates all three forecast error moments. As a consequence, even a belief mechanism with an arbitrarily small deviation from FIRE can match any empirical MZ, AC, and CG coefficient values.

Fortunately, there is a different moment of forecast errors which can discipline belief mechanisms across all possible ways that yields could be related to ℓ_{t+1}. The agent's forecasts can be characterized by a change of measure, where $S_{t+1} \equiv \frac{\pi_t^*(\ell_{t+1})}{\pi_t(\ell_{t+1})}$ is the likelihood ratio. Without needing to know anything about the relationship between yields and ℓ_{t+1}, other than the fact that their correlation cannot exceed 1, we show that the forecast error Sharpe ratio (FESR) for yields provides a lower bound for $Var(S_{t+1})$. We can then evaluate which belief mechanisms generate $Var(S_{t+1})$ high enough to exceed the empirically measured FESR. The logic behind this result comes directly from the original Hansen-Jagannathan bound and fits into the broader argument from Cui et al. (2026) that the change-of-measure tools from asset pricing can be used for much more than just pricing assets.[2]

[1] We give the full definition of persistently biased belief mechanisms in Section II, but the main intuition is that these are mechanisms in which biases do not immediately die out.

[2] Importantly, this result does not rely on the fact that our application is for forecasts of yields (i.e., a financial variable). Our belief-based Hansen-Jagannathan bound can be applied to non-financial variables such as GDP, inflation, unemployment, etc. The key is that the forecasts can be tied to a change of measure from FIRE.

Importantly, we are not arguing that the previous information in the MZ, AC, and CG coefficients should simply be tossed out. In fact, once we use the FESR approach, we can immediately show how information about forecast error predictability from these coefficients can be used to raise the empirical FESR and provide a higher value for the minimum required $Var\left(S_{t+1}\right)$. This is analogous to how evidence of return predictability can be used to raise financial Sharpe ratios which increases the lower bound for the volatility of the stochastic discount factor. Thus, our aim is to improve and generalize the existing tests of belief mechanisms so that they are not vulnerable to the problem of arbitrariness, rather than to dismiss them.

To demonstrate that this bound is useful, we measure the empirical FESR for the yield forecast and ask what this would require for Bayesian learning mechanisms. A commonly proposed explanation for patterns in forecast errors is that, while the patterns may be clear ex post, a Bayesian agent may have believed that the forecast errors were almost entirely noise in real time. For example, Veronesi (2000), Gilchrist and Saito (2008), Ai (2010), and Roussanov et al. (2021) provide examples of learning mechanisms in which the agent believes in real time that her forecast errors are 94-97% noise and only 3-6% incorrect beliefs about the underlying states.

Starting first with Bayesian learning in a linear, Gaussian setting, we show that $Var\left(S_{t+1}\right)$ only depends on two objects: the Kalman gain and the signal-to-noise ratio. All of the other details of the learning mechanism can be ignored. This provides a clear cut result where a high FESR can only be explained by a high Kalman gain. Specifically, the empirical FESR implies a minimum Kalman gain of 0.29. This means that we can rule out any mechanisms in which the agent thinks that her forecast errors are more than 71% noise. We then show that this same logic extends to a more generalized learning setting with non-linear Markov dynamics (e.g., regime-switching) and short samples. A high empirical FESR can only be explained if the Bayesian agent realizes in real time that at least 29% of her forecast error is due to incorrect beliefs about the hidden states. It cannot be that her mistake is largely

unnoticeable in real time.

We then apply our approach to behavioral mechanisms, specifically extrapolation, adaptive expectations, and over/underreaction. As with the Bayesian learning mechanisms, we show that $Var\left(S_{t+1}\right)$ can always be expressed in terms of just two parameters for each of these behavioral mechanisms. For extrapolation, $Var\left(S_{t+1}\right)$ solely depends on the FIRE persistence ρ and the believed persistence ρ^*. For any $\rho > 0$, we show that the empirical FESR requires a minimum $\rho^* - \rho$ of 0.40. In other words, the FESR directly tells us how much the believed parameter must deviate from FIRE, and the required deviation is fairly large. For context, we show a set of example extrapolation papers in which $\rho^* - \rho$ is below 0.1. An interesting exception is Barberis et al. (2015), which features agents who overstate the persistence of an endogenous sentiment variable and generates $\rho^* - \rho$ of 0.47.

For adaptive expectations, $Var\left(S_{t+1}\right)$ only depends on the gain (i.e., the constant that determines how much the agent updates her beliefs after observing her forecast error) and a parameter that captures the total persistence of ℓ_{t+1}. The empirical FESR requires a minimum gain of 0.40, which is substantially higher than the value of roughly 0.02 used in Orphanides and Williams (2005), Milani (2007), Hirshleifer et al. (2015), and Malmendier and Nagel (2016). The intuition for this result is the following. When the gain is low, the agent places almost equal weight on recent and distant observations, which means that noise in the observations largely averages out. Thus, while adaptive expectations are not FIRE, adaptive expectations are close to FIRE when the gain is low. In order to generate a high FESR, we require a large gain so that noise in recent observations substantially impacts the agent's beliefs and her expectations deviate sufficiently far from FIRE.

Finally, we consider over/underreaction in the form of diagnostic expectations. For these mechanisms, $Var\left(S_{t+1}\right)$ only depends on the objective informativeness of the shock and the amount that the agent exaggerates or understates recent information, θ. The empirical FESR implies a minimum required exaggeration/understatement of 50% which is easily satisfied by the typical value of 90% used in Bordalo et al. (2018), Bordalo et al. (2019), and Bordalo

et al. (2026). In other words, so long as the informativeness of the shock is moderately high, diagnostic expectations with $\theta = 0.90$ is a mechanism that clearly satisfies our bound.

Overall, this paper connects two separate literatures. The first studies predictable forecast errors and the implications for deviations from FIRE. Given the long history of this literature, we refer readers to the recent handbook chapters of Bachmann et al. (2023) for a comprehensive set of references.

A particularly relevant strand of this first literature are the papers proposing key moments of forecast errors. Mincer and Zarnowitz (1969) study the comovement of the forecast with the future realization. If we simply subtract 1 from the Mincer-Zarnowitz coefficient, this can be equivalently described as capturing the comovement of the forecast with future forecast errors. McNees (1978), Friedman (1980), and others later utilize the serial autocorrelation of forecast errors. More recently, Coibion and Gorodnichenko (2015) propose a new moment which utilizes the comovement of revisions in forecasts with the future forecast error.[3] We build on this agenda by proposing a new moment of forecast errors which can discipline belief mechanisms without needing to know the relationship between the forecasted variable and ℓ_{t+1}, which is particularly useful when ℓ_{t+1} may be a latent, unobservable variable. Further, we show how information from multiple previously studied moments can be easily incorporated into our proposed FESR to provide a stricter test that still relies on one summary moment.

Since our application utilizes forecasts of yields, another closely related strand is the recent work on yield forecast errors such as Chernov and Mueller (2012), Cieslak (2018), Bauer et al. (2024), Farmer et al. (2024), and Pang (2026). These papers document forecast errors for yields and propose explanations based on learning or behavioral biases. Rather than documenting wholly new forecast errors or a new explanation, our aim is to show a new methodology for evaluating these forecast errors and their implications. Using the "problem of arbitrariness," we stress that there are many possible mechanisms that can explain

[3]While our application is only to consensus expectations, Bordalo et al. (2020) demonstrate the informativeness of comparing individual-level and consensus-level Coibion-Gorodnichenko coefficients.

commonly studied moments of these forecast errors. This includes mechanisms that are arbitrarily close to FIRE. We then show that, by using the empirical FESR, we can establish a minimum required deviation from FIRE that holds for all potential mechanisms within a broad class, and that this required deviation has a straightforward interpretation for mechanisms of Bayesian learning, extrapolation, adaptive expectations, and over/underreaction.

The second broad literature is the asset pricing literature using volatility bounds to discipline stochastic discount factors (SDF) under minimal assumptions. Our approach is particularly inspired by Cochrane and Hansen (1992), who demonstrate the power of these types of bounds to discipline multiple classes of stochastic discount factors, and Hansen et al. (1995), who emphasize the usefulness of volatility bounds when the SDF potentially relies on variables that are unobservable to the researcher. More recently, Chen et al. (2020) establish how entropy bounds can be applied to asset prices to learn about investors' expectations, conditional on a preference specification. The key idea for connecting these bounds to forecasts is that these are fundamentally change-of-measure tools, rather than asset pricing tools. Even for forecasts of non-financial variables, these tools can still be applied so long as forecasts can be represented as a change of measure from FIRE.

In line with recent work by Bhandari et al. (2025) and Cui et al. (2026) using change-of-measure tools to understand subjective expectations, we demonstrate how these tools can be used to address the "problem of arbitrariness." Our approach is closely related to the problem studied in Chen et al. (2024), who take a fixed moment function and search for belief distortions within the researcher's chosen divergence bound. We instead take a specific belief mechanism (i.e., the distorted beliefs) as given and assess how this compares to a data-implied bound across all possible ways the forecasted variable could be related to the latent variable.

The rest of the paper is organized as follows. Section I provides the two assumptions that we use throughout the paper. Section II establishes the problem of arbitrariness. Section III shows a belief-based Hansen-Jagannathan bound. Section IV uses this bound to evaluate

mechanisms of Bayesian learning and behavioral biases. Section V concludes.

I. Overarching Assumptions

In this section, we lay out the two assumptions regarding subjective expectations that we will impose throughout the paper. In terms of notation, $E_t^*[\cdot]$ denotes the agent's subjective expectations at time t and $\pi_t^*(\cdot)$ denotes the agent's subjective probability distribution at time t.[4] All other operators use the Full-information Rational Expectations (FIRE) probability distribution. For example, $E_t[\cdot]$ denotes the conditional expectation under the FIRE probability distribution.

First, we assume the agent has non-FIRE beliefs about some variable ℓ_{t+1}: the FIRE probability distribution is $\pi_t(\ell_{t+1})$ and the subjective probability distribution is $\pi_t^*(\ell_{t+1})$, where the support of $\pi_t^*(\cdot)$ is a weak subset of the support of $\pi_t(\cdot)$.[5] Second, we assume that these non-FIRE beliefs about ℓ_{t+1} impact the agent's expectations of all other variables y_{t+1} following

$$E_t^*[y_{t+1}] = \int_{\ell_{t+1}} \pi_t^*(\ell_{t+1}) E_t[y_{t+1}|\ell_{t+1}] d\ell_{t+1} \tag{1}$$

where $E_t[y_{t+1}|\ell_{t+1}]$ is the conditional expectation of y_{t+1} given ℓ_{t+1} under the FIRE distribution. We use y_{t+1} to denote these other variables to align with our example of yields, but this could be any variable of interest.

Note that ℓ_{t+1} can be a vector that combines many different variables, which means these two assumptions can largely be thought of as definitional. If there is a variable y_{t+1} where $E_t^*[y_{t+1}]$ does not satisfy equation (1), then y_{t+1} can be included as part of the vector ℓ_{t+1}. We can repeat this process until we have defined ℓ_{t+1} such that assumption 2 is satisfied.

Intuitively, assumption 2 says that the agent agrees with FIRE about everything other

[4]For simplicity, we will focus on a single, representative agent, but the results can be easily extended to heterogeneous agents.

[5]This condition on the support of $\pi_t^*(\cdot)$ simply ensures that the agent does not assign positive probability to outcomes that are genuinely impossible. For all of the examples we study in Section IV, this condition trivially holds because normal shocks lead to full, unbounded support for $\pi_t(\cdot)$.

than the distribution of ℓ_{t+1}. The FIRE expectation for any variable y_{t+1} can always be written as

$$E_t\left[y_{t+1}\right] = \int_{\ell_{t+1}} \pi_t\left(\ell_{t+1}\right) E_t\left[y_{t+1}|\ell_{t+1}\right] d\ell_{t+1}. \tag{2}$$

Comparing this to equation (1), we see that assumption 2 says that the agent agrees with FIRE about how y_{t+1} is related to ℓ_{t+1} (i.e., $E_t\left[y_{t+1}|\ell_{t+1}\right]$) but the agent disagrees with FIRE about the distribution of future ℓ_{t+1}. For example, the agent may know how inflation impacts yields ($E_t\left[y_{t+1}|\ell_{t+1}\right]$) but not know the FIRE distribution for future inflation ($\pi_t\left(\ell_{t+1}\right)$).

Given that all of the non-FIRE aspects of the agent's expectations are captured in ℓ_{t+1}, we will refer to a "belief mechanism" as the FIRE processes for ℓ_{t+1} and $\pi_t^*\left(\ell_{t+1}\right)$. An example belief mechanism would be that ℓ_{t+1} is AR(1) with persistence ρ and conditional volatility σ and $\pi_t^*\left(\ell_{t+1}\right) = \pi_t\left(\ell_{t+1} - \left(\rho^* - \rho\right)\ell_t\right)$, i.e., the agent believes the persistence is ρ^* rather than ρ.

II. The Problem of Arbitrariness

Can commonly studied moments of forecast error predictability rule out a belief mechanism once we allow for latent variables? Unfortunately, for a large class of belief mechanisms, we will show that the answer is no. Suppose we have empirical data on some outcome $\hat{y}_{t+1}$ and survey expectations $\hat{E}_t^*\left[\hat{y}_{t+1}\right]$ and $\hat{E}_{t-1}^*\left[\hat{y}_{t+1}\right]$. We use the hat notation to distinguish these empirical objects from the analytical variables we will construct to match moments from the data. Let $\hat{f}_{t+1} \equiv \hat{y}_{t+1} - \hat{E}_t^*\left[\hat{y}_{t+1}\right]$ be the empirical forecast error.

When evaluating how these empirical forecast errors compare to FIRE, three of the most commonly studied moments are the Mincer-Zarnowitz coefficient, the serial autocorrelation of forecast errors, and the Coibion-Gorodnichenko coefficient. Conveniently, these three moments can each be expressed as a coefficient from a univariate regression predicting forecast errors $\hat{f}_{t+1}$ with the current forecast $\hat{E}_t^*\left[\hat{y}_{t+1}\right]$, the current forecast error $\hat{f}_t$, and the current

revision $\hat{E}_t^* [\hat{y}_{t+1}] - \hat{E}_{t-1}^* [\hat{y}_{t+1}]$, respectively,[6]

$$\hat{\beta}^{MZ} \equiv \frac{Cov\left(\hat{f}_{t+1}, \hat{E}_t^* [\hat{y}_{t+1}]\right)}{Var\left(\hat{E}_t^* [\hat{y}_{t+1}]\right)} \tag{3}$$

$$\hat{\beta}^{AC} \equiv \frac{Cov\left(\hat{f}_{t+1}, \hat{f}_t\right)}{Var\left(\hat{f}_t\right)} \tag{4}$$

$$\hat{\beta}^{CG} \equiv \frac{Cov\left(\hat{f}_{t+1}, \hat{E}_t^* [\hat{y}_{t+1}] - \hat{E}_{t-1}^* [\hat{y}_{t+1}]\right)}{Var\left(\hat{E}_t^* [\hat{y}_{t+1}] - \hat{E}_{t-1}^* [\hat{y}_{t+1}]\right)}. \tag{5}$$

We can also measure the autocorrelation of $\hat{y}_t$ to provide additional moments

$$\hat{\gamma}_j \equiv Corr(\hat{y}_t, \hat{y}_{t-j}) \tag{6}$$

for $j \leq n$ out to any desired n. Since $\hat{y}_{t+1}$ and $\hat{E}_t^* [\hat{y}_{t+1}]$ are empirical data, we will assume that the covariance matrix of $\left(\hat{y}_{t+1}\ \hat{y}_t ... \hat{y}_{t+1-n}\ \hat{E}_t^* [\hat{y}_{t+1}]\ \ \hat{E}_{t-1}^* [\hat{y}_t]\ ... \hat{E}_{t-n}^* [\hat{y}_{t+1-n}]\right)$ is positive definite.[7]

Now, suppose we are also given a belief mechanism $(\ell_{t+1}, \pi_t^* (\ell_{t+1}))$. Define $\varepsilon_{t+1} = \ell_{t+1} - E_t [\ell_{t+1}]$ as the FIRE shock and $\delta_t = E_t^* [\ell_{t+1}] - E_t [\ell_{t+1}]$ as the bias.[8] We will focus on the class of belief mechanisms where (i) all moments of ε_{t+1} are finite and ε_{t+1} has unbounded support and (ii) the bias can be expressed as

$$\delta_t = \sum_{k=0}^{\infty} \lambda_k \varepsilon_{t-k} + \eta_t$$

where $\lambda_k \neq 0$ for all k and η_t is independent of $\{\varepsilon_{t-k}\}_{k=0}^{\infty}$. We refer to this class as "persistently biased belief mechanisms." Section II.B discusses each of these conditions.

[6]The Mincer-Zarnowitz coefficient is technically the coefficient from regressing the future realization on the current forecast, i.e., $\frac{Cov\left(\hat{y}_{t+1}, \hat{E}_t^* [\hat{y}_{t+1}]\right)}{Var\left(\hat{E}_t^* [\hat{y}_{t+1}]\right)}$. We use $\frac{Cov\left(\hat{f}_{t+1}, \hat{E}_t^* [\hat{y}_{t+1}]\right)}{Var\left(\hat{E}_t^* [\hat{y}_{t+1}]\right)} = \frac{Cov\left(\hat{y}_{t+1}, \hat{E}_t^* [\hat{y}_{t+1}]\right)}{Var\left(\hat{E}_t^* [\hat{y}_{t+1}]\right)} - 1$ to highlight that this is just another measure of forecast error predictability, similar to the other two moments.

[7]By definition, the covariance matrix is always positive semi-definite. Assuming it is positive definite simply means we are assuming that none of the empirical variables are a perfect linear combination of the others.

[8]We use "bias" to denote any difference between subjective expectations and FIRE expectations. Thus, both behavioral mistakes and Bayesian learning would be included in the bias.

Proposition 1. *Given any persistently biased belief mechanism* $(\ell_{t+1}, \pi_t^*(\ell_{t+1}))$*, there exists*

$$y_{t+1} = a_t + b_t \ell_{t+1}$$

such that $\beta^{MZ} = \hat{\beta}^{MZ}$*,* $\beta^{AC} = \hat{\beta}^{AC}$*,* $\beta^{CG} = \hat{\beta}^{CG}$ *and* $\gamma_j = \hat{\gamma}_j$ *for all* $j \leq n$ *where*

$$\begin{aligned}
\beta^{MZ} &\equiv \frac{Cov\left(y_{t+1} - E_t^*[y_{t+1}], E_t^*[y_{t+1}]\right)}{Var\left(E_t^*[y_{t+1}]\right)} \\
\beta^{AC} &\equiv \frac{Cov\left(y_{t+1} - E_t^*[y_{t+1}], y_t - E_{t-1}^*[y_t]\right)}{Var\left(y_t - E_{t-1}^*[y_t]\right)} \\
\beta^{CG} &\equiv \frac{Cov\left(y_{t+1} - E_t^*[y_{t+1}], E_t^*[y_{t+1}] - E_{t-1}^*[E_t^*[y_{t+1}]]\right)}{Var\left(E_t^*[y_{t+1}] - E_{t-1}^*[E_t^*[y_{t+1}]]\right)} \\
\gamma_j &\equiv Corr\left(y_t, y_{t-j}\right).
\end{aligned}$$

All proofs are provided in Appendix B.

In words, for any belief mechanism that satisfies conditions (i)-(ii), there will always be a variable y_{t+1} that matches the empirical dynamics of $\hat{y}_{t+1}$ (i.e., $\gamma_j = \hat{\gamma}_j$ for all $j \leq n$) and matches all three moments of forecast error predictability β^{MZ}, β^{AC}, β^{CG}. Thus, we cannot reject that this belief mechanism explains the empirical ($\hat{\beta}^{MZ}$, $\hat{\beta}^{AC}$, $\hat{\beta}^{CG}$) for some latent component ℓ_{t+1} that affects y_{t+1}. By extension, this means that empirical ($\hat{\beta}^{MZ}$, $\hat{\beta}^{AC}$, $\hat{\beta}^{CG}$) can always be explained by mechanisms that only have minuscule differences from FIRE (e.g., mechanisms with $\sum_{k=0}^{\infty} \lambda_k$ arbitrarily close to 0).

In some of the later exercises, we will focus our attention on conditionally linear relationships between y_{t+1} and ℓ_{t+1}. This is why we show that Proposition 1 holds even when y_{t+1} is restricted to the form $a_t + b_t \ell_{t+1}$. Also, note that the two assumptions of Section I only specify the one-period subjective expectations of the agent. To calculate a Coibion-Gorodnichenko coefficient, we also need two-period subjective expectations so that we can calculate revisions, $E_t^*[y_{t+1}] - E_{t-1}^*[y_{t+1}]$. For simplicity, we define two-period subjective expectations using the law of iterated expectations, $E_{t-1}^*[y_{t+1}] = E_{t-1}^*[E_t^*[y_{t+1}]]$ in Proposition 1. In Section II.B, we discuss weaker versions of this condition that are sufficient for Proposition 1.

A. *Concrete example*

Consider a belief mechanism of AR(1) extrapolation. Specifically, ℓ_{t+1} is AR(1) with persistence $\rho = 0.9$ and standard normal shocks ε_{t+1} . Suppose the agent believes the persistence is $\rho^* = 0.95$. Normal shocks ε_{t+1} satisfy condition (i). The bias is $\delta_t = \sum_{k=0}^{\infty} (\rho^* - \rho) \rho^k \varepsilon_{t-k}$, so condition (ii) is satisfied. Thus, this is a persistently biased belief mechanism.

Suppose we collect expected yields $\hat{E}_t^* [\hat{y}_{t+1}]$ and realized yields $\hat{y}_{t+1}$ and calculate $\hat{\beta}^{MZ} = Cov\left(\hat{y}_{t+1} - \hat{E}_t^* [\hat{y}_{t+1}], \hat{E}_t^* [\hat{y}_{t+1}]\right) / Var\left(\hat{E}_t^* [\hat{y}_{t+1}]\right)$. We then want to know whether this $\hat{\beta}^{MZ}$ can rule out our AR(1) extrapolation belief mechanism. If we choose $y_{t+1} = \ell_{t+1}$, then $\beta^{MZ} = \frac{\rho - \rho^*}{\rho^*} \approx -0.05$ which means we cannot match any arbitrary value of $\hat{\beta}^{MZ}$ from the data. However, if we consider that the agent may not be extrapolating yields but instead is extrapolating some latent variable ℓ_{t+1} that is related to yields, $y_{t+1} = a_t + b_t \ell_{t+1}$, then we have much more freedom. As a simple example, if we consider $y_{t+1} = a\ell_t + \ell_{t+1}$, then we have

$$\beta^{MZ} = \frac{\rho - \rho^*}{a + \rho^*}. \tag{7}$$

By varying a from $-\infty$ to ∞, we can match any $\hat{\beta}^{MZ}$ from the data.

Importantly, we can do this for any $\rho^* \neq \rho$. If we changed our mechanism to $\rho^* = \rho + \epsilon$ for some tiny $\epsilon > 0$, we could still find $y_{t+1} = a\ell_t + \ell_{t+1}$ that matches any $\hat{\beta}^{MZ}$ from the data. Thus, while it is true that $\hat{\beta}^{MZ} \neq 0$ implies that the belief mechanism must deviate from FIRE (i.e., $\rho^* \neq \rho$), the value of $\hat{\beta}^{MZ}$ does not tell us whether the deviation $|\rho^* - \rho|$ needs to be large.

A similar point applies to the Coibion-Gorodnichenko coefficient where, under $y_{t+1} = a\ell_t + \ell_{t+1}$, we have

$$\beta^{CG} = \frac{1}{a + \rho^*} \frac{(\rho - \rho^*)(1 - \rho\rho^*)}{[1 - 2\rho\rho^* + \rho^{*2}]} \tag{8}$$

and by varying a from $-\infty$ to ∞, we can get any β^{CG}. The purpose of Proposition 1 is to show that, given the freedom to choose a_t and b_t, we can match $\hat{\beta}^{MZ}$ and $\hat{\beta}^{CG}$ simultaneously,

as well as $\hat{\beta}^{AC}$ and any number of desired autocorrelations of yields. Much of the complexity in proving Proposition 1 is showing that we can generate forecast error autocorrelation β^{AC} arbitrarily close to 1 and -1. But the core insight is that the same biased beliefs $E_t^* [\ell_{t+1}]$ will generate different moments for yield forecast error predictability depending on how yields relate to ℓ_{t+1}. This echoes and formalizes the point from Sargent (1993) that once we depart from FIRE, there are "problems of arbitrariness" as there are infinitely many ways one could choose to have extrapolation of a latent variable impact expectations of yields.

B. Discussion of conditions

As defined earlier, our class of belief mechanisms has two conditions. Section IV shows four example mechanisms which all satisfy these conditions. We view condition (ii) as the most economically meaningful aspect of this class, which is why we call the class "persistently biased belief mechanisms." Once biases in expectations of ℓ_{t+1} are persistent, one can generate rich dynamics for forecast errors of $y_{t+1} = a_t + b_t \ell_{t+1}$ by altering how a_t and b_t evolve over time. In comparison, we view condition (i) as more of a technicality to address $\hat{\beta}^{AC}$ values that are very close to 1. Note that condition (i) is satisfied if ε_{t+1} is normal or is the sum of a normal shock and some other shock.

In the statement of Proposition 1, we use $E_{t-1}^* [E_t^* [y_{t+1}]]$ as the two-period expectation for the revision used in the Coibion-Gorodnichenko coefficient. Assuming the law of iterated expectations would ensure that $E_{t-1}^* [y_{t+1}] = E_{t-1}^* [E_t^* [y_{t+1}]]$ for all possible variables y_{t+1}. This is sufficient for Proposition 1 but weaker assumptions will also suffice. In our proof of Proposition 1, the needed condition is only that $E_{t-1}^* [a_t + b_t (\ell_{t+1} - E_t^* [\ell_{t+1}])] = E_{t-1}^* [a_t]$ for any (a_t, b_t) processes. It is also worth noting that two-period expectations only matter for β^{CG} in Proposition 1. Changing the two-period expectations has no impact on Proposition 1's claims about β^{MZ}, β^{AC}, or γ_j. In this sense, the treatment of two-period expectations only applies to a limited portion of Proposition 1.

C. *Is the bar too high?*

An important point is that these moments ($\hat{\beta}^{MZ}$, $\hat{\beta}^{AC}$, $\hat{\beta}^{CG}$) were originally designed to only test mechanisms applied directly to the empirical variable, i.e., $y_{t+1} = \ell_{t+1}$. So, there is a natural question whether the bar is simply too high in Proposition 1. If one is allowed to choose any $y_{t+1} = a_t + b_t\ell_{t+1}$, then perhaps there is no moment of forecast errors for the empirical $\hat{y}_{t+1}$ that cannot be matched by a mechanism. In fact, it seems quite plausible that, with the infinitely many degrees of freedom that (a_t, b_t) provide, there is simply nothing that can be ruled out.

The goal of Section III is to demonstrate that the bar is not too high and to provide a constructive solution to the issue raised in Proposition 1. There is a fairly straightforward moment of forecast errors that can rule out a mechanism not only for all conditionally linear y_{t+1} (i.e., $y_{t+1} = a_t+b_t\ell_{t+1}$) but more broadly for all possible y_{t+1}. Additionally, Section III.B shows how we can incorporate information from ($\hat{\beta}^{MZ}$, $\hat{\beta}^{AC}$, $\hat{\beta}^{CG}$) to increase this moment and provide a stricter test. Thus, the aim of this paper is not to argue that we should throw out these previous tests. Instead, our goal is to improve and generalize these tests so that we can make statements about mechanisms that hold across all possible latent variables ℓ_{t+1}.

D. *Interpretation under Bayesian learning mechanisms*

There are two ways to break Full-information Rational Expectations. The first is that the agent has full information and does not have rational expectations. In this setting, the interpretation of subjective expectations of $a_t + b_t\ell_{t+1}$ is straightforward. For any (a_t, b_t) processes that the researcher chooses, a_t and b_t will be known to the agent at time t since she has full information. Thus, the subjective expectation is simply $E_t^*[a_t + b_t\ell_{t+1}] = a_t + b_tE_t^*[\ell_{t+1}]$.

The second way to break Full-information Rational Expectations is to have rational expectations but limited information. In this setting, the interpretation of subjective expectations of $a_t + b_t\ell_{t+1}$ is more subtle, as one might worry that the agent could learn from

(a_t, b_t) processes which could change her expectations of future ℓ_{t+1}. This is where the order of Proposition 1 becomes important. The researcher is given a belief mechanism, which is a full process for ℓ_{t+1} and $\pi_t^*(\ell_{t+1})$. This then defines the agent's subjective expectation $E_t^*\left[\tilde{a}_t + \tilde{b}_t \ell_{t+1}\right] = \tilde{a}_t + \tilde{b}_t E_t^*[\ell_{t+1}]$ following equation (1) for any $\tilde{a}_t$ and $\tilde{b}_t$. For example, the agent could report at time $t-1$ her expectation of $5+2\ell_t$, then at time t report her expectation of $-0.1-4\ell_{t+1}$, then at time $t+1$ report her expectation of $17-3\ell_{t+2}$. The researcher then gets to pick a particular (a_t, b_t) so that subjective expectations of $a_t + b_t\ell_{t+1}$ match the targeted moments. So, the agent does not learn from the researcher's choice of (a_t, b_t) because these are chosen after she has defined her subjective expectations for all possible $\tilde{a}_t + \tilde{b}_t \ell_{t+1}$.

The intuition of Proposition 1 under learning is easiest to understand when we focus on a finite sample. To make this concrete, suppose the agent observes $\ell_t = \mu_t + \varepsilon_t$ where μ_t is an unobserved AR(1) state and ε_t is observation noise. The agent's subjective expectation for ℓ_{t+1} will also define her subjective expectations for $\tilde{y}_{1,t+1} = -\ell_t + \ell_t\ell_{t+1}$ and for $\tilde{y}_{2,t+1} = \ell_t - \ell_t\ell_{t+1}$. If the agent's subjective expectations are generally too low over this finite sample, the researcher may choose to define the yield as $y_{t+1} = \tilde{y}_{1,t+1}$ to match the desired empirical targets. If the agent's subjective expectations are generally too high over this finite sample, the researcher may choose to define the yield as $y_{t+1} = \tilde{y}_{2,t+1}$ to match the empirical targets. This means that the researcher's choice of (a_t, b_t) in $y_{t+1} = a_t + b_t\ell_{t+1}$ does contain information about the underlying μ_t over this sample, but the agent does not learn from this information because the researcher makes her choice of (a_t, b_t) after the agent's expectations were defined for all possible $\tilde{y}_{t+1} = \tilde{a}_t + \tilde{b}_t\ell_{t+1}$.

The final Bayesian learning (i.e., rational expectations) detail to discuss is the two-period subjective expectations used in the Coibion-Gorodnichenko coefficient of Proposition 1. Note that this does not impact any of the other moments in Proposition 1. The logic is similar to the one-period expectations. The researcher is given a Bayesian learning belief mechanism. This then defines the agent's one-period subjective expectation

$E_t^*\left[\tilde{a}_t + \tilde{b}_t \ell_{t+1}\right] = \tilde{a}_t + \tilde{b}_t E_t^*\left[\ell_{t+1}\right]$ for any $\tilde{a}_t$ and $\tilde{b}_t$ and also defines her two-period expectations $E_t^*\left[\tilde{a}_t^- + \tilde{u}_{a,t+1} + \left(\tilde{b}_t^- + \tilde{u}_{b,t+1}\right)\ell_{t+2}\right] = \tilde{a}_t^- + E_t^*\left[\tilde{u}_{a,t+1}\right] + \tilde{b}_t^- E_t^*\left[\ell_{t+2}\right] + E_t^*\left[\tilde{u}_{b,t+1}\ell_{t+2}\right]$ for any $\left(\tilde{a}_t^-, \tilde{b}_t^-\right)$ and processes for $(\tilde{u}_{a,t+1}, \tilde{u}_{b,t+1})$. The researcher then gets to pick a particular (a_t, b_t), $\left(a_t^-, b_t^-\right)$, and processes for $(u_{a,t+1}, u_{b,t+1})$ so that $a_t = a_{t-1}^- + u_{a,t}$ and $b_t = b_{t-1}^- + u_{b,t}$ and the subjective expectations match the empirical targets. Again, the agent does not learn from the researcher's choices because these are chosen after she has defined her subjective expectations.

III. A Hansen-Jagannathan Bound for Beliefs

In this section, we provide a solution to the problem of arbitrariness. We propose a new moment of forecast errors: the forecast error Sharpe ratio (FESR), which is simply the average forecast error divided by the standard deviation of the forecast error. We show that the FESR can evaluate a mechanism for all possible latent ℓ_{t+1}. Further, we show that the empirical FESR provides a lower bound on the required deviation from FIRE.

Importantly, this section and all of the following sections only require the two assumptions of Section I and do not impose either of the conditions added in Section II. Under the two assumptions of Section I, we have that the subjective expectation for any variable y_{t+1} can be expressed as

$$E_t^*\left[y_{t+1}\right] \quad = \quad E_t\left[S_{t+1} y_{t+1}\right] \tag{9}$$

where $S_{t+1} \equiv \frac{\pi_t^*(\ell_{t+1})}{\pi_t(\ell_{t+1})}$ summarizes the change of measure from the FIRE distribution $\pi_t\left(\ell_{t+1}\right)$ to the subjective distribution $\pi_t^*\left(\ell_{t+1}\right)$. This is very similar to the change of measure approach in asset pricing and, following Cui et al. (2026), we will refer to S_{t+1} as the subjective belief factor (SBF) in an intentional nod to the stochastic discount factor (SDF) used in asset pricing. Conveniently, even if ℓ_{t+1} is a vector of variables, we can still summarize the

subjective expectations with a single scalar S_{t+1}. Since $S_{t+1} \equiv \frac{\pi_t^*(\ell_{t+1})}{\pi_t(\ell_{t+1})}$, we know that

$$E_t\left[S_{t+1}\right] \;=\; 1. \tag{10}$$

The fact that we can represent subjective expectations of y_{t+1} using an SBF does not rely on our choice to study subjective expectations of yields (i.e., a financial variable) in our application of Section IV. Equation (9) could just as easily be applied to subjective expectations of non-financial variables, e.g., inflation, GDP, or unemployment. A large goal of this paper is to demonstrate that the mathematical tools of asset pricing can be used for much more than pricing assets once we frame subjective expectations using a change of measure.

Let $f_{y,t+1} \equiv y_{t+1} - E_t^*\left[y_{t+1}\right]$ be the forecast error and note that $Std_t\left(f_{y,t+1}\right) = Std_t\left(y_{t+1}\right)$ since $E_t^*\left[y_{t+1}\right]$ is known at time t. We define the FESR as $\frac{E[f_{y,t+1}]}{Std(f_{y,t+1})}$ and the conditional FESR as $\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}$. Then, we have the following bound.

Lemma 1. *The variance of the SBF bounds the squared FESR and the average squared conditional FESR,*

$$Var\left(S_{t+1}\right) \geq \;\; E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \geq \;\; \left(\frac{E\left[f_{y,t+1}\right]}{Std\left(f_{y,t+1}\right)}\right)^2. \tag{11}$$

The proof follows the same logic as the original Hansen-Jagannathan bound, which relates the variance of the SDF to the squared Sharpe ratio for excess returns on assets. It is useful to see the main part of the proof for intuition. From equation (9), we have that

$$\begin{aligned} \frac{E_t\left[f_{y,t+1}\right]}{Std_t\left(f_{y,t+1}\right)} \;&=\; \frac{E_t\left[y_{t+1}\right] - E_t^*\left[y_{t+1}\right]}{Std_t\left(f_{y,t+1}\right)} = \frac{Cov_t\left(1 - S_{t+1}, y_{t+1}\right)}{Std_t\left(f_{y,t+1}\right)} \leq Std_t\left(S_{t+1}\right) \\ \Rightarrow E\left[\left(\frac{E_t\left[f_{y,t+1}\right]}{Std_t\left(f_{y,t+1}\right)}\right)^2\right] \;&\leq\; E\left[Std_t\left(S_{t+1}\right)^2\right] = Var\left(S_{t+1}\right). \end{aligned}$$

The first line uses the fact that the SBF allows us to turn the expected forecast error $E_t\left[f_{y,t+1}\right]$ into a covariance $Cov_t\left(1 - S_{t+1}, y_{t+1}\right)$ and then uses the fact that correlations are at most 1, so $Cov_t\left(1 - S_{t+1}, y_{t+1}\right) \leq Std_t\left(S_{t+1}\right) Std_t\left(f_{y,t+1}\right)$. The second line simply takes the first line, squares both sides, and takes the unconditional expectation. Since $E_t\left[S_{t+1}\right] = 1$

always, $E\left[Std_t\left(S_{t+1}\right)^2\right]$ simplifies to $Var\left(S_{t+1}\right)$. The proof in Appendix B shows the final step which is showing $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \geq \left(\frac{E[f_{y,t+1}]}{Std(f_{y,t+1})}\right)^2$. Intuitively, adding any information that changes $E_t\left[f_{y,t+1}\right]$ over time raises $E\left[\left(E_t\left[f_{y,t+1}\right]\right)^2\right]$ above $E\left[f_{y,t+1}\right]^2$ due to the Jensen's term and lowers $E\left[Std_t\left(f_{y,t+1}\right)^2\right]$ below $Std\left(f_{y,t+1}\right)^2$ by the law of total variance. Thus, the only detail we leave to the Appendix is handling the interaction of the numerator and denominator.

Equation (11) provides a powerful bound that allows us to evaluate a belief mechanism across all possible ways that y_{t+1} could be related to ℓ_{t+1}. To start, we will focus on the simpler version, $Var\left(S_{t+1}\right) \geq \left(\frac{E[f_{y,t+1}]}{Std(f_{y,t+1})}\right)^2$. Given a belief mechanism (i.e., the processes for ℓ_{t+1} and $\pi_t^*\left(\ell_{t+1}\right)$), we can calculate $Var\left(S_{t+1}\right) = Var\left(\frac{\pi_t^*(\ell_{t+1})}{\pi_t(\ell_{t+1})}\right)$. Given forecast error data for some variable y_{t+1}, we can calculate the empirical FESR $\frac{E[f_{y,t+1}]}{Std(f_{y,t+1})}$. If $Var\left(S_{t+1}\right) < \left(\frac{E[f_{y,t+1}]}{Std(f_{y,t+1})}\right)^2$, then we can conclude that this belief mechanism cannot explain the empirical FESR no matter how y_{t+1} is related to ℓ_{t+1}. Regardless of whether ℓ_{t+1} is a latent trend in y_{t+1}, ℓ_{t+1} is related to the growth in y_{t+1}, ℓ_{t+1} is related to a non-inflation component of y_{t+1}, etc., our conclusion would still hold. This is similar to how the original Hansen-Jagannathan bound allows researchers to use empirical Sharpe ratios for excess returns to reject SDF's even when the SDF's depend on latent, unobservable variables.

To highlight this usefulness, let's return to the concrete example of Section II.A. We have an extrapolative belief mechanism, where ℓ_{t+1} is AR(1) with persistence $\rho = 0.9$ and standard normal shocks and the agent believes the persistence is $\rho^* = 0.95$, so $\pi_t^*\left(\ell_{t+1}\right) = \pi_t\left(\ell_{t+1} - \left(\rho^* - \rho\right)\ell_t\right)$. Then, the SBF is $S_{t+1} = \exp\left(-\frac{1}{2}\left(\rho^* - \rho\right)^2 \ell_t^2 + \left(\rho^* - \rho\right)\ell_t \varepsilon_{t+1}\right)$ and $Var\left(S_{t+1}\right) = \left(1 - 2\frac{(\rho^* - \rho)^2}{1-\rho^2}\right)^{-1/2} - 1 \approx 0.013$. As shown in Section II.A, this belief mechanism can explain any Mincer-Zarnowitz coefficient or any Coibion-Gorodnichenko coefficient for forecast errors of y_{t+1} if we simply consider $y_{t+1} = a\ell_t + \ell_{t+1}$ for some a value. Proposition 1 shows that we can explain any Mincer-Zarnowitz coefficient, Coibion-Gorodnichenko coefficient, autocorrelation of forecast errors, and arbitrarily many autocorrelations of y_{t+1} if we allow $y_{t+1} = a_t + b_t\ell_{t+1}$. Yet, if the empirical squared FESR for y_{t+1} is above 0.013, then we

can reject this mechanism for all possible ways that y_{t+1} could be related to ℓ_{t+1}, including $y_{t+1} = a\ell_t + \ell_{t+1}$, $y_{t+1} = a_t + b_t\ell_{t+1}$, and any non-linear relationships. Thus, the empirical FESR gives us a way to test belief mechanisms that does not suffer from the problem of arbitrariness (i.e., changing how y_{t+1} relates to ℓ_{t+1} will not change the result of the test).

A. Minimum deviation from FIRE

For many researchers, the likely benefit from equation (11) is not that it can reject specific mechanisms (e.g., extrapolation with $\rho = 0.9$ and $\rho^* = 0.95$), but instead that it can be used to ask, "what does a belief mechanism need in order to not be rejected by the empirical FESR?" In the concrete example of extrapolation, we have $Var\left(S_{t+1}\right) = \left(1 - 2\frac{(\rho^*-\rho)^2}{1-\rho^2}\right)^{-1/2} - 1$. This means that, given an empirical FESR and some FIRE persistence ρ, we can calculate the minimum required $|\rho^* - \rho|$ for $Var\left(S_{t+1}\right)$ to exceed the squared empirical FESR. Thus, the empirical FESR provides a lower bound on the underlying mistake in the agent's beliefs (i.e., the gap between the FIRE ρ and the believed ρ^*). In contrast, Proposition 1 shows that the empirical Mincer-Zarnowitz, Coibion-Gorodnichenko, and autocorrelation coefficients only tell us that $|\rho^* - \rho| > 0$; these coefficients can always be matched by $|\rho^* - \rho| = \epsilon$ for arbitrarily small ϵ.

Intuitively, what is $Var\left(S_{t+1}\right)$? To answer this, we start with the conditional variance, $Var_t\left(S_{t+1}\right)$. Figure 2 provides a useful visualization. If the agent's beliefs exactly align with FIRE, then $S_{t+1} = \frac{\pi_t^*(\ell_{t+1})}{\pi_t(\ell_{t+1})}$ is 1 for all ℓ_{t+1}. Given that $E_t\left[S_{t+1}\right] = 1$ by definition, the conditional variance $Var_t\left(S_{t+1}\right)$ summarizes how much S_{t+1} tends to deviate from 1 (i.e., how much beliefs tend to deviate from FIRE). The unconditional variance is then simply the average of the conditional variance, $Var\left(S_{t+1}\right) = E\left[Var_t\left(S_{t+1}\right)\right]$, so the unconditional variance tells us, for an average period, how much the SBF tends to deviate from FIRE. For our concrete example of extrapolation, the conditional variance $Var_t\left(S_{t+1}\right) = \exp\left(\left(\rho^* - \rho\right)^2 \ell_t^2\right) - 1$ tells us how much beliefs deviate from FIRE each period depending on the current ℓ_t, and $Var\left(S_{t+1}\right) = \left(1 - 2\frac{(\rho^*-\rho)^2}{1-\rho^2}\right)^{-1/2} - 1$ tells us how much beliefs deviate from FIRE on average.

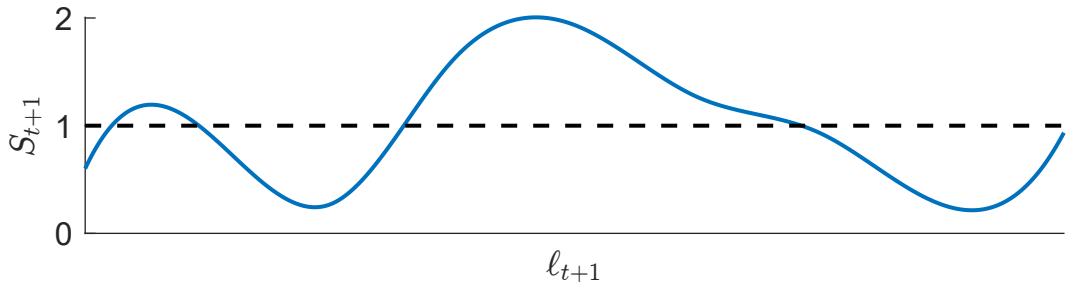


Figure 2. Variation in SBF. This figure provides a visualization of an example SBF $S_{t+1} = \frac{\pi_t^*(\ell_{t+1})}{\pi_t(\ell_{t+1})}$ for different ℓ_{t+1}. If ℓ_{t+1} is a vector of variables, then this would be a multidimensional figure. By construction, the mean of the SBF is 1. The variance of the SBF captures how much it deviates from 1.

B. Incorporating more information into the bound

The bound, $Var\left(S_{t+1}\right) \geq \left(\frac{E[f_{y,t+1}]}{Std(f_{y,t+1})}\right)^2$, discussed above applies for any relationship between y_{t+1} and ℓ_{t+1} and does not require us to take a stand on the FIRE process for forecast errors. This generality can be useful, but we can derive a tighter bound if we have more information. The subsections below discuss two important ways to tighten the bound.

B.1. Predictors of forecast errors

As mentioned in Section II.C, the goal of this paper is not to say that we should cease using Mincer-Zarnowitz, Coibion-Gorodnichenko, or autocorrelation coefficients. Instead we want to highlight that, on their own, these moments suffer from the problem of arbitrariness of Proposition 1 but, once combined with the FESR approach, these moments contain valuable information.

From equation (11), we see that a tighter version of the $Var\left(S_{t+1}\right) \geq \left(\frac{E[f_{y,t+1}]}{Std(f_{y,t+1})}\right)^2$ bound is $Var\left(S_{t+1}\right) \geq E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$. Thus, evidence of forecast error predictors (e.g., the current forecast, the current forecast error, or the current forecast revision) can be used to raise the value of the empirically measured $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$. This increases the required $Var\left(S_{t+1}\right)$, allowing us to reject more mechanisms and raising the required minimum deviation from FIRE.

Additionally, the average squared conditional FESR, $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$, shows how we

can distill information from multiple forecast error predictors into a single number which we can then compare to belief mechanisms. This method of incorporating multiple predictors of forecast errors is analogous to asset pricing work using predictors of excess returns to raise the maximum achievable Sharpe ratio for a portfolio. In this sense, the FESR approach provides a new framework to organize and interpret existing and future evidence of forecast error predictability.

B.2. Conditionally linear variables

The example above showed how we can tighten the bound by raising the empirical side of the bound. Now, we consider a separate approach which is to tighten the bound by lowering the mechanism side of the bound. The bound of equation (11) applies to all possible ways y_{t+1} could be related to ℓ_{t+1}. If we restrict our attention to conditionally linear relationships, then we can significantly tighten the bound.

We define y_{t+1} as being conditionally linear with ℓ_{t+1} if

$$E_t\left[y_{t+1}|\ell_{t+1}\right] \;=\; a_t + b_t\ell_{t+1}.$$

This obviously includes the case of $y_{t+1} = a_t + b_t\ell_{t+1}$, which we showed was sufficient for Proposition 1, but also includes any $y_{t+1} = a_t + b_t\ell_{t+1} + \eta_{t+1}$ where $E_t\left[\eta_{t+1}\right] = 0$ and η_{t+1} is independent of ℓ_{t+1}. Note that this can still include higher order moments of y_{t+1}. For example, suppose we have a vector $\ell_{t+1} = \left(\mu_{t+1}\ \sigma_{t+1}\right)'$ and $y_{t+1} = a_t + b_t\mu_{t+1} + c_t\sigma_{t+1}\varepsilon_{t+1}$ where ε_{t+1} is white noise. Then we have that $E_t\left[y_{t+1}|\ell_{t+1}\right] = a_t + \left(b_t\ 0\right)\ell_{t+1}$, so y_{t+1} would be conditionally linear with ℓ_{t+1}.

If y_{t+1} is conditionally linear with ℓ_{t+1}, then

$$Var\left(S_{t+1}^{(1)}\right) \geq \;\; E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \geq \;\; \left(\frac{E\left[f_{y,t+1}\right]}{Std\left(f_{y,t+1}\right)}\right)^2 \tag{12}$$

for $S_{t+1}^{(1)} = 1 + \left(E_t^*\left[\ell_{t+1}\right] - E_t\left[\ell_{t+1}\right]\right)'\Sigma_t^{-1}\varepsilon_{t+1}$ where $\Sigma_t = Cov_t\left(\ell_{t+1}, \ell_{t+1}\right)$ and $\varepsilon_{t+1} \equiv \ell_{t+1} - E_t\left[\ell_{t+1}\right]$. This is because this $S_{t+1}^{(1)}$ satisfies $E_t^*\left[y_{t+1}\right] = E_t\left[S_{t+1}^{(1)}y_{t+1}\right]$ for all conditionally linear y_{t+1}, which means we can repeat the same steps as Lemma 1 to derive the bound using

$Var\left(S_{t+1}^{(1)}\right)$.[9] In words, equation (12) says that we only need to consider variation coming from changes to the first moment of the ℓ_{t+1} distribution (i.e., $E_t^*\left[\ell_{t+1}\right] - E_t\left[\ell_{t+1}\right]$). Variation in S_{t+1} related to changes in higher order moments of ℓ_{t+1} will not matter for conditionally linear y_{t+1} and therefore can be removed from the mechanism side of the bound. We use the superscript to highlight that this alternative SBF $S_{t+1}^{(1)}$ only depends on the first-moment beliefs.

IV. Application to Yield Forecast Errors

Equations (11) and (12) are bounds. Like all bounds, they face the natural question of whether they are informative or trivial. To address this question, we study the empirical FESR of yield forecasts and show what this implies for mechanisms of Bayesian learning, extrapolation, adaptive expectations, and under/over-reaction (e.g., diagnostic expectations). For Bayesian learning, we will show that the bound is quite useful for y_{t+1} that are conditionally linear with ℓ_{t+1} and, for the behavioral mechanisms, we will show that the bound is useful for all possible relationships between y_{t+1} and ℓ_{t+1}.

A. Empirical FESR

As shown in the introduction, we focus on the ten-year maturity yield, which we denote as y_t. We collect one-year forecasts for the yield, $E_t^*\left[y_{t+1}\right]$, from the Blue Chip's mean professional forecast, which began covering the ten-year yield in 1989. We take the future realized values, y_{t+1}, from the Federal Reserve Bank of St. Louis. Our sample is 1989-2024. While forecasts are provided monthly, we only use annual observations to avoid overlapping observations.

To avoid having to take a stand on the FIRE process for forecast errors, we use the FESR $\frac{E[f_{y,t+1}]}{Std(f_{y,t+1})}$ rather than the conditional FESR. Over our sample, we estimate that the FESR

[9] There are infinitely many $\tilde{S}_{t+1}$ that satisfy $E_t^*\left[y_{t+1}\right] = E_t\left[\tilde{S}_{t+1}y_{t+1}\right]$ for conditionally linear y_{t+1}. So, equation (12) holds for all of these possible $\tilde{S}_{t+1}$. To make the bound as tight as possible, we choose the minimum variance $\tilde{S}_{t+1}$, which is $S_{t+1}^{(1)} = 1 + \left(E_t^*\left[\ell_{t+1}\right] - E_t\left[\ell_{t+1}\right]\right)' \Sigma_t^{-1} \varepsilon_{t+1}$.

is -0.64 (0.00014) where the HAC p-value is in parentheses. In the subsections below, we directly address finite sample implications for each mechanism. The FESR comes from an average forecast error of -0.52 p.p. and a standard deviation of 0.81 p.p.

To ensure that our choice to use the empirical $\frac{E[f_{y,t+1}]}{Std(f_{y,t+1})}$ rather than estimating an empirical $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ does not cause us to reject a mechanism, we will calculate the maximum average squared conditional FESR $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ in each mechanism. In other words, our underlying test will be whether the maximum $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ possible in each mechanism exceeds the empirical $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$. We then use the fact that $\left(\frac{E[f_{y,t+1}]}{Std(f_{y,t+1})}\right)^2 \leq E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ to say that the empirical $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ must be at least as large as the empirical $\left(\frac{E[f_{y,t+1}]}{Std(f_{y,t+1})}\right)^2$, meaning that we can reject a mechanism if the maximum possible $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ in the mechanism is lower than our empirical $\left(\frac{E[f_{y,t+1}]}{Std(f_{y,t+1})}\right)^2$. This allows us to reject a mechanism without needing to take a stand on the FIRE process for forecast errors.

B. Bayesian mechanisms

A commonly proposed explanation for patterns in empirical forecast errors is that, while the patterns in forecast errors are clear ex post, a Bayesian agent may have thought that the forecast errors were almost entirely random noise in real time. In the subsections below, we discuss increasingly generalized Bayesian learning mechanisms and show a consistent result that we can reject mechanisms in which the Bayesian agent believes, in real time, that her forecast errors are predominantly (e.g., 90%) noise. Rephrased, we can reject mechanisms in which, due to observation noise, the Bayesian agent largely does not realize that her priors are incorrect or is uncertain whether her priors are incorrect.

B.1. Linear Gaussian

We start with the classic setting of Bayesian learning with linear Gaussian relationships. The agent observes ℓ_t where

$$\ell_t = c'\alpha_t + \varepsilon_t \tag{13}$$

$$\alpha_t = F\alpha_{t-1} + \eta_t. \tag{14}$$

We allow α_t to be an arbitrarily long vector of hidden states. In this setting, ε_t is the observation noise, c is a vector that determines how ℓ_t depends on the hidden states, F is a matrix, and η_t is the innovation to the hidden states. We assume $\varepsilon_t \sim N(0, \sigma^2)$ and $\eta_t \sim N(0, Q)$.

In this setting, the subjective agent's beliefs about future ℓ_{t+1} will differ from the beliefs of a FIRE agent that knows α_t. The first thing to note is that $Var_t(S_{t+1})$ is potentially infinite for $S_{t+1} = \frac{\pi_t^*(\ell_{t+1})}{\pi_t(\ell_{t+1})}$, depending on the values of c, σ, F, and Q. This is because learning can generate fat tails for the subjective believed distribution. In fact, for parameters where $Var_t(S_{t+1})$ is infinite, Appendix A.1 shows that if we allow for non-linear relationships between y_{t+1} and ℓ_{t+1} (e.g., $y_{t+1} = a_t + b_t \mathbb{1}\{\ell_{t+1} > k\}$) then we can replicate any possible moment of survey expectations and forecast error predictability by focusing on the tails of the distribution. So, the fact that the empirical FESR cannot reject these mechanisms under non-linear relationships is not simply a limitation of the FESR; under non-linear relationships, there is no possible empirical moment that can reject the mechanism.

Fortunately, many learning models focus on conditionally linear relationships.[10] In this case, we have a straightforward bound.

Lemma 2. *$E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \leq Var\left(S_{t+1}^{(1)}\right) = \left(\frac{K}{1-K} - \phi^2\right)\frac{1}{\phi^2+1}$ for any y_{t+1} that is conditionally linear with ℓ_{t+1}. The bound depends on $\phi = \frac{\sqrt{c'Qc}}{\sigma}$, which is the signal-to-noise ratio,*

[10] Recall from Section III.B.2 that conditionally linear relationships can still include higher order moments of y_{t+1}. The previously given example is $\ell_{t+1} = (\mu_{t+1}\ \sigma_{t+1})'$ and $y_{t+1} = a_t + b_t\mu_{t+1} + c_t\sigma_{t+1}\varepsilon_{t+1}$ with white noise ε_{t+1}. This is still a conditionally linear relationship as $E_t[y_{t+1}|\ell_{t+1}] = a_t + (b_t\ 0)\,\ell_{t+1}$.

and K, which is the steady-state Kalman gain,

$$E_t^* \left[c' \alpha_t \right] \;=\; E_{t-1}^* \left[c' \alpha_t \right] + K \left(\ell_t - E_{t-1}^* \left[c' \alpha_t \right] \right).$$

We will focus on what the Kalman gain K means for Lemma 2. One could imagine a setting (i.e., a choice of c, σ, F, and Q) in which the subjective agent believes that her forecast errors are largely driven by the noisy shocks ε_t. This would mean that the Kalman gain K is low, as the subjective agent does not think that her forecast error $\ell_t - E_{t-1}^* \left[c' \alpha_t \right]$ is very informative about the underlying states $c' \alpha_t$.[11] Lemma 2 says that this type of setting cannot generate a large $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})} \right)^2 \right]$ even though $E_t \left[f_{y,t+1} \right]$ and $Std_t \left(f_{y,t+1} \right)$ use the FIRE beliefs.

Rephrased, to have a large $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})} \right)^2 \right]$, we must have that the Bayesian agent is aware in real time that her forecast errors are not primarily noise. To show this visually, Figure 3 shows the combinations of (K, ϕ) that satisfy our empirical lower bound for $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})} \right)^2 \right]$ of 0.40. All parameter combinations above the blue curve satisfy the bound, i.e., $\left(\frac{K}{1-K} - \phi^2 \right) \frac{1}{\phi^2+1} \geq 0.40$. We see that the lowest possible K that satisfies the bound is 0.29. So, in real time the agent must know that at least 29% of her forecast error represents mistakes about the underlying $c' \alpha_t$ and at most 71% of the forecast error is due to noise. This rules out mechanisms where the underlying states are stable and the Bayesian agent largely makes mistakes (relative to FIRE) because of noise. As we increase the signal-to-noise ratio ϕ, the required K increases further.

To give a sense of comparison, the black dashed line of Figure 3 shows the parameter values for a hypothetical FIRE agent, where K is 0 since the agent knows the underlying α_t, and the purple dots show example values of (K, ϕ) from Veronesi (2000), Gilchrist and Saito (2008), Ai (2010), and Roussanov et al. (2021). Importantly, these example papers do not attempt to explain forecast errors for yields, which means we are not saying that we reject the claims of these papers. Instead, our goal is to make the following two points.

[11] To state this formally, $1 - K$ is the fraction of $\ell_t - E_{t-1}^* \left[c' \alpha_t \right]$ variance that the agent believes is due to the noisy shock.

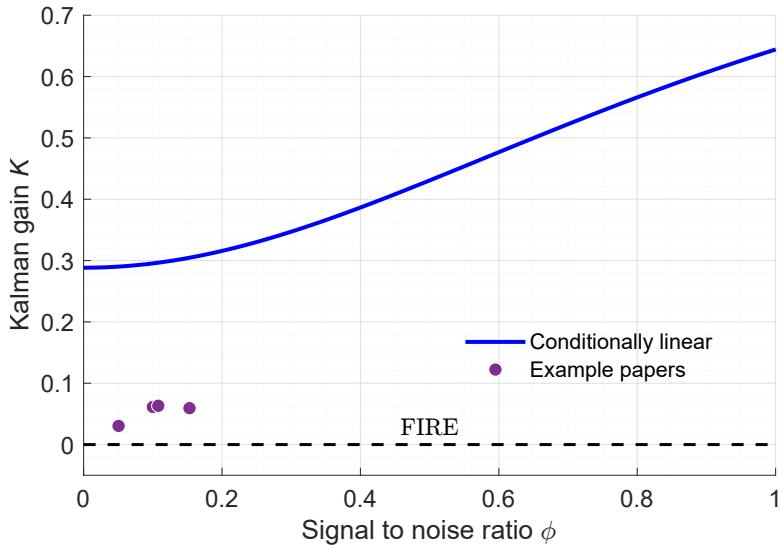


Figure 3. Bayesian learning (parameters above curve satisfy bound). This figure shows the minimum required deviation from FIRE to match the empirical FESR. For each value of the signal to noise ratio, the blue curve shows the minimum required Kalman gain in order to generate $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ as large as the squared empirical FESR for a conditionally linear y_{t+1}. The black dashed line shows the benchmark of FIRE, where the Kalman gain is 0. The purple dots show example paper values for comparison.

First, we know from Proposition 1 that one could always take the mechanisms for $E_t^*[\ell_{t+1}]$ from these papers and match the empirical $\hat{\beta}^{MZ}, \hat{\beta}^{AC}, \hat{\beta}^{CG}$ for yields for some $y_{t+1} = a_t + b_t\ell_{t+1}$. The fact that these papers were not attempting to explain yields might even be interpreted as further support for this mechanism, i.e., Bayesian learning about a persistent hidden component with large observation noise can explain moments of both stock returns (as shown in the original papers) and yield forecast errors. Second, in contrast to the first result, we know that there is no way to take these mechanisms and construct a y_{t+1} that is conditionally linear with ℓ_{t+1} that will explain the empirical yield FESR. A key benefit is that we can reach this conclusion without needing to individually check each possible $y_{t+1} = a_t + b_t\ell_{t+1} + \eta_{t+1}$ (i.e., every possible way yields could be related to a persistent component plus large observation noise). This is the benefit of having a lower bound on the required deviation from FIRE. While none of these mechanisms are FIRE, the blue curve of Figure 3 shows that they are still too close to FIRE to match the empirical FESR.

B.2. General Markov process

What if the agent is learning about a regime-switching process or other non-Gaussian underlying states? The linear Gaussian setting provides the main intuition, showing how the FESR is related to the Kalman gain and the signal-to-noise ratio. We now show that this result holds in a more general setting of Bayesian learning.

We assume the agent observes

$$\ell_t \quad = \quad c'\alpha_t + \varepsilon_t \tag{15}$$

where α_t is an arbitrarily long vector of states that follow a stationary Markov process (e.g., regime-switching) and $\varepsilon_t \sim N(0, \sigma^2)$ is the observation noise. Let $Q_t \equiv Cov_t(\alpha_{t+1}, \alpha_{t+1})$ be the covariance of innovations to α_t. Because the underlying states may be non-Gaussian, we no longer have that the Bayesian agent follows a linear updating rule. However, we can still define a generalized Kalman gain

$$\bar{K} \quad \equiv \quad \frac{Cov\left(E_t^*\left[c'\alpha_t\right] - E_{t-1}^*\left[c'\alpha_t\right], \ell_t - E_{t-1}^*\left[c'\alpha_t\right]\right)}{Var\left(\ell_t - E_{t-1}^*\left[c'\alpha_t\right]\right)} \tag{16}$$

which measures, on average, how much the agent updates her beliefs in response to her forecast error. Equivalently, we can say that the agent believes, on average, that $1 - \bar{K}$ of her forecast error is due to observation noise. Similarly, we can define a generalized signal-to-noise ratio,

$$\bar{\phi} \quad \equiv \quad \frac{\sqrt{c'\bar{Q}c}}{\sigma} \tag{17}$$

where $\bar{Q} \equiv E[Q_t]$ is the average covariance of the innovations. If the underlying states follow equation (14), then $\left(\bar{K}, \bar{\phi}\right)$ become the standard Kalman gain and signal-to-noise ratio of the previous subsection.

We have nearly the same result as the previous lemma.

Lemma 3. *In the generalized setting, if* $Cov\left(Var_t(\ell_{t+1}), \frac{(E_t[\ell_{t+1}] - E_t^*[\ell_{t+1}])^2}{Var_t(\ell_{t+1})}\right) \geq 0$, *then we have* $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \leq Var\left(S_{t+1}^{(1)}\right) \leq \left(\frac{\bar{K}}{1-\bar{K}} - \bar{\phi}^2\right)\frac{1}{\bar{\phi}^2+1}$ *for any* y_{t+1} *that is conditionally*

linear with ℓ_{t+1}.

Lemma 3 is identical to Lemma 2 except we replace (K, ϕ) with $\left(\bar{K}, \bar{\phi}\right)$. This means that the previous intuition still holds. We cannot generate a large $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ if the agent believes, in real time, that her forecast errors are 94-97% noise. The only requirement for this more generalized setting is that $Cov\left(Var_t\left(\ell_{t+1}\right), \frac{(E_t[\ell_{t+1}]-E_t^*[\ell_{t+1}])^2}{Var_t(\ell_{t+1})}\right) \geq 0$. If the conditional variance of innovations to α_t is constant over time, then this condition is immediately satisfied. If we allow the conditional variance of innovations to vary over time, we simply need that periods of volatile shocks to the underlying states also tend to be the periods in which a hypothetical FIRE agent can easily predict the forecast errors of the Bayesian agent (i.e., high $\frac{(E_t[\ell_{t+1}]-E_t^*[\ell_{t+1}])^2}{Var_t(\ell_{t+1})}$).

B.3. Finite samples and parameter learning

Given that our empirical estimates are based on a finite sample, there is a natural question of what these Bayesian mechanisms would predict over finite samples. A particularly relevant example would be parameter learning where some of the elements of α_t are constant over time. In steady state, the agent would perfectly learn those parameters but, over a finite sample, the agent will make mistakes about these parameters that can be large if she starts with a prior that is far from the actual parameter values. More broadly, suppose time $t = 0$ marks a large change in the hidden states or a change to parameter values. Then, a Bayesian agent who has been learning from prior observations will enter time 0 with a prior that is heavily biased relative to a hypothetical FIRE agent that knows the new values.

To address this question, we allow the agent to have a prior at time 0 which can differ from the steady-state prior. We do not require her to have the correct prior distribution or to know that a change has occurred at time 0. We only require that her prior is sufficiently diffuse that she is not overconfident.[12] To handle this added generality, we will assume a

[12]Specifically, we require that $E\left[E_t^*\left[f_{\ell,t+1}^2\right]\right] \geq E\left[f_{\ell,t+1}^2\right]$ for the finite sample $t = 0, 1, ..., T-1$.

slightly stricter condition than Section IV.B.2, which is that $Var_t(\ell_{t+1})$ is constant.[13] Let $\bar{P}_t \equiv E[Cov_t^*(\alpha_{t+1}, \alpha_{t+1})]$ reflect the agent's uncertainty about underlying states and let $Q = Cov_t(\alpha_{t+1}, \alpha_{t+1})$ reflect the FIRE uncertainty.

Note that we can consider an agent who has an overconfident prior but, from a practical perspective, this then becomes similar to a behavioral bias. Over a finite sample, there may only be a small distinction between a Bayesian agent who is wrong about the underlying parameters and clings to those beliefs because she is overconfident and a behavioral agent who is wrong about the underlying parameters (e.g., an extrapolative agent believes in an incorrect persistence parameter) and clings to those beliefs because she is irrational. Given that Section IV.C will cover behavioral agents who simply stick to their incorrect beliefs and will cover the finite sample implications, we think the more relevant Bayesian case to study is an agent who may have an incorrect prior but is at least aware of how wrong she may be about the underlying states/parameters (i.e., an agent whose prior is not overconfident).

For each period, the value

$$K_t = 1 - \frac{\sigma^2}{c'\bar{P}_t c + \sigma^2}, \tag{18}$$

captures how much the agent believes her forecast error reflects information about the states versus observation noise. Let $\bar{K}_T$ be a finite-sample version of $\bar{K}$ which solves

$$\frac{\bar{K}_T}{1-\bar{K}_T} = \frac{1}{T}\sum_{t=1}^{T}\frac{K_t}{1-K_t}. \tag{19}$$

The signal to noise ratio is $\phi = \frac{\sqrt{c'Qc}}{\sigma}$. Then, we have the following result.

Lemma 4. *In the finite sample setting, for any y_{t+1} that is conditionally linear with ℓ_{t+1}, we have $\frac{1}{T}\sum_{t=1}^{T} E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \leq \left(\frac{\bar{K}_T}{1-\bar{K}_T} - \phi^2\right)\frac{1}{\phi^2+1}$.*

Like with Lemma 3, this is the same bound as Lemma 2 except that K is replaced with a generalized finite sample $\bar{K}_T$. Thus, the insights from the blue curve of Figure 3 still hold, just with a more generalized version of the vertical axis. We cannot have a large

[13] Section IV.B.2 allows $Var_t(\ell_{t+1})$ to vary over time so long as $Cov\left(Var_t(\ell_{t+1}), \frac{(E_t[\ell_{t+1}]-E_t^*[\ell_{t+1}])^2}{Var_t(\ell_{t+1})}\right) \geq 0$.

$\frac{1}{T}\sum_{t=1}^{T} E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ if the agent believes in real time that her forecast errors are 94-97% noise. We must have $\bar{K}_T$ above 0.29.

For intuition, it is helpful to consider an example where the agent's prior distribution is very inaccurate. This could then generate a large $\frac{1}{T}\sum_{t=1}^{T} E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ because the hypothetical FIRE agent who knows α_t would be able to easily predict the Bayesian agent's forecast errors. However, because the subjective agent is Bayesian, she will notice that the realized ℓ_t is very unlikely based on her prior distribution and will substantially revise her expectations of the underlying states (i.e., she will generate a large finite-sample $\bar{K}_T$). So, the empirical FESR can still rule out a mechanism of regime shifts or parameter learning over a finite sample in which the change to the underlying states is largely undetected by a Bayesian agent in real time (i.e., low $\bar{K}_T$). For the shift to not be detected by a Bayesian agent, there must be a large amount of observation noise, which would lower $\frac{1}{T}\sum_{t=1}^{T} E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ since observation noise increases $Std_t\left(f_{y,t+1}\right)$.

C. Behavioral mechanisms

Now, we turn our attention to behavioral mechanisms. For Bayesian learning mechanisms, the broad takeaway was that we can reject mechanisms in which patterns in forecast errors are difficult for the agent to detect in real time and are only obvious ex post. As summarized in Figure 3 and the more general Lemmas 3 and 4, those mechanisms (i.e., those with low Kalman gains) do not deviate enough from FIRE to match the empirical FESR. For behavioral mechanisms, the broad takeaway is that the parameters controlling the degree of bias must be sufficiently large. For extrapolation, adaptive expectations, and over/underreaction, this means a large difference between the FIRE persistence and the subjective persistence, a large gain, and a large exaggeration/understatement of recent information, respectively.

Similar to the Bayesian learning mechanisms, the key question will not be whether these mechanisms deviate from FIRE but whether the deviation is sufficiently large to match the empirical FESR. Importantly, failing this test does not mean that the particular behavioral

mechanism is completely absent from the empirical subjective expectations. It simply says that the particular behavioral mechanism alone is not sufficient to explain the empirical FESR, and how far it falls below the bound tells us how much of the empirical FESR would need to be accounted for by a different belief mechanism.

C.1. Extrapolation

We start with extrapolation, which has been the concrete example used throughout the paper. There is an AR(1) process

$$\ell_{t+1} \quad = \quad \rho\ell_t + \sigma\varepsilon_{t+1}$$

with standard normal shocks ε_{t+1}. The subjective agent believes the persistence is $\rho^* > \rho$.

Lemma 5. *For any y_{t+1}, we have $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \leq Var\left(S_{t+1}\right) = \left(1 - 2\frac{(\rho^*-\rho)^2}{1-\rho^2}\right)^{-1/2} - 1$. We have a tighter bound of $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \leq Var\left(S^{(1)}_{t+1}\right) = \frac{(\rho^*-\rho)^2}{1-\rho^2}$ for any y_{t+1} that is conditionally linear with ℓ_{t+1}.*

Conveniently, both the unrestricted bound and the conditionally linear bound only depend on ρ and ρ^*.

The blue and red curves of Figure 4's left panel show the combinations of (ρ, ρ^*) which satisfy the bounds for our empirical squared FESR of 0.40. These curves show two main results. First, for many values of ρ, there is no value of $\rho^* \leq 1$ that satisfies the bound. If we allow for any possible y_{t+1}, then we can reject any mechanism with $\rho > 0.60$. If we restrict to conditionally linear y_{t+1}, then we can reject any mechanism with $\rho > 0.42$. Thus, without needing to calculate each example, we can already rule out any story where the yield FESR is explained by extrapolation of inflation, a persistent trend in the natural rate r*, slow-moving changes in productivity, or any other highly persistent variable.

Second, the curves show that $\rho^* - \rho$ must be large. The subjective agent's beliefs about the persistence cannot be slightly wrong; they must be substantially wrong. The dashed black line shows the FIRE benchmark, which is simply that $\rho^* = \rho$. For any possible y_{t+1},

the red solid curve shows that we must have $\rho^* - \rho > 0.4$. This demonstrates how the FESR provides a lower bound on the required deviation from FIRE that holds even if we (the researchers) do not observe ℓ_t and/or do not know how y_t is related to ℓ_t. If we add the restriction that y_{t+1} is conditionally linear with ℓ_{t+1}, then the required deviation becomes even larger, $\rho^* - \rho > 0.58$.

The purple dots in Figure 4's left panel show example values for (ρ, ρ^*) from Altı and Tetlock (2014), Barberis et al. (2015), Greenwood and Hanson (2015), and Rozsypal and Schlafmann (2023). Generally, these papers produce $\rho^* - \rho$ less than 0.1 which is noticeably lower than the required 0.40 or 0.58 minimum. Visually, we see that most of the purple dots are simply too close to the FIRE benchmark. One notable exception is Barberis et al. (2015), who study extrapolation of an endogenous sentiment variable rather than extrapolation of an exogenous process. This creates a feedback loop, wherein the level of sentiment impacts extrapolative expectations, which then endogenously impact the level of sentiment, and so on. Their model generates $(\rho, \rho^*) = (0.37, 0.84)$ when 75% of agents are extrapolators, which is their highest considered value, and this gap between ρ^* and ρ is large enough to satisfy the unrestricted bound.[14]

The right panel of Figure 4 considers finite samples. As discussed in Section IV.B.3, under Bayesian learning, the agent's updating (i.e., $\bar{K}_T$) is an endogenous object that depends on her initial prior. For behavioral mechanisms, the parameters controlling how the agent updates her beliefs are typically fixed which means we can calculate the entire distribution of results over finite samples. As an example, for extrapolation, the probability of a finite sample $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2$ greater than 0.40 is determined by a weighted chi-squared distribution for conditionally linear y_{t+1}. Appendix A.2 provides the full calculation for the finite sample probabilities.

The right panel of Figure 4 shows the values of (ρ, ρ^*) for which there is at least a 5%

[14]To calculate a representative ρ^*, we average across the two types of agents. 75% of agents believe the persistence of sentiment is 1 and 25% of agents believe the persistence is ρ. Given $\rho = 0.37$, this means the average believed persistence is 0.84.

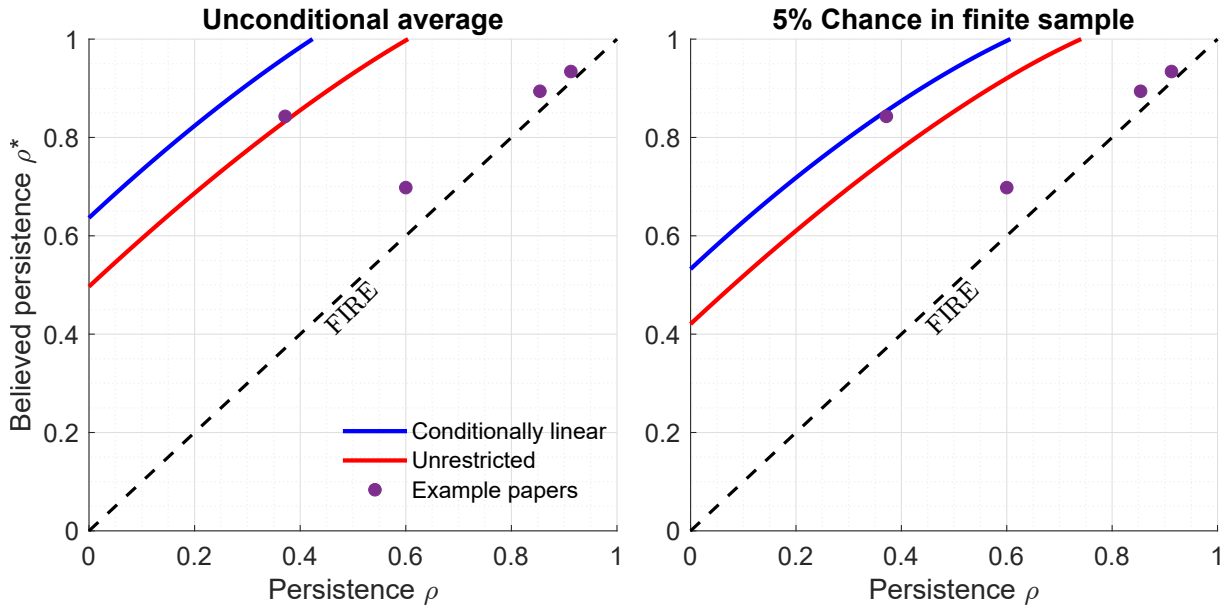


Figure 4. Extrapolation (parameters above curve satisfy bound). The left panel shows the minimum required deviation from FIRE to match the empirical FESR. For each value ρ for the persistence, the red curve shows the minimum required believed persistence ρ^* in order to generate $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ as large as the squared empirical FESR. The blue curve shows a higher required believed persistence when we restrict y_{t+1} to be conditionally linear with ℓ_{t+1}. The black dashed line shows the benchmark of FIRE, where $\rho^* = \rho$. The purple dots show example paper values for comparison. In the right panel, the red curve shows the minimum required believed persistence ρ^* for there to be a 5% chance of generating a finite sample $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2$ as large as the squared empirical FESR over a 34 period sample. The blue curve again shows a higher required ρ^* when we restrict y_{t+1} to be conditionally linear with ℓ_{t+1}. The black dashed line and the purple dots are exactly the same as the left panel.

chance of observing $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2$ greater than 0.40 over a 34 year period. While this relaxes the bound somewhat, we see that the main results of the left panel continue to hold. For unrestricted y_{t+1} and conditionally linear y_{t+1}, the minimum required $\rho^* - \rho$ is 0.26 and 0.40. As shown in the right panel, the required deviation from FIRE is still noticeably above the values produced in most of the example papers, with Barberis et al. (2015) again being the standout which satisfies the bound for unrestricted y_{t+1} and very nearly satisfies the bound for conditionally linear y_{t+1}.

C.2. Adaptive expectations

A long-studied deviation from FIRE is adaptive expectations, in which the agent forms her expectations following a constant-gain learning rule. Let ℓ_t be some AR(∞) time series,[15]

$$\ell_{t+1} \quad = \quad (1-\Gamma)\,\mu + \sum_{j=0}^{\infty} \gamma_j \ell_{t-j} + \sigma\varepsilon_{t+1} \tag{20}$$

where $\Gamma = \sum_{j=0}^{\infty} \gamma_j$ and ε_{t+1} is standard normal. While a FIRE agent would simply know the value of μ, our subjective agent forms her expectation of the mean using

$$\mu_t^* \quad = \quad \mu_{t-1}^* + g\left(\ell_t - E_{t-1}^*\left[\ell_t\right]\right) \tag{21}$$

where $E_t^*\left[\ell_{t+1}\right] = (1-\Gamma)\,\mu_t^* + \sum_{j=0}^{\infty} \gamma_j \ell_{t-j}$. So, the agent's believed distribution $\pi_t^*\left(\ell_{t+1}\right)$ is $N\left(E_t^*\left[\ell_{t+1}\right], \sigma^2\right)$.

Lemma 6. *For any y_{t+1}, we have $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \leq Var\left(S_{t+1}\right) = \left(1 - 2\frac{(1-\Gamma)g}{2-(1-\Gamma)g}\right)^{-1/2} - 1$. We have a tighter bound of $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \leq Var\left(S_{t+1}^{(1)}\right) = \frac{(1-\Gamma)g}{2-(1-\Gamma)g}$ for any y_{t+1} that is conditionally linear with ℓ_{t+1}.*

Like with extrapolation, the bounds show that we can ignore most details of the setting and solely focus on two parameters (Γ, g), rather than needing to know the entire $\{\gamma_j\}_{j=0}^{\infty}$.

Figure 5 shows an analogous version of Figure 4 but for adaptive expectations. In the left panel, the blue and red curves show the values of (Γ, g) that would be needed to match the empirical FESR. The dashed black line shows the FIRE benchmark, where g would be zero because the agent would already know μ. At $\Gamma = 0$, the minimum required g is 0.40 if we allow for any y_{t+1} and 0.58 if we restrict to conditionally linear y_{t+1}. This is substantially higher than the values typically used in the literature. The purple dots show values from Orphanides and Williams (2005), Milani (2007), Hirshleifer et al. (2015), and Malmendier and Nagel (2016). Because the blue and red curves are so far from the example paper values, we slightly shift the paper values so that the dots do not completely overlap in the figure.

[15] Any ARMA(p,q) time series process has an AR(∞) representation. We choose to use this representation as it allows us to easily define our object of interest Γ.

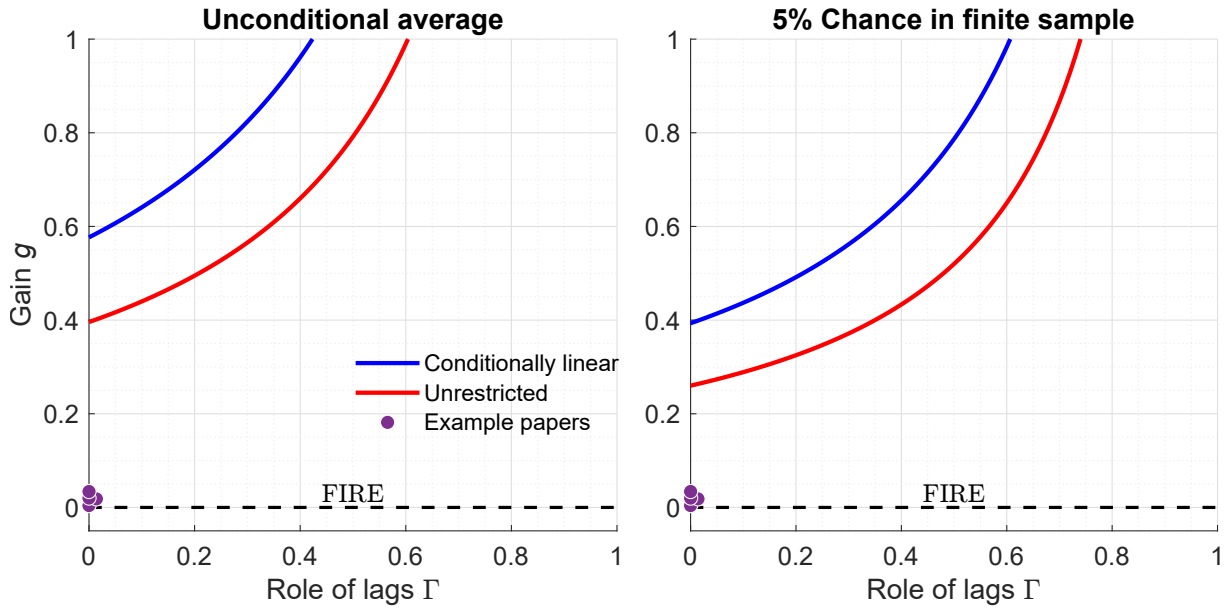


Figure 5. Adaptive expectations (parameters above curve satisfy bound). The left panel shows the minimum required deviation from FIRE to match the empirical FESR. For each value Γ of the total lag coefficients, the red curve shows the minimum required gain g in order to generate $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ as large as the squared empirical FESR. The blue curve shows a higher required gain when we restrict y_{t+1} to be conditionally linear with ℓ_{t+1}. The black dashed line shows the benchmark of FIRE, where $g = 0$. The purple dots show example paper values for comparison. In the right panel, the red curve shows the minimum required gain g for there to be a 5% chance of generating a finite sample $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2$ as large as the squared empirical FESR over a 34 period sample. The blue curve again shows a higher required gain when we restrict y_{t+1} to be conditionally linear with ℓ_{t+1}. The black dashed line and the purple dots are exactly the same as the left panel.

The example paper values all use $\Gamma = 0$ and $g \in (0.018, 0.02)$.[16]

For intuition on why we require a high g to match the empirical FESR, it is useful to focus on the case where ℓ_{t+1} is iid (i.e., $\gamma_j = 0$ for all j). In this case, the agent's expected mean μ_t^* is $g\sum_{j=0}^{\infty}(1-g)^j \ell_{t-j}$, which is simply a weighted average of past observed ℓ_t. For

[16]Because the example papers focus on iid ℓ_{t+1}, the bounds of Lemma 6 are conveniently immune to changing the horizon of the forecast. At time t, the FIRE belief is $\ell_{t+h} \sim N(\mu, \sigma^2)$ for any horizon h and the subjective belief is $\ell_{t+h} \sim N(\mu_t^*, \sigma^2)$ for any horizon h. This means that the maximum possible $E\left[\left(\frac{E_t[f_{y,t+h}]}{Std_t(f_{y,t+h})}\right)^2\right]$ is the same for all h. Thus, even though these example papers use quarterly models and our empirics use annual forecasts, the maximum possible $E\left[\left(\frac{E_t[f_{y,t+4}]}{Std_t(f_{y,t+4})}\right)^2\right]$ for annual forecasts in these models is the same as the maximum possible $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ for quarterly forecasts, which is $\left(1-2\frac{g}{2-g}\right)^{-1/2} - 1$ for unrestricted y_{t+4} and $\frac{g}{2-g}$ for conditionally linear y_{t+4}.

low values of g, the agent's μ_t^* is close to an equally weighted average of all past ℓ_{t-j} which would be the true value of μ. We need a high value of g so that μ_t^* largely depends on only a few recent observations and noise in those observations can push μ_t^* far away from μ. With low g, the random noise in each ℓ_{t-j} tends to average out because we are spreading the weight almost equally over all of the observations. Thus, while the agent's beliefs are non-FIRE for any $g > 0$, the FESR tells us how far g needs to be from zero (i.e., the minimum deviation from FIRE).

In the right panel, we account for finite samples. The blue and red curves show the values of (Γ, g) for which there is at least a 5% chance of observing $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2$ greater than 0.40 over a 34 year period. This lowers the required gain to 0.26 for unrestricted y_{t+1} and 0.39 for conditionally linear y_{t+1} but these values are still far above what is typically studied in the literature. Appendix A.2 provides the details but the overall approach closely mirrors the extrapolation result. For example, like with extrapolation, the probability of observing $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2$ greater than 0.40 is determined by a weighted chi-squared distribution for conditionally linear y_{t+1}.

C.3. Diagnostic expectations

As our final example, we consider diagnostic expectations. We have some MA(∞) time series,

$$\ell_{t+1} = \sum_{j=0}^{\infty} \lambda_j \sigma \varepsilon_{t-j} + \sigma \varepsilon_{t+1}.$$

As in the example of adaptive expectations, any ARMA(p,q) time series has an MA(∞) representation and an AR(∞) representation. We simply choose the representation that is notationally most convenient. The agent's believed distribution is

$$\pi_t^*\left(\ell_{t+1}\right) = \pi_t\left(\ell_{t+1} - \theta\lambda_0\sigma\varepsilon_t\right), \tag{22}$$

meaning that her subjective expectation is

$$E_t^* [\ell_{t+1}] = E_t [\ell_{t+1}] + \theta\lambda_0\sigma\varepsilon_t \quad (23)$$

$$= E_t [\ell_{t+1}] + \theta (E_t [\ell_{t+1}] - E_{t-1} [\ell_{t+1}]) . \quad (24)$$

The value of λ_0 determines how informative the recent shock $\sigma\varepsilon_t$ is for the future ℓ_{t+1} and the parameter $\theta > 0$ controls the degree to which the agent overreacts to this information. For generality, we also consider underreaction, i.e., $\theta < 0$.

Note that the results of this section only require the assumptions of Section I. However, if we also want to satisfy conditions (i)-(ii) of Section II, then we can consider the version of diagnostic expectations of Maxted (2024), in which $E_{t-1}^* [\ell_{t+1}] = E_{t-1}^* [E_t^* [\ell_{t+1}]]$ and $\pi_t^* (\ell_{t+1}) = \pi_t (\ell_{t+1} - \theta\lambda_0\mathcal{I}_t)$ where $\mathcal{I}_t = \sum_{j=0}^{\infty} \kappa^j \sigma\varepsilon_{t-j}$. Condition (i) is always satisfied, and for any $\kappa > 0$ condition (ii) is satisfied. We can then consider arbitrarily small κ which allows us to have subjective $\pi_t^* (\ell_{t+1})$ arbitrarily close to $\pi_t (\ell_{t+1} - \theta\lambda_0\sigma\varepsilon_t)$.

Given subjective beliefs (22), we have the following bound.

Lemma 7. *For any y_{t+1}, we have $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \leq Var (S_{t+1}) = \left(1 - 2 (\theta\lambda_0)^2\right)^{-1/2} - 1$. We have a tighter bound of $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \leq Var\left(S_{t+1}^{(1)}\right) = (\theta\lambda_0)^2$ for any y_{t+1} that is conditionally linear with ℓ_{t+1}.*

As in the other lemmas, the bound only depends on a few parameters (θ, λ_0) and does not require knowing the full details of $\{\lambda_j\}_{j=0}^{\infty}$.

The blue and red curves in the left panel of Figure 6 show the values of $(|\lambda_0|, |\theta|)$ that would be needed to match the empirical FESR. The dashed black line at $|\theta| = 0$ shows the FIRE benchmark when there is no overreaction or underreaction. For objective informativeness $|\lambda_0|$ near 1 (i.e., when ε_t matters as much for ℓ_{t+1} as ε_{t+1}), the required over/underreaction magnitude $|\theta|$ is 0.50 for unrestricted y_{t+1} and 0.64 for conditionally linear y_{t+1}. As the objective informativeness declines, the required over/underreaction magnitude increases in inverse proportion.

As shown in the purple dots of the left panel, the commonly used value of $\theta \approx 0.90$

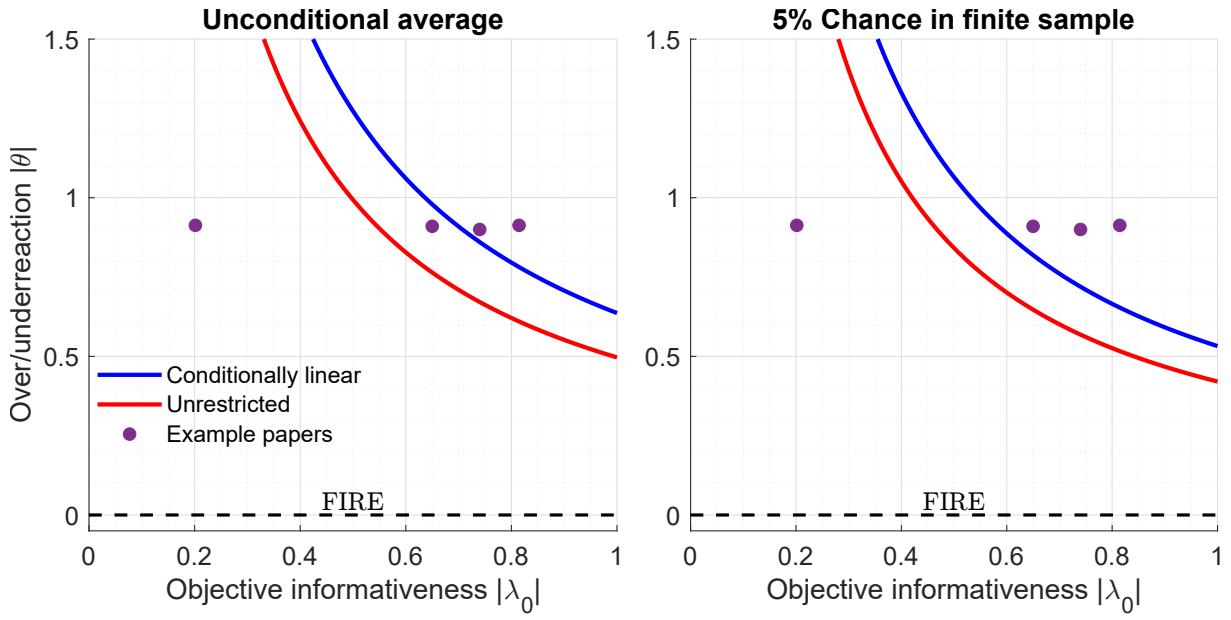


Figure 6. Over/under-reaction (parameters above curve satisfy bound). The left panel shows the minimum required deviation from FIRE to match the empirical FESR. For each absolute value $|\lambda_0|$ of the informativeness of the shock, the red curve shows the minimum required absolute over/underreaction parameter $|\theta|$in order to generate $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$ as large as the squared empirical FESR. The blue curve shows a higher required $|\theta|$ when we restrict y_{t+1} to be conditionally linear with ℓ_{t+1}. The black dashed line shows the benchmark of FIRE, where $\theta = 0$. The purple dots show example paper values for comparison. In the right panel, the red curve shows the minimum required $|\theta|$ for there to be a 5% chance of generating a finite sample $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2$ as large as the squared empirical FESR over a 34 period sample. The blue curve again shows a higher required $|\theta|$ when we restrict y_{t+1} to be conditionally linear with ℓ_{t+1}. The black dashed line and the purple dots are exactly the same as the left panel.

for diagnostic expectations is well above the required $|\theta|$ of 0.50 and 0.64.[17] So long as the objective informativeness $|\lambda_0|$ is not too low, a mechanism of diagnostic expectations with $\theta = 0.90$ satisfies both the unrestricted bound and the conditionally linear bound. The right panel of Figure 6 shows that when we only require a 5% chance of matching the empirical FESR in a finite sample, the minimum required $|\lambda_0|$ decreases to 0.47 for the unrestricted bound and 0.59 for the conditionally linear bound, assuming $\theta = 0.90$.[18] Thus, while diagnostic expectations is not the only mechanism that satisfies the bound, the degree to which it surpasses the bound is notable. Underreaction to the same degree (i.e.,

[17]The example papers are Bordalo et al. (2018), Bordalo et al. (2019), and Bordalo et al. (2026). The final paper appears as two dots in the figure, as the model studies diagnostic expectations of two separate processes.

[18]Appendix A.2 provides the details for the finite sample calculation.

$\theta = -0.90$) would also be just as successful in satisfying the bound.

V. Conclusion

Once a researcher departs from Full-information Rational Expectations (FIRE), there are infinitely many ways a belief mechanism could relate to a forecasted variable. We show that this "problem of arbitrariness" has an important consequence: commonly studied moments of forecast errors can be matched by any persistently biased belief mechanism applied to some latent component of the forecasted variable. This includes mechanisms arbitrarily close to FIRE.

We propose the forecast error Sharpe ratio (FESR) and a Hansen-Jagannathan bound for beliefs as a solution to this problem. Just as the original Hansen and Jagannathan (1991) bound allows researchers to reject stochastic discount factors without knowing how the SDF relates to observed assets, our bound allows researchers to reject belief mechanisms without knowing how the mechanism relates to the forecasted variable. Further, the empirical FESR provides a lower bound on the required deviation from FIRE.

For professional forecasts of yields, the empirical FESR requires a substantial deviation from FIRE. For Bayesian learning, the generalized Kalman gain must be at least 0.29. For extrapolation, the difference between the believed persistence and the actual persistence must be at least 0.40. For adaptive expectations, the gain must be at least 0.40. For over/underreaction, the agent must exaggerate or understate recent shocks by at least 50%. Of the example papers considered, extrapolation of an endogenous sentiment variable and diagnostic expectations with the commonly used overreaction parameter of 0.90 satisfy our bound. The FESR therefore distinguishes mechanisms that merely depart from FIRE from those that depart far enough to explain the empirical forecast errors.

References

Ai, Hengjie, 2010, Information quality and long-run risk: Asset pricing implications, *Journal of Finance* 65, 1333–1367.

Altı, Aydoğan, and Paul C. Tetlock, 2014, Biased beliefs, asset prices, and investment: A structural approach, *Journal of Finance* 69, 325–361.

Bachmann, Rüdiger, Giorgio Topa, and Wilbert van der Klaauw, eds., 2023, *Handbook of Economic Expectations* (Academic Press, London and San Diego, CA).

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

Bauer, Michael D., Carolin E. Pflueger, and Adi Sunderam, 2024, Perceptions about monetary policy, *Quarterly Journal of Economics* 139, 2227–2278.

Bhandari, Anmol, Jaroslav Borovička, and Paul Ho, 2025, Survey data and subjective beliefs in business cycle models, *Review of Economic Studies* 92, 1375–1437.

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

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer, 2020, Overreaction in macroeconomic expectations, *American Economic Review* 110, 2748–2782.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2018, Diagnostic expectations and credit cycles, *Journal of Finance* 73, 199–227.

Bordalo, Pedro, Nicola Gennaioli, Andrei Shleifer, and Stephen J. Terry, 2026, Real credit cycles, *American Economic Review* 116, 1274–1308.

Chen, Xiaohong, Lars Peter Hansen, and Peter G. Hansen, 2020, Robust identification of investor beliefs, *Proceedings of the National Academy of Sciences* 117, 33130–33140.

Chen, Xiaohong, Lars Peter Hansen, and Peter G. Hansen, 2024, Robust inference for moment condition models without rational expectations, *Journal of Econometrics* 243, 105653.

Chernov, Mikhail, and Philippe Mueller, 2012, The term structure of inflation expectations, *Journal of financial economics* 106, 367–394.

Cieslak, Anna, 2018, Short-rate expectations and unexpected returns in Treasury bonds, *Review of Financial Studies* 31, 3265–3306.

Cochrane, John H., and Lars Peter Hansen, 1992, Asset pricing explorations for macroeconomics, in Olivier Jean Blanchard, and Stanley Fischer, eds., *NBER Macroeconomics Annual 1992, Volume 7*, 115–165 (MIT Press, Cambridge, MA).

Coibion, Olivier, and Yuriy Gorodnichenko, 2015, Information rigidity and the expectations formation process: A simple framework and new facts, *American Economic Review* 105, 2644–2678.

Cui, Tingyue, Ricardo Delao, and Sean Myers, 2026, The subjective belief factor, *SSRN Working Paper* .

Farmer, Leland E., Emi Nakamura, and Jón Steinsson, 2024, Learning about the long run, *Journal of Political Economy* 132, 3334–3377.

Friedman, Benjamin M., 1980, Survey evidence on the "Rationality" of interest rate expectations, *Journal of Monetary Economics* 6, 453–465.

Gilchrist, Simon, and Masashi Saito, 2008, Expectations, asset prices, and monetary policy: The role of learning, in *Asset prices and monetary policy*, 45–102 (University of Chicago Press).

Greenwood, Robin, and Samuel G. Hanson, 2015, Waves in ship prices and investment, *Quarterly Journal of Economics* 130, 55–109.

Hansen, Lars Peter, John Heaton, and Erzo G. J. Luttmer, 1995, Econometric evaluation of asset pricing models, *Review of Financial Studies* 8, 237–274.

Hansen, Lars Peter, and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225–262.

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

Malmendier, Ulrike, and Stefan Nagel, 2016, Learning from inflation experiences, *Quarterly Journal of Economics* 131, 53–87.

Maxted, Peter, 2024, A macro-finance model with sentiment, *Review of Economic Studies* 91, 438–475.

McNees, Stephen K., 1978, The "Rationality" of economic forecasts, *American Economic Review* 68, 301–305, Papers and Proceedings.

Milani, Fabio, 2007, Expectations, learning and macroeconomic persistence, *Journal of Monetary Economics* 54, 2065–2082.

Mincer, Jacob A., and Victor Zarnowitz, 1969, The evaluation of economic forecasts, in Jacob A. Mincer, ed., *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, NBER Studies in Business Cycles, 3–46 (National Bureau of Economic Research, New York).

Orphanides, Athanasios, and John C. Williams, 2005, The decline of activist stabilization policy: Natural rate misperceptions, learning, and expectations, *Journal of Economic Dynamics and Control* 29, 1927–1950.

Pang, Hao, 2026, Forecast bias across horizons: Inflation expectations and the Treasury yields, *Journal of Financial Economics* 183, 104327.

Roussanov, Nikolai, Hongxun Ruan, and Yanhao Wei, 2021, Marketing mutual funds, *The Review of Financial Studies* 34, 3045–3094.

Rozsypal, Filip, and Kathrin Schlafmann, 2023, Overpersistence bias in individual income expectations and its aggregate implications, *American Economic Journal: Macroeconomics* 15, 331–371.

Sargent, Thomas J., 1993, *Bounded Rationality in Macroeconomics*, The Arne Ryde Memorial Lectures (Oxford University Press, Oxford).

Veronesi, Pietro, 2000, How does information quality affect stock returns?, *Journal of Finance* 55, 807–837.

Appendix

A. Additional analytical results

A.1. Matching conditional FESR means all moments can be matched

The paper primarily focuses on how the FESR can be used to reject belief mechanisms even when the Mincer-Zarnowitz coefficient, serial correlation of forecast errors, and Coibion-Gorodnichenko coefficient cannot reject these mechanisms. There is an interesting converse to this argument, which is that if a mechanism can match the empirical conditional FESR at each point in time, then there is no possible moment from the time series data of survey expectations or predictable forecast errors that can reject the mechanism.

Let $\hat{E}_t^* [\hat{y}_{t+1}]$ denote the empirical survey expectations and let $\hat{E}_t \left[\hat{f}_{t+1}\right]$ be statistical expectations of the future forecast error. These may be the true FIRE expectations or simply statistical expectations based on observable predictors of $\hat{f}_{t+1}$. Define the empirical shock as $\hat{\varepsilon}_{t+1} = \hat{f}_{t+1} - \hat{E}_t \left[\hat{f}_{t+1}\right]$.

Assume we have some belief mechanism $(\ell_{t+1}, \pi_t^* (\ell_{t+1}))$ where $\varepsilon_{t+1} = \ell_{t+1} - E_t [\ell_{t+1}]$. Our goal is to create a process y_{t+1} such that $E_t^* [y_{t+1}]$, and $E_t [f_{y,t+1}]$ perfectly match $\hat{E}_t^* [\hat{y}_{t+1}]$, and $\hat{E}_t \left[\hat{f}_{t+1}\right]$ at every point in time. Consider $y_{t+1} = a_t + b_t \ell_{t+1} + \eta_{t+1}$ where

$$\begin{aligned} b_t &= \frac{-\hat{E}_t \left[\hat{f}_{t+1}\right]}{E_t^* [\ell_{t+1}] - E_t [\ell_{t+1}]} \\ a_t &= \hat{E}_t^* [\hat{y}_{t+1}] - b_t E_t^* [\ell_{t+1}] . \end{aligned}$$

This gives $E_t [f_{y,t+1}] = b_t (E_t [\ell_{t+1}] - E_t^* [\ell_{t+1}])$. We are free to choose any η_{t+1} process that is mean zero and independent of ε_{t+1}. This y_{t+1} will satisfy our goal of exactly matching $\hat{E}_t^* [\hat{y}_{t+1}]$, and $\hat{E}_t \left[\hat{f}_{t+1}\right]$ at every point in time. The only condition we have to verify is that a choice for η_{t+1} exists, i.e., that there is a process for η_{t+1} with nonnegative variance.

We have that $Var_t (\eta_{t+1}) \geq 0$ is equivalent to $\hat{Var}_t (\hat{\varepsilon}_{t+1}) \geq b_t^2 Var_t (\varepsilon_{t+1})$ which is equivalent to $\frac{1}{\hat{Var}_t(\hat{\varepsilon}_{t+1})} \leq \frac{1}{b_t^2 Var_t(\varepsilon_{t+1})}$. Since $\hat{Var}_t (\hat{\varepsilon}_{t+1}) = \hat{Var}_t \left(\hat{f}_{t+1}\right)$ and $Var_t (\varepsilon_{t+1}) =$

$Var_t\left(f_{\ell,t+1}\right)$, the condition $\frac{1}{\hat{Var}_t(\hat{\varepsilon}_{t+1})} \leq \frac{1}{b_t^2 Var_t(\varepsilon_{t+1})}$ can be rearranged to

$$\frac{\hat{E}_t\left[\hat{f}_{t+1}\right]^2}{\hat{Var}_t\left(\hat{f}_{t+1}\right)} \leq \frac{\left(E_t^*\left[\ell_{t+1}\right] - E_t\left[\ell_{t+1}\right]\right)^2}{Var_t\left(f_{\ell,t+1}\right)}.$$

So, as long as the conditional squared FESR for ℓ_{t+1} is weakly greater than the empirical conditional squared FESR, this construction is possible and we can perfectly replicate the entire empirical time series.

If we allow for non-linear relationships, then the conditions become even weaker. Consider $y_{t+1} = a_t + b_t z_{t+1} + \eta_{t+1}$ for some $z_{t+1} = g_t\left(\ell_{t+1}\right)$ where

$$\begin{aligned} b_t &= \frac{-\hat{E}_t\left[\hat{f}_{t+1}\right]}{E_t^*\left[z_{t+1}\right] - E_t\left[z_{t+1}\right]} \\ a_t &= \hat{E}_t^*\left[\hat{y}_{t+1}\right] - b_t E_t^*\left[z_{t+1}\right]. \end{aligned}$$

Again, we are free to choose any η_{t+1} process we wish that is independent of $\varepsilon_{z,t+1} \equiv z_{t+1} - E_t\left[z_{t+1}\right]$. To ensure this construction is possible, we again just need that $\hat{Var}_t\left(\hat{\varepsilon}_{t+1}\right) \geq b_t^2 Var_t\left(\varepsilon_{z,t+1}\right)$, which is equivalent to

$$\frac{\hat{E}_t\left[\hat{f}_{t+1}\right]^2}{\hat{Var}_t\left(\hat{f}_{t+1}\right)} \leq \frac{\left(E_t^*\left[z_{t+1}\right] - E_t\left[z_{t+1}\right]\right)^2}{Var_t\left(f_{z,t+1}\right)}.$$

This directly connects to the discussion in Section IV.B.1 about non-linear relationships for learning problems. If $Var_t\left(S_{t+1}\right)$ is infinite, then we can always choose some threshold indicator $z_{t+1} = \mathbb{1}\left\{\ell_{t+1} > k\right\}$ with k sufficiently large that $\frac{\left(E_t^*[z_{t+1}] - E_t[z_{t+1}]\right)^2}{Var_t(f_{z,t+1})}$ exceeds $\frac{\hat{E}_t\left[\hat{f}_{t+1}\right]^2}{\hat{Var}_t\left(\hat{f}_{t+1}\right)}$.

A.2. Finite sample distributions for behavioral mechanisms

Rather than studying each of the three behavioral mechanisms individually, we consider a generalized behavioral mechanism that nests the mechanisms studied in Section IV.C. The FIRE distribution is $\ell_{t+1} \sim N\left(\mu_t, \sigma^2\right)$ and the believed distribution is $\ell_{t+1} \sim N\left(\mu_t^*, \sigma^2\right)$ where $\mu_t^* - \mu_t = \alpha z_t$ and z_t is a normally distributed AR(1) process with persistence ρ and variance $\frac{\sigma^2}{1-\rho^2}$. For extrapolation, ρ would be the FIRE persistence of ℓ_t and α would be

$\rho^* - \rho$. For adaptive expectations, ρ would be $1-(1-\Gamma)\,g$ and α would be $(1-\Gamma)\,g$. For diagnostic expectations, ρ would be 0 and α would be $\theta\lambda_0$.

In this generalized behavioral mechanism, $S_{t+1} = \frac{\pi_t^*(\ell_{t+1})}{\pi_t(\ell_{t+1})} = \exp\left(-\frac{1}{2}\left(\frac{\alpha z_t}{\sigma}\right)^2 + \frac{\alpha z_t}{\sigma}\varepsilon_{t+1}\right)$ is conditionally log-normal. This means that $Var_t\left(S_{t+1}\right) = \exp\left(\left(\frac{\alpha z_t}{\sigma}\right)^2\right) - 1$. We also have $S_{t+1}^{(1)} = 1 + \frac{\alpha z_t}{\sigma}\varepsilon_{t+1}$, so $Var_t\left(S_{t+1}^{(1)}\right) = \frac{\alpha^2 z_t^2}{\sigma^2}$. Following the steps from Section III, the belief-based Hansen-Jagannathan bound holds period by period. Specifically, we know that $E_t^*\left[y_{t+1}\right] = E_t\left[S_{t+1}y_{t+1}\right]$, so

$$\frac{E_t\left[f_{y,t+1}\right]}{Std_t\left(f_{y,t+1}\right)} \;=\; \frac{E_t\left[y_{t+1}\right] - E_t^*\left[y_{t+1}\right]}{Std_t\left(f_{y,t+1}\right)} = \frac{Cov_t\left(1-S_{t+1}, y_{t+1}\right)}{Std_t\left(f_{y,t+1}\right)} \leq Std_t\left(S_{t+1}\right)$$

which implies $\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2 \leq Var_t\left(S_{t+1}\right) = \exp\left(\left(\frac{\alpha z_t}{\sigma}\right)^2\right) - 1$. We can repeat the process to show $\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2 \leq Var_t\left(S_{t+1}^{(1)}\right) = \frac{\alpha^2 z_t^2}{\sigma^2}$ for conditionally linear y_{t+1}.

For conditionally linear y_{t+1}, the probability of $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2$ greater than 0.40 is at most the probability of $\frac{1}{T}\sum_{t=1}^{T}\alpha^2 z_t^2$ greater than 0.40, which is determined by a weighted chi-squared distribution. The probability is $P\left(\sum_{j=1}^{T}\lambda_j X_j^2 \geq 0.4T\frac{\left(1-\rho^2\right)}{\alpha^2}\right)$ where X_j^2 are chi-squared variables and λ_j are the eigenvalues of the matrix

$$\begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{bmatrix}.$$

Thus, for any given (ρ, α) values, we can calculate an upper bound for the probability of $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2$ greater than 0.40.

For unrestricted y_{t+1}, the probability of $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2$ greater than 0.40 is at most the probability of $\frac{1}{T}\sum_{t=1}^{T}\left[\exp\left(\left(\frac{\alpha z_t}{\sigma}\right)^2\right) - 1\right]$ greater than 0.40. Because of the nonlinearity, $\frac{1}{T}\sum_{t=1}^{T}\left[\exp\left(\left(\frac{\alpha z_t}{\sigma}\right)^2\right) - 1\right]$ does not have a simple closed-form distribution but it can be easily estimated with Monte Carlo simulations. So, for any given (ρ, α) values, we can calculate an upper bound for the probability of $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2$ greater than 0.40 for

all possible y_{t+1} (i.e., we can do a single estimation, rather than individually evaluating every possible way that y_{t+1} could be related to ℓ_{t+1}).

B. Proofs

B.1. Proof of Proposition 1

We provide a constructive proof that covers all possible values for the targets $\hat{\beta}^{MZ}, \hat{\beta}^{AC}, \hat{\beta}^{CG}$, and $\hat{\gamma}_j$ for $j \leq n$. Importantly, this is simply one solution for $y_{t+1} = a_t + b_t \ell_{t+1}$ that matches these moments. There may be infinitely many solutions. Further, we choose to provide a single construction method that can match all possible targets, including extreme cases, rather than different construction methods for different ranges of target values. Much of the complexity of the proof is to accommodate all possible targets, particularly $\hat{\beta}^{AC}$ arbitrarily close to 1 or -1. For a single set of target values, there may be much simpler solutions for the a_t and b_t processes.

Let $\hat{q}_t \equiv \hat{E}_t^*[\hat{y}_{t+1}]$ denote the subjective expectation. WLOG, we normalize $Var(\hat{y}_t)$ to 1 since this does not affect any of the three target regression coefficients or the target autocorrelations. Let

$$\widehat{C}_j \equiv Cov\left(\begin{pmatrix} \hat{f}_{t+1} \\ \hat{q}_t \end{pmatrix}, \begin{pmatrix} \hat{f}_{t-j+1} \\ \hat{q}_{t-j} \end{pmatrix}\right)$$

denote the empirical covariance of the forecast errors and subjective expectations and let $\iota \equiv (1\ 1)'$. We then have that

$$\begin{aligned} (\widehat{C}_1)_{11} &= \hat{\beta}^{AC}(\widehat{C}_0)_{11}, \\ (\widehat{C}_0)_{12} &= \hat{\beta}^{MZ}(\widehat{C}_0)_{22}, \\ \iota'\widehat{C}_j\iota &= \hat{\gamma}_j, \qquad j = 0, \ldots, n. \end{aligned}$$

We also have that the block covariance matrix generated by $\widehat{C}_0, \ldots, \widehat{C}_n$ is strictly positive definite.

Our first step is to change the construction of a_t and b_t to a construction of q_t and b_t. Given any admissible q_t and b_t, set $a_t \equiv q_t - b_t E_t^*[\ell_{t+1}]$. This guarantees that $E_t^*[y_{t+1}] = a_t + b_t E_t^*[\ell_{t+1}] = q_t$. Further, since $\ell_{t+1} = E_t[\ell_{t+1}] + \varepsilon_{t+1}$ and $\delta_t = E_t^*[\ell_{t+1}] - E_t[\ell_{t+1}]$, the forecast error is

$$f_{t+1} \equiv y_{t+1} - E_t^*[y_{t+1}] = b_t\left(\varepsilon_{t+1} - \delta_t\right).$$

We will construct $q_t = q_{t-1}^B + q_t^C$, where q_{t-1}^B is realized at time $t-1$ and q_t^C is ζ_t times a variable realized at time $t-1$ for an iid mean-zero, unit variance ζ_t. This means that $E_{t-1}^*\left[E_t^*\left[y_{t+1}\right]\right] = E_{t-1}^*\left[q_t\right] = q_{t-1}^B$ since ζ_t is independent of ℓ_t. So, the forecast revision for the Coibion-Gorodnichenko coefficient is

$$E_t^*\left[y_{t+1}\right] - E_{t-1}^*\left[E_t^*\left[y_{t+1}\right]\right] = q_t^C.$$

The second step is to set aside a small part of the contemporaneous variance for matching the Coibion-Gorodnichenko coefficient. Choose $\omega > 0$ and define

$$\Sigma^C \equiv \omega \begin{pmatrix} 1 + 2\left(\hat{\beta}^{CG}\right)^2 & \hat{\beta}^{CG} \\ \hat{\beta}^{CG} & 1 \end{pmatrix}.$$

Since $\det(\Sigma^C) = \left[1 + \left(\hat{\beta}^{CG}\right)^2\right]\omega^2 > 0$, we know that Σ^C is positive definite. Let

$$\begin{aligned} C_0^B &\equiv \widehat{C}_0 + \omega \begin{pmatrix} -2\left[1 + \left(\hat{\beta}^{CG}\right)^2 + \hat{\beta}^{MZ}\right] & \hat{\beta}^{MZ} - \hat{\beta}^{CG} \\ \hat{\beta}^{MZ} - \hat{\beta}^{CG} & 0 \end{pmatrix}, \\ C_1^B &\equiv \widehat{C}_1 + \omega\hat{\beta}^{AC}\left(1 + 2\hat{\beta}^{MZ}\right)\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \\ C_j^B &\equiv \widehat{C}_j \text{ for } j = 2, \ldots, n. \end{aligned}$$

This ensures that

$$\begin{aligned}(C_1^B)_{11} &= \hat{\beta}^{AC}\left\{(C_0^B)_{11} + \left[1 + 2\left(\hat{\beta}^{CG}\right)^2\right]\omega\right\},\\ (C_0^B)_{12} + \hat{\beta}^{CG}\omega &= \hat{\beta}^{MZ}\left[(C_0^B)_{22} + \omega\right],\\ \iota' C_j^B \iota &= \hat{\gamma}_j \text{ for } j = 1, \ldots, n.\end{aligned}$$

Thus, if the B-component has covariance matrices C_j^B and the C-component has contemporaneous covariance Σ^C, the desired AC, MZ, and autocorrelation targets will all hold. Every adjustment above is proportional to ω. Therefore, because the empirical covariance matrix is strictly positive definite, the block covariance matrix associated with $C_0^B, \ldots, C_n^B$ remains strictly positive definite for sufficiently small ω.

Next, we convert these finite-lag covariances into a finite collection of response profiles. Choose a sufficiently large integer $L > n$. We know that there exists a positive-definite $2L \times 2L$ matrix $P = [P_{kh}]_{k,h=0}^{L-1}$ partitioned into 2×2 blocks, such that

$$\sum_{k=j}^{L-1} P_{k,k-j} = C_j^B \text{ for } j = 0, \ldots, n.$$

Specifically, we can use $P_{k,k-j} = \frac{C_j^B}{L-j}$ for $j = 0, ..., n$. The positive-definite completion theorem for band matrices then supplies the unspecified blocks of P. As L becomes large, every fully specified principal block converges, after multiplication by L, to a principal block of the strictly positive-definite covariance matrix associated with $C_0^B, \ldots, C_n^B$. By construction, summing along the j-th block diagonal gives C_j^B.

We then show that we can implement these response profiles. Fix a tail threshold K_B and let $p_B \equiv \text{Prob}\left(|\varepsilon_t| > K_B\right)$ be the probability that the shock exceeds the threshold in absolute value. Then let $J_t \equiv \frac{\mathbf{1}\{|\varepsilon_t| > K_B\}}{\sqrt{p_B}}$ and $D_t \equiv \frac{J_t}{\varepsilon_t}$, where D_t is set to 0 whenever J_t is 0. For a lag date $t - r$ and $k = 0, \ldots, L-1$, define

$$Y_{t-r+k+2,k} \equiv \frac{D_{t-r}}{\lambda_{k+1}}\left(\delta_{t-r+k+1} - \varepsilon_{t-r+k+2}\right).$$

The moving-average representation for the bias gives

$$\delta_{t-r+k+1} \;=\; \lambda_{k+1}\varepsilon_{t-r} + V_{t-r+k+1,k},$$

where $V_{t-r+k+1,k}$ is square-integrable and independent of ε_{t-r}. Consequently,

$$Y_{t-r+k+2,k} \;=\; J_{t-r} + \frac{D_{t-r}}{\lambda_{k+1}}\left(V_{t-r+k+1,k} - \varepsilon_{t-r+k+2}\right).$$

Suppressing the dependence of J, D, and Y on K_B, fix L. Since $|D_{t-r}| \leq J_{t-r}/K_B$, for every $k = 0, \ldots, L-1$, we have that

$$E\left[(Y_{t-r+k+2,k} - J_{t-r})^2\right] \;\leq\; \frac{E[(V_{t-r+k+1,k} - \varepsilon_{t-r+k+2})^2]}{\lambda_{k+1}^2 K_B^2}$$

where the RHS converges to 0 as $K_B \to \infty$. Thus, holding L fixed, the $\mathcal{L}^2$ distance between $Y_{t-r+k+2,k}$ and J_{t-r} converges to zero as $K_B \to \infty$ for every $k = 0, \ldots, L-1$. Let

$$H_B \;\equiv\; E\left[\begin{pmatrix} Y_{t-r+2,0} \\ J_{t-r} \\ \vdots \\ Y_{t-r+L+1,L-1} \\ J_{t-r} \end{pmatrix}\begin{pmatrix} Y_{t-r+2,0} \\ J_{t-r} \\ \vdots \\ Y_{t-r+L+1,L-1} \\ J_{t-r} \end{pmatrix}'\right].$$

As $K_B \to \infty$, every entry of H_B converges to one. For sufficiently large K_B, the matrix defined entry by entry by $\widetilde{P}_{ij} \equiv \frac{P_{ij}}{(H_B)_{ij}}$ is therefore positive definite. Factor this matrix as $\widetilde{P} = XX'$ and denote the m-th column of X as $(c_{m0}\ q_{m0}\ \ldots\ c_{m,L-1}\ q_{m,L-1})'$.

Let $\{\xi_{t,m}\}$ be independent Rademacher variables, independent of the belief mechanism. Define

$$\begin{aligned} b_{t-1}^B \;&\equiv\; -\sum_{m=1}^{2L}\sum_{k=0}^{L-1}\frac{c_{mk}}{\lambda_{k+1}}\xi_{t-k-1,m}D_{t-k-1}, \\ q_{t-1}^B \;&\equiv\; \sum_{m=1}^{2L}\sum_{k=0}^{L-1} q_{mk}\xi_{t-k-1,m}J_{t-k-1}. \end{aligned}$$

The forecast error generated by this component is

$$f_{t+1}^B \;=\; b_{t-1}^B\left(\varepsilon_{t+1}-\delta_t\right)=\sum_{m=1}^{2L}\sum_{k=0}^{L-1}c_{mk}\xi_{t-k-1,m}Y_{t+1,k}.$$

The independent Rademacher variables eliminate all covariance terms coming from different lag dates or different columns. For a given lag date, the second-moment matrix is $H_B\circ XX'=P$, where $\circ$ denotes entry-by-entry multiplication. This ensures that

$$Cov\left(\begin{pmatrix}f_{t+1}^B\\ q_{t-1}^B\end{pmatrix},\begin{pmatrix}f_{t-j+1}^B\\ q_{t-j-1}^B\end{pmatrix}\right) \;=\; \sum_{k=j}^{L-1}P_{k,k-j}=C_j^B$$

for $j=0,\ldots,n$. Also note that q_{t-1}^B is realized at time $t-1$, so this component generates no forecast revision. This means we can achieve the desired covariance C_j^B for all $j\le n$.

The last remaining challenge is to construct the component that matches $\hat{\beta}^{CG}$. Let $I_t\equiv\mathbf{1}\{|\varepsilon_t|>K_C\}$, $U_{t+1}\equiv I_{t-1}\left(\varepsilon_{t+1}-\delta_t\right)$, and $Z_{t-1}\equiv I_{t-1}\varepsilon_{t-1}$. Since

$$\varepsilon_{t+1}-\delta_t \;=\; -\lambda_1\varepsilon_{t-1}+\widetilde{V}_{t+1},$$

where $\widetilde{V}_{t+1}$ is square-integrable and independent of ε_{t-1}, tail observations make U_{t+1} arbitrarily close to $-\lambda_1 Z_{t-1}$. Specifically, we have that $\frac{(E[U_{t+1}Z_{t-1}])^2}{E[U_{t+1}^2]E[Z_{t-1}^2]}$ converges to 1 as $K_c\to\infty$. This means we can choose K_C sufficiently large so that $\frac{(E[U_{t+1}Z_{t-1}])^2}{E[U_{t+1}^2]E[Z_{t-1}^2]}>\frac{\left(\hat{\beta}^{CG}\right)^2}{1+2\left(\hat{\beta}^{CG}\right)^2}$ since the RHS is strictly less than 1. Then, define $h_1\equiv\left(\frac{\omega[1+2(\hat{\beta}^{CG})^2]}{E[U_{t+1}^2]}\right)^{0.5}$, $h_2\equiv\frac{\hat{\beta}^{CG}\omega}{h_1E[U_{t+1}Z_{t-1}]}$, and $h_3\equiv\left(\omega-h_2^2E[Z_{t-1}^2]\right)^{0.5}$. We can always choose K_C so that h_3 is real. Let $\{\zeta_t\}$ and $\{\nu_t\}$ be independent mean-zero, unit-variance shocks, independent of the belief mechanism and of the Rademacher variables. Define

$$\begin{aligned}b_t^C &\equiv \zeta_t h_1 I_{t-1},\\ q_t^C &\equiv \zeta_t\left(h_2 I_{t-1}\varepsilon_{t-1}+h_3\nu_{t-1}\right).\end{aligned}$$

Then

$$f_{t+1}^C \;=\; b_t^C(\varepsilon_{t+1}-\delta_t)=\zeta_t h_1 U_{t+1},$$

and direct calculation gives

$$\begin{aligned} Var\left(f_{t+1}^{C}\right) &= [1+2\left(\hat{\beta}^{CG}\right)^{2}]\omega, \\ Cov\left(f_{t+1}^{C}, q_{t}^{C}\right) &= \hat{\beta}^{CG}\omega, \\ Var\left(q_{t}^{C}\right) &= \omega. \end{aligned}$$

The mean-zero shock ζ_t makes all positive-lag covariances of this component zero and makes this component uncorrelated with the B-component. Moreover, $E_{t-1}^{*}\left[q_{t}^{C}\right] = 0$, so q_t^C is exactly the forecast revision generated by this component.

To review and combine the results, set

$$\begin{aligned} b_t &\equiv b_{t-1}^{B} + b_{t}^{C}, \\ q_t &\equiv q_{t-1}^{B} + q_{t}^{C}, \end{aligned}$$

and choose

$$a_t \equiv q_t - b_t E_t^{*}[\ell_{t+1}].$$

Then,

$$\begin{aligned} E_t^{*}[y_{t+1}] &= a_t + b_t E_t^{*}[\ell_{t+1}] = q_t \\ y_{t+1} - E_t^{*}[y_{t+1}] &= b_t\left(\varepsilon_{t+1} - \delta_t\right) = f_{t+1}^{B} + f_{t+1}^{C} \\ E_t^{*}[y_{t+1}] - E_{t-1}^{*}[E_t^{*}[y_{t+1}]] &= q_t^{C}. \end{aligned}$$

The equations defining C_0^B, C_1^B, and Σ^C ensure

$$\begin{aligned} \beta^{AC} &= \frac{(C_1^B)_{11}}{(C_0^B)_{11} + \left[1+2\left(\hat{\beta}^{CG}\right)^{2}\right]\omega} = \hat{\beta}^{AC} \\ \beta^{MZ} &= \frac{(C_0^B)_{12} + \hat{\beta}^{CG}\omega}{(C_0^B)_{22} + \omega} = \hat{\beta}^{MZ} \\ \beta^{CG} &= \frac{\hat{\beta}^{CG}\omega}{\omega} = \hat{\beta}^{CG}. \end{aligned}$$

Finally,

$$Corr\left(y_t, y_{t-j}\right) \;=\; \iota' C_j^B \iota = \hat{\gamma}_j \text{ for } j \leq n.$$

For the fixed thresholds K_B and K_C, the process $\{D_t\}$ is bounded and J_t has unit second moment. The constructed processes are finite sums of stationary square-integrable variables and independent finite-variance auxiliary shocks. Therefore the forecast, forecast error, forecast revision, and y_t have finite, time-invariant second moments, so every covariance and regression coefficient above is well defined.

B.2. Proofs for lemmas

Proof of Lemma 1: Section III shows the entire proof for $Var\left(S_{t+1}\right) \geq E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right]$. The only remaining step is to show that $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \geq \left(\frac{E[f_{y,t+1}]}{Std(f_{y,t+1})}\right)^2$. We have that

$$\begin{aligned} E\left[f_{y,t+1}\right]^2 &= E\left[E_t\left[f_{y,t+1}\right]\right]^2 = E\left[\frac{E_t\left[f_{y,t+1}\right]}{Std_t\left(f_{y,t+1}\right)} Std_t\left(f_{y,t+1}\right)\right]^2 \\ &\leq E\left[\left(\frac{E_t\left[f_{y,t+1}\right]}{Std_t\left(f_{y,t+1}\right)}\right)^2\right] E\left[Std_t\left(f_{y,t+1}\right)^2\right] \\ &\leq E\left[\left(\frac{E_t\left[f_{y,t+1}\right]}{Std_t\left(f_{y,t+1}\right)}\right)^2\right] Std\left(f_{y,t+1}\right)^2 \end{aligned}$$

where the second line uses the Cauchy-Schwarz inequality and the third line uses the law of total variance $Std\left(f_{y,t+1}\right)^2 \geq E\left[Std_t\left(f_{y,t+1}\right)^2\right]$. This implies $\frac{E[f_{y,t+1}]^2}{Std(f_{y,t+1})^2} \leq E\left[\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right]^2$.

Proof of Lemma 2: Lemma 2 is a special case of Lemma 3, so we defer to the proof of Lemma 3.

Proof of Lemma 3: Equation (12) ensures that $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \leq Var\left(S_{t+1}^{(1)}\right) = E\left[\left(\frac{E_t\left[f_{\ell,t+1}\right]}{Std_t\left(f_{\ell,t+1}\right)}\right)^2\right]$. In other words, instead of worrying about all possible conditionally linear y_{t+1}, we just need to prove that forecast errors for ℓ_{t+1} satisfy $E\left[\left(\frac{E_t\left[f_{\ell,t+1}\right]}{Std_t\left(f_{\ell,t+1}\right)}\right)^2\right] \leq$

$\left(\frac{\bar{K}}{1-\bar{K}} - \bar{\phi}^2\right)\frac{1}{\bar{\phi}^2+1}$. Since $Cov\left(Var_t\left(\ell_{t+1}\right), \frac{(E_t[\ell_{t+1}]-E_t^*[\ell_{t+1}])^2}{Var_t(\ell_{t+1})}\right) \geq 0$, we know that $E\left[\left(\frac{E_t\left[f_{\ell,t+1}\right]}{Std_t\left(f_{\ell,t+1}\right)}\right)^2\right] \leq \frac{E\left[E_t\left[f_{\ell,t+1}\right]^2\right]}{E\left[Var_t\left(f_{\ell,t+1}\right)\right]}$, so we will prove that $\frac{E\left[E_t\left[f_{\ell,t+1}\right]^2\right]}{E\left[Var_t\left(f_{\ell,t+1}\right)\right]} = \left(\frac{\bar{K}}{1-\bar{K}} - \bar{\phi}^2\right)\frac{1}{\bar{\phi}^2+1}$.

Let $\delta_t \equiv E_t^*\left[\ell_{t+1}\right] - E_t\left[\ell_{t+1}\right]$. Since the observation noise ε_{t+1} and the shocks to α_{t+1} are uncorrelated with each other and with δ_t, we know that $Var\left(f_{\ell,t+1}\right) = \sigma^2 + E\left[\delta_t^2\right] + c'\bar{Q}c$ and $Cov\left(f_{\ell,t+1}, f_{\ell,t+1} - \varepsilon_{t+1}\right) = E\left[\delta_t^2\right] + c'\bar{Q}c$. So, the generalized Kalman gain is

$$\bar{K} = \frac{E\left[\delta_t^2\right] + c'\bar{Q}c}{\sigma^2 + E\left[\delta_t^2\right] + c'\bar{Q}c}.$$

Given that the generalized signal to noise ratio is $\bar{\phi} = \frac{\sqrt{c\bar{Q}c'}}{\sigma}$, we have that

$$\frac{\bar{K}}{1-\bar{K}} = \bar{\phi}^2 + \frac{E\left[\delta_t^2\right]}{\sigma^2}$$

which means

$$\begin{aligned}\frac{E\left[E_t\left[f_{\ell,t+1}\right]^2\right]}{E\left[Var_t\left(f_{\ell,t+1}\right)\right]} &= \frac{E\left[\delta_t^2\right]}{\sigma^2 + c\bar{Q}c'} = \left(\frac{\bar{K}}{1-\bar{K}} - \bar{\phi}^2\right)\frac{\sigma^2}{\sigma^2 + c'\bar{Q}c} \\ &= \left(\frac{\bar{K}}{1-\bar{K}} - \bar{\phi}^2\right)\frac{1}{\bar{\phi}^2+1}.\end{aligned}$$

Proof of Lemma 4: Following the steps from Section III, the belief-based Hansen-Jagannathan bound holds period by period. Specifically, for conditionally linear y_{t+1}, we know that $E_t^*\left[y_{t+1}\right] = E_t\left[S_{t+1}^{(1)} y_{t+1}\right]$, which means

$$\frac{E_t\left[f_{y,t+1}\right]}{Std_t\left(f_{y,t+1}\right)} = \frac{E_t\left[y_{t+1}\right] - E_t^*\left[y_{t+1}\right]}{Std_t\left(f_{y,t+1}\right)} = \frac{Cov_t\left(1 - S_{t+1}^{(1)}, y_{t+1}\right)}{Std_t\left(f_{y,t+1}\right)} \leq Std_t\left(S_{t+1}^{(1)}\right).$$

So, for any fixed t, $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \leq E\left[Var_t\left(S_{t+1}^{(1)}\right)\right] = \frac{E\left[\delta_t^2\right]}{c'Qc+\sigma^2}$ where $\delta_t = E_t^*\left[\ell_{t+1}\right] - E_t\left[\ell_{t+1}\right]$ even when the distribution of δ_t changes depending on t (i.e., we are not in a steady-state distribution and the distribution of δ_1 differs from the distribution of δ_2 as the agent learns over the finite sample). Because the agent's prior is sufficiently diffuse that she is not overconfident, we have that $E\left[E_t^*\left[f_{\ell,t+1}^2\right]\right] \geq E\left[f_{\ell,t+1}^2\right]$ which means $c'\bar{P}_t c + \sigma^2 \geq c'Qc + \sigma^2 + E\left[\delta_t^2\right]$. Combining these two inequalities, we have $E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \leq \frac{c'\left(\bar{P}_t - Q\right)c}{c'Qc+\sigma^2}$.

From the definitions of K_t, ϕ, and $\bar{K}_T$, we have that

$$\begin{aligned}\left(\frac{\bar{K}_T}{1-\bar{K}_T}-\phi^2\right)\frac{1}{\phi^2+1} &= \left(\left[\frac{1}{T}\sum_{t=1}^{T}\frac{K_t}{1-K_t}\right]-\phi^2\right)\frac{1}{\phi^2+1} \\ &= \left(\frac{1}{T}\sum_{t=1}^{T}\frac{c'\left(\bar{P}_t-Q\right)c}{\sigma^2}\right)\frac{1}{\phi^2+1} \\ &= \frac{1}{T}\sum_{t=1}^{T}\frac{c'\left(\bar{P}_t-Q\right)c}{c'Qc+\sigma^2}.\end{aligned}$$

This means that $\frac{1}{T}\sum_{t=1}^{T}E\left[\left(\frac{E_t[f_{y,t+1}]}{Std_t(f_{y,t+1})}\right)^2\right] \leq \frac{1}{T}\sum_{t=1}^{T}\frac{c'\left(\bar{P}_t-Q\right)c}{c'Qc+\sigma^2} = \left(\frac{\bar{K}_T}{1-\bar{K}_T}-\phi^2\right)\frac{1}{\phi^2+1}$.

Proof of Lemmas 5, 6, and 7: We consider a generalized behavioral mechanism that nests the mechanisms studied in Lemmas 5-7. The FIRE distribution is $\ell_{t+1}\sim N(\mu_t,\sigma^2)$ and the believed distribution is $\ell_{t+1}\sim N(\mu_t^*,\sigma^2)$ where $\mu_t^*-\mu_t=\alpha z_t$ and z_t is a normally distributed AR(1) process with persistence ρ and variance $\frac{\sigma^2}{1-\rho^2}$. Extrapolation translates to ρ being the FIRE persistence of ℓ_t and $\alpha=\rho^*-\rho$. Adaptive expectations translates to $\rho=1-(1-\Gamma)g$ and $\alpha=(1-\Gamma)g$. Diagnostic expectations translates to $\rho=0$ and $\alpha=\theta\lambda_0$.

In this generalized setting, $S_{t+1}=\frac{\pi_t^*(\ell_{t+1})}{\pi_t(\ell_{t+1})}=\exp\left(-\frac{1}{2}\left(\frac{\alpha z_t}{\sigma}\right)^2+\frac{\alpha z_t}{\sigma}\varepsilon_{t+1}\right)$ which is conditionally log-normal. This means that $Var_t(S_{t+1})=\exp\left(\left(\frac{\alpha z_t}{\sigma}\right)^2\right)-1$. The unconditional variance is then $Var(S_{t+1})=E[Var_t(S_{t+1})]=\left(1-2\frac{\alpha^2}{1-\rho^2}\right)^{-1/2}-1$. Similarly, we have $S_{t+1}^{(1)}=1+\frac{\alpha z_t}{\sigma}\varepsilon_{t+1}$ and $Var\left(S_{t+1}^{(1)}\right)=\frac{\alpha^2}{1-\rho^2}$.