Under- and Overreaction in Yield Curve Expectations*

Chen Wang†

October 5, 2021

Abstract

I document a robust pattern in how Treasury market participants' yield curve expectations respond to new information: forecasts for short-term rates underreact to news while forecasts for long-term rates overreact. I propose a new explanation of this based on "autocorrelation averaging," whereby, due to limited processing capacity, forecasters' estimate of the autocorrelation of a given process is biased toward the average autocorrelation of all related processes. Consistent with this view, forecasters *overestimate* the autocorrelation of the less persistent term-premium component of interest rates and *underestimate* the autocorrelation of the more persistent short-rate component; a calibrated model quantitatively matches the documented pattern of misreaction. Moreover, banks' allocations to Treasuries vary *positively* with their expectations of bond returns and misreaction proxies can strongly predict future short- and long-term bond returns, respectively.

*I am deeply grateful to the members of my committee—Nicholas Barberis, Stefano Giglio, Kelly Shue, and Tobias Moskowitz—for their invaluable guidance. I also thank James Choi, Shane Corwin, Zhi Da, Eduardo Dávila, Xavier Gabaix, Paul Goldsmith-Pinkham, Jon Ingersoll, Lawrence Jin, Bryan Kelly, Ye Li, Yueran Ma, Cameron Peng, Michael Sockin, Ken Singleton, Neng Wang, and seminar participants at Cheung Kong Graduate School of Business, Chinese University of Hong Kong, Cornerstone Research, Hong Kong University of Science and Technology, Michigan Ross, National University of Singapore, Notre Dame, University of Florida, University of Hong Kong, Yale SOM, and WFA Annual Meeting 2021 for helpful discussions and comments. I acknowledge financial support from a Whitebox Advisors research grant from the International Center for Finance at Yale School of Management. All errors are my own. The Internet Appendix can be found here. First version: September 30, 2019.

†Mendoza College of Business, University of Notre Dame. E-mail: chen.wang@nd.edu

1 Introduction

The US Treasury bond market is uniquely situated at the intersection of financial markets and the macroeconomy: the short-term interest rate is an important monetary policy instrument for achieving macro-finance stability; and the yield curve is a fundamental building block for pricing numerous financial assets. Investors' beliefs in this market, in principle, reveal their perceptions of the economy and bond risk premia. For decades, the predominant approach to modeling beliefs has been the rational expectations (RE) framework. Under the canonical full-information rational expectations (FIRE), investors instantaneously incorporate new information in their yield curve expectations, and their beliefs can be accurately recovered from observed yields, leaving little room for studying actual belief formation.

However, a growing body of recent research—using survey forecasts made by professional economists, financial analysts, and individual households—has documented systematic departures from rational expectations. These departures, in the form of predictable forecast errors, show that people's beliefs appear to underreact to news for some variables (e.g., many macroeconomic and near-term firm fundamentals), while overreacting for others (e.g., asset prices and long-term firm fundamentals). Researchers have proposed disparate explanations for belief misreaction in different settings. Unfortunately, these explanations seem to offer unclear guidance on how (or if) beliefs in Treasury yields deviate from rational expectations—the Treasury yield curve spans both the short and long term, and it relates to both asset prices and macroeconomic variables. Hence, yield curve expectations, in addition to informing us of people's expectations of monetary policy and risk compensation, present a unique setting to answer key questions concerning belief formation: How do people react to new information in an integrated market with a term structure? If belief misreaction changes along the yield curve, can we offer a unified explanation?

In this paper, I aim to answer these questions by studying measured beliefs from professional forecasts of interest rates across the entire yield curve and examining the term structure of under- and overreaction to new information. I document a novel pattern that, when updating beliefs about future interest rates, professional forecasters react very differently across maturities: they underreact for short-maturity and overreact for long-maturity interest rates. Specifically, I apply a methodology developed by Coibion and Gorodnichenko (2015, CG)—which assesses rationality by investigating the predictability of forecast errors (FE) from forecast revisions (FR)—to quantify misreaction to information for individual- and consensus-level forecasts. In particular, the FE-on-FR regression coefficients, as a function of maturity, are downward-sloping and cross zero at around the two-year maturity. Furthermore, this downward-sloping term structure of under- and overreaction is evident in both the individual and consensus forecasts, indicating a similar departure from rationality at both levels; it is robust to a battery of robustness checks concerning issues such as small sample bias and measurement errors. These findings present a challenge for existing models of expectation formation that deviate from full-information rational expectations. In fact, neither rational Bayesian learning nor any of the commonly used models of biased beliefs—such as sticky, extrapolative, or diagnostic expectations—can, in their standard formulations, generate the documented pattern[12].

To jointly explain this pattern of under- and overreaction, I propose a simple bounded-rationality framework based on "autocorrelation averaging." Investors, exposed to many time series in real time, may not have the cognitive processing capacity to learn the true autocorrelation of each one.[3] Instead, when forecasting, they may use something closer to an average of the true autocorrelations of the series that they are forecasting. Simply put, if the series that investors are working with have true autocorrelations ranging from 0.7 to 0.9, the investors may instead forecast all variables using an autocorrelation in the neighborhood of 0.8. An immediate consequence of this is that they will *over*-react to information about variables with less persistent processes—for example, those with autocorrelations close to 0.7—but will *under*-react to information about variables with more persistent processes—for example, those with autocorrelations closer to 0.9.

I bring these ideas to the context of the yield curve. The yield on a bond has two components: one that is an average of expected short rates over the life of the bond (the expectations-hypothesis, or EH, component), and one that captures the term premium (the TP component). Suppose that the true autocorrelation of the EH component exceeds that of the TP component. Suppose also that, due to bounded rationality, investors forecast both components using an intermediate, average autocorrelation. As described above, this "autocorrelation averaging" means that investors will underreact to news about the EH component but overreact

[1]Refer to Section 3 for a detailed discussion of various models of expectations.

[2]The pattern of under- and overreaction remains stable for each forecaster over time, indicating that Bayesian learning may not play a significant role in determining how forecasters respond differently to information at various maturities.

[3]We can define processing capacity with regard to the forecaster's innate cognitive ability or to institutional resources.

to news about the TP component. Since, for short-maturity bonds, the EH component is more important than the TP component in driving the yield variation, this predicts underreaction to information about short-term bond yields; and since, for long-maturity bonds, the TP component is more important, this predicts overreaction to information about long-term bond yields. This is precisely the pattern I document in the data.

I present several pieces of evidence to show that this parsimonious model of "autocorrelation averaging" can qualitatively and quantitatively unify the pattern of under- and overreaction. First, I compute the sample autocorrelations of the short rate and the term premia across maturities—the short rate is very persistent, with an autocorrelation close to 0.97, while the term premia are less persistent, with autocorrelations close to 0.75.[4] I then use the measured beliefs from surveys to structurally estimate forecasters' perceived autocorrelations of the short rate and term premia. I find that they are very close—in the range of 0.92 to 0.96—and that they lie between the estimated true autocorrelations of 0.75 and 0.97. These estimates are consistent with "autocorrelation averaging": instead of using the true autocorrelations of the time series, forecasters are using an average autocorrelation. This, in turn, generates underreaction for short-rate and overreaction for term-premium components. Moreover, I calibrate my autocorrelation-averaging model with the estimated true and perceived autocorrelations. Though the model makes no specific assumption about the relative importance of the short rate and term premia, it generates the downward-sloping term structure of misreaction, with the FE-on-FR coefficients statistically close to the empirical estimates.

One may naturally wonder why professional forecasters and investors—whose career advancement and financial compensation depend on accurately capturing yield curve dynamics—do not seem to fully learn the true autocorrelations of the short-rate and term-premium components over time. It is worth stressing that while belief distortion from "autocorrelation averaging" goes a long way towards explaining the term structure of misreaction, the departure from rational expectations is small in absolute magnitude. To be precise, forecasters only slightly deviate from the true autocorrelations, especially for the directly observable short rate.[5] Moreover, forecasters do learn and improve over time. For example, more experienced forecasters have more accurate, albeit still distorted, perceived autocorrelation.

[4] All autocorrelation coefficients are computed quarterly.

[5] It is true that the series with different levels of persistence have distinguishable half-lives when one observes them long enough. In a real-world forecasting and investing environment where the persistence of various series is likely to be time-varying and forecasters face constraints on their processing capacity, it is much harder to do so.

One possible psychological driver of "autocorrelation averaging" is limited attention. When investors' attention varies, so will their forecast misreaction. In other words, when people's processing capacity is more constrained relative to the difficulty of their forecasting task, their learning is more impaired, and they are more prone to "autocorrelation averaging." By relating each forecaster's time-varying subjective autocorrelation to various characteristics, I show that forecasters tend to make bigger mistakes in autocorrelations following recessions, large monetary policy shocks, or heightened economic uncertainty. This additional body of evidence further supports a bounded-rationality interpretation of the results.

Next, I show that the pattern of belief misreaction has direct implications for asset prices. One prerequisite for such pricing impact from distorted beliefs is that the market participants should *act* in accordance with their stated beliefs—i.e., they are willing to put their money behind these numbers. Notice that forecasters in the sample, whose identities are known, are either major players in the US Treasury market—for example, 17 of them are primary dealers of the New York Fed—or likely to influence important market participants through various client services. I match a subset of Blue Chip forecasters—which are banks—to their balance sheet information from the Call Reports. I find that, consistent with Merton's (1969) model of portfolio choice, banks' allocations to Treasuries vary positively and significantly with their subjective expectations of bond returns at the corresponding maturities. The magnitude is economically significant: a one-standard-deviation increase in forecasts of next year's 10-year yield leads to a $1.64 billion decrease in Treasuries (maturity between 5 and 15 years) held by an average bank, representing a sizable 40% decrease.

With this key prerequisite fulfilled, I investigate some implications of overreaction for long-term bond prices: when investors overreact to the arrival of new information by raising their forecasts of the long-bond yield too much, they are likely to push the bond price down too low (and yields too high); however, when this price pressure gradually subsides, the bond price will correct to a sensible level. Therefore, an increase in the forecast revision should forecast higher bond returns in the future. Since forecast revision is primarily driven by the *lagged* forecast error (realized at time t), I turn this into a prediction that is easier to test: lagged forecast errors should *positively* predict subsequent bond returns. This prediction is strongly confirmed in the data—a single ex-ante measurable variable, the lagged forecast error for the 10-year Treasury yield, predicts excess bond returns in and out of sample, subsumes several commonly used predictors and strongly rejects the "spanning hypothesis" using the robust procedure of Bauer and Hamilton

(2017). Analogously, I also confirm the opposite prediction of underreaction in short-term bond prices and Federal funds futures prices.

Related literature. The main contributions of this paper are fourfold: (a) by using individual and consensus-level yield curve expectations, I document a robust pattern of how beliefs in the integrated US Treasury market depart from rational expectations—a downward-sloping term structure of under- and overreaction to new information. (b) I present and empirically test a new, parsimonious model of distorted beliefs based on "autocorrelation averaging" that unifies the pattern of belief misreaction. (c) I systematically connect banks' yield curve expectations to their Treasuries portfolios and show that banks allocate funds per their stated beliefs. (d) I uncover strong predictability in bond and interest rate futures prices as predicted by both under- and overreaction in beliefs.

There has recently been renewed interest in studying the rationality of beliefs using micro-level survey data. This paper, together with recent work such as Bouchaud et al. (2019) and Bordalo et al. (2020, BGMS), follows a methodology developed by Coibion and Gorodnichenko (2015), which quantifies departures from RE through the lens of predictability of forecast errors from forecast revisions. Most studies, except BGMS, document one form of misreaction—either under- or overreaction—for various series.[6] BGMS, in analyzing professional forecasts of 22 macroeconomic variables, find that individual forecasts tend to overreact to information while consensus forecasts tend to underreact. They combine a new model of expectations, developed by Bordalo et al. (2018), called "diagnostic expectations"—which generates individual overreaction—and rational inattention—which generates consensus underreaction—to explain their specific pattern of misreaction.[7] Distinct from previous papers, I organize the test of rationality along the Treasury yield curve, where interest rates are disciplined by no-arbitrage conditions, and establish a clear downward-sloping term structure of belief misreaction. The pattern—underreaction for short rates and increasing overreaction for long rates—is robust at

[6]In an experimental study of expectation formation, Afrouzi et al. (2020) document both under- and overreaction in individual forecasts, and find that overreaction, in the form of extrapolative expectations, predominates.

[7]BGMS focus on the distinct patterns of forecaster-level overreaction and consensus-level underreaction. Among many variables they examine, they find underreaction for FFR and 3-month T-bill rate and overreaction for 10-year Treasury yield. These findings are not fully explained by their diagnostic expectations framework. In another contemporaneous work, d'Arienzo (2020) focuses on the greater overreaction of forecasts for long-term yields. To rationalize this phenomenon, he follows BGMS and builds a model of diagnostic expectations in which long-term outcomes have higher uncertainty, making investors overreact more aggressively for longer-term variables. This "representativeness"–based mechanism, in both BGMS and d'Arienzo (2020), is different from mine, which has its root in bounded rationality.

both individual and consensus levels. This finding is at odds with the interpretation from BGMS and from models that generate only one form of misreaction. Recently, in a setting similar to mine where people are estimating/learning about the time-varying parameters governing bond yield dynamics, Singleton (2021) shows that beliefs from a Bayesian econometrician can generate CG regression coefficients close to the empirically estimated ones.[8]

Models that depart from full rationality have been developed to explain under- and overreaction in other settings. For macroeconomic variables and short-term earnings forecasts, rational inattention and information rigidities have been the main frameworks to explain underreaction.[9] Finance models, often aiming to explain overreaction from a behavioral perspective, are built on psychological foundations such as representativeness and overconfidence;[10] existing models of under- and overreaction include those of Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), Rabin (2002), and BGMS. Different from the approaches based on representativeness, this paper presents a model of distorted beliefs based on "autocorrelation averaging." The model is akin to many bounded-rationality models, such as Gabaix (2019), in which the decision-maker, with limited processing capacity, anchors her perception of key structural parameters on some default values. In particular, I present—and offer direct empirical support for—a unified framework with a bounded-rationality origin for under- and overreaction.

Finally, a growing body of research studies investors' subjective beliefs in the Treasury bond market (see Singleton, 2021 for a thorough review). Previous studies have explored asset pricing implications of distorted beliefs in particular sections of the yield curve (e.g., Cieslak, 2018 and Xu, 2020 in short rates and d'Arienzo, 2020 in long rates), or of heterogeneous beliefs (and disagreement) about interest rates among forecasters (e.g., Buraschi, Piatti, and Whelan, 2018; Giacoletti, Laursen, and Singleton, 2021). Related to the findings on perceived autocorrelations in this article, Piazzesi, Salomao, and Schneider (2015) find that bond risk premia implied by survey

[8] While I take a "behavioral" approach and offer evidence supporting the bounded rationality origin of misreaction, my "autocorrelation-averaging" mechanism is not inconsistent with a learning framework—in Section A.2 of the Internet Appendix, I show that "autocorrelation-averaging" can be motivated by a learning problem where the forecaster puts a heavy weight on her priors.

[9] See Sims (2003), Woodford (2003), Carroll (2003), Mankiw and Reis (2002), and Gabaix (2014) for discussions of underreaction for macroeconomic variables.

[10] In the behavioral finance literature, overreaction to news is often modeled using extrapolative beliefs, which has been developed in generations of models such as De Long et al. (1990), Frankel and Froot (1990), Barberis, Shleifer, and Vishny (1998), Hong and Stein (1999) and, more recently, Barberis et al. (2015, 2018) and Glaeser and Nathanson (2017). Greenwood and Shleifer (2014) find strong evidence of extrapolative expectations in several surveys of stock market returns.

forecasts are more persistent and less volatile than those obtained from bond yields. I complement these papers in three ways. First, I focus on beliefs about the entire term structure of interest rates and document a robust pattern of under- and overreaction in both individual and consensus forecasts. Second, I manually link survey forecasts of banks to their Treasuries portfolios and establish a significant, positive relationship between banks' allocations to Treasuries and their subjective expected bond returns.[11] Third, I provide strong statistical evidence of "excess" bond return predictability induced by under- and overreaction in beliefs; the direction of the predictability depends on how investors misreact to news across maturities. This "excess" return predictability is similar, in spirit, to that of Cieslak (2018), which also stems from distorted beliefs, but is distinct from the conventional interpretations based on hidden factors or measurement error.[12]

The rest of the paper is organized as follows. Section 2 describes the Blue Chip Financial Forecasts data and examines the accuracy of the survey forecasts. Section 3 documents how forecasters react to new information for short- and long-maturity interest rates. Section 4 builds a bounded-rationality model, empirically estimates the actual and perceived autocorrelations, and calibrates the model to match the term structure of under- and overreaction. Section 5 explores the implications of under- and overreaction for asset prices and portfolio allocation. Section 6 concludes.

2 Data

2.1 Survey expectation data

The primary dataset used in this paper is the survey forecasts of US Treasury bond yields across maturities and other interest rates, which I obtain from the Blue Chip Financial Forecasts (BCFF) survey. This monthly survey maintains a stable and large panel of professional forecasters and has a long sample dating back to the 1980s. Among the various datasets of professional forecasts,

[11] A series of recent papers, starting from Manski (2004), has linked stated beliefs to investment behavior and portfolio choices (e.g., Vissing-Jorgensen, 2003; Dominitz and Manski, 2007; Kézdi and Willis, 2009; Amromin and Sharpe, 2014; Drerup, Enke, and von Gaudecker, 2017; Choi and Robertson, 2020; Giglio et al., 2021). While most of the literature focuses on the equity market and individual investors, this paper is the first to systematically link survey forecasts of banks to their Treasuries portfolios.

[12] A related strand of literature has found excess sensitivity of long-term claims, such as Stein (1989) in volatility; Giglio and Kelly (2018) in various cash flow streams; and Gürkaynak, Sack, and Swanson (2005), Hanson and Stein (2015), Hanson, Lucca, and Wright (2018) and Brooks, Katz, and Lustig (2019) in interest rates. The overreaction for long-rate expectations, documented in this paper, complements these findings in bond prices.

it is especially suitable for studying expectation formation and asset prices.

Each month, the BCFF survey collects forecasts from a panel of, on average, over 40 economists from leading financial institutions and economic consulting firms. They are asked to provide point forecasts of future financial and macroeconomic variables at horizons from the current quarter ("nowcast") to four quarters ahead (five quarters since January 1997). The forecasts are collected over a two-day period, usually between the 23rd and 27th of each month, and published on the first day of the following month. A sample BCFF survey questionnaire is presented in the Internet Appendix.

Variables. To study the subjective expectations of bond yields across the entire yield curve and of other interest rates, I require that the forecasts have reasonably long and continuous time series. Specifically, I focus on the forecasts of the following interest rate variables: Treasury bills with maturities of three months and one year ($tb3m$ and $tb1y$), Treasury notes and bonds with maturities of 2, 5, 10, and 30 years ($tn2y$, $tn5y$, $tn10y$, and $tn30y$), Federal Funds Rate (ffr), one-month commercial paper rate ($cp1m$), prime bank rate (pr), three-month LIBOR rate ($libor$), Aaa and Baa corporate bond rates (aaa and baa), and home mortgage rate (hmr). For each target variable at each forecast horizon, I obtain both forecasts of individual forecasters and the consensus forecast (cross-sectional mean).

Forecasters. One of the advantages of the BCFF survey is that it includes each forecaster's name and affiliated institution.[13] This feature allows us to keep track of the time series of each firm's forecasts and hence make the BCFF forecasts a panel dataset. However, the institution names change from time to time due to reasons such as mergers and acquisitions. To make the panel of forecasts as balanced as possible, I manually check the name changes of the forecasters using the information provided by the Federal Financial Institutions Examinations Council (FFIEC) and concatenate the observations belonging to the same entity. This manual match gives us 86 unique forecasters with more than 60 monthly forecasts, among which 26 are banks, 15 are broker-dealers, and 17 are primary dealers of the Federal Reserve Bank of New York. Table A.1 in the Internet Appendix provides a full list of institutions that participate in the BCFF survey, grouped by type of institution.

[13]As the forecasts mostly reflect collective expectations of the institutions, for the rest of the paper, I use "forecaster" to refer to the institution.

The final BCFF survey forecast dataset has a sample period from January 1988 to December 2018. I choose the start date such that the forecasts of all Treasury yields that I study are available.

2.2 Realized interest rates and macro data

The nominal zero-coupon Treasury yields are mainly from two sources. I obtain the US Treasury yields and forward rates from the fitted Treasury yield curve of Gürkaynak, Sack, and Wright (2006, GSW). The GSW data is updated regularly and available on the Federal Reserve Board website. Because it contains only bonds with maturities from one year to 30 years, I use constant-maturity Treasury (CMT) yields from the H.15 statistical release of the Federal Reserve for Treasury bills with maturities shorter than one year.[14] Additionally, I use Fama and Bliss bond yields and forward rates from the Center for Research in Securities Prices (CRSP) to reproduce the Cochrane and Piazzesi (2005) bond return predictor (CP). To complement the bond returns constructed from the fitted Treasury yields of GSW, I obtain from CRSP the returns of the Fama Maturity Portfolios, which are actual coupon bond portfolios sorted by maturity.

The realized values of other interest rates, such as the effective Federal Funds Rate, are obtained from the St. Louis Fed FRED database. I sample the daily data at a monthly (quarterly) frequency by using the last observation of each month (quarter). The interest rates are expressed in percent per annum. I also obtain macroeconomic and aggregate financial variables from FRED.

2.3 Notation

Bond notation. I follow the standard notation in the bond literature, in which the maturity is given in parentheses as a superscript. Assume that interest rates are continuously compounded. The log price and the yield of an n-year bond at time t are denoted $p_t^{(n)}$ and $y_t^{(n)} = -\frac{1}{n}p_t^{(n)}$, respectively. The h-year holding period return on an n-year zero-coupon bond is defined as the change in the log price:

$$r_{t+h}^{(n)} \equiv ny_t^{(n)} - (n-h)y_{t+h}^{(n-h)}.$$

In the asset pricing analyses in this paper, I follow the literature and focus on the one-year holding period excess return:

$$rx_{t+1}^{(n)} \equiv ny_t^{(n)} - (n-1)y_{t+1}^{(n-1)} - y_t^{(1)}.$$

[14] I use the GSW zero-coupon yields in order to follow the bond return predictability literature. The results in this paper are robust if I instead use the Fed's H.15 statistical release for all analyses.

Expectation notation. There are several different forms of expectations studied in this paper. With a slight abuse of notation, I denote the rational expectation by the expectation operator $\mathbb{E}(.)$. The BCFF survey-based subjective expectation and the expectation using the econometrician's real-time information set are denoted as $\mathbb{E}^S(.)$ and $\widehat{\mathbb{E}}(.)$ respectively.

2.4 Properties of survey expectations of Treasury yields

I first examine the accuracy, in the form of forecast errors and out-of-sample predictability, of the survey-based expectations of interest rates. Forecast errors are the differences between the ex-post realized values and the forecasts. For an n-year Treasury bond, I define the h-period-ahead forecast error (FE) as[15]

$$FE_t\left(y_{t+h}^{(n)}\right) = y_{t+h}^{(n)} - \mathbb{E}_t^S\left(y_{t+h}^{(n)}\right). \tag{1}$$

A non-zero forecast error may stem from a shock between the time a forecast is made and the outcome is realized or from a systematic departure from rationality. Panel A of Table 1 reports summary statistics for the individual forecast errors for FFR and Treasury yields across maturities. The forecasts are pooled across horizons h.[16] As in Cieslak's (2018) study on consensus short-rate forecasts, the average and median forecast errors for interest rates across maturities are negative and small in magnitude: The means are less than 0.4% in absolute value, and the standard deviations are around 1%, indicating that the professional forecasters regularly overestimate the future Treasury yields to a moderate degree. Interest rates with different maturities have quantitatively similar forecast errors on average. The forecast errors normalized by contemporaneous realized interest rates also have similar medians. The statistics are similar, albeit less variable, when we look at the consensus-level forecasts, as reported in the Internet Appendix.

Apart from checking the average accuracy, I use the out-of-sample (OOS) R^2 to examine the predictive power of the individual- and consensus-level survey forecasts against several

[15] BCFF panelists forecast various interest rates in the current and next four to five quarters; and they report the quarterly average level. To match this survey convention, I average the monthly interest rates within a quarter when calculating the realized interest rates.

[16] Summary statistics of individual forecasts of each horizon h are reported in the Internet Appendix.

alternative statistical models:

$$R^2_{OOS,i} = 1 - \frac{\sum_t \left(y^{(n)}_{t+1} - \mathbb{E}^S_t \left(y^{(n)}_{t+1} \right) \right)^2}{\sum_t \left(y^{(n)}_{t+1} - \mathbb{E}^i_t \left(y^{(n)}_{t+1} \right) \right)^2}, \tag{2}$$

where $\mathbb{E}^i_t \left(y^{(n)}_{t+1} \right)$ is the prediction from model i at time t. The OOS R^2 contrasts errors of the subjective survey forecasts with those from an alternative model. A positive OOS R^2 indicates superior predictive power of the survey forecasts over the alternative model, and vice versa. I consider several commonly used statistical models: moving average (Mean), AR(1), AR(p), and ARIMA(1,1,0).[17] The moving average is a default alternative model in the forecast evaluation and asset pricing literature and ARIMA-class models are known to model the level of the interest rate well with an R^2 close to 1.[18]

Panels C and D of Table 1 summarize the results for all interest rates by tabulating the median individual OOS R^2 and consensus OOS R^2. The results for the ARIMA-class models offer a mixed picture. While professional forecasters make better predictions than statistical models at short maturities, they perform poorly at longer maturities in general, even though the statistical fitness of the ARIMA models are comparable at short- and long-maturity interest rates. The moving average is too smooth to capture the business-cycle frequency fluctuation of interest rates; thus, we see that survey forecasts perform much better, on average, at both the individual and consensus levels.

Overall, professional forecasters, on average, overestimate interest rates moderately across maturities. The accuracy of the short- and long-maturity forecasts diverges when compared to predictions from commonly used statistical models. Potentially due to the sophistication of the survey participants, the short-rate forecasts are reasonably accurate; however, those for long-maturity rates are much less so.

3 Misreaction to Information across Maturities

In this section, I formally test whether the professional forecasts of interest rates across maturities are rational and, more precisely, whether the deviation from rationality stems from under- or

[17]To prevent the alternative models from using future information, I estimate AR(1), AR(p), and ARIMA(1,1,0) recursively with rolling windows.

[18]See Goyal and Welch (2008) and Clark and McCracken (2013) for reviews of forecast evaluation.

overreaction to the new information that forecasters receive in real time. To do this, I follow a methodology developed by Coibion and Gorodnichenko (2015), which assesses under- and overreaction by examining the predictability of forecast errors from forecast revisions and has been used extensively in recent work to study the dynamics of expectations. CG use forecast revisions to capture the new information available to forecasters; this circumvents the problem that the real-time information set of the forecasters is not observable to the econometrician ex-post. CG derive the predicted relationship between ex-post forecast errors and ex-ante forecast revisions at the consensus level from sticky-information and noisy-information models. However, their methodology can also be applied to individual-level forecasts, as in Bouchaud et al. (2019) and Bordalo et al. (2020).

Consider a target variable x_t. Formally, the forecast revision (FR) at time t is defined as the difference between the time t forecast for x_{t+h} and the forecast for the same quantity made at time $t-k$:

$$FR_t(x_{t+h}) \equiv \mathbb{E}_t^S\left(x_{t+h}\right) - \mathbb{E}_{t-k}^S\left(x_{t+h}\right). \tag{3}$$

The extent to which individual and average professional forecasters under- or overreact to news can be evaluated by estimating the following regression

$$FE_{i,t}\left(x_{t+h}\right) = \alpha_i + \beta FR_{i,t}\left(x_{t+h}\right) + \epsilon_{i,t,h}, \tag{4}$$

where x_t is the underlying interest rate variable and h denotes the specific forecast horizon, ranging from one (next quarter) to four quarters ahead. As noted by Coibion and Gorodnichenko (2012, 2015), this structure accommodates patterns of both under- and overreaction. When $\beta > 0$, the forecaster insufficiently incorporates new information into her forecasts, indicating underreaction and expectation stickiness. When $\beta < 0$, the forecaster responds too much to new information, indicating overreaction. When $\beta = 0$, the subjective forecasts are consistent with rational expectations.

Several details warrant emphasis regarding the above regression. First, the regression is estimated at the quarterly frequency for ease of interpretation; $k = 1Q$ indicates a one-quarter forecast revision. In the Internet Appendix, I estimate Equation (4) at the monthly frequency and with one-month forecast revisions; the results are similar. Second, I run the above regression for each interest rate for both *individual*-level and *consensus*-level forecasts. When applied to the individual level, the regression includes forecaster (institution) fixed effects to control for cross-sectional unobserved heterogeneity across forecasters. This individual-level regression specification is different from the main test in Coibion and Gorodnichenko (2015); they assume rational Bayesian learning of private information at the individual level, which predicts $\beta = 0$ at the individual level. Third, when estimating Equation (4) for each interest rate at both the individual and consensus levels, I pool the observations across different forecast horizons to make full use of all information from the subjective expectations and hence increase the statistical power of the test.[19] Previous studies mostly focus on one specific forecast horizon (e.g., $h = 3Q$ in CG and BGMS) due, in some cases, to data availability. In the baseline results, I include the Federal Funds Rate and Treasury yields for maturities of 3 months, and 1, 2, 5, 10, and 30 years as the underlying interest rates.

Main results. Panel A of Figure 1 presents the regression coefficients $\hat{\beta}$ for each interest rate from the individual-level panel regressions. The dots depict the point estimates and the range of each bar represents the 95%-confidence interval of each estimate. Standard errors are clustered by forecaster and time. A clear pattern emerges as maturity increases. The term structure of the FE-on-FR regression coefficients is downward-sloping. Short-maturity interest rates (less than two years) have $\hat{\beta} > 0$ while long-maturity interest rates (greater than two years) have $\hat{\beta} < 0$. Moreover, I show that the same pattern of misreaction is also manifested in the consensus-level regressions. Panel B of Figure 1 plots regression coefficients estimated using consensus forecasts. The standard errors are calculated as in Driscoll and Kraay (1998), which allows for cross-sectional and serial correlations and for heteroskedasticity in the errors.

The pattern of the coefficients, at both the individual and consensus levels, indicates that individual forecasters *underreact* to new information about short-maturity interest rates and *overreact* to new information about long-maturity interest rates. This pattern of disparate misreaction is striking, especially given that the US Treasury market is largely integrated and that a strong factor structure in the yield curve induces high correlations among all interest rates. Table 2 reports the details of the regression results at both the individual and consensus levels. In Panel A, except for the two-year Treasury note that is situated in between short- and long-maturity interest rates, all β estimates are statistically and economically significant. Take the 10-year Treasury yield as an example: a one-percentage-point increase in the past forecast revision indicates that the future realized yields are, on average, 0.23 percentage points lower

[19] Coibion and Gorodnichenko (2015) also find very little cross-horizon heterogeneity in inflation expectations.

than the previous forecasts. Panel B reports results for consensus-level regressions. The sign and statistical significance of the β estimates are similar to those in the individual-level regressions. There are two differences from Panel A: the absolute values of $\hat{\beta}$ across maturities are higher in Panel B, especially at the short end of the yield curve and the $\hat{\beta}$ of the two-year Treasury note is statistically significant.

The downward-sloping pattern in FE-on-FR regression coefficients is robust to several alternative specifications, especially at the individual level. In the Internet Appendix, I report results for Equation (4) estimated (a) using all available forecast data from 1982, (b) without firm fixed effects at the individual level, (c) at a monthly frequency using one-month forecast revisions, and (d) separately for each forecast horizon.[20] All alternative specifications preserve the downward-sloping term structure of under- and overreaction and the beta coefficients cross zero at around the two-year maturity.

To examine more extensively the contrasting reactions to news for short- and long-maturity rates, I estimate Equation (4) for a few additional interest rates, which I divide, broadly, into two groups based on their maturities. The new interest-rate variables include one-month commercial paper rate ($cp1m$), prime bank rate (pr), three-month LIBOR rate ($libor$), Aaa and Baa corporate bond rates (aaa and baa), and home mortgage rate (hmr).[21] Though these interest rates contain additional risks, such as default risk and prepayment risk, compared with Treasury yields, they correlate strongly with Treasury yields across all maturities. However, the correlations of forecast errors/revisions are high only within maturity groups and low across groups.[22] I focus on the individual-level regressions. Figure 2 plots the β estimates of the additional short-maturity (Panel A) and long-maturity (Panel B) interest rates and Table 3 reports the details of the corresponding regressions. The dichotomy of forecasters' reactions to new information for the extended set of short and long rates is evident in the figure. Moreover, the downward-sloping pattern of coefficients is largely preserved. The point estimates are positive for all short-term interest rates and negative for all long-term interest rates. All but the two-year Treasury yield are statistically

[20] The underlying interest rate variables were introduced in the survey in a staggered fashion. Some interest rates, such as the Federal Funds Rate, appear in the survey earlier than 1988. Alternative specification (1) uses all available forecasts for each interest rate.

[21] Both Aaa and Baa indexes are calculated based on corporate bonds with maturities of 20 years and above. The home mortgage rate has a maturity of 30 years.

[22] In the Internet Appendix, I plot pairwise correlations of the level, one-year changes, forecast errors, and forecast revisions of different interest rates. The interest rates include the Treasury and additional interest rates. The level of all interest rates correlates strongly, while there is a clear two-group structure in the correlations of forecast errors and forecast revisions of different interest rates.

different from zero. The regression results are also robust if I restrict the sample to the forecasters who make forecasts for the complete set of interest rates.[23]

In summary, I document a robust downward-sloping term structure of under- and overreaction to new information in the professional forecasts of interest rates. This pattern is evident at both individual and consensus levels.

Commonly-used models of expectations. How does the above evidence square with commonly used models of expectations? In what follows, I consider several candidate models in addition to the full-information rational expectation benchmark. I briefly discuss the implications of each model and explain why none of the commonly used models, at least in their standard form, can deliver the under- and overreaction pattern that I document. I relegate all related derivations to the Internet Appendix.

- The benchmark FIRE model posits that the forecast error is noise, orthogonal to any information known to the forecaster, and is therefore not predictable; thus there should be no relationship between an individual's past forecast revisions and subsequent forecast errors. FIRE therefore is at odds with either form of misreaction to information. Relatedly, models of rational Bayesian learning, which relaxes the full information aspect of FIRE, have been proposed to explain belief formation in financial markets (e.g., Singleton, 2021). However, as I show in the Internet Appendix, the pattern of misreaction remains stable for each forecaster across different subperiods. This indicates that a Bayesian learner has to learn very slowly despite that the interest rate environment has changed significantly over the past decades.

- Sticky expectations capture the idea that the forecaster is sluggish when updating her beliefs, so that the current forecast gives significant weight to lagged beliefs. Coibion and Gorodnichenko (2015) apply sticky expectations to the aggregate consensus forecast; such expectations can be derived from infrequent information updating (Mankiw and Reis, 2002) or signal extraction from heterogeneous signals. When applied at the individual level, as

[23] Measurement error in forecasts may mechanically lead to negative predictability of forecast errors from forecast revisions. However, additional robustness checks indicate that measurement error may not be a big concern. First, because actual interest rates are measured with little noise and the forecast errors are small on average, the noise in forecasts should be small. Second, I follow the procedure proposed by Bordalo et al. (2020) to regress forecast errors at a certain horizon on forecast revisions for a different horizon. Given that misreaction is positively correlated for forecasts at different horizons, this specification yields CG coefficients with the same sign while avoiding the mechanical measurement error problem of overlapping left- and right-hand-side variables.

in the setting of Equation (4), sticky expectations predict a positive relationship between forecast errors and forecast revisions, indicating underreaction to information across all maturities. Hence, sticky expectations cannot explain the evidence that I document.

- Extrapolative expectations, in the functional form reviewed by Barberis (2018), model the current forecast as a weighted average of past realizations. The weights are exponentially decaying and higher for the more recent past. Extrapolative expectations typically generate overreaction in the stock market when people form expectations about non-persistent stock returns. In the Internet Appendix, I show that, when the underlying process is persistent enough, as is the case for *all* interest rates, underreaction to information prevails and the coefficient in the FE-on-FR regression is always positive. Another frequently used form of extrapolative expectations, as surveyed by Afrouzi et al. (2020), is a backward-looking extrapolative expectation in which forecasts are determined by the current outcome and the recent one-period trend: $\mathbb{E}_t^S(x_{t+h}) = x_t + \theta(x_t - x_{t-1})$. As shown in the Internet Appendix, this framework nests both underreaction (when $\theta < 0$) and overreaction (when $\theta > 0$). However, it is difficult to argue that people have drastically different extrapolative parameters θ—with opposite signs—for highly correlated processes such as short- and long-maturity interest rates.

- Diagnostic expectations, formalized by Bordalo et al. (2018, 2019), incorporate a belief distortion rooted in the concept of representativeness first introduced by Kahneman and Tversky (1972, 1973). Under diagnostic expectations, individual forecasts overweight future outcomes in light of incoming data; the individual-level FE-on-FR regression coefficient therefore is always negative. BGMS extend the framework to allow for imperfectly correlated heterogeneous private signals. The new elements of the model can generate underreaction to aggregate (average) information when applied to consensus-level forecasts. However, neither the original nor the richer diagnostic expectations model can simultaneously deliver underreaction for short rates and overreaction to long rates at both the individual and consensus levels.

- Natural expectations posit that forecasters obtain their expectations by taking an average of expectations under both the true model and a more parsimonious but misspecified intuitive model. Under the specification formalized in Fuster, Laibson, and Mendel (2010), the true data-generating process is a stationary AR(2), while the forecasters' intuitive model contains a unit root in the first lag. This deviation from rationality gives rise to overreaction. Therefore, the natural expectations framework cannot generate underreaction without making additional assumptions about the true and intuitive models of expectations.

4 A Model Based on "Autocorrelation Averaging"

In this section, I propose a simple bounded-rationality framework for understanding the above results. The framework is based on the distorted subjective perception of autocorrelations. A boundedly-rational forecaster has limited working memory or finite processing capacity for carrying out complex calculations.[24] When facing multiple time series with different autocorrelations, she may not attend promptly to all relevant information to correctly estimate each autocorrelation. Instead, she uses something closer to an average autocorrelation of all the processes to which she is exposed and only imperfectly adjusts toward the true autocorrelations. For example, if the series have true autocorrelations ranging from 0.7 to 0.9, the forecaster may instead forecast all variables using an autocorrelation in the neighborhood of 0.8. An immediate consequence of this is that she will overreact to information about less persistent variables—for example, those with autocorrelations close to 0.7—but will underreact to information about more persistent variables—those with autocorrelations closer to 0.9. Formally, I term this behavior "autocorrelation averaging."

In the context of interest rates, the yield on a bond has two components: one that is an average of expected short rates over the life of the bond (the expectations-hypothesis, or EH, component) and one that captures the term premium (TP). For each maturity n, the forecaster needs to estimate separate autocorrelations for the EH and TP components. Suppose for a moment that the true autocorrelation of the EH component exceeds that of the TP component. Also, suppose that, due to bounded rationality, the forecaster forecasts both components using an intermediate, average autocorrelation. Given a large number of autocorrelations to estimate, she anchors her subjective autocorrelation to the simple average autocorrelation across all processes. As described above, this means that she will underreact to news about the EH component but overreact to news about the TP component. Since, for short-maturity bonds, the EH component is the more important, this predicts under-reaction to information about short-term bond yields. Since, for long-maturity bonds, the TP component is the more important, this

[24] See Gabaix (2019) for a detailed review.

predicts overreaction to information about long-term bond yields. This prediction is precisely the pattern of belief misreaction.

In the remainder of this section, I first formalize the predictions for belief formation in a model in which the boundedly-rational forecaster uses an average autocorrelation to forecast the interest rate components. I then empirically estimate the subjective and actual autocorrelations of the EH and TP components to corroborate the model's critical assumption. Last, I show that the "autocorrelation averaging" model, calibrated to the estimated autocorrelations, can quantitatively match the downward-sloping term structure of under- and overreaction.

4.1 Subjective autocorrelation

I start by exploring the effect of biased perceived autocorrelations on the forecaster's reaction to news—specifically, in the form of the FE-on-FR regression coefficient. Suppose that an underlying variable z_t follows an AR(1) process:

$$z_{t+1} = \rho z_t + \varepsilon_{t+1}, \quad \varepsilon_t \sim \mathcal{N}\left(0, \sigma^2\right).$$

A boundedly-rational forecaster uses her subjectively perceived autocorrelation ($\rho^s \neq \rho$) of the underlying process. She forms her subjective expectation of the h-period-ahead value of z_t as

$$\mathbb{E}_t^S\left(z_{t+h}\right) = \rho^s z_t,$$

so that

$$\mathbb{E}_t^S\left(z_{t+h}\right) = \left(\frac{\rho^s}{\rho}\right)^h \mathbb{E}_t\left(z_{t+h}\right).$$

Following the previous definitions, her forecast error and one-period forecast revision are

$$\begin{aligned} FE_t\left(z_{t+h}\right) &= z_{t+h} - \mathbb{E}_t^S\left(z_{t+h}\right) = \left(\rho - \rho^s\right) z_t + \varepsilon_{t+h}, \\ FR_t\left(z_{t+h}\right) &= \mathbb{E}_t^S\left(z_{t+h}\right) - \mathbb{E}_{t-1}^S\left(z_{t+h}\right) = \rho^s\left(\rho - \rho^s\right) z_{t-1} + \rho^s \varepsilon_t. \end{aligned}$$

The covariance between forecast error and past forecast revision, the numerator of the FE-on-FR regression coefficient, can be derived as

$$\operatorname{Cov}\left(FE_t\left(z_{t+h}\right), FR_t\left(z_{t+h}\right)\right) = \frac{\rho^s\left(\rho - \rho^s\right)\left(1 - \rho^s \rho\right)\sigma^2}{1 - \rho^2}, \tag{5}$$

which gives us the following proposition about the sign of the FE-on-FR coefficient and the subjective autocorrelation ρ^s:

Proposition 1 *As long as* $\frac{\rho^s(1-\rho^s\rho)}{1-\rho^2} > 0$*, which is satisfied when both actual and subjective autoregressive processes are persistent and stationary, we have:*

1. *Under rational expectations* ($\rho^s = \rho$)*, the subjective autocorrelation equals the true autocorrelation and the FE-on-FR regression coefficient is zero.*
2. *When* $\rho^s > \rho$*, the FE-on-FR regression coefficient is negative, indicating overreaction to new information.*
3. *When* $\rho^s < \rho$*, the FE-on-FR regression coefficient is positive, indicating underreaction to new information.*

4.2 The FE-on-FR regression coefficient for interest rates

I now apply the bounded-rationality framework to the context of interest rates across maturities. Denote the one-period nominal short rate as i_t. We can iterate the definition of holding period excess returns forward and obtain an identity that relates the long rate to the short rate:

$$\begin{aligned} rx_{t+1}^{(n)} &= ny_t^{(n)} - (n-1)\, y_{t+1}^{(n-1)} - i_t \\ y_t^{(n)} &= \underbrace{\frac{1}{n}\mathbb{E}_t\left(\sum_{i=0}^{n-1} i_{t+i}\right)}_{\text{Expectations hypothesis,} eh_t^{(n)}} + \underbrace{\frac{1}{n}\mathbb{E}_t\left(\sum_{i=0}^{n-2} rx_{t+i+1}^{(n-i)}\right)}_{\text{Term premium,} tp_t^{(n)}}. \end{aligned} \tag{6}$$

The current n-year bond yield is the sum of the expected future average short rate and the average expected excess returns earned over the life of the bond. The identity, similar to the Campbell and Shiller (1988) identity in the aggregate stock market, decomposes the current bond yield into an expectations hypothesis (EH) component and a term premium (TP) component. Notice that the above identity holds both ex-ante and ex-post, so we can apply an arbitrary expectation operator to both sides, including the forecaster's subjective expectation. Empirically, the autocorrelations of the short rate and the term premia are different.

Suppose that the underlying processes of the short rate (i_t) and n-year term premium ($tp_t^{(n)}$) are both AR(1) with autocorrelations ρ_1 and $\rho_p^{(n)}$, respectively, and that term premium shocks are uncorrelated with short-rate shocks[25]

$$
\begin{aligned}
i_{t+1} &= \rho_1 i_t + \varepsilon_{t+1}, \quad \varepsilon_t \sim \mathcal{N}\left(0, \sigma_1^2\right) && (7)\\
tp_{t+1}^{(n)} &= \rho_p^{(n)} tp_t^{(n)} + \upsilon_{t+1}, \quad \upsilon_t \sim \mathcal{N}\left(0, \sigma_p^2\right). && (8)
\end{aligned}
$$

For exposition purposes, I suppress the superscript in the term premium autocorrelation and simply denote it as ρ_p. Under these assumptions, I can specify the EH component in Equation (6) using the dynamics of the short rate i_t. I thus rewrite the decomposition of the n-year long rate, using the true short-rate autocorrelation, as

$$
y_t^{(n)} = \frac{1}{n}\frac{1-\rho_1^n}{1-\rho_1} i_t + tp_t^{(n)}, \tag{9}
$$

or, using the subjective autocorrelation, as

$$
y_t^{(n)} = \frac{1}{n}\frac{1-\left(\rho_1^s\right)^n}{1-\rho_1^s} i_t + tp_t^{s(n)}, \tag{10}
$$

where $tp_t^{(n)}$ and $tp_t^{s(n)}$ represent the n-year actual and subjective TP components, respectively. Since TP components are not observable, we can recover them using Equations (9) and (10). I derive the covariance between forecast errors and forecast revisions of an n-year bond by applying the results of Proposition 1:

$$
\begin{aligned}
&\operatorname{Cov}\left(FE_t\left(y_{t+1}^{(n)}\right), FR_t\left(y_{t+1}^{(n)}\right)\right)\\
&\quad= \frac{1}{n}\frac{1-\left(\rho_1^s\right)^n}{1-\rho_1}\operatorname{Cov}\left(FE_t\left(i_{t+1}\right), FR_t\left(i_{t+1}\right)\right) + \operatorname{Cov}\left(FE_t\left(tp_{t+1}^{(n)}\right), FR_t\left(tp_{t+1}^{(n)}\right)\right) \qquad (11)\\
&\quad= \frac{1}{n}\frac{1-\left(\rho_1^s\right)^n}{1-\rho_1^s}\frac{\rho_1^s\left(\rho_1-\rho_1^s\right)\left(1-\rho_1^s\rho_1\right)\sigma_1^2}{1-\rho_1^2} + \frac{\rho_p^s\left(\rho_p-\rho_p^s\right)\left(1-\rho_p^s\rho_p\right)\sigma_p^2}{1-\rho_p^2}.
\end{aligned}
$$

I assume for now—and corroborate in the next subsection—that the short rate (equivalently, the EH component) is more persistent than the TP components.

Assumption 1 (autocorrelation of EH and TP components) *The true autocorrelations satisfy* $\rho_{EH} = \rho_1 > \rho_p^{(n)}$ *for all* n.

The forecaster deals with many EH and TP processes with various autocorrelations, each

[25] The assumption of no correlation between the EH and TP components is empirically verified in the results of Cieslak and Povala (2016) and Duffee (2018).

demanding allocation of her limited processing power. Define the average autocorrelation across all components as

$$\bar{\rho} = \frac{1}{2N}\left(\sum_{n=1}^{N}\rho_{EH}^{(n)} + \sum_{n=1}^{N}\rho_{p}^{(n)}\right) = \frac{1}{2N}\left(N\rho_1 + \sum_{n=1}^{N}\rho_{p}^{(n)}\right), \tag{12}$$

where N is number of interest rates to which the forecaster is exposed. For a given process, she only imperfectly perceives the true autocorrelation and anchors her subjective autocorrelation on the average autocorrelation $\bar{\rho}$. Therefore, "autocorrelation averaging," under a similar formulation as in Gabaix (2019), dictates that her perceived autocorrelation is a weighted average:

$$\rho_i^s = (1-m)\rho_i + m\bar{\rho}, \quad i \in \{1, p\}, \tag{13}$$

where the weight m measures the strength of "autocorrelation averaging." When $m = 1$, she uses the average autocorrelation for all processes. There has been some recent empirical evidence of the averaging behavior in beliefs in other settings, such as DellaVigna and Pollet (2007), Enke and Zimmermann (2019), and Matthies (2018), that is also consistent with a bounded-rationality interpretation.[26] "Autocorrelation averaging", together with the assumption in (13), has the following direct implications for under- and overreaction.

Proposition 2 (under/overreaction of EH and TP) *Forecasters underreact to information in the expectations hypothesis (EH) component ($\rho_1^s < \rho_1$) and overreact to information in the term premium (TP) component ($\rho_p^s > \rho_p$).*

Since short rates have only EH components, we immediately obtain underreaction for short rates. Long rates have both EH and TP components; the following lemma establishes the condition under which the forecaster overreacts for long rates.

Lemma 1 *If the overreaction effect from the term-premium component dominates the underreaction effect from the short-rate component, i.e.*

$$\left|\frac{\rho_p^s\left(\rho_p-\rho_p^s\right)\left(1-\rho_p^s\rho_p\right)\sigma_p^2}{1-\rho_p^2}\right| > \left|\frac{1}{n}\frac{1-\left(\rho_1^s\right)^n}{1-\rho_1^s}\frac{\rho_1^s\left(\rho_1-\rho_1^s\right)\left(1-\rho_1^s\rho_1\right)\sigma_1^2}{1-\rho_1^2}\right|,$$

[26] Alternatively, I can motivate the "autocorrelation averaging" behavior with "slow" learning. Suppose that the forecaster holds a prior that autocorrelations are drawn from a distribution with mean $\bar{\rho}$. As long as the forecaster learns slowly—either because she penalizes heavily any deviation from her prior or because the noise-to-signal ratio is too low—her perceived autocorrelations of the EH and TP components stay close to the mean autocorrelation $\bar{\rho}$, and I obtain the same "autocorrelation averaging" behavior. See the Internet Appendix for more details.

the coefficient in the FE-on-FR regression for long-maturity interest rates is negative.

Intuitively, this condition is more likely to hold as maturity n increases or when the subjective autocorrelation of TP deviates more from the objective one.

4.3 Estimation of the subjective autocorrelations

The simple model shows that "autocorrelation averaging" can potentially explain the documented pattern of under- and overreaction in beliefs. The goals of the empirical exercise hereafter are to (a) empirically verify "autocorrelation averaging" and (b) show that this model can quantitatively match the term structure of misreaction. To achieve (a), I empirically estimate the actual and subjective autocorrelations of interest rate components. The estimation of actual autocorrelations follows the standard AR(1) model estimation procedure. To obtain subjective autocorrelations of the *short rate*, I rely on the relationships of interest rate forecasts between different maturities and across different horizons. For term premia, I recover the subjective autocorrelations using only the relationships of the forecasts across horizons.

Specifically, I allow the AR(1) process for interest rates and their components to have non-zero long-run means to better match the dynamics of interest rates. The process for the short rate i_t is

$$i_{t+1} = (1-\rho_1)\,\bar{i} + \rho_1 i_t + \varepsilon_{t+1}$$

and its subjective expectation is

$$\mathbb{E}_t^S\left(i_{t+h}\right) = \left(1-\left(\rho_1^s\right)^h\right)\bar{i} + \left(\rho_1^s\right)^h i_t. \tag{14}$$

Substituting Equation (14) into the decomposition in Equation (6), I obtain

$$\mathbb{E}_t^S\left(y_{t+h}^{(n)}\right) - \bar{i} = \frac{1}{n}\frac{1-\rho_1^n}{1-\rho_1}\left[\mathbb{E}_t^S\left(i_{t+h}\right)-\bar{i}\right] + tp_t^{s(n)}, \tag{15}$$

where $tp_t^{s(n)}$ denotes subjective term premium with maturity n. Given the empirically consistent assumption that the short-rate and term-premium shocks are uncorrelated, the presence of TP does not bias the estimation of the short-rate subjective autocorrelation.

GMM moment conditions. Since there are two sets of moment conditions concerning the perceived short-rate autocorrelation ρ_1^s, I estimate ρ_1^s using the generalized method of moments

(GMM). The advantage of using such an over-identified GMM estimation over a nonlinear regression to recover the subjective autocorrelation is that GMM moment conditions take into account both the time-series (Equation (16)) and cross-sectional (Equation (17)) restrictions. Therefore, GMM estimation imposes the sensible requirement that the forecaster uses the same short-rate subjective autocorrelation ρ_1^s in all relevant forecasts. The two sets of moment conditions are as follows:

1. (*Forecasts of short rates at different horizons*) For forecast horizon $h \in \{1, 2, 3, 4\}$ quarters, the loadings of the short-rate forecasts on the current short rate follow a geometric progression as implied by the AR(1) structure:[27]

$$\mathbb{E}_t^S \left(i_{t+h} \right) - \left(1 - \left(\rho_1^s \right)^h \right) \bar{i} - \left(\rho_1^s \right)^h i_t = 0. \tag{16}$$

2. (*Cross-maturity relationship*) For maturity $n \in \{2, 5, 10\}$ years and forecast horizon $h \in \{1, 2, 3, 4\}$ quarters, the yield decomposition equation binds long-maturity yields to the short rate[28]:

$$\mathbb{E}_t^S \left(y_{t+h}^{(n)} \right) - \bar{i} - \frac{1}{n} \frac{1 - \left(\rho_1^s \right)^n}{1 - \rho_1^s} \left[\mathbb{E}_t^S \left(i_{t+h} \right) - \bar{i} \right] - c = 0. \tag{17}$$

In total, there are 16 moment conditions and ρ_1^s is the parameter of interest to be estimated.

Short-rate subjective autocorrelation estimation. I estimate the model using monthly observations, forecaster-by-forecaster, and with rolling windows of 10 years (120 months). With the monthly frequency, I use all the available forecasts, which increases the sample size and thus the accuracy of the rolling GMM estimation. Next, estimating ρ_1^s for each forecaster allows for cross-sectional heterogeneity in forecasting methodology, which may originate from the different levels of "autocorrelation averaging" across forecasters. Moreover, as the data sample spans over three decades, macroeconomic conditions and inflation regimes change, and the economists who lead the forecasting effort at each institution also change. Hence it is reasonable to allow the subjective autocorrelation of the short rate to fluctuate over time. The 10-year window length is close to a typical forecasting economist's tenure at a firm in the sample.

[27] Notice that the forecast horizons are in quarters, so the estimated autocorrelations are quarterly.

[28] The free parameter c accounts for non-zero term premium and is estimated by the GMM procedure.

To estimate this over-identified system, I use Hansen's (1982) two-step GMM procedure, with the efficient weighting matrix used in the second step. Following Cieslak (2018), I choose the Federal Funds Rate (ffr) as the primary measure for the short rate, but the results are quantitatively similar if I use the one-year Treasury bill yield instead. I use the average short rate, estimated using an expanding window and available to the forecasters in real time, as the empirical measure of the long-run mean of the short rate $\bar{i}$.[29] Fixing the long-run mean reduces the degree of freedom and enables a more accurate estimation of the parameter of interest, ρ_1^s.

As reported in Panel A of Table 4, the GMM estimation generates 11,106 and 314 valid subjective autocorrelations at the individual and consensus level, respectively. The GMM procedure almost always produces statistically significant estimates with p-values close to zero. I therefore focus on the point estimates and omit the statistical significance from the table. The mean ρ_1^s estimates for the individual- and consensus-level forecasts are 0.92 and 0.89, respectively, which are slightly lower than the mean ρ_1 estimate of 0.97 for the realized short rate[30][31]. This slight deviation indicates that "autocorrelation averaging"-prone forecasters do not make significant mistakes for a fundamental interest rate such as the Fed Funds Rate and that it will take the forecasters a long time to spot and self-correct errors in their forecasting models. As a consequence of such deviations, the forecasters underperceive the persistence of the short rate. As a consequence, this empirical relationship between the subjective and actual autocorrelations of the short rate guarantees that short-rate expectations underreact to information at both the individual and consensus levels, as prescribed in Proposition 1. Figure 3 plots the histogram of ρ_1^s estimates: Panel A pools estimates across time and forecasters and Panel B plots the median estimate for each forecaster. In both, the majority of ρ_1^s estimates lie between 0.75 and 1, indicating that the subjective short-rate autocorrelations, though lower than their actual counterpart, do not deviate much from the actual autocorrelation. With the time-varying ρ_1^s estimates, I further show that the EH components, which are linear transformations of the short rates for a given ρ_1^s, also have lower subjective autocorrelations than their objective counterparts.

[29] Since the full history of the short rate is available to the forecasters, I start the expanding window at the beginning of the Federal Funds Rate data, namely, July 1954.

[30] In comparison, the autocorrelation estimated using the longest available sample of the Federal Funds Rate is 0.97.

[31] The median ρ_1^s estimates for the individual- and consensus-level forecasts—0.95 and 0.96, respectively—are even closer to the median true autocorrelation estimates of the short rate.

Subjective autocorrelation of the term- premium component. With the subjective and actual autocorrelation estimates of the short rate in hand, I can recover both the subjective and actual TP components by inverting Equations (9) and (10), respectively. There is only one set of moment conditions regarding TP components. Since an n-year TP follows an AR(1) process, regressing the h-quarter-ahead TP forecasts ($tp^{s(n)}_{t+h}$) on those with horizon $h-1$ quarters ($tp^{s(n)}_{t+h-1}$) reveals the subjective autocorrelation of the TP component:

$$tp^{s(n)}_{t+h} = \alpha + \rho^{s(n)}_{p} tp^{s(n)}_{t+h-1} + \varepsilon_t. \tag{18}$$

Similarly, this regression is estimated separately for each forecaster and for TP with each maturity.

Statistics for estimated TP autocorrelations are reported in Panels B to D of Table 4. Several findings are worth highlighting. First, the mean and median estimates of the subjective TP autocorrelations $\rho^{s(n)}_{p}$, at both individual and consensus levels (Panels B and C), range from 0.91 to 0.96 across maturities; they are much higher than their actual counterparts, which are around 0.75 (Panel D). Second, the estimated actual TP autocorrelations are lower than the actual short-rate autocorrelation: $\rho^{(n)}_{p} < \rho_1, \forall n$. This relationship corroborates the critical Assumption 1 that the short rate is more persistent than the term premia. Third, the subjective autocorrelations of the short-rate (EH) and TP components are close, located between the disparate actual autocorrelations of EH and TP. This is precisely "autocorrelation averaging" that is theorized in the previous section. Finally, the dispersion of the subjective TP autocorrelation estimates, measured by standard deviations, is much smaller than that of the EH components, suggesting that forecasters disagree less in their subjective autocorrelations of term premia.

Figure 4 summarizes the relationship between the subjective and actual autocorrelations of the short rate and term premia across maturities. It plots the median autocorrelation coefficients from rolling-window estimations and includes autocorrelation estimates for all individual forecasters, the consensus forecast, and the econometrician (actual autocorrelation estimated from realized series). Each blue circle represents a forecaster's median estimates of subjective autocorrelations ρ^{s}_{1} and $\rho^{s(n)}_{p}$; the size of the circle corresponds to the number of monthly surveys this forecaster participates in. The orange diamond and green square, respectively, represent the median autocorrelations estimated from the consensus forecasts and the realized series. The boundaries of the dashed box are determined by actual short-rate and TP autocorrelations. "Autocorrelation averaging" implies that all subjective autocorrelations should be within the box. As is evident in all panels of the plot, across maturities, the majority of the individual blue dots and the consensus orange diamond are located within the box (i.e., $\rho_p < \rho_1^s, \rho_p^{s(n)} < \rho_1$) and a significant mass of dots, gauged by size of the dots, centers around the consensus estimate. Admittedly, the result is slightly weaker in Panel D where some forecasters' subjective autocorrelations of 30-year TP are too high to be bounded by the dashed box. In sum, the relative positions of subjective and actual autocorrelations in Figure 4 strongly support the "autocorrelation-averaging" hypothesis.

Under- and overreaction for EH and TP components. A direct implication of Proposition 1 is that forecasters underreact to new information in EH components and overreact to new information in TP components in FE-on-FR regressions. I formally confirm this prediction in Table 5 at both individual (Panel A) and consensus (Panel B) levels. Compared with Table 2, there are fewer observations, as theye estimation of ρ_1^s and $\rho_p^{s(n)}$ requires 120 months of data to begin with. In Panel A, I obtain positive and significant regression coefficients for the EH components of 2-, 5-, and 10-year bonds; the coefficients are slightly higher than those of the short rates in Table 2. The coefficients of TP regressions across all maturities are significant and negative, exhibiting stronger overreaction than in Table 2. In Panel B, the sign and statistical significance of the coefficient estimates largely persist. The consensus-level regressions show that the FE-on-FR regression coefficients are greater in magnitude for EH components, and smaller for TPs, than the individual-level ones. Consistent with the relationship between *average* subjective and actual autocorrelations, the average forecaster underreacts significantly to the news in EH and overreacts significantly to the news in TP.

4.4 Model calibration

To quantitatively evaluate the performance of the model of "autocorrelation averaging," I calibrate it with the estimated subjective and actual autocorrelations and compare the model-generated FE-on-FR regression coefficients with the empirical ones. Because the estimated subjective autocorrelations are, on average, close at the individual and consensus levels, I use the individual-level median autocorrelation estimates as inputs to the model. Also, because the estimation starts in 1998, I reestimate the FE-on-FR regressions from 1998—which generates slightly different regression coefficients from those in Table 2—to compare with the model predictions. Consistent with the choice made in the autocorrelation estimations, I use the FFR as the short rate and focus on the FE-on-FR regression coefficients for FFR, 2-, 5-, 10-, and 30-year bonds. On the other hand, I compute the model-implied regression coefficients based on Equation (11) and on the realized variance of forecast revisions. Table 6 reports the parameter values and results from the calibration exercise. FE-on-FR coefficients from the model and the data are close. They are positive for FFR and 2-year bond and negative for the rest. We cannot reject that empirical and model-generated coefficients are statistically different for FFR, 10 and 30-year yields.

Figure 5 depicts the calibrated FE-on-FR coefficients against the individual- and consensus-level regression estimates across maturities. The blue circles and orange triangles are the individual- and consensus-level regression coefficients from the data and the green squares are the calibrated counterparts. It is clear from the figure that the calibration generates the same downward-sloping pattern of FE-on-FR regression coefficients. Moreover, except for the 5-year yield, the calibration-generated coefficients are mostly within the 95%-confidence interval of those estimated in the data. The calibration exercise makes no specific assumption about the value of autocorrelations or the relative importance of the EH and TP components for each interest rate. Using the discrepancy between subjective and actual autocorrelations extracted the survey data, it successfully matches the downward-sloping term structure of under- and overreaction and generates FE-on-FR regression coefficients quantitatively close to those from regressions. This exercise highlights the quantitative success of the "autocorrelation averaging" mechanism.

4.5 Determinants of subjective autocorrelations

The underlying psychology of "autocorrelation averaging" posits that professional forecasters and investors are boundedly rational—they have limited cognitive and institutional processing capacity, especially when facing many demanding tasks; therefore, they learn the true time-varying autocorrelations slowly and rely on something closer to an average autocorrelation for their forecasts. An immediate implication of such an interpretation is that when a forecaster's processing capacity is more constrained, one would expect her subjective autocorrelations to deviate further from the true autocorrelations and her misreaction to new information to be exacerbated.

To jointly measure the deviations of subjective autocorrelations of EH and TP components from their respective actual counterparts, I define at time t a *signed* distance, $\|\rho\|_{it}$, between forecaster i's subjective autocorrelations and the true autocorrelations for EH and TP components:

$$\|\rho\|_{it} = \begin{cases} \sqrt{(\rho^s_{1,it} - \rho_{1,t})^2 + (\rho^s_{p,it} - \rho_{p,t})^2}, & \text{if } \rho_{p,t} < \rho^s_{1,it}, \rho^s_{p,it} < \rho_{1,t} \\ -\sqrt{(\rho^s_{1,it} - \rho_{1,t})^2 + (\rho^s_{p,it} - \rho_{p,t})^2}, & \text{otherwise.} \end{cases} \tag{19}$$

I represent the short-rate (equivalently EH) and TP components with FFR and 10-year term premium, respectively. The time variation of $\|\rho\|_{it}$ stems from the rolling-window estimation. The distance is positive when forecaster i conforms to "autocorrelation averaging" and negative otherwise. Thus, a higher value of $\|\rho\|_{it}$ reflects a larger deviation from true autocorrelations in the direction that is consistent with "autocorrelation averaging."

Next, I regress the signed distance $\|\rho\|_{it}$ on forecasters' characteristics and aggregate time-series variables:

$$\|\rho\|_{it} = \alpha_i + \beta X_{i,t} + \gamma Z_t + \epsilon_{it}, \tag{20}$$

where X_{it} are forecaster-level characteristics such as the number of years a forecaster has been participating in the survey (Forecaster Experience) and Z_t includes aggregate variables such as 5-year cumulative absolute monetary policy shocks constructed by Swanson (2021) (Cum MP Shock)[32]; number of months in recession during the past five years (Recession Past 5Y); the 5-year average volatility of 10-year Treasury yields (Avg Yield Volatility)[33]; 5-year average economic policy uncertainty (Avg EPU); and numbers of scheduled/unscheduled Fed meetings and special programs during the past five years (Fed Meetings/Programs).[34] I include forecaster fixed effects to absorb unobserved firm-level heterogeneity in forecasting.

Table 7 reports results on the determinants of forecaster-level autocorrelation deviations. The signed distance—forecasters' deviation from the true autocorrelations—tends to decrease with the forecaster's years of experience (Column 1); it tends to increase during periods when there are bigger monetary policy shocks (Column 2, when the economy is in recession (Column 3), when there is high economic uncertainty (Column 5), and when the Federal Reserve has more unscheduled meetings and special programs (Column 6). Moreover, the signed distance decreases

[32]The absolute size of the monetary policy shock represents uncertainty about monetary policy: the higher the uncertainty, the more constrained a forecaster's processing capacity. In this exercise, the sign of the monetary policy shock is not needed.

[33]The yield volatility is first computed using daily observations within a year and then averaged within a five-year window to obtain Avg Yield Volatility. The regression results remain stable if we instead compute yield volatility using daily observations of 5 years.

[34]I use information during the past five years to account for the fact that the estimation of autocorrelations uses overlapping rolling windows.

when the recent bond yield volatility is high; this relationship, which merits further study, may be due to heightened attention induced by significant yield movements. All of the above effects are strongly statistically significant. Column 7 runs a "kitchen-sink" regression with all determinant variables, where monetary policy shock, average yield volatility, and economic policy uncertainty remain significant, although the length of recent recession flips sign. Taken together, these regression results support a bounded-rationality interpretation—forecasters are more engaged in "autocorrelation averaging" when their processing capacity or attention is limited relative to the difficulty of the forecasting task—and are broadly consistent with findings in previous studies where forecasters make bigger mistakes in recessions (Cieslak, 2018), and allocate resources to different tasks across the business cycles (Kacperczyk, Van Nieuwerburgh, and Veldkamp, 2014).

Given the various determinants of the subjective autocorrelations, we can go one step further and test the bounded-rationality mechanism of "autocorrelation averaging" by showing that the strength of the forecast misreaction depends on these determinants. To do this, I split the sample into two parts based on the binary values (indicator variables such as recessions and 2007-09 financial recessions) or median values (monetary policy shocks, EPU, and yield volatility) and run the previous FE-on-FR regressions for each subsample. Figure A.4 in the Internet Appendix plots the term structure of FE-on-FR regression coefficients with five sets of split samples. The downward-sloping term structure of misreaction—underreaction for short-term rates and overreaction for long-term rates—is evident in all subsamples; however, the strength of misreaction varies. In times when forecasters are more prone to "autocorrelation averaging," namely recession/financial crisis and periods with big cumulative monetary policy shocks, high EPU, or low yield volatility, the absolute values of the CG coefficients are almost always larger than those during the opposite periods, indicating stronger misreactions. These conditional FE-on-FR regressions, though not statistically significant in all specifications, provide suggestive evidence that tightens the link between bounded rationality and forecast misreaction—both under- and overreaction are stronger when the forecaster's processing capacity is more constrained.

5 Belief Misreaction, Portfolio Allocation, and Asset Prices

The pattern of misreaction to new information, as shown in the previous sections, is evident not only at the individual forecaster level but also at the *aggregate* consensus level. BCFF forecasters, whose identities are known, are either major players in the US Treasury market—for example, 17 of them are primary dealers of the New York Fed—or are likely to influence significant market participants through various client services. Given these features, one would naturally expect professional forecasts to represent the beliefs of major Treasury market participants and the pattern of misreaction to manifest itself in related asset prices. In this section, I explore this possibility by examining the impact of forecast misreaction on bond and interest rate futures prices.

5.1 Stated Beliefs and Bank Treasury Portfolios

One prerequisite for such pricing impact from distorted beliefs is that the market participants should *act* in accordance with their stated beliefs by putting their money behind these numbers. To clear this significant hurdle, I start by matching a subset of Blue Chip forecasters—which are banks—to their balance sheet information from the Call Reports. The details of the bank balance sheet data and the matching procedure can be found in the Internet Appendix, but I also describe them briefly here. The variables of interest are each bank's holdings of Treasuries and other fixed-income securities of various maturities. The Call Reports do not record the holdings of individual securities on banks' balance sheets; instead, they group securities by asset class and by maturity range.[35] I match these ranges to the yield forecasts with the closest maturity—for instance, holdings of 5- to 15-year securities are matched with forecasts of 10-year yields.

To guide our intuition for the relationship between yield curve expectations and holdings of Treasuries, I use Merton's (1969) model of portfolio choice as a benchmark. For a power-utility investor in the Treasury bond market, her portfolio weight in the risky bond with n-maturity is determined as

$$w_{i,t} = \frac{1}{\gamma}\frac{\mathbb{E}^S_{i,t}[R_{t+1}] - R_f}{\mathrm{Var}_i[R_t]} = \frac{1}{\gamma}\frac{\mathbb{E}^S_{i,t}\left[ny_t^{(n)} - (n-1)y_{t+1}^{(n-1)} - y_t^{(1)}\right]}{\mathrm{Var}_i[R_t]}, \tag{21}$$

where γ is the coefficient of relative risk aversion and $\mathbb{E}^S_{i,t}\left[y_{t+1}^{(n-1)}\right]$ is her yield forecast for maturity $n-1$. The above formula assumes that there is only a risky bond and a risk-free asset to choose from, which is permissible given the high correlations between Treasury yields. The simple frictionless benchmark implies that, all else equal, the portfolio weight or allocation is

[35] The maturity ranges in the Call Reports are: less than 3-month, 3-month to 1-year, 1- to 3-year, 3- to 5-year, 5- to 15-year and beyond 15-year.

negatively related to the subjective expectation of the future yield

$$w_{i,t} \propto -\mathbb{E}_t^S(y_{t+1}^{(n)}). \tag{22}$$

To test the relationship between portfolio allocations and stated beliefs, given data availability, I use the allocations to Treasuries at a certain maturity as the main holdings measure and estimate the following regression:

$$Treasury(n)_{i,t} = \alpha_i + \beta\mathbb{E}_{i,t}^S(y_{t+h}^{(n)}) + \gamma \cdot X_{i,t} + \varepsilon_{i,t}, \tag{23}$$

where the dependent variable is the dollar amount of Treasuries with maturity close to n years. β is the coefficient of interest; it measures the sensitivity of portfolio allocations to beliefs, i.e., how much an individual bank's allocation changes in dollars at the end of the quarter for each percentage point change in the one-year expected yield. To account for unobserved heterogeneity at the bank level, I include bank fixed effects in the main specification. $X_{i,t}$ is a vector of control variables. The regression is estimated monthly, in which the quarterly allocation data is assigned to all three months within a quarter. I include all the maturity ranges that are greater than one year.

The results are reported in Table 8. Panel A, which fixes the forecast horizon h to 4 quarters for ease of interpretation, directly tests Equation (21). Panel B pools yield expectations across forecast horizons (1-4Q) and obtains slightly stronger results. A key finding emerges from these regressions: the β estimates are negative across all maturities and, except for the 30-year maturity in Panel A, also statistically significant, confirming the prediction from Equation 22. To highlight the economic magnitude of the estimated β-coefficient, I take the 10-year Treasuries (Column 3, Panel A) as an example: a one-standard-deviation increase in the expected one-year-ahead 10-year Treasury yield leads to a \$1.64 billion decrease in the 5- to 15-year Treasuries held by an average bank in the sample, which—given that an average bank holds around \$4 billion in Treasuries in this maturity range—represents a sizable 40% decrease. The negative link between yield forecasts and portfolio allocation also persists in other fixed-income securities where interest rate risk at various maturities is an integral component: In Table A.19 of the Internet Appendix, I provide additional consistent evidence using allocations to tradable securities, residential mortgage-back securities (RMBS), and the total fixed-income assets.

To sum up, I show that banks' allocations to Treasuries and other related fixed-income securities vary positively and significantly with their subjective expectations of bond returns at the corresponding maturities. The effect is substantial for an average bank, meaning that banks take these forecasts seriously and put real money behind them. This robust link between stated beliefs and asset allocations, apart from validating the survey forecasts, offers crucial justification for the potential asset pricing implications and is analogous to evidence from other groups of investors (e.g., Giglio et al., 2021)

5.2 Return predictability from overreaction

The term structure of forecast misreaction has implications for both short-term and long-term bond prices. Recall that belief overreaction is the dominant feature for bonds with a maturity greater than two years, and these are also the bonds that most studies of Treasury bond returns have focused on. I start by testing the following implication of overreaction: when investors revise upwards their forecasts of the long-bond yield following the arrival of new information, they push the bond price down too low (and bond yields too high); subsequently, this price pressure, due to overreaction, gradually subsides and the price is likely to correct to a sensible level. Therefore, this predictable price movement implies that a positive forecast revision of long rates predicts higher bond returns in the future. I empirically confirm this return predictability in Table A.9 in the Internet Appendix.

Furthermore, we can simplify the prediction. Decomposing the variance of forecast revisions reveals that the lagged forecast error (realized at time t) is an important, perhaps the most important, contributor to the variation in forecast revisions. For instance, the lagged forecast errors for 10-year Treasury yields explain 61% of the variation of the forecast revision at the consensus level for the one-quarter horizon and 36% for the one-year horizon:

$$FR_t\left(y_{t+1Q}^{(10)}\right) = -0.11 + \underset{(t=14.10)}{0.56} \times FE_{t-1Q}\left(y_t^{(10)}\right) + \varepsilon_t, \quad R^2 = 0.61.$$

Table A.10 in the Internet Appendix reports the contemporaneous relationship between forecast revisions and lagged forecast errors across maturities for the one-quarter and one-year forecast horizons. Across all interest rates, lagged forecast errors explain a significant portion of the variation in forecast revisions.[36]

There are two reasons that the lagged forecast error may be preferred to forecast revision

[36] Empirically, forecast errors are negatively autocorrelated; this is consistent with overreaction.

as a predictor of bond returns. First, the bond literature focuses on the one-year holding period return for ease of construction and interpretation. Matching the one-year holding period requires survey forecasts with a horizon longer than one year to calculate the forecast revision, which significantly shortens the available sample for the return predictability exercise. Second, the interval at which the forecast revision is calculated is arbitrary; however, this additional degree of freedom is not essential to establish return predictability from overreaction, as other intervals also pick up the same overreaction dynamic. The lagged forecast error does not suffer from these deficiencies. It offers a simpler measure of information updates and retain a long sample that is crucial for statistical inferences in return predictability exercises. I therefore choose the lagged forecast errors as the main predictor and turn the prediction into one that is easier to test:

Prediction 1 (overreaction and subsequent bond returns) *Lagged forecast errors for the long rate should positively predict future Treasury bond returns.*

To test this prediction, I pick one interest rate—the 10-year Treasury yield—as the workhorse long rate, given the strong one-factor structure in Treasury bond yields and returns (Cochrane and Piazzesi, 2009). The 10-year T-note rate is a benchmark bond yield extensively used by both market participants and researchers.[37] Formally, the overreaction-motivated return predictor is the difference between the 10-year Treasury yield at time t and its consensus forecast from one year ago: $FE_{t-1Y}\left(y_t^{(10)}\right)$. For brevity, I label it as $FE10Y_t$. The forecast horizon of one year is picked to match the typically used one-year holding period bond returns. The time-series dynamics of $FE10Y_t$ are plotted in Panel A of Figure 6. The lagged forecast errors for the 10-year Treasury yield exhibit strong countercyclical fluctuations over time at a business-cycle frequency.

I continue by running simple return predictive regressions at the monthly frequency with overlapping one-year holding period excess returns as the dependent variable:

$$rx_{t+1}^{(n)} = \alpha + \beta FE10Y_t + \gamma \cdot X_t + \varepsilon_{t+1}, \tag{24}$$

where $rx_{t+1}^{(n)}$ is the excess return of an n-year bond; $FE10Y_t$ is the main overreaction-motivated return predictor; and X_t is a vector of control variables—such as the first three principal components (PC) of the yield curve—and other bond predictors. To take into account the

[37] For example, Brooks and Moskowitz (2017) use the 10-year yield as an empirical proxy for the yield curve level factor and show that it traces the statistical level factor well.

heteroskedasticity and autocorrelation in the error terms and the overlapping structure of the excess returns on the left-hand side, I construct two robust standard errors: Newey and West (1987) standard errors with 12 lags and Hodrick (1992) standard errors which retain correct size in small samples for overlapping returns. I follow the implementation of Wei and Wright (2013) to calculate the Hodrick standard error by running the "reverse regressions."[38]

Table 9 reports the results of estimating Equation (24) with no control variables (Panel A) and with the first three PCs of the yield curve as controls (Panel B). Specifically, I use $FE10Y$ to predict the future one-year holding period excess returns of Treasury bonds with 2, 3, 5, 7, 20, and 30-year maturities. The dependent variable in the last column is an average excess return weighted by the inverse of bond maturities $\overline{rx}_{t+1} \equiv (1/\sum \frac{1}{n}) \sum \frac{1}{n} rx_{t+1}^{(n)}$ that does not overweight long maturity returns.[39] As is clearly shown in Panel A of Table 9, $FE10Y$ positively and significantly predicts the future excess returns across all maturities, consistent with the prediction of overreaction. The coefficients are all statistically significant under the more stringent Hodrick (1992) standard errors. In particular, $FE10Y$ predicts the average excess return $\overline{rx}_{t+1}$ with an R^2 of 0.25, which means that, in economic magnitude, a one-standard-deviation increase in $FE10Y$ is associated with a 3.38% (98% of the unconditional mean of $\overline{rx}$) increase in the average return of Treasury bonds in the coming year. The statistical significance and magnitude barely change in Panel B when I control for the first three PCs. The noticeable difference is that the two-year regression coefficient becomes less significant, which is mostly consistent with the weaker (or absent) overreaction at the two-year maturity in Section 3. Under the null hypothesis of no additional predictive power beyond the information contained in yields (the "spanning hypothesis"), the three yield curve PCs should span all auxiliary bond return predictors. However, the evidence in Panel B strongly rejects the null of no incremental predictability from $FE10Y$. Bauer and Hamilton (2017) recently point out significant small-sample distortions in the test of the "spanning hypothesis" and cast serious doubt on the additional predictive power of many auxiliary bond return predictors. To address this small-sample issue, I follow their proposed parametric bootstrap procedure to test the null hypothesis of no additional predictability. The test, detailed in Section A.4 of the Internet Appendix, strongly rejects the null hypothesis and confirms the robust predictive power of $FE10Y$.

[38]The reverse regressions use the one-month excess returns on Treasury bonds as the left-hand-side variable. I obtain them by interpolating the available yields using the cubic spline method to get the yield of a bond with maturity $(n - 1/12)$.

[39]Alternatively, constructing $\overline{rx}$ as a simple cross-sectional average generates very similar results.

Excess predictability from overreaction. Next, I show that the "excess" predictability from $FE10Y$ stems from overreaction to information. The underlying mechanism shares some similarities with the one in Cieslak (2018), who finds that a wedge between investors' perceived and realized dynamics of the *short rate* has predictive power for bond returns. Though both papers explore the information embedded in subjective expectations, I focus on investors' overreaction to information in the long rate or, more precisely, the term-premium component. To support this interpretation, I construct the lagged forecast error for the 10-year term premium as $FE_{t-1}\left(tp_t^{(10)}\right)$ using TP constructed in the consensus-level yield decomposition in Section 4.3. Panel A of Table 10 reports the results of the predictive regressions using forecast errors for TP. The sample is shorter due to the rolling-window estimation of the subjective autocorrelations. Nonetheless, the predictive power from TP forecast errors is comparable to that from $FE10Y$, confirming the overreaction origin of return predictability. However, the point estimates are less significant due to the noise introduced in the autocorrelation estimation. As a placebo test, I examine the predictive power of the lagged forecast error for the short rate (e.g., FFR, three-month, or one-year T-bill rates) and the EH components. Panel B reports the results using FFR, where there is barely any predictability. Unreported results show that the same conclusion holds for other short-rate variables.

Robustness. I perform several additional robustness checks concerning the predictive power from overreaction. First, the return predictability captured by $FE10Y$ is distinct from other existing bond return predictors, including a linear combination of forward rates proposed by Cochrane and Piazzesi (2005); the cycle factor (cf) from Cieslak and Povala (2015); a growth factor (GRO) and an inflation factor ($INFL$) from Joslin, Priebsch, and Singleton (2014); and the eight principal components of a large set of macroeconomic variables from Ludvigson and Ng (2009). The magnitude and statistical significance of $FE10Y$ barely change when these alternative bond predictors are added to the predictive regressions. The correlation between $FE10Y$ and the additional predictors and results from the multivariate "horse race" predictive regressions are reported in the Internet Appendix. Second, $FE10Y$ can predict excess returns of coupon bonds across various maturity brackets. Table A.13 in the Internet Appendix reports the prediction results using the CRSP Fama bond portfolios. The results are comparable to those using zero-coupon bonds. Last, the survey-based forecast errors $FE10Y$ contain unique information regarding future bond prices. In Tables A.16 and A.17 in the Internet Appendix, I show that the predictability cannot be replicated with other related forecast error measures such as an econometrician's forecast errors obtained from a forecasting system using macro, monetary, and financial variables (Cieslak, 2018).

5.3 Underreaction and asset prices

I complete this section by testing the implications of underreaction for asset prices. The symmetry in the mechanism of "autocorrelation averaging" imples that, for short-term bonds, investors insufficiently adjust their beliefs (and move prices) in light of new information, leading to a subsequent continuation of price movement in the same direction. This underreaction naturally generates the opposite prediction for short-term bond prices. That is, forecast revision and lagged forecast errors for the short rate should *negatively* predict future short-term bond returns. Calculating the holding period return on a very short-term zero-coupon bond requires additional assumptions regarding interpolation, which may induce additional measurement error. To circumvent this issue, I instead use the changes in yields of six-month and one-year T-bills—which closely resemble short-term bond returns—as the dependent variables. Additionally, I pick coupon bonds with maturities of less than 12 and 24 months from the Fama Maturity Portfolios and test the prediction using short-term coupon bond excess returns. Consistent with the previous sections, I use the Federal Funds Rate (FFR) as the benchmark short rate and define the lagged forecast errors accordingly. To accommodate the short maturities of the bonds, I use the one-quarter horizon to calculate the forecast errors, yield changes, and excess returns. In Table 11, I regress future inverse yield changes (Columns 1 and 2)[40] and the excess returns on the two short-term bond portfolios (Columns 3 and 4) on the time-t lagged forecast errors for FFR. The regressions are estimated at the monthly frequency. The coefficients across the four columns are negative and statistically significant using Newey-West standard errors with three lags. These results confirm the predictions of underreaction.

Underreaction in Federal funds futures. Another natural candidate for testing the predictions of underreaction is the Federal funds futures.[41] The Federal funds futures, one of the most actively traded interest rate futures contracts, are known to well reflect investors'

[40]The inverse yield change $y_t - y_{t+1}$ has the same sign as the bond returns.

[41]Notice that the Treasury bond futures contracts have hypothetical coupon rates; the futures implied yields are different from both the survey forecasts and zero-coupon yields. The Federal funds futures does not have this feature and therefore is suitable for testing the prediction.

information and expectations.[42] Underreaction in beliefs implies that investors only partially incorporate news—which is captured by forecast revisions—into the futures prices. Since the payoff of Federal funds futures is determined by the average effective FFR in a given month, I define the futures-based forecast error for FFR as

$$FE_t^{FUT}\left(ffr_{t+h}\right) = \overline{ffr}_{t+h} - \mathbb{E}_t^{FUT}\left(ffr_{t+h}\right),$$

where $\overline{ffr}_{t+h}$ is the average FFR and $\mathbb{E}_t^{FUT}\left(ffr_{t+h}\right)$ is the end-of-month futures-implied FFR.[43] To be consistent with the previous exercises, I use the survey-based forecast revisions to capture information updates[44] and test underreaction in the Federal funds futures by estimating the following FE-on-FR regression:

$$FE_t^{FUT}\left(ffr_{t+h}\right) = \alpha + \beta FR_t^S\left(ffr_{t+h}\right) + \epsilon_{t,h}, \tag{25}$$

where $FE_t^{FUT}\left(ffr_{t+h}\right)$ is the futures-based forecast error and $FR_t^S\left(ffr_{t+h}\right)$ is the *survey*-based forecast revision constructed using consensus forecasts. Table 12 reports the findings. I estimate Regression (25) at a quarterly frequency by pooling across forecast horizons (Column 1) and separately for horizons from one to four quarters (Columns 2–5). The coefficients across columns are positive and statistically significant, which strongly support the prediction of underreaction in the Federal funds futures market. The point estimates of the FE-on-FR coefficient are larger in absolute terms than those from the individual- and consensus-level forecasts in part because the futures forecast errors are much more volatile than those of survey forecasts.

6 Conclusion

In this paper, I investigate how measured beliefs from professional forecasts of interest rates across the entire yield curve respond to new information. I document a robust downward-sloping term structure of misreaction in both individual- and consensus-level forecasts: forecasters

[42] The futures data are available via Datastream, where I obtain daily settlement prices and resample them to the monthly and quarterly frequencies. Federal funds futures have one contract for each month and liquidity concern is minimized for contracts that expire within a year.

[43] Consistent with Cieslak (2018), the statistical properties of futures- and survey-based expectations are similar for FFR—the correlation between survey and futures-based forecast errors is around 0.6. The complete summary statistics are reported in the Internet Appendix.

[44] Defining forecast revision using Federal funds futures yields similar results.

underreact to news for short-term bonds and overreact for long-term bonds. I present a bounded-rationality model of belief formation based on "autocorrelation averaging." It is rooted in a key observation that short rate and term premia, building blocks of the Treasury yield curve, exhibit different levels of persistence cross-sectionally and over time. Investors, facing these disparate time series in real time, may lack the cognitive processing capacity or institutional resources to learn the true autocorrelation of each series. Instead, when forecasting, they use perceived autocorrelations that are closer to an average of the true autocorrelations of all the series they are exposed to. In doing so, they overreact to less persistent term-premium processes and underreact to more persistent short-rate processes. My theoretical explanation is strongly supported in the data, where investors' perceived autocorrelations, structurally estimated from their forecasts, are compressed toward the average. This departure from rational expectations is parsimonious and small in absolute magnitude, yet when calibrated to the estimated perceived autocorrelations, the model quantitatively matches the downward-sloping term structure of misreaction.

Since professional forecasts are likely to represent the beliefs of market participants, under- and overreaction in interest rate expectations have immediate predictions for asset prices: an overreaction-motivated predictor, the lagged forecast errors for 10-year Treasury yields, robustly forecasts future excess bond returns. I also confirm the analogous prediction of underreaction for short-term bonds and Federal funds futures prices.

The "autocorrelation averaging" behavior has its root in limited attention. I show that, in line with the inattention interpretation, when people's processing capacity is more constrained relative to the difficulty of their forecasting task—such as during recessions or periods with substantial monetary policy shocks—they are more prone to "autocorrelation averaging" and the pattern of belief misreaction is considerably stronger. The origin of such inattention could be psychological or institutional. My model of distorted beliefs is "portable" and can be taken to other contexts where investors, who are boundedly rational, form expectations about different time series. Here, I focus on misreaction in interest rate forecasts across maturities, but "autocorrelation averaging" may provide useful guidance for reconciling similar phenomena in other asset markets. For example, it could be applied to the aggregate stock market, unifying underreaction in short-term earnings forecasts and overreaction in return expectations.

Table 1 Summary statistics and out-of-sample performance of the interest rate forecasts

Panels A and B report summary statistics of the individual-level forecast errors and forecast revisions. The results are pooled across forecast horizons h. The last rows of Panels A and B are medians of forecast errors and forecast revisions, normalized by the contemporaneous interest rates. Panels C and D evaluate the out-of-sample (OOS) performance of survey forecasts against alternative models, including moving average ($Mean$), $AR(1)$, $AR(p)$, and ARIMA(1,1,0) ($ARIMA$). Panel C reports median OOS R^2 of the individual forecasts and Panel D reports OOS R^2 of the consensus forecasts. The underlying variables are the Federal Funds Rate (ffr) and the Treasury bill, note, and bond yields with maturities of 3 months, 1, 2, 5, 10 and 30 years ($tb3m$, $tb1y$, $tn2y$, $tn5y$, $tn10y$ and $tn30y$). The data cover 1988–2018.

	ffr	$tb3m$	$tb1y$	$tn2y$	$tn5y$	$tn10y$	$tn30y$
Panel A: Individual forecast errors $FE_{i,t}\,(x_{t+h})$, pooled across horizons							
Count	23005	22836	21392	23041	22966	23203	22361
Mean	-0.26	-0.29	-0.31	-0.40	-0.36	-0.19	-0.18
SD	0.90	0.93	0.98	0.93	0.84	0.77	0.71
Min	-5.07	-4.95	-4.66	-4.71	-3.59	-3.71	-3.86
p25	-0.52	-0.64	-0.76	-0.88	-0.87	-0.66	-0.61
p50	-0.06	-0.12	-0.17	-0.27	-0.32	-0.21	-0.17
p75	0.10	0.13	0.18	0.10	0.14	0.27	0.25
Max	6.22	4.36	3.89	3.60	3.45	4.60	6.39
p50 (Normalized)	-0.05	-0.10	-0.08	-0.10	-0.08	-0.04	-0.03
Panel B: Individual forecast revisions $FR_{i,t}\,(x_{t+h})$, pooled across horizons							
Count	20852	20406	18831	20440	20346	20613	19821
Mean	-0.14	-0.15	-0.15	-0.16	-0.15	-0.14	-0.12
SD	0.60	0.59	0.61	0.61	0.59	0.53	0.49
Min	-6.00	-5.50	-4.80	-4.30	-5.10	-6.10	-6.00
p25	-0.34	-0.38	-0.40	-0.41	-0.42	-0.40	-0.40
p50	0.00	-0.02	-0.06	-0.09	-0.10	-0.10	-0.10
p75	0.10	0.10	0.14	0.15	0.20	0.15	0.15
Max	6.30	6.00	2.50	2.80	5.89	5.60	5.20
p50 (Normalized)	0.00	-0.02	-0.03	-0.03	-0.03	-0.02	-0.02
Panel C: Median OOS R^2 of individual forecasts							
$R^2_{OOS,Mean}$	0.72	0.73	0.69	0.73	0.78	0.86	0.72
$R^2_{OOS,AR(1)}$	0.16	-0.07	-0.05	-0.15	-0.23	-0.15	-0.28
$R^2_{OOS,AR(p)}$	0.27	-0.03	0.02	-0.13	-0.19	-0.15	-0.28
$R^2_{OOS,ARIMA}$	0.04	0.04	-0.11	-0.16	-0.33	-0.44	-0.85
Panel D: OOS R^2 of consensus forecasts							
$R^2_{OOS,Mean}$	0.86	0.81	0.83	0.84	0.87	0.90	0.79
$R^2_{OOS,AR(1)}$	0.23	-0.03	0.06	-0.04	-0.12	0.00	-0.12
$R^2_{OOS,AR(p)}$	0.34	0.00	0.13	-0.02	-0.06	0.00	-0.12
$R^2_{OOS,ARIMA}$	0.11	0.06	-0.02	-0.13	-0.22	-0.24	-0.65

Table 2 Forecast error on forecast revision regression results for interest rates across maturities

This table reports the coefficients from the forecast error on the forecast revision regression of Coibion and Gorodnichenko (2015) for each interest rate:

$$FE_{i,t}\left(x_{t+h}\right)=\alpha_i+\beta FR_{i,t}\left(x_{t+h}\right)+\epsilon_{i,t,h},$$

where the forecasts are pooled across horizon h. Panel A reports the baseline results using individual-level forecasts. Standard errors are clustered by both forecaster and time; forecaster fixed effects are included. Panel B reports the results using consensus-level forecasts. Standard errors are calculated following Driscoll and Kraay (1998). The underlying variables are the Federal Funds Rate (ffr), Treasury bill, note and bond yields with maturities of 3 months, 1, 2, 5, 10 and 30 years ($tb3m$, $tb1y$, $tn2y$, $tn5y$, $tn10y$ and $tn30y$). The data are quarterly and cover 1988–2018. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

	Dependent variable: $FE_{i,t}\left(x_{t+h}\right)$						
	ffr	$tb3m$	$tb1y$	$tn2y$	$tn5y$	$tn10y$	$tn30y$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Individual forecasts							
$FR_{i,t}\left(x_{t+h}\right)$	0.30***	0.25***	0.16**	0.03	-0.17**	-0.23***	-0.26***
	(0.07)	(0.07)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)
N	20,603	20,406	18,831	20,440	20,346	20,613	19,821
R^2	0.09	0.08	0.07	0.04	0.07	0.10	0.13
Panel B: Consensus forecasts							
$FR_{i,t}\left(x_{t+h}\right)$	0.64***	0.63***	0.48***	0.26***	-0.04	-0.15***	-0.19***
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.05)	(0.03)
Constant	-0.17***	-0.20***	-0.26***	-0.37***	-0.38***	-0.22***	-0.22***
	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)	(0.07)	(0.06)
N	453	459	456	456	456	456	459
R^2	0.15	0.12	0.07	0.02	-0.002	0.01	0.01

Table 3 Forecast error on forecast revision regression results for extended short and long rates at the individual level

This table reports the coefficients from the forecast error on the forecast revision regression of Coibion and Gorodnichenko (2015) for each interest rate:

$$FE_{i,t}\left(x_{t+h}\right) = \alpha_i + \beta FR_{i,t}\left(x_{t+h}\right) + \epsilon_{i,t,h},$$

where the individual-level forecasts are pooled across horizon h, standard errors are clustered by both forecaster and time, and forecaster fixed effects are included. Panel A reports results for short-maturity interest rates. Panel reports results for long-maturity interest rates. The underlying variables are the Federal Funds Rate (ffr), Treasury bill, note and bond yields with maturities of 3 months, 1, 2, 5, 10 and 30 years ($tb3m$, $tb1y$, $tn2y$, $tn5y$, $tn10y$ and $tn30y$), one-month commercial paper rate ($cp1m$), prime bank rate (pr), three-month LIBOR rate ($libor$), Aaa and Baa corporate bond yields (aaa and baa) and home mortgage rate (hmr). The data are quarterly and cover 1988–2018. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

	Dependent variable: $FE_{i,t}\left(x_{t+h}\right)$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Short-maturity interest rates							
	ffr	$tb3m$	$tb1y$	$tn2y$	$cp1m$	pr	$libor$
$FR_{i,t}\left(x_{t+h}\right)$	0.30***	0.25***	0.16**	0.03	0.33***	0.26***	0.23***
	(0.07)	(0.07)	(0.08)	(0.07)	(0.09)	(0.07)	(0.08)
N	20,603	20,406	18,831	20,440	12,384	20,068	18,522
R^2	0.09	0.08	0.07	0.04	0.08	0.08	0.06
Panel B: Long-maturity interest rates							
	$tn5y$	$tn10y$	$tn30y$	aaa	baa	hmr	
$FR_{i,t}\left(x_{t+h}\right)$	-0.17**	-0.23***	-0.26***	-0.20***	-0.24***	-0.19***	
	(0.07)	(0.07)	(0.07)	(0.06)	(0.07)	(0.06)	
N	20,346	20,613	19,821	18,184	10,925	19,160	
R^2	0.07	0.10	0.13	0.13	0.11	0.11	

Table 4 Summary statistics of short-rate and term-premium autocorrelation estimates

This table reports subjective and actual autocorrelation estimates of the short rate and term premia. Panel A summarizes short-rate subjective autocorrelations (ρ_1^s) at the individual and consensus levels (lines 1 and 2), and actual short-rate autocorrelations ρ_1 (line 3). Panels B and C summarize estimates of the subjective autocorrelation of the term premia (ρ_p^s) at the individual (Panel B) and consensus (Panel C) levels. Panel D summarizes estimates of the actual autocorrelation of the term premia (ρ_p). Term premia have maturities of 2, 5, 10, and 30 years. Details of the estimation are in Section 4.3.

	Count	Mean	SD	Min	p25	p50	p75	Max
Panel A: Short rate autocorrelation estimates ρ_1^s								
ρ_1^s Individual	11106	0.92	0.17	-1.16	0.93	0.95	0.97	1.10
ρ_1^s Consensus	314	0.89	0.23	-0.05	0.94	0.96	0.96	1.01
ρ_1 Actual	314	0.97	0.06	0.64	0.96	0.98	1.00	1.16
Panel B: Term premium autocorrelation estimates: Individual subjective $\rho_{p,i}^s$								
2Y	6499	0.91	0.08	-2.03	0.88	0.92	0.95	2.18
5Y	6499	0.93	0.07	0.22	0.90	0.93	0.96	2.59
10Y	6499	0.94	0.08	-2.14	0.91	0.94	0.97	1.63
30Y	6499	0.95	0.07	-0.82	0.93	0.96	0.99	1.54
Panel C: Term premium autocorrelation estimates: Consensus subjective $\rho_{p,con}^s$								
2Y	314	0.91	0.02	0.81	0.90	0.91	0.92	0.95
5Y	314	0.93	0.02	0.90	0.92	0.92	0.93	1.21
10Y	314	0.94	0.03	0.90	0.93	0.94	0.95	1.24
30Y	314	0.96	0.06	0.90	0.95	0.96	0.97	1.95
Panel D: Term premium autocorrelation estimates: Actual ρ_p								
2Y	309	0.69	0.13	0.00	0.69	0.73	0.75	0.80
5Y	309	0.74	0.13	0.01	0.71	0.78	0.82	0.85
10Y	309	0.77	0.16	0.02	0.73	0.83	0.86	0.90
30Y	309	0.77	0.18	0.02	0.74	0.83	0.88	0.93

Table 5 Forecast error on forecast revision regression results for expectations hypothesis (EH) and term premium (TP) components

This table reports coefficients from the forecast error on the forecast revision regression of Coibion and Gorodnichenko (2015) for expectations hypothesis (EH) and term premium (TP) components:

$$FE_{i,t}\left(x_{t+h}\right) = \alpha_i + \beta FR_{i,t}\left(x_{t+h}\right) + \epsilon_{i,t,h},$$

where the forecasts are pooled across horizon h. Panel A reports the results using individual-level forecasts. Standard errors are clustered by both forecaster and time, and forecaster fixed effects are included. Panel B reports the results using consensus-level forecasts. Standard errors are calculated following Driscoll and Kraay (1998). The underlying variables are expectations hypothesis (EH) and term premium (TP) components with maturities of 2, 5, 10, and 30 years. For each maturity n, forecasts of EH and TP are constructed by decomposing yield forecasts using the estimated subjective short-rate autocorrelation, and realized EH and TP are constructed by decomposing realized yields using the estimated actual short-rate autocorrelation. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

	Dependent variable: $FE_{i,t}\left(x_{t+h}\right)$							
	$eh^{(2)}$	$eh^{(5)}$	$eh^{(10)}$	$eh^{(30)}$	$tp^{(2)}$	$tp^{(5)}$	$tp^{(10)}$	$tp^{(30)}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Individual level EH and TP								
$FR_{i,t}\left(x_{t+h}\right)$	0.31***	0.27***	0.22**	0.11	-0.21***	-0.20***	-0.17**	-0.20***
	(0.10)	(0.09)	(0.09)	(0.08)	(0.06)	(0.07)	(0.07)	(0.07)
N	11,628	11,628	11,628	11,628	11,439	11,340	11,427	10,703
R^2	0.08	0.08	0.08	0.10	0.05	0.04	0.03	0.04
Panel B: Consensus level EH and TP								
$FR_t\left(x_{t+h}\right)$	0.56***	0.52***	0.42***	0.32***	-0.05	-0.07**	-0.06**	-0.10***
	(0.05)	(0.05)	(0.05)	(0.06)	(0.03)	(0.03)	(0.03)	(0.04)
Constant	-0.01	-0.02	-0.04	-0.04**	-0.004	-0.02	-0.03	-0.05**
	(0.07)	(0.06)	(0.04)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
N	371	371	371	371	371	371	371	371
R^2	0.08	0.07	0.05	0.03	-0.002	-0.001	-0.001	0.001

Table 6 Model calibration with individual-level forecasts

This table reports the results from a calibration exercise of the "autocorrelation averaging" model. The model is calibrated with individual-level median estimates of the subjective and actual autocorrelations (reported in the left two sections). The empirical and model-generated (Coibion and Gorodnichenko, 2015) coefficients for each interest rate are reported in the right section, where the last column presents p-values of a test that the empirical and model-generated coefficients are statistically different. The calibration uses the Federal funds rate (ffr) as short rate and Treasury yields with maturities of 2, 5, 10 and 30 years ($tn2y$, $tn5y$, $tn10y$ and $tn30y$) as long rates.

	Subjective autocorrelation		Actual autocorrelation		FE-on-FR coefficients		
	Short rate ρ_1^s	Term premium $\rho_p^{s(n)}$	Short rate ρ_1	Term premium ρ_p	Data	Model	$p(\text{Data} \neq \text{Model})$
ffr	0.95		0.98		0.36	0.40	0.67
$tn2y$		0.92		0.73	0.05	0.28	0.03
$tn5y$		0.93		0.78	-0.21	0.07	0.00
$tn10y$		0.94		0.83	-0.30	-0.15	0.12
$tn30y$		0.96		0.83	-0.32	-0.42	0.23

Table 7 Determinants of time-varying subjective autocorrelations

This table relates forecasters' time-varying subjective autocorrelations to various forecaster-level and aggregate-level characteristics. The dependent variable is the signed distance $\|\rho\|_{it}$ between forecaster i's subjective autocorrelations and the true autocorrelations for FFR and 10-year term premium:

$$\|\rho\|_{it} = \begin{cases} \sqrt{(\rho^s_{1,it} - \rho_{1,t})^2 + (\rho^s_{p,it} - \rho_{p,t})^2}, & \text{if } \rho_{p,t} < \rho^s_{1,it}, \rho^s_{p,it} < \rho_{1,t} \\ -\sqrt{(\rho^s_{1,it} - \rho_{1,t})^2 + (\rho^s_{p,it} - \rho_{p,t})^2}, & \text{otherwise.} \end{cases}$$

The explanatory variables include forecaster experience (in years); 5-year cumulative absolute monetary policy shocks constructed by Swanson (2021) (Cum MP Shock); number of months in recession during the past 5 years (Recession Past 5Y); 5-year average volatility of 10-year Treasury yields (Avg Yield Volatility); 5-year average economic policy uncertainty (Avg EPU); and numbers of scheduled/unscheduled Fed meetings and special programs during the past 5 years (Fed Meetings/Programs). Standard errors are clustered by both forecaster and time, and forecaster fixed effects are included. The data are monthly and cover 1993–2018. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

	Dependent variable: Signed distance $\|\rho\|_{it}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Forecaster Experience	-0.01***						0.01
	(0.003)						(0.004)
Cum MP Shock		0.53***					0.63***
		(0.17)					(0.15)
Recession Past 5Y			0.01***				-0.00**
			(0.001)				(0.001)
Avg Yield Volatility				-9.50***			-41.40***
				(2.4)			(9.1)
Avg EPU					0.11**		1.00***
					(0.05)		(0.11)
Fed Meetings/Programs						0.01***	0.00
						(0.0007)	(0.002)
Forecaster FE	✓	✓	✓	✓	✓	✓	✓
R^2	0.40	0.20	0.31	0.32	0.29	0.36	0.61
N	5,569	5,023	5,569	5,569	5,569	5,373	5,023

Table 8 Subjective beliefs and banks' Treasury portfolio allocation: Regression by maturity

This table reports results from regressing banks' Treasury allocations on survey forecasts at the monthly frequency.

$$\text{Treasury}(n)_{i,t} = \alpha_i + \beta \mathbb{E}^S_{i,t}(y^{(n)}_{t+h}) + \gamma X_{i,t} + \varepsilon_{i,t}$$

The dependent variables are bank i's dollar allocations to the US Treasury with maturities of 1-3 years, 3-5 years, 5-15 years, and over 15 years. The independent variables are bank i's yield forecasts with the closest maturities. Monthly forecasts within each quarter are matched with quarter-end allocations. Panel A fixes the forecast horizon to 4 quarters, and Panel B pools across forecast horizons. Standard errors are clustered by firm and month. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

	Dependent variable:			
	Treasury$(1-3Y)$	Treasury$(3-5Y)$	Treasury$(5-15Y)$	Treasury$(>15Y)$
	(1)	(2)	(3)	(4)
Panel A: $h = 4Q$				
$tn2y$	-0.40**			
	(0.18)			
$tn5y$		-0.50*		
		(0.30)		
$tn10y$			-1.31**	
			(0.64)	
$tn30y$				-1.21
				(0.78)
Firm FE	✓	✓	✓	✓
N	2,581	2,570	2,583	2,371
R^2	0.61	0.52	0.57	0.64
Panel B: $h = 1, 2, 3, 4Q$				
$tn2y$	-0.77*			
	(0.46)			
$tn5y$		-0.67**		
		(0.33)		
$tn10y$			-2.27*	
			(1.17)	
$tn30y$				-1.99*
				(1.05)
Firm FE	✓	✓	✓	✓
N	13,342	13,290	13,365	12,289
R^2	0.48	0.55	0.57	0.57

Table 9 Predicting one-year excess bond returns with overreaction-motivated predictor $FE10Y$

This table presents results of the predictive regressions of one-year bond excess returns on the overreaction-motivated predictor $FE10Y$:

$$rx_{t+1}^{(n)} = \alpha + \beta FE10Y_t + \gamma \cdot X_t + \varepsilon_{t+1},$$

where $rx_{t+1}^{(n)}$ is the one-year holding period excess return of an n-year bond and $\overline{rx}_{t+1}$ is the average excess return weighted by the inverse of bond maturities. Panel A includes no additional independent variable and Panel B includes the first three yield curve principal components (PCs). T-statistics are reported for two types of standard errors: Newey and West (1987) standard errors with 12 lags (in parentheses) and Hodrick (1992) standard errors obtained from reverse regressions (in brackets). The data are monthly and cover 1988–2018. The results for the intercept are omitted.

	$rx_{t+1}^{(2)}$	$rx_{t+1}^{(3)}$	$rx_{t+1}^{(5)}$	$rx_{t+1}^{(7)}$	$rx_{t+1}^{(10)}$	$rx_{t+1}^{(20)}$	$rx_{t+1}^{(30)}$	$\overline{rx}_{t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: No controls								
$FE10Y_t$	0.55	1.11	2.12	3.04	4.33	7.82	10.23	3.81
	(4.68)	(5.14)	(5.57)	(5.72)	(5.68)	(5.15)	(4.36)	(5.69)
	[4.01]	[3.80]	[3.51]	[3.41]	[3.35]	[3.13]	[2.80]	[3.21]
N	348	348	348	348	348	348	348	348
R^2	0.14	0.16	0.18	0.20	0.22	0.25	0.20	0.25
Panel B: Controlling for PCs								
$FE10Y_t$	0.23	0.60	1.42	2.36	3.92	8.88	12.94	3.78
	(1.78)	(2.47)	(3.33)	(4.11)	(5.19)	(6.45)	(6.04)	(5.90)
	[1.87]	[2.10]	[2.38]	[2.67]	[3.03]	[3.39]	[3.22]	[3.35]
PCs	✓	✓	✓	✓	✓	✓	✓	✓
N	348	348	348	348	348	348	348	348
R^2	0.27	0.28	0.33	0.38	0.43	0.45	0.35	0.42

Table 10 Predicting one-year excess bond returns with lagged forecast errors for short rates and 10-year term premium

This table presents results of the predictive regressions of one-year bond excess returns on the short rate and 10-year term premium:

$$rx_{t+1}^{(n)} = \alpha + \beta FE_{t-1}(z_t) + \gamma' PC_t + \varepsilon_{t+1}, \quad z \in \{tp^{(10)}, ffr\},$$

where $rx_{t+1}^{(n)}$ is the one-year holding period excess return of an n-year bond and $\overline{rx}_{t+1}$ is the average excess return weighted by the inverse of bond maturities. Panels A and B report results of the 10-year term premium and the Federal Funds Rate, respectively. The first three yield curve principal components (PCs) are included in both panels. T-statistics are reported for two types of standard errors: Newey and West (1987) standard errors with 12 lags (in parentheses) and Hodrick (1992) standard errors obtained from reverse regressions (in brackets). The data are monthly and cover 1988–2018. The results for the intercept are omitted.

	$rx_{t+1}^{(2)}$	$rx_{t+1}^{(3)}$	$rx_{t+1}^{(5)}$	$rx_{t+1}^{(7)}$	$rx_{t+1}^{(10)}$	$rx_{t+1}^{(20)}$	$rx_{t+1}^{(30)}$	$\overline{rx}_{t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: 10-year term premium								
$FE_{t-1}\left(tp_t^{(10)}\right)$	0.42	0.94	1.94	2.91	4.22	6.75	7.15	3.34
	(3.65)	(4.03)	(4.64)	(5.08)	(5.25)	(3.89)	(2.14)	(4.53)
	[2.72]	[2.64]	[2.49]	[2.37]	[2.20]	[1.78]	[1.36]	[1.84]
PCs	✓	✓	✓	✓	✓	✓	✓	✓
N	273	273	273	273	273	273	273	273
R^2	0.10	0.14	0.18	0.22	0.24	0.21	0.10	0.23
Panel B: Federal Funds Rate								
$FE_{t-1}(ffr_t)$	-0.18	-0.38	-0.70	-0.91	-1.05	-0.65	0.34	-0.60
	(-1.37)	(-1.52)	(-1.69)	(-1.70)	(-1.56)	(-0.59)	(0.19)	(-1.06)
	[-1.09]	[-1.11]	[-1.05]	[-0.97]	[-0.83]	[-0.32]	[0.12]	[-0.35]
PCs	✓	✓	✓	✓	✓	✓	✓	✓
N	360	360	360	360	360	360	360	360
R^2	0.04	0.05	0.05	0.04	0.03	0.00	0.00	0.02

Table 11 Predicting future short-maturity bond returns with underreaction-motivated predictor

This table presents results from testing the asset pricing prediction of underreaction:

$$Z_{t+1Q} = \alpha + \beta FE_{t-1Q}(ffr_t) + \varepsilon_{t+1Q},$$

where the left-hand-side variables in Columns 1 and 2 are the one-quarter yield changes of one-year and six-month T-bills and in Columns 3 and 4 are one-quarter holding period excess returns of the coupon bond portfolios with less than 12-month and 24-month maturities, respectively. The underreaction-motivated predictor is the lagged forecast error for the Federal Funds Rate $FE_{t-1Q}(ffr_t)$. T-statistics based on Newey and West (1987) standard errors with 3 lags are reported in parentheses. The data are monthly and cover the period 1982 to 2018. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

$Z_{t+1Q} =$	$y_t^{(1)} - y_{t+1Q}^{(1)}$	$y_t^{(6m)} - y_{t+1Q}^{(6m)}$	$rx < 12m$	$rx < 24m$
	(1)	(2)	(3)	(4)
$FE_{t-1Q}(ffr_t)$	-0.19***	-0.23***	-0.05*	-0.19**
	(-2.64)	(-3.33)	(-1.76)	(-2.01)
Constant	0.02	0.02	0.11***	0.27***
	(0.56)	(0.48)	(6.03)	(4.48)
N	428	428	428	428
R^2	0.05	0.08	0.02	0.03

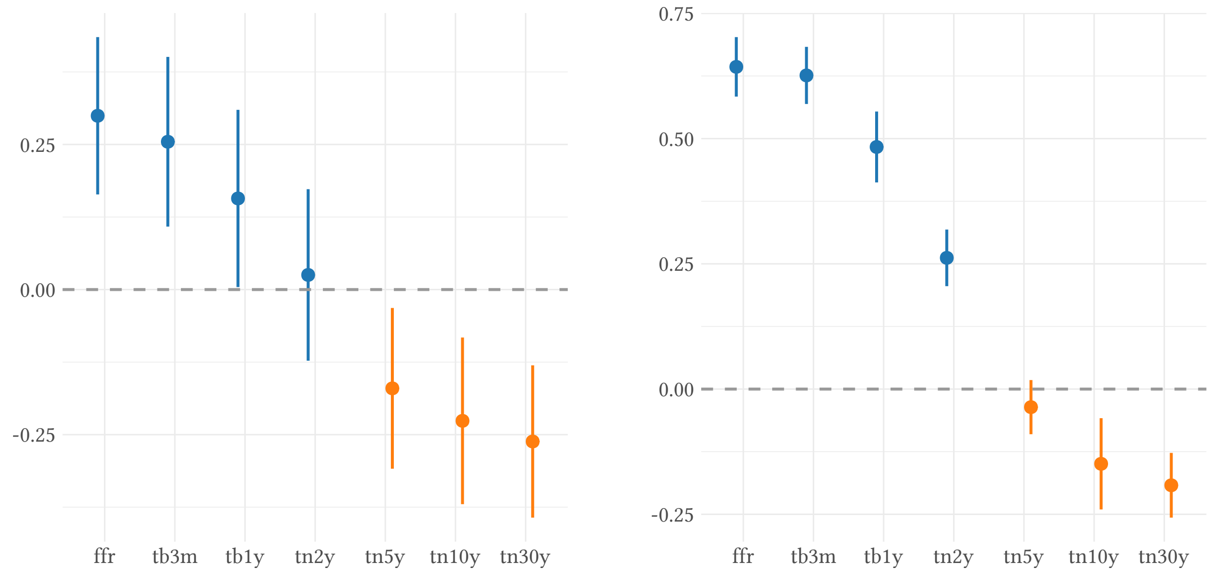
Table 12 Predicting futures-based forecast errors with survey-based forecast revisions

This table presents results of the predictive regressions of futures-based forecast errors on survey-based forecast revisions of the Federal Funds Rate:

$$FE_t^{FUT}(ffr_{t+h}) = \alpha + \beta FR_t^S(ffr_{t+h}) + \epsilon_{t,h}.$$

The regressions use consensus-level forecasts. In Column 1, observations are pooled across horizon h. Columns 2 to 5 report results for each forecast horizon. The futures-based forecast error for FFR is defined as $FE_t^{FUT}(ffr_{t+h}) \equiv \overline{ffr}_{t+h} - \mathbb{E}_t^{FUT}(ffr_{t+h})$, where $\overline{ffr}_{t+h}$ is the within-month average FFR and $\mathbb{E}_t^{FUT}(ffr_{t+h})$ is the end-of-month futures implied FFR. Standard errors are calculated following Driscoll and Kraay (1998) and reported in parentheses. The data are quarterly and cover 2002–2018. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

	Dependent variable: $FE_t^{FUT}(x_{t+h})$				
	Pooled	$h = 1Q$	$h = 2Q$	$h = 3Q$	$h = 4Q$
	(1)	(2)	(3)	(4)	(5)
$FR_t^S(x_{t+h})$	1.24**	1.19**	0.68**	1.45***	1.49**
	(0.49)	(0.60)	(0.29)	(0.56)	(0.65)
Constant	-0.22	-0.19	-0.24	-0.14	-0.29
	(0.14)	(0.13)	(0.15)	(0.12)	(0.18)
N	258	68	66	63	61
R^2	0.14	0.12	0.03	0.21	0.15



A. Individual level

B. Consensus level

Figure 1 Forecast error on forecast revision regression coefficients for short- and long-maturity interest rates: Baseline evidence

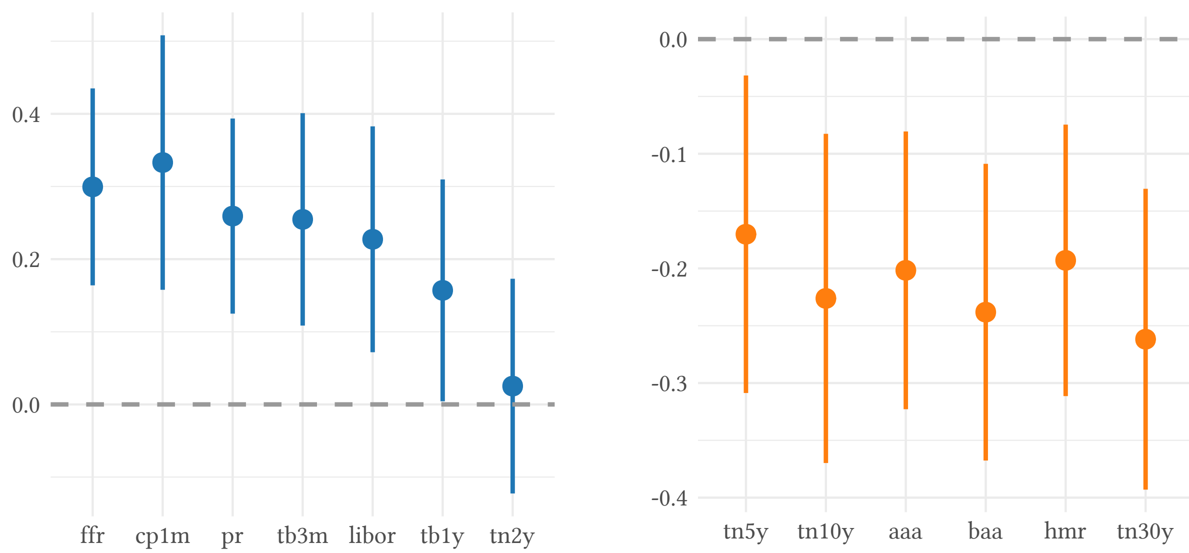
Panel A plots the coefficients from the forecast error on the forecast revision regression of Coibion and Gorodnichenko (2015, CG) for each interest rate using individual-level forecasts:

$$FE_{i,t}\left(x_{t+h}\right) = \alpha_i + \beta FR_{i,t}\left(x_{t+h}\right) + \epsilon_{i,t,h},$$

where the forecasts are pooled across horizon h, standard errors are clustered by both forecaster and time, and forecaster fixed effects are included. Panel B plots FE-on-FR regression coefficients using consensus-level forecasts:

$$FE_t\left(x_{t+h}\right) = \alpha + \beta FR_t\left(x_{t+h}\right) + \epsilon_{t,h},$$

where the forecasts are pooled across horizon h and standard errors are calculated following Driscoll and Kraay (1998). The underlying variables are the Federal Funds Rate (ffr) and the Treasury bill, note and bond yields with maturities of 3 months, 1, 2, 5, 10 and 30 years ($tb3m$, $tb1y$, $tn2y$, $tn5y$, $tn10y$ and $tn30y$). The data are quarterly and cover 1988–2018. The range of each whisker depicts the 95%-confidence interval.



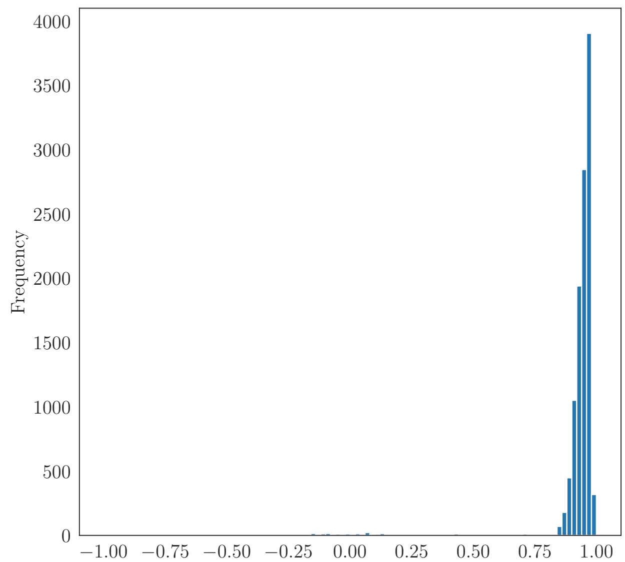
A. Short Rates **B.** Long Rates

Figure 2 Forecast error on forecast revision regression coefficients for short- and long-maturity interest rates: Additional individual-level evidence

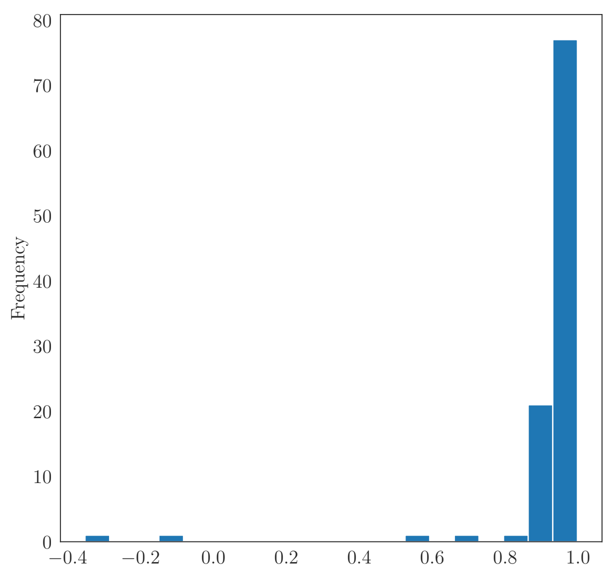
This figure plots the coefficients from the forecast error on the forecast revision regression of Coibion and Gorodnichenko (2015) for an extended list of interest rates using individual-level forecasts:

$$FE_{i,t}\left(x_{t+h}\right) = \alpha_i + \beta FR_{i,t}\left(x_{t+h}\right) + \epsilon_{i,t,h},$$

where the forecasts are pooled across horizon h, standard errors are clustered by both forecaster and time, and forecaster fixed effects are included. The underlying variables are divided into short rates (Pane A) and long rates (Panel B) and ordered by their maturities in each panel. Short rates include Federal Funds Rate (ffr) and Treasury yields with maturities of 3-month and 1-, 2-years ($tb3m$, $tb1y$ and $tn2y$), one-month commercial paper rate ($cp1m$), prime bank rate (pr), three-month LIBOR rate ($libor$). Long rates include Treasury yields with maturities of 5, 10, and 30 years ($tn5y$, $tn10y$ and $tn30y$), Aaa and Baa corporate bond yields (aaa and baa), and home mortgage rate (hmr). The range of each whisker depicts the 95%-confidence interval.



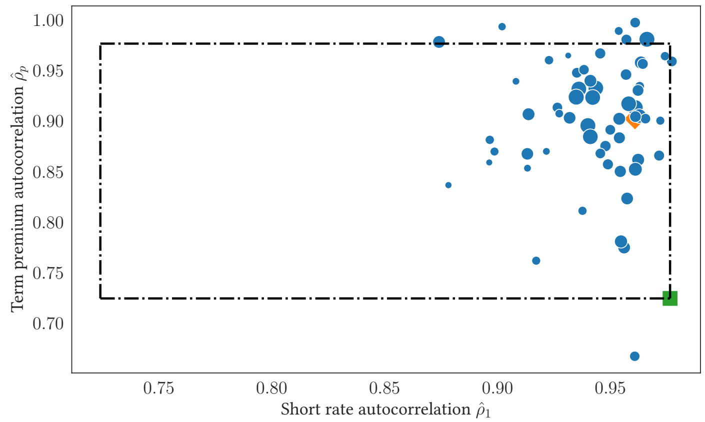
A. All ρ_1^s estimates



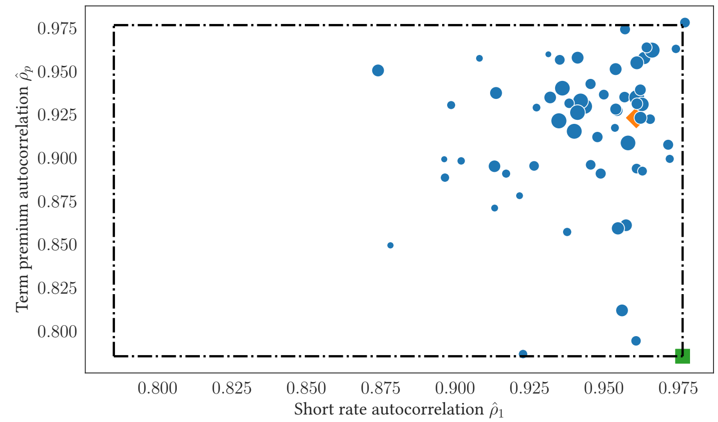
B. Forecaster-level median ρ_1^s estimates

Figure 3 Histograms of the short-rate subjective autocorrelation ρ_1^s estimates

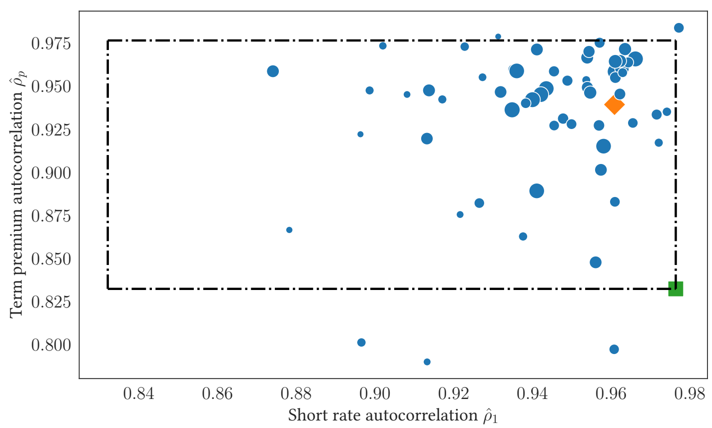
This figure summarizes GMM estimation results of the short-rate subjective autocorrelation ρ_1^s. Subjective autocorrelations are estimated forecaster-by-forecaster and on a 120-month rolling basis. Panel A plots all ρ_1^s estimates across time and forecasters. Panel B plots forecaster-level median ρ_1^s estimates. The details of the estimation are in Section 4.3.



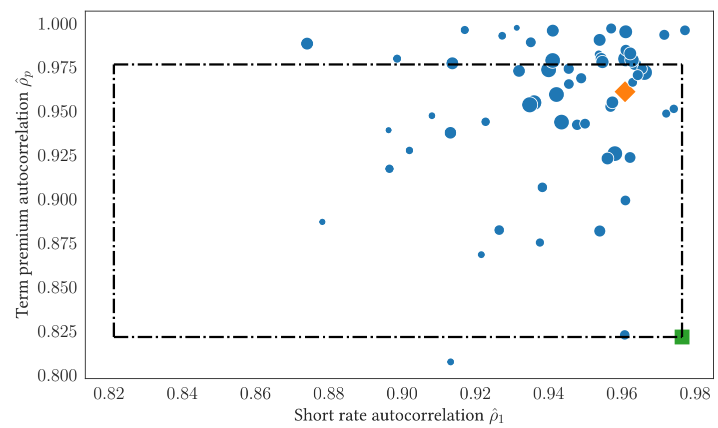
A. 2-year term premium



B. 5-Year term premium



C. 10-Year term premium



D. 30-Year term premium

Figure 4 Forecaster-level median autocorrelation estimates of the short rate and term premia

This figure plots the median autocorrelation estimates of the short rate and term premia. It includes both subjective and actual autocorrelation estimates. Each blue circle represents a forecaster's median subjective autocorrelation estimates of short rate ρ_1^s and term premium $\rho_p^{s(n)}$. The size of the circle corresponds to the number of this forecaster's valid autocorrelation estimates. The orange diamond represents the median subjective autocorrelation estimates for the consensus forecasts. The green square at the bottom-right corner represents the median actual autocorrelation estimates of short rate ρ_1 and term premium ρ_p. "Autocorrelation averaging" implies that all subjective autocorrelations should be within the dashed box (i.e., $\rho_p < \rho_1^s, \rho_p^{s(n)} < \rho_1$). The details of the estimation are in Section 4.3.

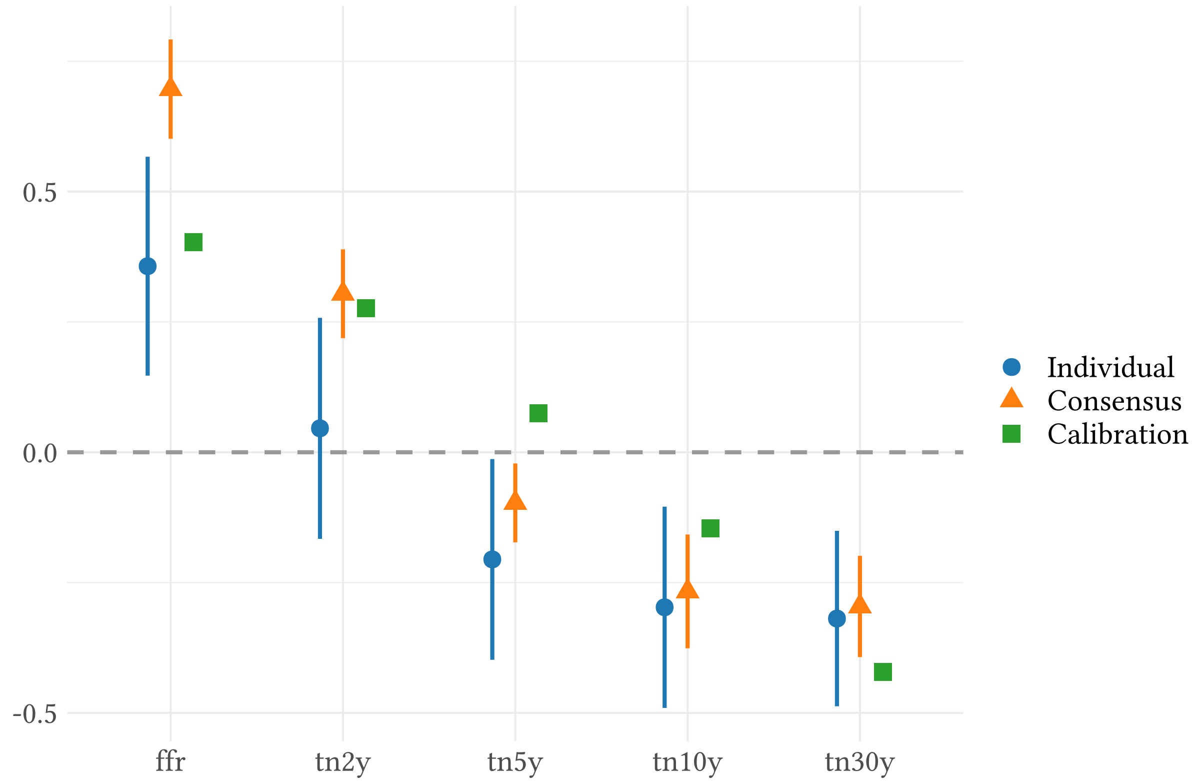


Figure 5 Calibrated vs. empirically estimated FE-on-FR regression coefficients at the individual level

This figure plots the Coibion and Gorodnichenko (2015) regression coefficients estimated from both individual- and consensus-level forecasts and from the calibrated model. The model is calibrated with individual-level average estimates of the subjective and actual autocorrelations. The blue dots and orange triangles represent empirically estimated coefficients at the individual and consensus level, respectively. The green squares represent model-generated coefficients. The calibration exercise uses the Federal funds rate (ffr) as short rate and Treasury yields with maturities of 2, 5, 10 and 30 years ($tn2y$, $tn5y$, $tn10y$ and $tn30y$) as long rates.

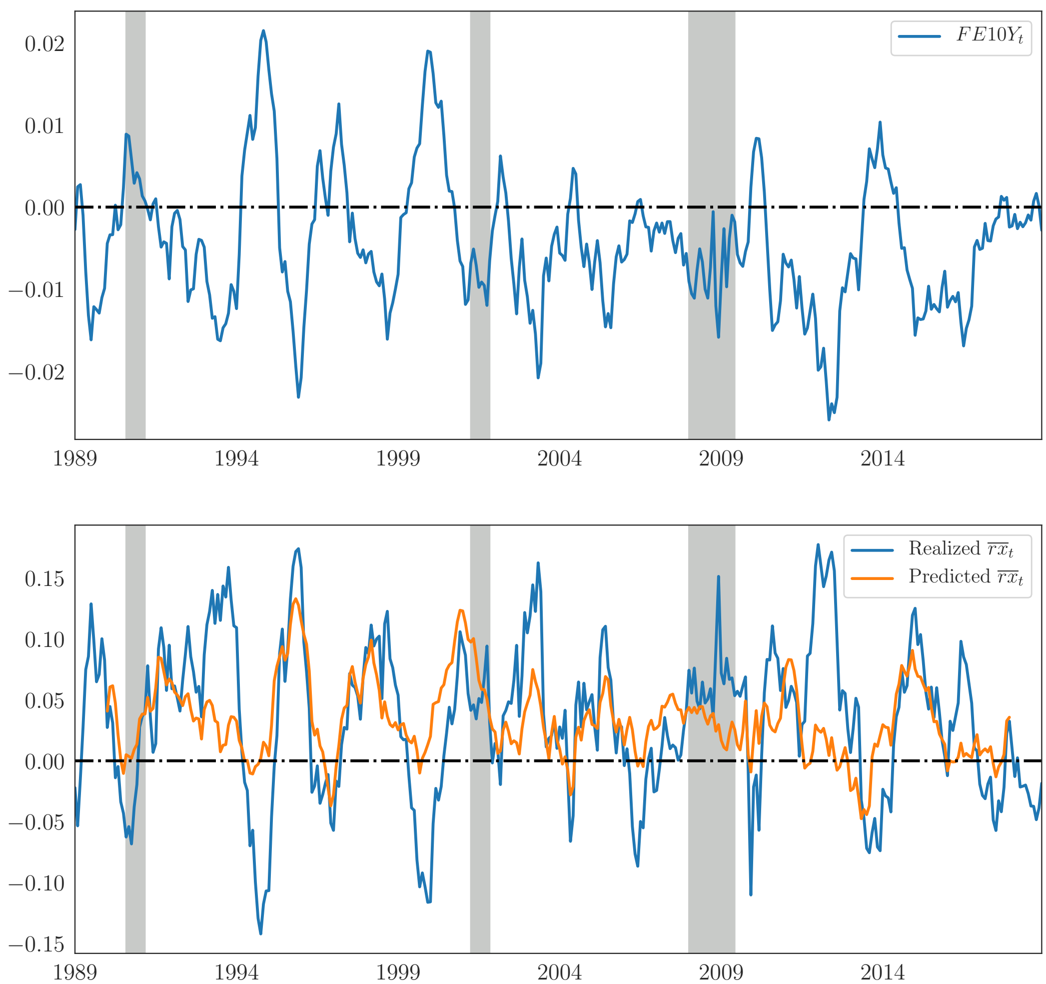


Figure 6 The time series of return predictor $FE10Y$ and its predicted average bond excess returns

Panel A plots the monthly time-series of the return predictor $FE10Y_t$, defined as the lagged forecast error for the 10-year Treasury yield. Panel B plots the realized and in-sample fitted average one-year excess bond returns. The fitted values are from a univariate predictive regression using $FE10Y_t$. The grey shaded areas are NBER-dated recessions.

References

Afrouzi, Hassan, Spencer Kwon, Augustin Landier, Yueran Ma, and David Thesmar, 2020, Overreaction and Working Memory, Working Paper.

Amromin, Gene, and Steven A. Sharpe, 2014, From the horse's mouth: Economic conditions and investor expectations of risk and return, *Management Science* 60, 845–866.

Barberis, Nicholas, 2018, Psychology-Based Models of Asset Prices and Trading Volume, in B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds., *Handbook of Behavioral Economics - Foundations and Applications 1*, volume 1, 79–175 (North-Holland).

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.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2018, Extrapolation and Bubbles, *Journal of Financial Economics* 2, 203–227.

Barberis, Nicholas, Andrei Shleifer, and Robert W. Vishny, 1998, A Model of Investor Sentiment, *Journal of Financial Economics* 49, 307–343.

Bauer, Michael D., and James D. Hamilton, 2017, Robust Bond Risk Premia, *Review of Financial Studies* 31, 399–448.

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.

Bouchaud, Jean-Philippe, Philipp Kruger, Augustin Landier, and David Thesmar, 2019, Sticky Expectations and the Profitability Anomaly, *Journal of Finance* 74, 639–674.

Brooks, Jordan, Michael Katz, and Hanno N. Lustig, 2019, Post-FOMC Announcement Drift in U.S. Bond Markets, Working Paper.

Brooks, Jordan, and Tobias J. Moskowitz, 2017, Yield Curve Premia, Working Paper.

Buraschi, Andrea, Ilaria Piatti, and Paul Whelan, 2018, Rationality and Subjective Bond Risk Premia, Working Paper.

Campbell, John Y., and Robert J. Shiller, 1988, The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, *Review of Financial Studies* 1, 195–228.

Carroll, Christopher, 2003, Macroeconomic Expectations of Households and Professional Forecasters, *Quarterly Journal of Economics* 118, 269–298.

Choi, James J., and Adriana Z. Robertson, 2020, What Matters to Individual Investors? Evidence from the Horse's Mouth, *Journal of Finance* 75, 1965–2020.

Cieslak, Anna, 2018, Short-Rate Expectations and Unexpected Returns in Treasury Bonds, *Review of Financial Studies* 31, 3265–3306.

Cieslak, Anna, and Pavol Povala, 2015, Expected Returns in Treasury Bonds, *Review of Financial Studies* 28, 2859–2901.

Cieslak, Anna, and Pavol Povala, 2016, Information in the Term Structure of Yield Curve Volatility, *Journal of Finance* 71, 1393–1436.

Clark, Todd, and Michael McCracken, 2013, Advances in Forecast Evaluation, in Graham Elliott, and Allan Timmermann, eds., *Handbook of Economic Forecasting*, volume 2 of *Handbook of Economic Forecasting*, 1107 – 1201 (Elsevier).

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

Cochrane, John H., and Monika Piazzesi, 2009, Decomposing the Yield Curve, Working Paper.

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

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.

Daniel, Kent D., David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor Psychology and Security Market Under- and Overreactions, *Journal of Finance* 53, 1839–1885.

d'Arienzo, Daniele, 2020, Maturity Increasing Overreaction and Bond Market Puzzles, Working Paper.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Positive Feedback Investment Strategies and Destabilizing Rational Speculation, *Journal of Finance* 45, 379–395.

DellaVigna, Stefano, and Joshua M. Pollet, 2007, Demographics and Industry Returns, *American Economic Review* 97, 1667–1702.

Dominitz, Jeff, and Charles F. Manski, 2007, Expected equity returns and portfolio choice: Evidence from the Health and Retirement Study, *Journal of the European Economic Association* 5, 369–379.

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl, 2017, The Deposits Channel of Monetary Policy, *Quarterly Journal of Economics* 132, 1819–1876.

Drerup, Tilman, Benjamin Enke, and Hans-Martin von Gaudecker, 2017, The precision of subjective data and the explanatory power of economic models, *Journal of Econometrics* 200, 378 − 389.

Driscoll, John C., and Aart C. Kraay, 1998, Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data, *Review of Economics and Statistics* 4, 549–560.

Duffee, Gregory R., 2018, Expected Inflation and Other Determinants of Treasury Yields, *Journal of Finance* 73, 2139–2180.

English, William B., Skander J. Van den Heuvel, and Egon Zakrajšek, 2018, Interest Rate Risk and Bank Equity Valuations, *Journal of Monetary Economics* 98, 80 − 97.

Enke, Benjamin, and Florian Zimmermann, 2019, Correlation Neglect in Belief Formation, *Review of Economic Studies* 86, 378 − 389.

Frankel, Jeffrey A., and Kenneth A. Froot, 1990, Chartists, fundamentalists, and trading in the foreign exchange market, *The American Economic Review* 80, 181–185.

Fuster, Andreas, David Laibson, and Brock Mendel, 2010, Natural Expectations and Macroeconomic Fluctuations, *Journal of Economic Perspectives* 24, 67–84.

Gabaix, Xavier, 2014, A Sparsity-Based Model of Bounded Rationality, *Quarterly Journal of Economics* 129, 1661–1710.

Gabaix, Xavier, 2019, Behavioral Inattention, in B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds., *Handbook of Behavioral Economics - Foundations and Applications 2*, volume 2, 261 − 343 (North-Holland).

Giacoletti, Marco, Kristoer T. Laursen, and Kenneth J. Singleton, 2021, Learning From Disagreement in the U.S. Treasury Bond Market, *Journal of Finance* 76, 395–441.

Giglio, Stefano, and Bryan T. Kelly, 2018, Excess volatility: Beyond discount rates, *Quarterly Journal of Economics* 133, 71–127.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Steve Utkus, 2021, Five Facts about Beliefs and Portfolios, *American Economic Review* 111, 1481–1522.

Glaeser, Edward L, and Charles G Nathanson, 2017, An Extrapolative Model of House Price Dynamics, *Journal of Financial Economics* 126, 147–170.

Goyal, Amit, and Ivo Welch, 2008, A Comprehensive Look at The Empirical Performance of Equity Premium Prediction, *Review of Financial Studies* 21, 1455–1508.

Greenwood, Robin, and Andrei Shleifer, 2014, Expectations of Returns and Expected Returns, *Review of Financial Studies* 27, 714–746.

Gürkaynak, Refet S., Brian Sack, and Eric Swanson, 2005, The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models, *American Economic Review* 95, 425–436.

Gürkaynak, Refet S., Brian P. Sack, and Jonathan H. Wright, 2006, The U.S. Treasury Yield Curve: 1961 to the Present, *Journal of Monetary Economics* 54, 2291–2304.

Hansen, Lars Peter, 1982, Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica* 50, 1029–1054.

Hanson, Samuel G., David O. Lucca, and Jonathan H. Wright, 2018, The Excess Sensitivity of Long-Term Rates: A Tale of Two Frequencies, Working Paper.

Hanson, Samuel G., and Jeremy C. Stein, 2015, Monetary Policy and Long-Term Real Rates, *Journal of Financial Economics* 115, 429–448.

Hodrick, Robert J., 1992, Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement, *Review of Economic Dynamics* 5, 357–386.

Hong, Harrison, and Jeremy C. Stein, 1999, A Unified Theory of Underreaction , Momentum Trading , and Overreaction in Asset Markets, *Journal of Finance* 54, 2143–2184.

Joslin, Scott, Marcel Priebsch, and Kenneth J. Singleton, 2014, Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks, *Journal of Finance* 69, 1197–1233.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2014, Time-Varying Fund Manager Skill, *Journal of Finance* 69, 1455–1484.

Kahneman, Daniel, and Amos Tversky, 1972, Subjective Probability: A Judgment of Representativeness, *Cognitive Psychology* 3, 430–454.

Kahneman, Daniel, and Amos Tversky, 1973, On the psychology of Prediction, *Psychological Review* 80, 237–251.

Kézdi, Gábor, and Robert J. Willis, 2009, Stock Market Expectations and Portfolio Choice of American Households, Working Paper.

Ludvigson, Sydney C., and Serena Ng, 2009, Macro Factors in Bond Risk Premia, *Review of Financial Studies* 22, 5027–5067.

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

Manski, Charles F., 2004, Measuring Expectations, *Econometrica* 72, 1329–1376.

Matthies, Ben, 2018, Biases in the Perception of Covariance, Working Paper.

Merton, Robert C., 1969, Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case, *Review of Economics and Statistics* 51, 247–257.

Newey, Whitney K., and Kenneth D. West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

Piazzesi, Monika, Juliana Salomao, and Martin Schneider, 2015, Trend and Cycle in Bond Premia, Working Paper.

Rabin, Matthew, 2002, Inference by Believers in the Law of Small Numbers, *Quarterly Journal of Economics* 117, 775–816.

Sims, Christopher, 2003, Implications of Rational Inattention, *Journal of Monetary Economics* 50, 665–690.

Singleton, Kenneth J., 2021, Presidential Address: How Much "Rationality" Is There In Bond Market Expectations, *Journal of Finance* 76, 1611–1654.

Stein, Jeremy C., 1989, Overreactions in the Options Market, *Journal of Finance* 44, 1011–1023.

Swanson, Eric, 2021, Measuring the Effects of Federal Reserve Forward Guidance and Asset Purchases on Financial Markets, *Journal of Monetary Economics* 118, 32–53.

Vissing-Jorgensen, Annette, 2003, Perspectives on Behavioral Finance: Does" Irrationality" Disappear with Wealth? Evidence from Expectations and Actions, *NBER Macroeconomics Annual* 139–194.

Wei, Min, and Jonathan H. Wright, 2013, Reverse Regressions and Long-Horizon Forecasting, *Journal of Applied Econometrics* 28, 353–371.

Woodford, Michael, 2003, *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, chapter Imperfect Common Knowledge and the Effects of Monetary Policy, 25–58 (Princeton University Press).

Xu, Zhengyang, 2020, Expectation Formation in the Treasury Bond Market, Working Paper.