The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

Check for updates

The Journal of Finance®

Kenneth J. Singleton
The President of the American Finance Association 2020

Presidential Address: How Much "Rationality" Is There in Bond-Market Risk Premiums?

KENNETH J. SINGLETON

ABSTRACT

Beliefs of professional forecasters are benchmarked against those of a Bayesian econometrician $\mathcal{BE}$ who is learning about the unknown dynamics of the bond risk factors. Consistent with rational Bayesian learning, the forecast errors of individual professionals and $\mathcal{BE}$ are comparably predictable over the business cycle. The secular and cyclical patterns of professionals' forecasts relative to those of $\mathcal{BE}$ are explored in depth. Inconsistent with many models with belief dispersion, the relationship between professionals' yield disagreement and their *matched* disagreements about macroeconomic fundamentals is very weak.

IN THIS PAPER, I REFLECT on two broad questions related to the nature of belief formation in bond markets: $\mathcal{Q}1$—what guidance does theory provide for assessing the rationality of beliefs in bond markets? and $\mathcal{Q}2$—what guidance do historical bond yields and professional surveys provide for modeling investors' belief formation?

This exploration is premised on three considerations. (i) In general, dispersion of beliefs is a source of risk in bond markets and *bonds are not priced as if* one is in an economy of homogeneous agents holding an average belief (Jouini and Napp (2007)). (ii) Outside of knife-edge cases, the beliefs of each individual investor will differ from those of an observer econometrician who is modeling the data-generating process (DGP) for equilibrium bond prices (e.g., Xiong and Yan (2009)). (iii) While surveys provide an incomplete picture of the beliefs of major market participants, forecasts and measured disagreement from surveys are informative about bond-market risk premiums (Giacoletti, Laursen, and Singleton (2021, GLS). Building on this theory and evidence, I explore

Kenneth J. Singleton is at the Graduate School of Business, Stanford University, and NBER. A very special thank you to Marco Giacoletti, Anh Le, and Albert Marcet for their collaboration on ongoing research on learning in financial markets, and for the insightful comments they shared while I was preparing this address. Helpful comments were also received from Anna Cieslak, Stefan Nagel, Paul Whelan, Wei Xiong, and seminar participants at LBS and UCL. No third party had the right to review this paper prior to its circulation and submission. Singleton has nothing to disclose in regard to *The Journal of Finance* disclosure policy.

Correspondence: Kenneth J. Singleton, Graduate School of Business, Stanford University, and NBER; e-mail: kenneths@stanford.edu.

DOI: 10.1111/jofi.13062

© 2021 the American Finance Association

what historical and survey data reveal about the structure and "rationality" of beliefs in the U.S. Treasury bond markets.

Although $\mathcal{Q}1$ and $\mathcal{Q}2$ are intimately connected, the literature has largely bifurcated its answers to these questions along two distinct conceptual paths. Building on David (2008) and Xiong and Yan (2009), $\mathcal{Q}1$ is typically addressed using a class of equilibrium models in which agents optimally filter for an unknown inflation or output growth rate, while "agreeing to disagree" about the *known* values of the parameters governing the state process.[1] This model design induces speculative trading and variation in relative wealths that affect risk premiums on bonds. Effectively, investors disagree about expected excess returns on bonds because they disagree about the underlying inflation/output fundamentals that determine bond prices.[2]

Motivated by Coibion and Gorodnichenko (2012, 2015, C-G), much of the empirical work addressing $\mathcal{Q}2$ explores the nature of the serial dependence of the survey respondents' forecast errors. For their "noisy-signal" environment, C-G argue that optimal filtering gives rise to serially independent forecast errors at the individual level but serially correlated errors by the cross-sectional average forecaster. Since empirical evidence on forecast errors of economic time series often shows serial dependence at the *individual* level, this has led some to explore behavioral frictions as explanations for this seemingly nonrational behavior of professional forecasters (e.g., Bordalo et al. (2020), d'Arienzo (2020), Wang (2020)).

Informed by considerations (i) and (ii) above, I argue in Section I that optimal learning in the presence of (typical) informational frictions gives rise to serially correlated forecast errors at the *individual* level. (In this case the errors of the cross-sectional average forecaster simply reflect this individual-level serial dependence.) In such settings, serial dependence of individual-level forecast errors in and of itself is silent about the rationality of investors' beliefs. Accordingly, assessing an individual's rationality using this econometric criterion would seem to require benchmarking her errors against the degree of serial dependence arising within a reference rational learning rule. Furthermore, the noisy-signal friction, while plausible in some macroeconomic settings, does not seem like the most natural reference friction for bond markets. Not only are bond prices readily observed, but investors can effectively infer many of the parameters governing the pricing distribution of bonds from the cross-section of yields on traded bonds (Duffee (2011), GLS). More likely to be unknown to professionals are the parameters governing the dynamics of the risk factors underlying variation in bond yields, especially when these parameters shift over time owing to stochastic trends or changes in global economic conditions. Accordingly, the benchmark I use is the learning rule of the "outside observer" Bayesian econometrician $\mathcal{BE}$ introduced in GLS.

[1] Buraschi and Whelan (2016) and Ehling et al. (2018), among others, extend the Xiong-Yan model to constant relative risk aversion (CRRA) preferences and preference shocks.

[2] Buraschi and Whelan (2016) and Andrade et al. (2016), among others, study disagreement about "fundamentals" and the correlations of these disagreements with bond yields.

An initial look in Section II at the properties of the forecast errors of the professional forecasters comprising the Blue Chip Financial Forecasters (BCFF) survey reveals three striking patterns. First, there is little persistence in the relative positioning of professional forecasters in terms of either their yield forecasts or the accuracy of these forecasts. To a reasonable approximation, those forecasters in the top 20% of yield forecasts today end up being randomly distributed across the support of the forecast distribution one year hence. Second, there is considerable cross-sectional heterogeneity in the correlations of professionals' yield forecast errors with lagged revisions in their expectations about the future level ($PC1$) and slope ($PC2$) of the yield curve. The frequency distributions of responses to expectation revisions for $PC1$ become increasingly negative as the maturity of the underlying bond increases. Third, the cross-sectional median responses of the BCFF professionals to forecast revisions are approximately the same as the responses of the benchmark learner $\mathcal{BE}$. In this (specific) sense, the BCFF professionals appear to respond "on average" to new information as would a Bayesian who is learning from the history of the yield curve.

Pursuing this last finding, I decompose bond yields into their expectations and term-premium components in Section III. I show that the increasingly leftward shift of $\mathcal{BE}$'s sensitivities to forecast revisions for $PC1$ is intimately connected to the larger role of shocks to term premiums in the overall volatility of yields on longer term bonds. More broadly, there are interesting cyclical patterns of persistence in forecast errors associated with the low-order principal components of yields (PCs), measures of disagreement, and forecast revisions. Once again, the sensitivities of $\mathcal{BE}$'s forecast errors to lagged information are similar to the median sensitivities across BCFF forecasters. Key exceptions are the differences in their responses to the yield curve slope and to revisions in their forecasts of future PCs.

Unanswered by my analysis up to this point is the question of why professional forecasters disagree. A strong implication of the equilibrium models in which agents "agree to disagree" is that yield disagreement is directly linked to disagreement about inflation or output growth (the central state variable observed with noise). I show in Section V that, in fact, there is essentially no connection at all between professionals' disagreement about yields and their matched disagreements about inflation or output growth. This finding leads to some reflections on what might be missing from our equilibrium models of bond yield determination in the presence of belief dispersion.

In Section VI, I take a deeper dive into the rationality of professional forecasters and highlight several secular and cyclical reasons why the forecasts of the typical BCFF professional are less accurate than those of the Bayesian $\mathcal{BE}$.

Throughout this analysis, when I use the term *rational* I mean *internal rationality* in the sense of Adam and Marcet (2011). Agents make fully optimal, dynamically consistent decisions given the informational frictions they face and their subjective beliefs about the state of the economy that is outside of their control. There is no presumption, and indeed it is generally not the case, that each agent's subjective beliefs correspond to the objective probability associated with the DGP for bond yields.

I. Benchmarking Beliefs in Bond Markets

We know very little about professionals' conditioning information sets or their underlying forecasting models, and (perhaps) only slightly more about the incentives that might guide construction of their forecasts. Fortunately, for an interesting array of informational frictions, economic theory provides substantial guidance about the properties of forecasts and forecast errors that arise in learning environments. In this section, I review some of the categories of agent-forecasters within equilibrium models with heterogeneous beliefs, and then set forth the benchmark learning model that I will use in subsequent empirical assessments of the properties of professional yield forecasts.

A. Disagreement When Agents Agree to Disagree

The influential model of Xiong and Yan (2009) has two (classes of) agents, neither of whom observes the long-run mean θ_t of inflation π_t. Both investors observe a signal $\mathcal{S}_t$ about θ_t that in reality is pure noise, but they both believe that this signal is correlated with θ_t and thus informative about bond prices. Each agent interprets the signal differently, and they agree to disagree about its correlation with θ_t. This disagreement gives rise to speculative trading on the part of the two agents that impacts equilibrium bond prices. Specifically, letting ω_t^i denote group i's wealth share, the yield on an m-period zero-coupon bond takes the form

$$y_t^m = A_m + B_m(\pi)\pi_t - \frac{1}{m}\log\left[\omega_t^a e^{-B_m(\theta)\hat{\theta}_t^a} + \omega_t^b e^{-B_m(\theta)\hat{\theta}_t^b}\right] \tag{1}$$

$$\approx A_m + B_m(\pi)\pi_t - B_m(\theta)\left(\omega_t^a\hat{\theta}_t^a + \omega_t^b\hat{\theta}_t^b\right) \tag{2}$$

for constant loadings $(A_m, B_m(\pi), B_m(\theta))$, where $\hat{\theta}_t^i$ is group i's optimally filtered value of the unknown θ_t based on knowledge of $(\pi_t, \mathcal{S}_t)$ and their assumed correlation between $(\theta_t, \mathcal{S}_t)$. Disagreement between their internally rational forecasts, $(\hat{\theta}_t^a - \hat{\theta}_t^b)$, drives bond yields directly through the weighted average beliefs in (1) and indirectly through the wealth weights (ω_t^a, ω_t^b) (see Xiong and Yan (2009), Theorem 1).[3]

Now suppose that an observer Bayesian econometrician $\mathcal{BE}$ in this setting understands that $\mathcal{S}_t$ is pure noise and, as such, she ignores $\mathcal{S}_t$ an optimal filter for θ_t as the mean $\hat{\theta}_t^{\mathcal{E}}$ of her posterior distribution $\theta_t|\{\pi_\tau\}_{\tau=0}^t$. Although θ_t is unobservable to all three agents (a, b) and $\mathcal{BE}$, their belief systems are in general different. Consequently, when $\mathcal{BE}$ studies the forecasts of bond yields under this

[3] Relaxing the assumption of logarithmic utility introduces a third, direct channel through which $(\hat{\theta}_t^a - \hat{\theta}_t^b)$ impact bond yields, namely, the induced hedging demands studied by Buraschi and Whelan (2016).

θ-observability friction, she concludes that the forecast errors $(y^m_{t+h} - E^i_t(y^m_{t+h}))$ are serially dependent, to differing degrees across $i \in (a, b)$ because they have different subjective beliefs.

Learning in these endowment economies impacts equilibrium bond prices, but of course it does not impact the exogenously given inflation and consumption/endowment processes. An omniscient observer econometrician is one who observes the true long-run mean θ_t of inflation.[4] Through the lens of this observer's eyes, the yield forecast errors of the agents (a, b) *and* the observer $\mathcal{BE}$ will show serial dependence. This is true in steady state since the disparate beliefs of $(\mathcal{BE}, a, b)$ do not converge in large samples to beliefs conditioned on θ_t.

B. Persistent Forecast Errors Induced by Noisy Signals

Empirical studies of persistence in forecast errors are increasingly stepping back from these theoretical foundations and are framing their null hypotheses around the inference strategy proposed by Coibion and Gorodnichenko (2015) for their *noisy-signal* environment. Of particular relevance here, C-G argue that "the predictability of the average ex post forecast errors across agents ... is an emergent property of the aggregation across individuals, not a property of the individual forecasts" (p. 2652). Similarly, in their application of the C-G framework, Bordalo et al. (2020) state that "Tests of individual beliefs are informative about departures from rationality. Tests of consensus forecasts yield additional information about the role of information frictions" (p. 3). How are these quotes reconciled with theoretical models of belief dispersion in which *forecast errors of individual forecasters are serially correlated*?

To revisit the implications of noisy signals for the persistence of forecast errors in bond markets, suppose that bond yields follow a linear model in terms of the risk factors $\mathcal{P}_t$:

$$y^m_t = A_m + B_m \mathcal{P}_t. \tag{3}$$

Le, Singleton, and Dai (2010) show that (3) is a good approximation in a representative agent model with habit-formation preferences, and Xiong and Yan (2009) find that the affine representation (2) is an accurate approximation to the actual pricing relation for their model with belief heterogeneity. Without loss of generality (see Joslin, Singleton, and Zhu (2011)), we can rotate any affine factor model (including (3) and (2)) to set $\mathcal{P}$ to the principal components

[4] In a related environment in which agents are learning about the unobserved stochastic long-run mean of the consumption/endowment process, Buraschi and Whelan (2016) show that an omniscient agent's (instantaneous) expected excess return on an m-period bond is a weighted average of the forecast errors of their two agent-groups. This illustrates the impact of disagreement on equilibrium bond yields and risk premiums.

(PCs) of yields.[5] We also assume that $\mathcal{P}$ follows the Markov process

$$\mathcal{P}_t = \kappa_0 + \kappa_{\mathcal{P}} \mathcal{P}_{t-1} + \epsilon_{\mathcal{P}t}, \qquad \epsilon_{\mathcal{P}t} \sim N(0, \Sigma_{\mathcal{P}}). \tag{4}$$

Initially, for ease of notation, I focus on the case of scalar $\mathcal{P}$ and set A_m in (3) to zero.

Suppose that a Bayesian agent $\mathcal{BA}$ does not observe the risk factor $\mathcal{P}$ directly, but she receives a noisy signal $\mathcal{S}$ about $\mathcal{P}$:

$$\mathcal{S}_t = \mathcal{P}_t + \omega_t, \qquad \omega_t \sim N(0, \sigma_\omega^2), \tag{5}$$

where ω_t is i.i.d. and independent of $\mathcal{P}_s$ for all t and s. $\mathcal{BA}$'s steady-state Kalman filter for this setting has the updating rule

$$E^{\mathcal{BA}}[\mathcal{P}_t|\mathcal{S}_0^t] = (1-G)E^{\mathcal{BA}}[\mathcal{P}_t|\mathcal{S}_0^{t-1}] + G\mathcal{S}_t, \qquad G = \frac{\hat{\sigma}^2}{\hat{\sigma}^2 + \sigma_\omega^2}, \tag{6}$$

where $\mathcal{S}_0^t = \{\mathcal{S}_0, \ldots, \mathcal{S}_t\}$ and $\hat{\sigma}^2$ is the steady-state variance of the forecast error $\mathcal{P}_{t+1} - E^{\mathcal{BA}}[\mathcal{P}_{t+1}|\mathcal{S}_0^t]$. This is the starting point of C-G's analysis.

Proceeding at the level of the individual forecaster $\mathcal{BA}$, the definition of $\mathcal{S}_t$ in (5) and the fact that $E^{\mathcal{BA}}[\mathcal{P}_{t+1}|\mathcal{S}_0^t] = \kappa_{\mathcal{P}} E^{\mathcal{BA}}[\mathcal{P}_t|\mathcal{S}_0^t]$ together imply that

$$\mathcal{P}_{t+1} - E^{\mathcal{BA}}[\mathcal{P}_{t+1}|\mathcal{S}_0^t] = \kappa_{\mathcal{P}}(1-G)\left(\mathcal{P}_t - E^{\mathcal{BA}}[\mathcal{P}_t|\mathcal{S}_0^{t-1}]\right) + \left(\epsilon_{\mathcal{P},t+1} - \kappa_{\mathcal{P}} G\omega_t\right). \tag{7}$$

Moreover, using the proportionality of y_t^m and $\mathcal{P}_t$, it follows immediately that[6]

$$\begin{aligned} &y_{t+1}^m - E^{\mathcal{BA}}[y_{t+1}^m|\mathcal{S}_0^t] \\ &\quad = \kappa_f(1-G)\left(y_t^m - E^{\mathcal{BA}}[y_t^m|\mathcal{S}_0^{t-1}]\right) + B_m\left(\epsilon_{\mathcal{P},t+1} - \kappa_{\mathcal{P}} G\omega_t\right) \qquad (8) \\ &\quad = \frac{1-G}{G}\left(E^{\mathcal{BA}}[y_{t+1}^m|\mathcal{S}_0^t] - E^{\mathcal{BA}}[y_{t+1}^m|\mathcal{S}_0^{t-1}]\right) + B_m G\left(\epsilon_{\mathcal{P},t+1} - \kappa_{\mathcal{P}}\omega_t\right), \qquad (9) \end{aligned}$$

with a version of the last expression holding for forecasts of y_{t+h}^m, $h > 1$. Thus, under C-G's framing of the noisy-signal friction, forecast errors at the individual level are serially dependent. Moreover, the persistence in forecast errors for the cross-sectional average forecaster is simply inherited from (9) by integrating out ω through cross-agent averaging.[7]

[5] It is true that low-dimensional factor models do not perfectly price the entire yield curve—there are small pricing or measurement errors. Yet, Joslin, Le, and Singleton (2013b) show that estimates of dynamic term structure models based on the Kalman filter that accounts for such errors give virtually identical results to those obtained when one assumes that the low-order PCs comprising $\mathcal{P}$ are priced perfectly.

[6] There are differences under this noisy-signal friction on how the persistence in forecast errors should be interpreted at the individual level versus the cross-sectional average (*consensus*) level, and on how best to estimate G in these different settings.

[7] As many have observed, this framing of noisy information implies that $G \in [0, 1]$, so a finding of a negative loading on revisions in expectations would amount to prima facie evidence against this friction.

The optimality of $\mathcal{BA}$'s filter is not inconsistent with serial dependence in $\mathcal{P}_{t+1} - \hat{\mathcal{P}}^{\mathcal{BA}}_{t+1,t}$, because her filter is designed to "whiten" the errors in forecasting $\mathcal{S}_{t+1}$; it is computing the objective expectation $E[\mathcal{S}_{t+1}|\mathcal{S}_0, \ldots, \mathcal{S}_t]$. Under $\mathcal{BA}$'s belief system, her forecast errors are serially independent. However, by assumption, $\mathcal{BA}$ never sees the true realization of $\mathcal{P}_t$ (she is incapable of estimating the projection in (7) to inform her beliefs). In contrast, an econometrician who observes $\{\mathcal{P}_t\}$ and $\mathcal{BA}$'s survey forecasts, and thus can estimate the linear projections (7) or (8), will find that $\mathcal{BA}$'s errors exhibit serial dependence under the DGP. This is a direct consequence of the econometrician—just as with the researchers using C-G's framework—being endowed with more information than $\mathcal{BA}$.

It bears emphasizing that the roles of noisy signals in the equilibrium analysis of Section I.A and C-G's reduced-form derivations are very different. Without loss of generality, Xiong and Yan's (2009) linearized relationship (2) can be rotated to express the entire term structure of zero yields as a factor model $y^m_t = \tilde{A}_m + \tilde{B}_m \mathcal{P}_t$, where $\mathcal{P}'_t = (PC1_t, PC2_t)$. The $(y^m_t, \mathcal{P}_t)$ are *fully observable* to all agents and the econometrician alike.[8] In contrast, C-G directly posit that their variable of interest $\mathcal{P}_t$ is observable only up to a noisy signal $\mathcal{S}_t$. This framing does not seem well matched to the learning challenges in bond markets, because yields are observed essentially without error. As such, $\sigma^2_\omega = 0$ and $G = 1$, which gives the counterfactual implication of *no persistence* in the forecast errors (see (7)).

In sum, within C-G's noisy signal framework, evidence of serial dependence at the individual level in and of itself is silent about whether agents are forming internally rational forecasts of bond yields. We should expect this friction to give rise to serial dependence in the forecast errors of individual Bayesian forecasters. Thus, to draw conclusions about the rationality of survey forecasts from forecast error persistence, we should be benchmarking the characteristics of survey data against forecasts by an internally rational agent facing frictions that give rise to persistent forecast errors. The next section sets forth what I believe is an empirically more relevant friction impacting respondents to surveys about yields.

C. Learning About the Dynamics of Risk Factors in Bond Markets

The benchmark agent that I focus on is a Bayesian econometrician $\mathcal{BE}$ learning about the unknown dynamics of the risk factors $\mathcal{P}$ while pricing bonds with a dynamic, arbitrage-free term structure model. This blending of learning under an informational friction with a formal term structure model brings realism—it is not unlike models used by traders for pricing and risk management. Accordingly, when assessing the rationality of the BCFF professionals, I often proceed as if these forecasters and $\mathcal{BE}$ share a common pricing problem. A drawback of the complexity of $\mathcal{BE}$'s learning/pricing problem is that

[8] This is not an "affine" term structure model, because the implied conditional distribution of $\mathcal{P}_t$ is nonaffine.

attribution of persistence in her forecast errors to specific frictions is obscured. The patterns in the survey forecasts documented below are robust to how BCFF survey respondents actually form their beliefs.

Suppose as before that yields satisfy the factor model (3) and, following GLS, all market participants have full knowledge of the factor loadings (A_m, B_m). Professional traders typically adopt a pricing model that shares this factor structure, and they treat the loadings (A_m, B_m) as both known and fixed when computing the prices of zero-coupon bonds. This assumption provides a plausible simplification for our discussion, since these loadings are inferred from the cross-section of yields using a spline or a parsimoniously parameterized arbitrage-free pricing model (see also Duffee (2011)).[9] If the factor loadings are known to traders, then their forecasts of yields and of the low-order *PC*s should also satisfy a factor model with the same loadings. GLS show that this is indeed the case, a strong finding in light of the substantial disagreement across the deciles of ordered professional forecasts.

With full knowledge of the pricing distribution, the informational challenge faced by market participants is that of learning the parameters governing the dynamics of the risk factors $\mathcal{P}$. It will be instructive to develop the implications of $\mathcal{BE}$'s learning problem for the persistence of her forecast errors in three steps: (i) time-varying factor dynamics, (ii) overidentifying restrictions implied by no-arbitrage pricing, and (iii) $\mathcal{BE}$'s beliefs compared to those implied by the DGP (an observer-econometrician's beliefs). Suppose that $\mathcal{BE}$ believes that the DGP for a *scalar* $\mathcal{P}$ has an autoregressive structure except that, instead of (4), $\mathcal{BE}$ believes that $\mathcal{P}_t$ follows the process

$$\mathcal{P}_{t+1} = \kappa_{0t} + \kappa_{\mathcal{P}t}\mathcal{P}_t + \epsilon_{\mathcal{P},t+1} \sim N(0,\, \Sigma_{\mathcal{P}}), \tag{10}$$

with unknown $\Theta_t' = (\kappa_{0t}, \kappa_{\mathcal{P}t})$ that evolves according to

$$\Theta_t = \Theta_{t-1} + Q_{t-1}^{1/2}\eta_t, \qquad \eta_t \overset{\text{iid}}{\sim} \text{Normal}(0, I), \tag{11}$$

where $Q_{t-1} = \text{Var}(\Theta_t|\Theta_{t-1})$ and η_t is independent of all past and future $\epsilon_{\mathcal{P}t}$. $\mathcal{BE}$ computes filtered estimates of Θ_t conditional on her knowledge of (A_m, B_m) for all m.

Adopting a Gaussian prior on Θ_0 leads to a posterior distribution for Θ_t that is also Gaussian, $\Theta_t|\mathcal{P}_1^t \sim \text{Normal}(\hat{\Theta}_t, P_t)$. So, letting $\mathcal{X}_t = (1, \mathcal{P}_t)$, $\mathcal{BE}$'s posterior mean and variance follow the recursions

$$\hat{\Theta}_t = \hat{\Theta}_{t-1} + P_{t-1}\mathcal{X}_{t-1}'(\mathcal{P}_t - \mathcal{X}_{t-1}\hat{\Theta}_{t-1})/\big(\mathcal{X}_{t-1}P_{t-1}\mathcal{X}_{t-1}' + \Sigma_{\mathcal{P}}\big), \tag{12}$$

$$P_t = P_{t-1} + Q_{t-1} - P_{t-1}\mathcal{X}_{t-1}'\mathcal{X}_{t-1}P_{t-1}/\big(\mathcal{X}_{t-1}P_{t-1}\mathcal{X}_{t-1}' + \Sigma_{\mathcal{P}}\big). \tag{13}$$

Using her posterior mean $\hat{\Theta}_t$, $\mathcal{BE}$'s forecast of the future yield y_{t+1}^m is

$$E^{\mathcal{BE}}[y_{t+1}^m|\mathcal{P}_0^t] = A_m + B_m E^{\mathcal{BE}}[\mathcal{P}_{t+1}|\mathcal{P}_0^t] = A_m + B_m\mathcal{X}_t\hat{\Theta}_t^{\mathcal{BE}}, \tag{14}$$

[9] Trading desks typically parameterize the pricing distribution of their risk factors $\mathcal{P}$ and treat the (A_m, B_m) as known functions of the parameters of this distribution.

which gives the one-period ahead forecast error $(y^m_{t+1} - (A_m + B_m \mathcal{X}_t \hat{\Theta}_t))$. Similar recursions are obtained for the more general case (considered below) of vector $\mathcal{P}_t$ with $\Theta_t' = (\kappa_{0t}', \text{vec}(\kappa_{\mathcal{P}})')$.

Forecasts from a dynamic term structure model potentially involve three ingredients: the dynamics of $\mathcal{P}$ under the historical DGP used directly to forecast future $\mathcal{P}$, the dynamics of $\mathcal{P}$ under the pricing distribution ($\mathbb{Q}$), and constraints on the market prices of risk (MPR) that link the DGP and $\mathbb{Q}$ distributions. I assume that, under $\mathbb{Q}$, $\mathcal{P}_t$ follows a Gaussian vector autoregression (VAR) with fixed autoregressive parameters $\Theta^{\mathbb{Q}} = (\kappa_0^{\mathbb{Q}}, \kappa_{\mathcal{P}}^{\mathbb{Q}})$. This is in line with my assumption that $\mathcal{BE}$ knows the (A_m, B_m) because these loadings are fully determined by $\Theta_0^{\mathbb{Q}}$ and the innovation variance $\Sigma_{\mathcal{P}}$ (the latter being quantitatively negligible for the maturity range considered here). GLS lend empirical support to the assumption of constant $\Theta^{\mathbb{Q}}$ as, even when $\mathcal{BE}$ was allowed to update her estimates of $\Theta^{\mathbb{Q}}$ every month, she held them nearly constant over time.

Do the no-arbitrage restrictions of a dynamic term structure model impact $\mathcal{BE}$'s learning? Joslin, Singleton, and Zhu (2011) show within a canonical Gaussian model with fixed parameters Θ_0 and no constraints on the MPR that the no-arbitrage structure of a dynamic term structure model has no impact on the estimation of the VAR parameters Θ_0. That is, the maximum likelihood estimates of Θ_0 within the pricing model are identical to the ordinary least squares estimates of the VAR for $\mathcal{P}$. I am breaking this irrelevance (in following GLS) by imposing constraints on the MPR. Specifically, $\mathcal{BE}$ initializes her learning rule by fitting a Gaussian model (absent learning) over a 10-year training period and "zeroing out" parameters of the MPR that are statistically insignificant. She then maintains these restrictions when updating $\hat{\Theta}_t$. In this manner, the known parameters of the $\mathbb{Q}$ distribution of $\mathcal{P}_t$ constrain some of the VAR parameters under the DGP to be zero.[10]

To complete $\mathcal{BE}$'s Bayesian learning problem, I characterize $\mathcal{BE}$'s beliefs about the conditional variance $\boldsymbol{Q}_{t-1}$ of Θ_t in (11) by the recursion

$$\boldsymbol{Q}_{t-1}^{-1} = \gamma \boldsymbol{Q}_{t-2}^{-1} + \frac{\gamma}{1-\gamma} \mathcal{X}_{t-1}' \Sigma_{\mathcal{P}}^{-1} \mathcal{X}_{t-1} = \frac{\gamma}{1-\gamma} \sum_{i=0}^{t-1} \gamma^i \mathcal{X}_{t-1-i}' \Sigma_{\mathcal{P}}^{-1} \mathcal{X}_{t-1-i}, \tag{15}$$

for known $\gamma \in (0, 1)$ and $\Sigma_{\mathcal{P}}$. As shown in GLS, this assumption implies that $\mathcal{BE}$'s Bayesian updating rule simplifies to constant-gain learning with gain parameter γ. It seems plausible that, even as far back as the 1980s, (sophisticated) market participants could have been pricing bonds and forecasting yields within a dynamic term structure model that embedded constant-gain learning from $\{\mathcal{P}_1^t\}$. Thus, the "bar of accuracy" set by this variant of $\mathcal{BE}$'s Bayesian learning rule serves as a plausible and interpretable benchmark against which BCFF professionals' forecasts are to be evaluated.

There is an important difference between $\mathcal{BE}$'s learning rule under the informational friction set forth above and the case of constant, and still

[10] Importantly, GLS found that updating the constraints on the MPR as new data became available to $\mathcal{BE}$ led to substantially inferior out-of-sample forecast accuracy.

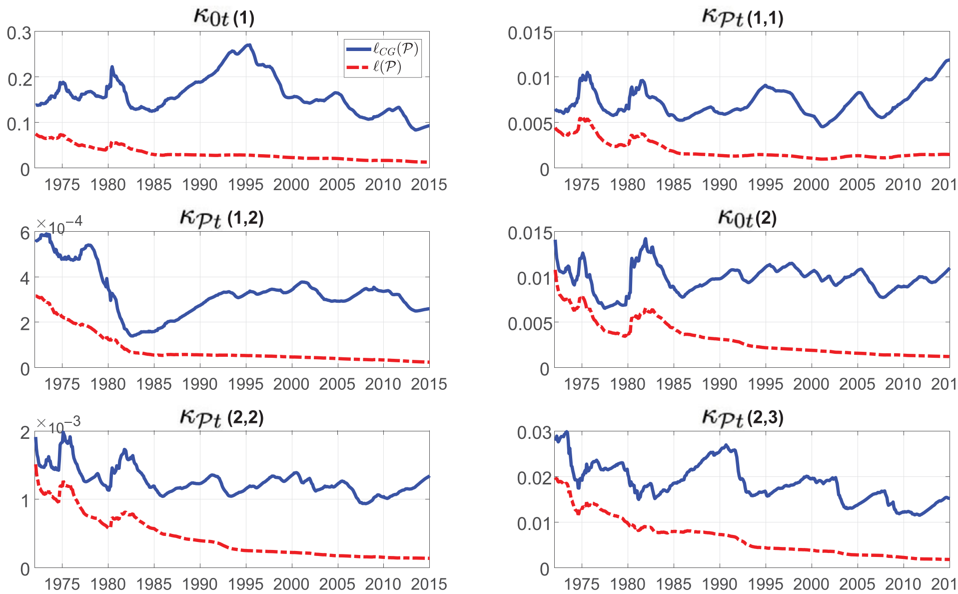


Figure 1. Posterior variances of unknown parameters in $(\kappa_{0t}, \kappa_{\mathcal{P}t})$ from the Bayesian learning rule in GLS. The dashed lines correspond to the learning rule where $\mathcal{BE}$ treats Θ_0 as constant, and the solid lines to the case in which $\mathcal{BE}$ believes that Θ_t follows a random walk. (Color figure can be viewed at wileyonlinelibrary.com)

unknown, autoregressive parameters Θ_0. Adopting a Gaussian prior on Θ_0 once again leads to a posterior distribution for Θ_t that is also Gaussian, $\Theta_t|\mathcal{P}_1^t \sim \text{Normal}(\hat{\Theta}_t, P_t)$. However, the assumption that $Q = 0$ implies that $\gamma = 1$, since $Q_{t-1} = (1-\gamma)R_t^{-1}/\gamma$. This gives the recursive (expanding-window) least squares estimator of Θ_0. The associated posterior variance of Θ_t now becomes

$$P_t = P_{t-1} - P_{t-1}\mathcal{X}'_{t-1}\mathcal{X}_{t-1}P_{t-1}/\big(\mathcal{X}_{t-1}P_{t-1}\mathcal{X}'_{t-1} + \Sigma_{\mathcal{P}}\big). \tag{16}$$

$\mathcal{BE}$ perceives herself as learning the true value of Θ and, as seen from (16), her posterior variance of Θ_t shrinks to zero over time.

This is illustrated by the Bayesian agent $\mathcal{BE}$ in GLS, who are learning in real time about the parameters governing the autoregressive process for the first three *PC*s of bond yields. From Figure 1 we can see that the posterior variances for the case in which agents perceive themselves as learning about constant parameters converge to zero over time. In contrast, when agents believe that there are permanent shifts in Θ_t, their posterior variances fluctuate over time at levels that do not converge toward zero.

There are compelling motivations for random parameters in (10). Dating back at least to Kozicki and Tinsley (2001), many studies have documented that capturing "shifting endpoints"—stochastic long-run means—of bond yields materially improves forecast accuracy in time-series models of

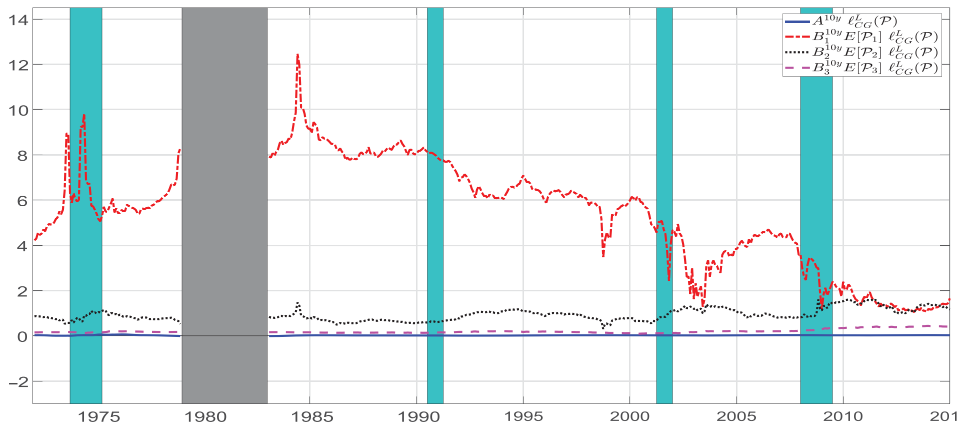


Figure 2. Decomposition of the "long-run means" of the 10-year yield y_t^{10} implied by rule $\ell^L_{CG}(\mathcal{P})$. The mean of $\mathcal{P}_t$ in month t is set to $(I - \kappa_{\mathcal{P}t})^{-1}\kappa_{0t}$ and contribution of each PCj is its loading B_j^{10y} multiplied by its pseudo long-run mean. (Color figure can be viewed at wileyonlinelibrary.com)

bond yields.[11] Based on historical relationships between spot and forward bond yields during his sample period, Fama (2006) argues that the long up and then downward trends in rates are the result of permanent shocks to the long-run means of bond yields.

Permanent-level shifts are accommodated in $\mathcal{BE}$'s learning rule through κ_{0t} in (10). My framing of shifting endpoints differs in two important respects from most prior studies of bond yields. First, the typical assumption is that the risk factors share a single, common stochastic trend. This introduces the possibility that the level and slope of the yield curve are constrained to share a trend. In contrast, the long-run means of $PC1$ and $PC2$ are different linear combinations of the vector $\hat{\kappa}_{0t}$. Second, prior studies typically presume that their common stochastic trend is observable to the econometrician and it is often proxied by a weighted average of past yield or macroeconomic information (e.g., a trend inflation series). Instead of linking κ_{0t} to specific, observed macroeconomic time series, $\mathcal{BE}$ treats these shifting endpoints as latent, and constructs filtered $\hat{\kappa}_{0t}$ based on the past history of the yield PCs. Richer conditioning information could easily be accommodated as in GLS.

If, at any date t, $\mathcal{BE}$ believed that the factor dynamics would remain fixed at the current value $\hat{\Theta}_t$, then she would compute the long-run means for the factors $\mathcal{P}_t$ as $(I - \hat{\kappa}_{\mathcal{P}t})^{-1}\hat{\kappa}_{0t}$. Figure 2 (taken from GLS) displays the time series of these pseudo long-run means for y^{10y} as $\mathcal{BE}$ updates her filtered estimates of $(\kappa_{0t}, \kappa_{\mathcal{P}t})$. The contribution of the mean of $PC1$ to the (pseudo) mean of y^{10y} shows a persistent downward drift from the mid 1980s until the last part of

[11] See also van Dijk et al. (2014), Cieslak and Povala (2015), and Bauer and Rudebusch (2020).

the sample.[12] At the same time, the pseudo long-run mean of $PC2$ in Figure 2 cycles around a roughly fixed value, consistent with it being trend free as seen through $\mathcal{BE}$'s eyes.

Of course another critical ingredient of $\mathcal{BE}$'s learning rule is her view about the dynamics of $\kappa_{\mathcal{P}t}$ as this autoregressive matrix (in the multifactor setting) governs the dynamics of the risk factors. Indeed, as we have just seen, $\hat{\kappa}_{\mathcal{P}t}$ is a key component of $\mathcal{BE}$'s shifting long-run means. I defer until later sections further discussion of the role of stochastic shifts in $\kappa_{\mathcal{P}t}$ and their relevance for benchmarking the beliefs of the BCFF professionals.

D. Outside/Inside Observers and Serial Dependence of Forecast Errors

$\mathcal{BE}$'s informational friction implies that her forecast errors depend on $\Theta_t - \hat{\Theta}_t$ as follows:

$$y^m_{t+1} - E^{\mathcal{BE}}[y^m_{t+1}|\mathcal{P}^t_0] = B_m[(\kappa_{0t} - \hat{\kappa}_{0t}) + (\kappa_{\mathcal{P}t} - \hat{\kappa}_{\mathcal{P}t})\mathcal{P}_t] + B_m\sigma_f\epsilon_{f,t+1}. \tag{17}$$

The difference $\Theta_t - \hat{\Theta}_t$ is time-varying, does not converge to zero as $t \to \infty$, and is dependent (through $\hat{\Theta}_t$) on $\mathcal{BE}$'s information set. To provide a foundation for subsequent comparisons of the properties of $\mathcal{BE}$'s forecasts to those of the professional BCFF professionals, is useful to expand on $\mathcal{BE}$'s position as an outside-observer econometrician and on potential sources of serial dependence in her forecast errors.

D.1. Outside/Inside Observers

An undercurrent that runs throughout the literature is the connection between the survey respondents who form the basis of empirical work and the agents whose dispersion of beliefs in theory impact equilibrium bond prices. Within a noisy-signal framework, Patton and Timmerman (2010) argue that disagreement about macroeconomic fundamentals is better explained by heterogeneous priors versus heterogeneity in forecasters' information signals, although the admissible priors did not encompass nonnested autoregressive dynamics. Bacchetta and Wincoop (2006) present evidence that disparate information sets across market participants explain several anomalies of exchange rate dynamics.

$\mathcal{BE}$'s benchmark rule that I use for assessing the rationality of survey respondents is that of an "outside econometrician." There is no assumption, one way or the other, about whether there exists a professional trader in the U.S. Treasury bond market who holds the beliefs of $\mathcal{BE}$, and no survey information is used in calibrating $\mathcal{BE}$'s learning rule. Nor should we necessarily expect that $\mathcal{BE}$'s beliefs should converge to those of an individual or consensus BCFF

[12] The period of the Federal Reserve monetary "experiment" from 1979 through 1982 is grayed out in Figure 2 owing to the sharp movements in the pseudo means that do not fit on the scale of this graph.

forecaster as time passes.[13] This rule is simply one of many rules that we could assign to an outside observer. A more flexible nonlinear rule could be captured by machine learning algorithms (e.g., Bianchi, Ludvigson, and Ma (2021)). I focus on this simpler, and yet still quite revealing, case of $\mathcal{BE}$.

Arguably, the survey professionals are also outside observers. Most are not themselves traders in U.S. Treasury bonds, but rather are providing forecasts on behalf of the advisory institution they represent. Their roles may be less impartial than $\mathcal{BE}$, however. While ideally (for matching to theory) forecast accuracy is in the mutual best interest of survey respondents and their firms, there may be conflicts of interest. Lamont (2002) argues that principal-agent problems incentivize older and more established forecasters to produce more radical forecasts. Similarly, Laster, Bennett, and Geoum (1999) argue that professional forecasters behave strategically, with those whose wages are tied to publicity-giving forecasts that are relatively far from the consensus (average). While prior studies focus on macroeconomic forecasts, one might reasonably wonder about conflicts in reporting yield forecasts. In moving forward with these data, I am reassured by the findings that measures of yield disagreement are similar across categories of professionals (see Figure 3 below) and that these yield disagreements are highly informative about risk premiums in bond markets.

To reiterate, $\mathcal{BE}$ should not be viewed as a "representative agent." As illustrated in Section I.A, the beliefs of the representative agent in environments with heterogeneous beliefs (the focus of subsequent empirical explorations) cannot be characterized as emerging from a Bayesian learning rule like the one I attribute to $\mathcal{BE}$. The outside observer $\mathcal{BE}$ presumes that the outcome of the equilibrium price setting process for bonds is a factor model for yields that is well described by (10) and (11).

D.2. Serial Dependence in Forecast Errors Under the DGP

In the noisy-signal setting of Section I.B, it is the strong informational hierarchy (the econometrician observes bond yields while the Bayesian $\mathcal{BA}$ does not) that gives rise to serial dependence. When $\mathcal{BA}$ is restricted to only seeing the noisy signal $\mathcal{S}_t$ about $\mathcal{P}_t$, her optimal filtering for $\mathcal{P}_t$ based on the past history $\mathcal{S}_0^t$ gives rise to serial dependence in the forecast errors $\mathcal{P}_t - E^{\mathcal{BA}}[\mathcal{P}_t|\mathcal{S}_0^t]$, which she never observes.

There is also an informational hierarchy–based source of serial dependence when agents and an econometrician ($\mathcal{OE}$) are learning about the dynamics of $\mathcal{P}$ (as in Section I.C) even though everyone observes yields perfectly. The optimal Kalman filter gives rise to serial correlation in $\Theta_t - \hat{\Theta}_t$ that, as is seen from (17), determines the serial dependence in $y_{t+1}^m - E^{\mathcal{BE}}[y_{t+1}^m|\mathcal{P}_0^t]$. Neither $\mathcal{BE}$ nor

[13] The issue of whether the beliefs of Bayesian learners converge asympotitcally—of whether disagreement vanishes—even when they agree on the basic structure an economy is explored in Acemoglu, Chernozhukov, and Yildiz (2016). Lahiri and Sheng (2008) document empirically that heterogeneous priors is a major contributor to persistent disagreement about the macroeconomy.

$\mathcal{OE}$ observe $\{\Theta_t\}$. The informational hierarchy in this case arises from the fact that $\mathcal{OE}$ is conditioning on $\{\mathcal{P}_t\}_{t=1}^{T}$, where T is sample size, and $\mathcal{BE}$ only knows $\{\mathcal{P}_t\}_{t=1}^{t}$ when she updates her beliefs at date t. Importantly, this informational advantage of $\mathcal{OE}$ over $\mathcal{BE}$ does not vanish over time, because Θ_t is changing randomly. So $\mathcal{OE}$ may find that the differences $\{E[\Theta_t|\mathcal{P}_0^T] - \hat{\Theta}_t\}_{t=1}^{T}$ are predictable based on $\mathcal{P}_0^t$ as a consequence of the retrospective Kalman smoother giving systematically more accurate filtered values for Θ_t.

Of potentially equal importance for the time-series properties of the forecast errors are the transient effects from learning. These effects arise even for the simplest case in which Θ is a constant unknown to $\mathcal{BE}$. In this case $Q = 0$ and Bayesian learning is recursive least squares (RLS) learning ($\gamma = 1$, and $\mathcal{OE}$ and $\mathcal{BE}$ have identical $\hat{\Theta}$ at date T). With a long enough sample and under the usual regularity conditions, $\mathcal{OE}$ can be thought of as knowing Θ. Earlier in the sample, when $\mathcal{BE}$ is learning about Θ from the subsample $\mathcal{P}_0^t$, her forecast error—representable as $E_t^{\mathcal{OE}}[y_{t+s}^m] - E_t^{\mathcal{BE}}[y_{t+s}^m]$ for $s > 0$—may be correlated with date t information. However, any such predictability is transient; for large T, $\mathcal{BE}$ also knows the true Θ.

Time-varying Θ_t brings the added complexity of $\mathcal{BE}$ never learning the true Θ_t at any t. This introduces Q_{t-1} into $\mathcal{BE}$'s posterior variance P_t of Θ_t in (13), which, in turn, affects the dynamics of $\hat{\Theta}_t$ in (12). Moreover, the equivalence between Bayesian and constant-gain learning relies on $\Sigma_{\mathcal{P}}$ being known and Q_{t-1}^{-1} following (15). Departures of $\mathcal{BE}$'s beliefs from the true DGP for Θ_t could, at the very least, lead to transient departures of her constant-gain rule from the efficient Bayesian rule.[14]

There are numerous examples in the literature of situations in which transient effects can persist over long periods. Collin-Dufresne, Johannes, and Lochstoer (2016) find this in a model with rare disasters. Complementary research in progress by Farmer, Nakamura, and Steinsson (2021) documents persistent effects of learning on the biases and serial dependencies of consensus survey forecasts. Giacoletti, Laursen, and Singleton (2021) illustrate slow convergence within a constant-parameter version of the three-factor model adopted by $\mathcal{BE}$ (see their Figure 1). Notably, the differences between the one-quarter-ahead forecasts of the 10-year yield based on full-sample estimates of $\kappa_{\mathcal{P}}$ versus those from RLS learning are typically positive and show strong cyclical patterns. Relative to the full-sample forecasts of $\mathcal{OE}$, the RLS learner systematically predicts lower excess returns following NBER recessions even though $\mathcal{OE}$ and the RLS learner have fully convergent forecasts at T.

That $\mathcal{BE}$ faces a confluence of informational frictions and transient dynamics when learning is a plus in my view, as any professional forecasting bond yields would likely face similar issues. No attempt is made to attribute the serial dependence in $\mathcal{BE}$'s forecast errors documented below to these various possible

[14] In a different economic setting in which an agent misperceives the true DGP, and her constant Q^{-1} is proportional to the second moment of $\mathcal{X}$, Sargent and Williams (2005) show equivalence in steady state. Prior to convergence to the steady-state Kalman filter, the constant-gain and Bayesian rules may differ.

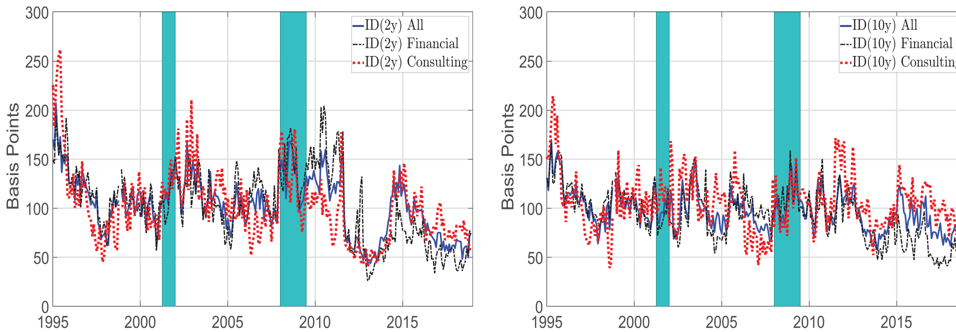


Figure 3. Interdecile disagreement about one-year-ahead two- and 10-year bond yields for "All," "Financial," and "Consulting" professionals in the BCFF survey. (Color figure can be viewed at wileyonlinelibrary.com)

sources. Instead, I focus on interpreting the properties of the forecasts of the BCFF professionals relative to those of $\mathcal{BE}$. In large part because of this *relative* focus, I also do not examine in more depth the consequences of $\mathcal{BE}$ misperceiving the DGP for $\mathcal{P}_t$ compared to its true objective DGP.

II. Belief Heterogeneity among BCFF Forecasters

I study the BCFF survey of yield and macroeconomic forecasts over the period from January 1985 through December 2018, with the start date determined by data availability. Of the 194 forecasters, 115 are categorized as financial institutions, 48 are consulting firms, and 31 represent other types of institutions. The forecasters submit forecasts of investment yields on U.S. Treasury bonds that have maturities of six months and one, two, three, five, seven, and 10 years. From these raw data, survey forecasts of zero-coupon bond yields of matching maturities are constructed as in Le and Singleton (2012). Survey data are released monthly at the beginning of the following month (usually the first business day), based on information collected over a two-day period (typically scheduled between the 20$^{\text{th}}$ and the 26$^{\text{th}}$ of the month). Disagreement is measured as the difference between the 90$^{\text{th}}$ and 10$^{\text{th}}$ percentiles of the cross-sectional distribution of BCFF zero-coupon yield forecasts.[15]

There is substantial disagreement among market professionals about future paths of bond yields. Figure 3 displays the time series of disagreements over a one-year horizon on two- and 10-year bond yields—$ID(2y)$ and $ID(10y)$—for all respondents, those from financial institutions, and those from consulting firms. Disagreement tends to be cyclical, rising during recessions, especially

[15] Typically about 45 forecasters submit forecasts in any given month. Our results are robust to measuring dispersion in beliefs as the cross-sectional (point-in-time) volatility of professional forecasts (Patton and Timmerman (2010)) or the cross-sectional mean absolute deviation in forecasts (Buraschi and Whelan (2016)), and our measure is similar to that used by Andrade et al. (2016).

for $ID(2y)$.[16] Interestingly, the respondents from financial and consulting firms show comparable magnitudes of disagreement as in the full set of survey respondents.[17] This is reassuring with regard to possible within-group incentives distorting my measures of investor disagreement. By construction, the *ID*s trim out the most extreme forecasters.

A. Are the Beliefs of Professional Forecasters Persistent?

Are professional forecasters persistently optimistic or pessimistic about bond yields? Cao et al. (2020) define investor groups as those that lie within the top or bottom 10% of forecasted yields, while Buraschi, Piatti, and Whelan (2021) focus on the top quartile of professional forecasters in terms of accuracy. Are their selection mechanisms (roughly) choosing a fixed group of forecasters? Or do the compositions of the investors in their groups, as well as those who determine $ID(2y)$ and $ID(10y)$, change substantially over time?

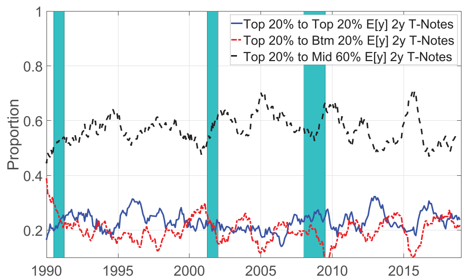
To explore questions about the persistence of the relative rankings of the forecasters' beliefs, we examine their positioning in the top 20%, middle 60%, or bottom 20% based on the construct of interest. For example, the solid line in Figure 4, Panel A, shows the probability of starting out in the top 20% of forecasted levels of the two-year bond yield and ending up in this same group of top 20% of forecasters 12 months later. The dashed line shows the probability of starting out in the top 20% of yield forecast levels and transitioning to the middle 60% of forecasted levels one year hence, etc. A 20%/60%/20% split would indicate random assignment of forecasters one year out. While there is some month-to-month variation in the proportions, it is striking that the pattern is roughly random.[18]

The transitions based on forecast accuracy are somewhat more nuanced (bottom row of Figure 4). Roughly 60% of the time the most accurate forecaster transitions to the middle category, as implied by random assignment. However, there is a strikingly strong negative correlation between staying in the top 20% accuracy or experiencing a complete collapse into the bottom 20% accuracy range 12 months later, particularly in the case of five- and 10-year bond yields and after the global financial crisis (GFC). For the shorter maturity bonds post-GFC, periods when high-accuracy forecasters remain in the lowest or middle categories of absolute forecast errors tend to be followed by reversions to roughly random assignment.

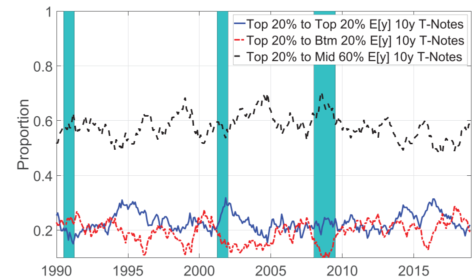
[16] Nieuwerburgh and Veldkamp (2010) discuss this cyclicality more broadly, and Cujean and Hasler (2017) connect it with higher return (in my setting, excess return) predictability in bad times.

[17] I confirmed this by projecting realized excess returns onto the first three yield *PC*s and either the two *ID*s for all firms or the four *ID*s for the financial and consulting firm groups. The incremental explanatory power from separating firms into the financial/consulting groupings in these regressions was small.

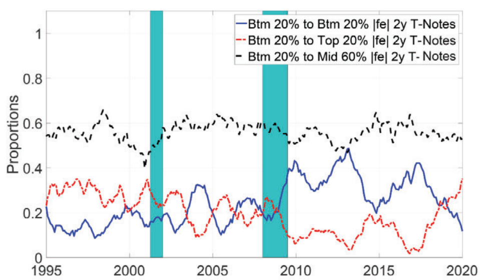
[18] The only notable departure from randomness, specific to the six-month maturity (not shown), occurs after 2015 when short Treasury bill rates were pegged at near-zero yields.



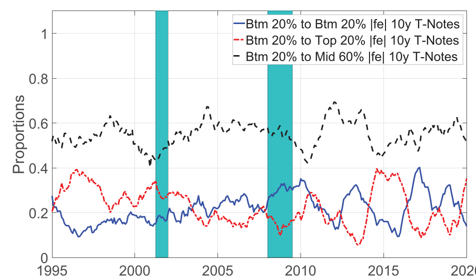
Panel A: Two-Year Yield Forecasts



Panel B: Ten-Year Yield Forecasts



Panel C: Absolute Forecast Errors y^{2y}



Panel D: Absolute Forecast Errors y^{10y}

Figure 4. Fraction of respondents in the bottom 20% who transition to the top 20%, middle 60%, or bottom 20% ranges one year hence. The upper two graphs correspond to yield forecasts and the lower two graphs to absolute forecast errors. (Color figure can be viewed at wileyonlinelibrary.com)

Buraschi, Piatti, and Whelan (2021) find somewhat more persistence in both beliefs and accuracy, more in line with the accuracies I document for y^{2y} after 2010. Relative to Figure 4, these authors focus on a subset of long-participating institutions (minimum of five years versus three years in my analysis). About 60% of their selected forecasters in the top/bottom *quartiles* transition to different quartiles one year later, whereas roughly 80% of my subset of forecasters transition to different *quintiles*. Institutions enter and leave the survey and, for some, participation is short-lived. These differences are likely indicative of some confounding of changes in beliefs and changes in the composition of the survey respondents.

Caution is therefore in order when matching of percentiles of the cross-sectional distributions of forecasts to specific agent-groups in theoretical models. For example, in Cao et al. (2020) two groups of investors (A, B) hold different beliefs about an unobserved, stochastic long-run mean of $PC1$. They assign the average beliefs about the monetary policy rate of the top 10 and bottom 10 survey respondents to groups A and B, respectively. Similarly, outside of a pricing model, Bianchi, Ludvigson, and Ma (2021) match the percentiles of the forecast distribution implied by their machine learning algorithms to the corresponding percentiles of the forecast distributions. An attractive feature of such matchings is they circumvent the formidable problem whereby the identities of the professionals reporting monthly change over time. Yet, we may well be studying pseudo-forecasters that bear little resemblance to actual (groups of) forecasters.

B. Persistence of Forecast Errors: A First Look

Initially, I explore the relationship between the errors of individual forecasters and their lagged revisions in expectations, though with a very different conceptual motivation.[19] When an agent $\mathcal{L}$ is learning about shifting risk-factor dynamics, the components of her forecast revisions are

$$E_t^{\mathcal{L}}[\mathcal{P}_{t+1}] = \hat{\kappa}_{0t}^{\mathcal{L}} + \hat{\kappa}_{\mathcal{P}t}^{\mathcal{L}}\mathcal{P}_t,$$
$$E_{t-1}^{\mathcal{L}}[\mathcal{P}_{t+1}] \approx \left(\hat{\kappa}_{0,t-1}^{\mathcal{L}} + \hat{\kappa}_{\mathcal{P},t-1}^{\mathcal{L}}\hat{\kappa}_{0,t-1}^{\mathcal{L}}\right) + (\hat{\kappa}_{\mathcal{P},t-1}^{\mathcal{L}})^2\mathcal{P}_{t-1},$$

where I assume that $\mathcal{L}$ holds the VAR parameters of $\mathcal{P}$ fixed at her current subjective values when forecasting future $\mathcal{P}$ (hence the approximation). Defining the forecast revision over k periods for a forecast horizon of h periods by

$$\mathcal{FR}_t(f; h, k) \equiv E[f_{t+h}|\mathcal{P}_0^t]) - E[f_{t+h}|\mathcal{P}_0^{t-k}], \tag{18}$$

$\mathcal{L}$'s forecast revision can be expressed as

$$\mathcal{FR}_t^{\mathcal{L}}(\mathcal{P}; 1, 1) \approx \left(\hat{\kappa}_{0t}^{\mathcal{L}} - \hat{\kappa}_{0,t-1}^{\mathcal{L}}\right) + \hat{\kappa}_{\mathcal{P}t}^{\mathcal{L}}\left(\mathcal{P}_t - E_{t-1}^{\mathcal{L}}[\mathcal{P}_t]\right) + \left(\hat{\kappa}_{\mathcal{P}t}^{\mathcal{L}} - \hat{\kappa}_{\mathcal{P},t-1}^{\mathcal{L}}\right)E_{t-1}^{\mathcal{L}}[\mathcal{P}_t]. \tag{19}$$

These revisions in PC forecasts embody several features of $\mathcal{L}$'s learning problem that connect back to the discussion of persistent forecast errors in Section I.C.

Note that the factor representation of y_t^m offers the possibility of deeper insights into the sources of predictability of investors' forecast errors relative to studies of yield forecasts alone. Specifically, revisions in forecasts of future yields can be represented in terms of revisions in forecasts of the risk factors. To set notation, I examine revisions in expectations of the "level" and "slope" of the yield curve by setting $f_t' = (PC1_t, PC2_t)$ in the projections

$$y_{t+h}^m - E^{\mathcal{BE}}[y_{t+h}^m] = \alpha_0^{mh} + \beta_f^{mh\prime}\mathcal{FR}_t^{\mathcal{BE}}(f; h, k) + \xi_{t+h}^{mh}. \tag{20}$$

[19] Though not explicit in my equations, I modify the projections (20) to accommodate the structure of the BCFF survey data. Specifically, the monthly BCFF survey asks for *average* forecasts over future calendar quarters. I am interested in monthly, constant-maturity forecasts over future quarterly intervals. To arrive at the latter from the reported data, I follow GLS and use weighted interpolations of the reported forecasts.

Implicit in the β_f^{mh}'s are the factor loadings B_m, so they contribute to the cross-maturity variation in the signs and magnitudes of the β^{mh} (typically symmetric about zero for $PC2$).

My benchmark learner $\mathcal{BE}$ follows the Bayesian rule $\ell_{CG}(\mathcal{P})$ from GLS that has her updating views on the dynamics of the risk factors $\mathcal{P}_t' = (PC1_t, PC2_t, PC3_t)$ using the history $\mathcal{P}_1^t$. I initialize $\mathcal{BE}$'s Θ at estimates from a standard Gaussian affine term structure model. I then treat the first 10 years of my sample as the burn-in period. Although an explicit role for disagreement as a risk factor, as in the difference-of-opinion models of Section I.A, is absent from this reduced-form benchmark, the $\mathcal{P}_t$ implicitly embody the role of disagreement as a priced risk. Approximately 50% of the variation in the BCFF-based measures of disagreement is accounted for by variation in $\mathcal{P}$ (GLS).

Since survey forecasts are at quarterly intervals out to one year, I examine nine-month (three-quarter)-ahead forecasts and use revisions of forecasts over a one-quarter interval of forecasts nine and 12 months ahead with $f_t' = (PC1_t, PC2_t)$:

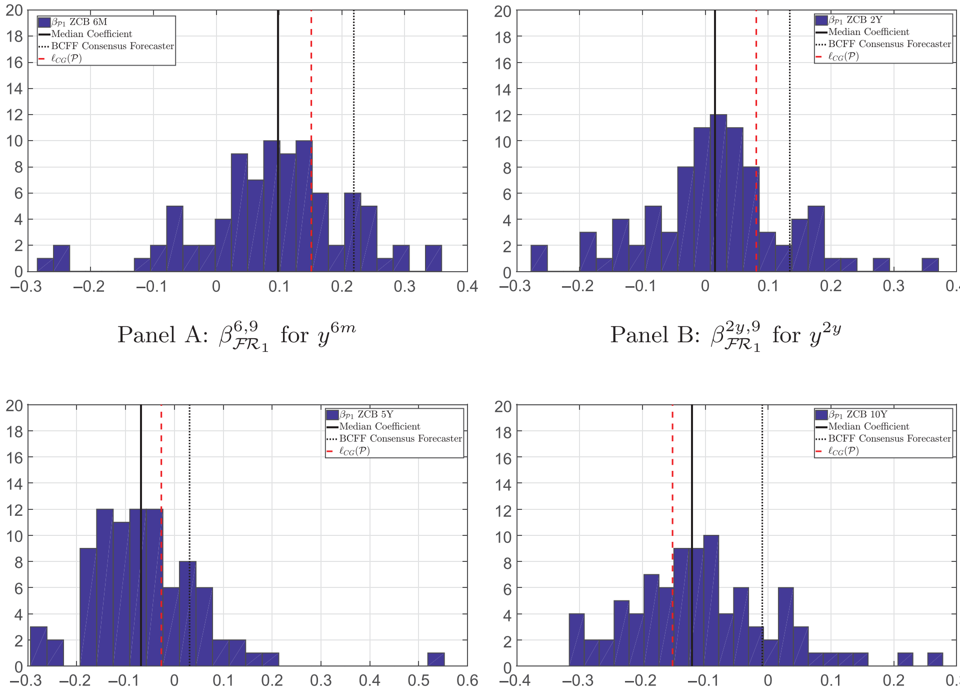
$$y_{t+9}^m - E_t^{\mathcal{L}}[y_{t+9}^m] = \alpha_0^{m9} + \beta_{\mathcal{FR}}^{m9\prime}\mathcal{FR}_t^{\mathcal{L}}(f; 9, 3) + \xi_{t+9}^{mh}, \tag{21}$$

where $\mathcal{L}$ indexes the BCFF forecasters. The frequency distribution for the estimates of $\beta_{\mathcal{FR}_1}^{m9}$ for $PC1$ forecast revisions are displayed in Figure 5 for the bonds of maturities six months and two, five, and 10 years to maturity, along with the median $\beta_{\mathcal{FR}_1}^{m9}$ across all BCFF forecasters (solid vertical line). I benchmark these estimates against the corresponding estimates of $\beta_{\mathcal{FR}}^{m9}$ for $\mathcal{BE}$'s learning rule (dashed vertical line).

Figure 5 shows that the BCFF forecasters are implicitly using different learning rules, as sensitivities $\beta_{\mathcal{FR}_1}^{m9}$ for the level factor show a wide range of positive and negative values.[20] Notwithstanding this heterogeneity, the median BCFF estimate of $\beta_{\mathcal{FR}_1}^{m9}$ and the estimate implied by $\mathcal{BE}$'s rule are comparable in magnitude for all maturities. This is a notable finding as there is no direct connection whatsoever between the BCFF forecasts and those of $\mathcal{BE}$. Moreover, $\mathcal{BE}$ is learning from the past history of $\mathcal{P}_t$ alone.

Another striking aspect of these distributions is that the levels of the loadings on $\mathcal{FR}(PC1)$ for all forecasters systematically shift to the left as the maturities of the bonds increase. The BCFF reactions tend to be positive for the six-month maturity and become largely negative at the 10-year maturity. This leftward shift is equally true of $\mathcal{BE}$'s estimates. The pattern of the loadings B_{m1} on $PC1$ is not the source of this leftward shift, as they are roughly flat across maturities m (see GLS). Rather, its source must be the dynamics of the forecast revisions as they relate to the predictability of forecast errors. Some have interpreted these patterns as evidence of "behavioral" underreactions to news at short maturities and overreactions to news at longer maturities (e.g., Bordalo et al. (2020), d'Arienzo (2020), Wang (2020)). By benchmarking against $\mathcal{BE}$, we

[20] There are forecasters who appear to have "outlier-like" responses. I have not omitted these forecasters as these outliers have only minor effects on the medians.



Panel A: $\beta^{6,9}_{\mathcal{FR}_1}$ for y^{6m}

Panel B: $\beta^{2y,9}_{\mathcal{FR}_1}$ for y^{2y}

Panel C: $\beta^{5y,9}_{\mathcal{FR}_1}$ for y^{5y}

Panel D: $\beta^{10y,9}_{\mathcal{FR}_1}$ for y^{10y}

Figure 5. Frequency distributions of the estimates of $\beta^{m9}_{\mathcal{FR}_1}$ for all BCFF forecasters. The solid vertical line corresponds to the median $\beta^{m9}_{\mathcal{FR}_1}$, while the dashed line to the estimate for the Bayesian $\mathcal{BE}$. (Color figure can be viewed at wileyonlinelibrary.com)

see that this shift is consistent (at least on average) with optimal Bayesian learning in the face of unknown and shifting dynamics of the risk factors.

The corresponding reaction coefficients for revisions in expectations about the slope of the yield curve are displayed in Figure 6. These frequency distributions are comparable across maturities, with the median response being slightly positive and once again quite close to $\mathcal{BE}$'s response (even more so than for $PC1$).

Finally, defining the realized excess return $xr^{m,h}_{t+h}$ over h periods as

$$hxr^{m,h}_{t+h} = -(m-h)y^{m-h}_{t+h} + my^m_t - hy^h_t, \tag{22}$$

errors in forecasting yields and excess returns are connected by the equality

$$-(m-h)\Big(y^{m-h}_{t+h} - E_t[y^m_{t+h}]\Big) = h\Big(xr^{m,h}_{t+h} - E_t[xr^{m,h}_{t+h}]\Big). \tag{23}$$

Thus, all of my evidence on yield forecast errors can be reinterpreted as evidence about errors in forecasting realized excess returns on bonds.

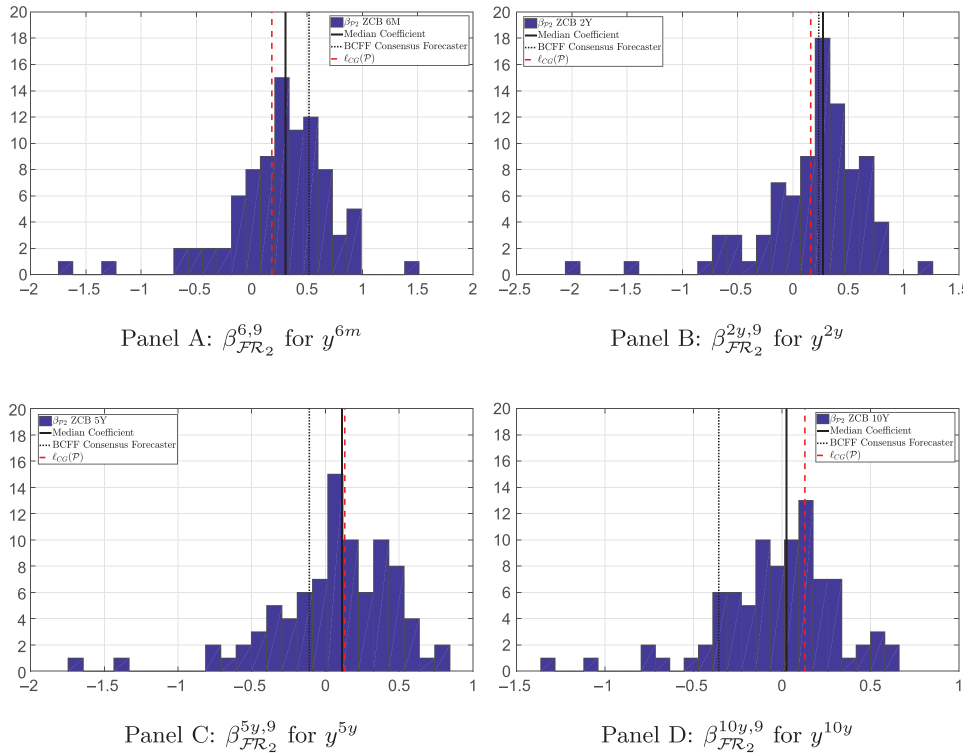


Figure 6. Frequency distributions of the estimates of $\beta^{m9}_{\mathcal{FR}_2}$ for all BCFF forecasters. The solid vertical line corresponds to the median $\beta^{m9}_{\mathcal{FR}_2}$, while the dashed line to the estimate for the Bayesian $\mathcal{BE}$. (Color figure can be viewed at wileyonlinelibrary.com)

C. Why Are the $\beta^{m9}_{\mathcal{FR}_1}$'s Increasingly Negative as Maturity Increases?

To address this question I start by decomposing yields into their expectations ($\mathcal{EH}$) and term premium ($\mathcal{TP}$) components using the six-month yield as the spot rate. Viewing (22) as a recursion and iterating forward gives the following tautology for the m-month yield that holds *for any belief system underlying* $\boldsymbol{E}^{\mathcal{L}}_t$ (e.g., Dai and Singleton (2002)):

$$y^m_t = \frac{6}{m}\sum_{j=1}^{m/6} \boldsymbol{E}^{\mathcal{L}}_t(y^6_{t+(j-1)6}) + \frac{6}{m}\sum_{j=1}^{m/6} \boldsymbol{E}^{\mathcal{L}}_t\left(xr^{m-(j-1)6}_{t+j6}\right) \equiv \mathcal{EH}^m_t + \mathcal{TP}^m_t. \qquad (24)$$

Using (24), the six-month-ahead forecast error for y^m_t decomposes as

$$\underbrace{y^m_{t+6} - \boldsymbol{E}^{\mathcal{L}}_t(y^m_{t+6})}_{\text{Yield Errors}} = \underbrace{\left(\mathcal{EH}^m_{t+6} - \boldsymbol{E}^{\mathcal{L}}_t(\mathcal{EH}^m_{t+6})\right)}_{\text{Expectations Errors}} + \underbrace{\left(\mathcal{TP}^m_{t+6} - \boldsymbol{E}^{\mathcal{L}}_t(\mathcal{TP}^m_{t+6})\right)}_{\text{Term Premium Errors}}. \qquad (25)$$

One of my motivations for examining this decomposition is the fact that time-series innovations in $\mathcal{TP}^m$ account for a higher percentage of the variance of innovations in y_t^m as m increases.

It is at this stage that I use a formal dynamic term structure model to evaluate $\mathcal{EH}_t^m$ and $\mathcal{TP}_t^m$. Under the assumption that $\mathcal{BE}$ knows the parameters of the Gaussian pricing distribution for $\mathcal{P}_t$, bond yields—and in particular all future y_{t+j}^6—are affine functions of the risk factors (Duffie and Kan (1996), Dai and Singleton (2000)). Following GLS, the expectations components $\mathcal{EH}_t^m$ are then computed under the additional simplifying assumption that, at each date t, $\mathcal{BE}$ holds the parameters governing her perceived historical distribution fixed for the purpose of forecasting future y_{t+j}^6. As time passes, $\mathcal{BE}$ updates $(\hat{\kappa}_{0t}, \hat{\kappa}_{ft})$ according to her Bayesian learning rule and, as such, her perceived DGP for $\mathcal{P}_t$ is a nonlinear process dependent on the full history $\mathcal{P}_0^{t-1}$.

Using the implied monthly time series $\mathcal{EH}_t^m$ and $\mathcal{TP}_t^m$, I compute linear projections of the errors in (25) onto $\mathcal{BE}$'s six-month revisions in her forecasts of $PC1$ and $PC2$, $\mathcal{FR}_t^{\mathcal{BE}}(f; 6, 6)$.[21] The estimates flagged with a $*$ associated with $PC1$ in the first column of Table I replicate the pattern of $\mathcal{BE}$'s forecasts in Figure 5: They start positive at the one-year maturity and then decline monotonically to -0.11 at the 10-year maturity. Reading across the rows, each number with a $*$ is the sum of the entries in columns (2) and (3), owing to (25). The matched sensitivities for $\mathcal{EH}^m$ in the middle column vary little with maturity. Thus, it is the impact of $\mathcal{FR}^{\mathcal{BE}}(PC1; 6, 6)$ on errors in forecasting $\mathcal{TP}^m$, combined with the increasing contribution of $\mathcal{TP}^m$ to variation in y^m, that drives the pronounced leftward shifts in Figure 5.

III. Cyclical Patterns in $\mathcal{BE}$'s Forecast Errors

In this section I dig deeper into three questions about learning, informational frictions, and the persistence of forecast errors ($\mathcal{Q}2$ in the opening paragraph). Specifically, (i) what are the cyclical patterns in the predictable components of $\mathcal{BE}$'s forecast errors; (ii) are forecast revisions especially revealing about informational frictions, or do other factors dominate; and (iii) building on GLS, are disagreements $(ID(2y), ID(10y))$ also informative about $\mathcal{BE}$'s forecast errors? Throughout this section $\mathcal{FR}_t(f)$ refers to $\mathcal{FR}_t(f; 6, 6)$.

A. Interpreting Forecast-Error Projections

To frame this inquiry, I begin by elaborating on my choices of conditioning information and on the interpretation of forecast error projections through the lens of the learning model in Section I. Realized forecast errors—the difference between the realized future bond yield and today's yield forecast—are not

[21] These are the same projections underlying the frequency distributions in Figures 5 and 6, except that now the forecast horizon is six months rather than nine months. As I explain below, using a six-month horizon allows me to compare my findings for $\mathcal{BE}$ to those for $\mathcal{CP}$ for the one-year yield y_t^{1y}.

Table I

Projections onto Forecast Revisions

This table presents projections of $\mathcal{BE}$'s errors from six-month-ahead forecasts of y_t^m, $\mathcal{EH}^m$, and $\mathcal{TP}^m$ onto revisions in her expectations about $PC1$ and $PC2$, $m = 1, 2, 5, 10$ years. Robust t-statistics are in brackets.

	(1)	(2)	(3)
	$y_t^{1y} - E_{t-6}^{\mathcal{BE}}(y_t^{1y})$	$\mathcal{EH}_t^{1y} - E_{t-6}^{\mathcal{BE}}(\mathcal{EH}_t^{1y})$	$\mathcal{TP}_t^{1y} - E_{t-6}^{\mathcal{BE}}(\mathcal{TP}_t^{1y})$
$\mathcal{FR}_t^{\mathcal{BE}}(PC1; 6, 6)$	0.0316*	0.0580	−0.0265
	[0.554]	[1.066]	[−2.439]
$\mathcal{FR}_t^{\mathcal{BE}}(PC2; 6, 6)$	0.2034	0.1995	0.0039
	[1.342]	[1.370]	[0.177]
Adj R^2	2.72%	3.87%	4.61%
	$y_t^{2y} - E_{t-6}^{\mathcal{BE}}(y_t^{2y})$	$\mathcal{EH}_t^{2y} - E_{t-6}^{\mathcal{BE}}(\mathcal{EH}_t^{2y})$	$\mathcal{TP}_t^{2y} - E_{t-6}^{\mathcal{BE}}(\mathcal{TP}_t^{2y})$
$\mathcal{FR}_t^{\mathcal{BE}}(PC1; 6, 6)$	−0.0087*	0.0647	−0.0733
	[−0.149]	[1.203]	[−3.031]
$\mathcal{FR}_t^{\mathcal{BE}}(PC2; 6, 6)$	0.1626	0.2262	−0.0636
	[1.206]	[1.587]	[−1.331]
Adj R^2	1.07%	4.82%	11.06%
	$y_t^{5y} - E_{t-6}^{\mathcal{BE}}(y_t^{5y})$	$\mathcal{EH}_t^{5y} - E_{t-6}^{\mathcal{BE}}(\mathcal{EH}_t^{5y})$	$\mathcal{TP}_t^{5y} - E_{t-6}^{\mathcal{BE}}(\mathcal{TP}_t^{5y})$
$\mathcal{FR}_t^{\mathcal{BE}}(PC1; 6, 6)$	−0.0732*	0.0589	−0.1321
	[−1.311]	[1.306]	[−3.057]
$\mathcal{FR}_t^{\mathcal{BE}}(PC2; 6, 6)$	0.0999	0.2660	−0.1661
	[0.866]	[2.089]	[−1.550]
Adj R^2	1.45%	6.82%	13.17%
	$y_t^{10y} - E_{t-6}^{\mathcal{BE}}(y_t^{10y})$	$\mathcal{EH}_t^{10y} - E_{t-6}^{\mathcal{BE}}(\mathcal{EH}_t^{10y})$	$\mathcal{TP}_t^{10y} - E_{t-6}^{\mathcal{BE}}(\mathcal{TP}_t^{10y})$
$\mathcal{FR}_t^{\mathcal{BE}}(PC1; 6, 6)$	−0.1121*	0.0424	−0.1544
	[−2.309]	[1.089]	[−3.067]
$\mathcal{FR}_t^{\mathcal{BE}}(PC2; 6, 6)$	0.0790	0.2895	−0.2105
	[0.743]	[2.434]	[−1.707]
Adj R^2	4.30%	8.40%	12.22%

directly informative about informational frictions or behavioral biases that may be affecting forecasters' beliefs. These errors confound (i) serial dependence induced by multiperiod forecasting; (ii) informational frictions, as illustrated by $\mathcal{BE}$'s learning in the presence of unknown factor dynamics; and (iii) any "behavioral" frictions guiding the beliefs of professional forecasters.

Ingredients (ii) and (iii) can be teased apart from (i) with projections. This is illustrated by the one-period-ahead forecasts (17): Projections of $(y_{t+1}^m - E_t[y_{t+1}^m])$ onto information $\mathcal{I}_t$ known to the forecaster at date t isolate the impact of learning—the term $Proj[(\kappa_{0t} - \hat{\kappa}_{0t}) + (\kappa_{ft} - \hat{\kappa}_{ft})\mathcal{P}_t|\mathcal{I}_t]$—from the orthogonal innovation $\epsilon_{f,t+1}$ (given that $Proj_t[\epsilon_{f,t+1}|\mathcal{I}_t] = 0$, provided $\mathcal{P}_0^t \in \mathcal{I}_t$). Similarly, for multiperiod forecasts, these projections (approximately) isolate the impact of learning from the induced moving-average errors.[22] Benchmarking these projections against the learning framework in Section I informs us about the roles of (ii) versus (iii).

[22] I say "approximately" because, strictly speaking, multiperiod forecasts involve nonlinearities in the presence of learning. The forecasts that I attribute to $\mathcal{BE}$ here circumvent this nonlinearity because she holds the filtered parameters fixed at the date she forms forecasts of future yields.

Table II

Projections onto Forecast Revisions, *PC*s, and Disagreement

This table presents projections of six-month-ahead forecast errors of $\mathcal{BE}$ for y_t^m, $\mathcal{EH}^m$, and $\mathcal{TP}^m$ onto $\mathcal{FR}_t(f)$ (forecast revisions), $\mathcal{P}$ (yield *PC*s), and H (disagreement) for $m = 1y, 10y$ yields. Robust t statistics are in brackets.

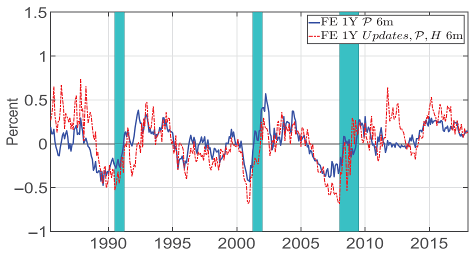
	$y_t^m - E_{t-6}^{\mathcal{BE}}(y_t^m)$		$\mathcal{EH}_t^m - E_{t-6}^{\mathcal{BE}}(\mathcal{EH}_t^m)$		$\mathcal{TP}_t^m - E_{t-6}^{\mathcal{BE}}(\mathcal{TP}_t^m)$	
m	1y	10y	1y	10y	1y	10y
$\mathcal{FR}_t(PC1)$	−0.019	−0.138	0.003	−0.002	−0.022	−0.136
	[−0.345]	[−2.859]	[0.067]	[−0.052]	[−2.320]	[−2.713]
$\mathcal{FR}_t(PC2)$	0.165	0.031	0.173	0.292	−0.009	−0.261
	[1.084]	[0.283]	[1.197]	[2.425]	[−0.394]	[−2.203]
$PC1$	−0.085	−0.038	−0.076	−0.023	−0.009	−0.015
	[−3.000]	[−1.608]	[−2.889]	[−1.265]	[−1.555]	[−0.714]
$PC2$	−0.001	0.120	−0.072	−0.150	0.071	0.270
	[−0.014]	[1.510]	[−0.763]	[−2.328]	[3.684]	[3.548]
$PC3$	0.859	0.224	0.865	0.521	−0.006	−0.297
	[2.286]	[0.867]	[2.421]	[1.998]	[−0.095]	[−1.185]
$ID(2y)$	−0.079	−0.368	−0.078	−0.170	−0.002	−0.198
	[−0.314]	[−1.711]	[−0.314]	[−0.668]	[−0.030]	[−0.760]
$ID(10y)$	0.677	0.724	0.615	0.337	0.061	0.386
	[2.354]	[2.967]	[2.234]	[1.381]	[1.220]	[1.450]
Adj R^2	14.76%	12.05%	18.26%	19.67%	17.14%	23.69%

The yield forecast errors over a six-month horizon ($h = 6$) are projected onto the risk factors $\mathcal{P}$ (first three *PC*s of bond yields), the revisions $\mathcal{FR}_t(f)$ in forecasts of f (the first two yield *PC*s), and the yield disagreement measures $H' = (ID(2y), ID(10y))$. The projection onto $\mathcal{P}_t$ serves as a useful reference, because $\mathcal{BE}$'s learning rule uses the history $\mathcal{P}_0^t$ to update her posterior views on $(\kappa_{0t}, \kappa_{ft})$. From (19) it can be seen that $\mathcal{FR}^{\mathcal{BE}}(f)$ is informative about informational frictions faced by $\mathcal{BE}$ and therefore may have predictive power for her errors even when $\mathcal{P}$ is the only relevant information for $\mathcal{BE}$'s learning problem.

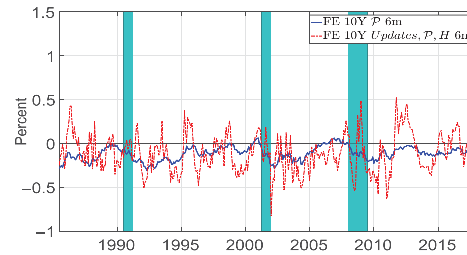
The role of disagreement H_t in these diagnostics is different. Within $\mathcal{BE}$'s learning rule, as I formulate it here, H cannot directly affect $\mathcal{BE}$'s expected excess returns, because her learning rule is conditioned only on the history $\mathcal{P}_0^t$. Accordingly, evidence that H has predictive power for forecast errors (over and above $\mathcal{P}_t$) would suggest that $\mathcal{BE}$ could have improved the accuracy of her forecasts by conditioning on $(\mathcal{P}_t, H_t)$.

Table II illustrates the richness of the serial dependence of forecast errors for the one- and 10-year yields using projections onto $(\mathcal{P}_t, H_t, \mathcal{FR}_t(f))$. Inclusion of $(\mathcal{P}_t, H_t)$ substantially increases the adjusted R^2s relative to their counterparts in Table I. Moreover, the R^2s from conditioning on $\mathcal{FR}(f)$ alone and those from conditioning on $(\mathcal{P}, H)$ (not shown) approximately sum to the adjusted R^2s in Table II. This observation explains the similar loadings on the forecast revisions across Tables I and II. Disagreement H_t adds explanatory power for all four yields, particularly for $\mathcal{EH}_t^m - E_{t-6}^{\mathcal{BE}}(\mathcal{EH}_t^m)$ and for $\mathcal{TP}_t^m - E_{t-6}^{\mathcal{BE}}(\mathcal{TP}_t^m)$ at the intermediate two- and five-year maturities.

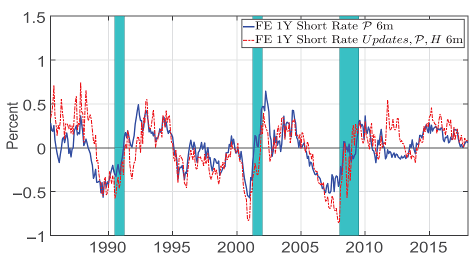
Also notable is the statistical significance of *PC*2 in the $\mathcal{TP}$ projections for all maturities. A strong connection between the slope of the yield curve and risk premiums has been widely documented in the term structure literature. In a



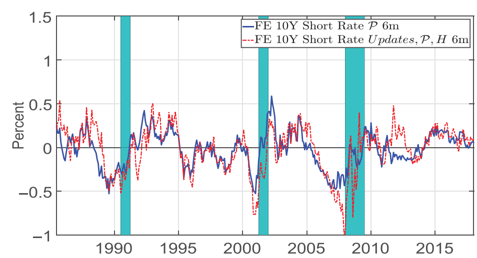
Panel A: $Proj\left(y^{1y}_{t+6} - E^{\mathcal{BE}}_t(y^{1y}_{t+6})|\mathcal{I}_t\right)$



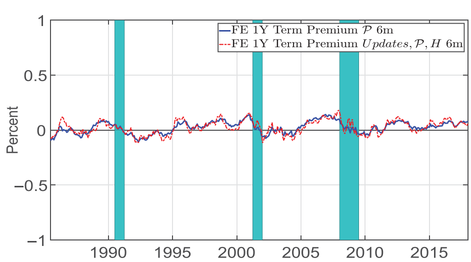
Panel B: $Proj\left(y^{10y}_{t+6} - E^{\mathcal{BE}}_t(y^{10y}_{t+6})|\mathcal{I}_t\right)$



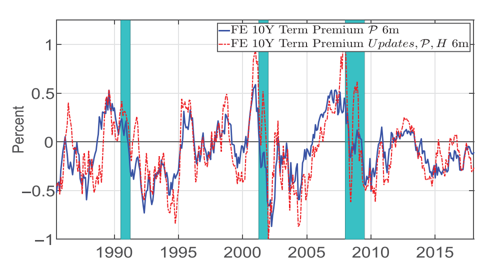
Panel C: $Proj\left(\mathcal{EH}^{1y}_{t+6} - E^{\mathcal{BE}}_t(\mathcal{EH}^{1y}_{t+6})|\mathcal{I}_t\right)$



Panel D: $Proj\left(\mathcal{EH}^{10y}_{t+6} - E^{\mathcal{BE}}_t(\mathcal{EH}^{10y}_{t+6})|\mathcal{I}_t\right)$



Panel E: $Proj\left(\mathcal{TP}^{1y}_{t+6} - E^{\mathcal{BE}}_t(\mathcal{TP}^{1y}_{t+6})|\mathcal{I}_t\right)$



Panel F: $Proj\left(\mathcal{TP}^{10y}_{t+6} - E^{\mathcal{BE}}_t(\mathcal{TP}^{10y}_{t+6})|\mathcal{I}_t\right)$

Figure 7. Fitted projections of $\mathcal{BE}$'s forecast errors. Panels A to F display the fitted projections of $\mathcal{BE}$'s forecast errors onto $\mathcal{I}_t = \mathcal{P}_t$ or $\mathcal{I}_t = (\mathcal{P}_t, H_t, \mathcal{FR}^{\mathcal{BE}}_t(f))$ for the one-year (left) and 10-year (right) bond yields. (Color figure can be viewed at wileyonlinelibrary.com)

learning setting, such as that faced by $\mathcal{BE}$, it would therefore be natural for $PC2$ to have predictive power for the forecast errors of $\mathcal{TP}$.

B. Cyclical Patterns in Fitted Forecast Errors

To shed more light on the properties of $\mathcal{BE}$'s forecasts, I examine the cyclical patterns of the fitted projections of the three forecast errors comprising (25). Figure 7 shows the projections for the one-year yield y^{1y}_{t+6} (left column) and the 10-year yield y^{10y}_{t+6} (right column) for two sets of conditioning variables: $\mathcal{P}_t$ for the baseline case and all right-hand side conditioning variables from Table II. Fitted values are displayed as of the dates of the conditioning information.[23]

Focusing first on the one-year results, one immediately sees the cyclicality of the fitted forecast errors (Figure 7, Panel A). Prior to every NBER-declared recession, the fitted errors turn negative ($\mathcal{BE}$'s forecast is systematically above the future realized yield). The fitted errors then shift signs from trough to peak during each recession, after which they tend to stay positive for several years until this cycle repeats itself. The source of this cyclical pattern is the forecast error for $\mathcal{EH}^{1y}_{t+6}$, as Figure 7, Panel C, reveals that the fitted values of these errors mirror almost exactly the pattern of the yield forecast errors. A very similar pattern emerges for the fitted forecast errors for the expectations component $\mathcal{EH}^{10y}_{t+6}$ of the 10-year yield (Figure 7, Panel D).

This pattern for $\text{Proj}[\mathcal{EH}^{1y}_{t+6} - E^{\mathcal{BE}}_t(\mathcal{EH}^{1y}_{t+6})|\mathcal{I}_t]$ seems plausible in the context of $\mathcal{BE}$'s learning problem. As economic activity weakens, $\mathcal{BE}$ receives market-based signals $\mathcal{P}$ that confound innovations to $\mathcal{P}$ and shocks to the parameters governing the dynamics of $\mathcal{P}$. Heading into recessions, this confounding induces systematic overestimation of average future short rates relative to realizations six months later. As $\mathcal{BE}$ achieves greater clarity about the degree of economic weakness (subsequently confirmed by NBER), $\mathcal{BE}$ adjusts her views and her errors in forecasting $\mathcal{EH}$ revert toward zero. Subsequently, she must judge how quickly the economy would emerge from recession. $\mathcal{BE}$'s learning leads to persistent underestimates of the rise in average short rates relative to her revised views six months later (positive fitted forecast errors subsequent to recessions).

Comparing Figure 7, Panels E and F, the variation in the fitted projection for $\mathcal{TP}^{1y}$ is much smaller than its counterpart for $\mathcal{TP}^{10y}$. These fitted forecast errors tend to peak just prior to or during NBER recessions, precisely when market risk premiums are large (countercyclical risk premiums). Evidently, it is during these periods of high-risk premiums that learning induces high degrees of predictability in $\mathcal{BE}$'s errors in forecasting future term premiums. Given the modest variation of $\text{Proj}(y^{10y}_{t+6} - E^{\mathcal{BE}}_t[y^{10y}_{t+6}]|\mathcal{P}_t)$ in Figure 7, Panel B, compared to the cyclicality of the fitted errors for $\mathcal{EH}^{10y}$ in Figure 7, Panel D, errors for $\mathcal{EH}^{10y}$ and $\mathcal{TP}^{10y}$ are necessarily (through (24)) negatively correlated.

To a large extent, the incremental predictive power of disagreement for $\mathcal{BE}$'s forecast errors is manifested through her expectations components $\mathcal{EH}^m$. The

[23] Note that $\mathcal{EH}^m_{t+6} - E^{\mathcal{BE}}[\mathcal{EH}^m_{t+6}|\mathcal{P}_t]$ is a weighted linear combination of forecast revisions based on $\mathcal{P}_t$ alone:

$$\mathcal{EH}^m_t - E^{\mathcal{BE}}(\mathcal{EH}^m_t|\mathcal{P}_{t-6}) = B_6 \frac{6}{m} \sum_{j=1}^{m/6} \Big(E^{\mathcal{BE}}[\mathcal{P}_{t+(j-1)6}|\mathcal{P}_t] - E^{\mathcal{BE}}[\mathcal{P}_{t+(j-1)6}|\mathcal{P}_{t-6}] \Big).$$

While revisions in beliefs about near-dated short rates are always material, revisions for far-dated short rates also showed sizable moves during the dotcom bust and the GFC. Furthermore, each difference in forecasts is based on filtered $\hat{\Theta}_t$ computed at dates t and $t-j$. Thus, the terms in the summation share revisions in the filtered trend $\hat{\kappa}_{0t}$, and this brings commonality to the predicted forecast errors for $\mathcal{EH}^m$ across all m.

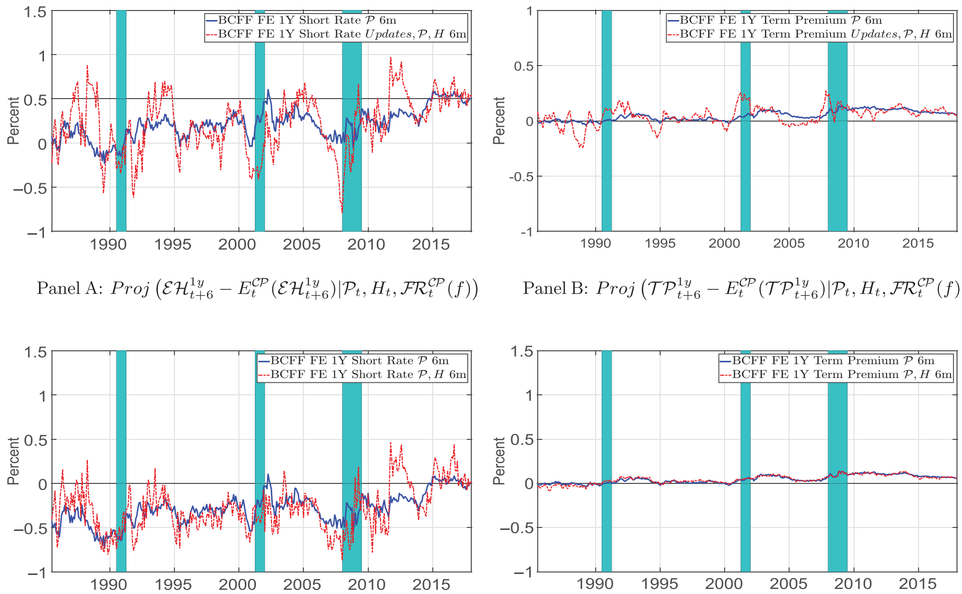
forecast revisions $\mathcal{FR}_t(f)$ bring material explanatory power over and above $(\mathcal{P}_t, H_t)$ to fitted forecast errors for $\mathcal{EH}^m$ and $\mathcal{TP}^m$. For yields of all maturities, the incremental predictive power of $\mathcal{FR}_t(f)$ is most pronounced near peaks and troughs or other sharp movements in the forecast errors for these components.

IV. A "Representative Consensus" BCFF Professional?

The vast majority of studies involving BCFF survey data construct a "consensus" forecaster as the average or median forecast of the variables of interest. Two properties of this pseudo-forecaster—henceforth referred to as the "consensus professional" ($\mathcal{CP}$)—warrant special emphasis for my analysis. First, $\mathcal{CP}$ is in general not the "representative investor" that is well defined in some economies under belief dispersion (Jouini and Napp (2007)). As Xiong and Yan (2009) illustrate, the beliefs of the representative investor are typically a weighted average (in their case, a wealth-weighted average) of the beliefs of the individual agent/groups. No wealth-weighing goes into the construction of $\mathcal{CP}$'s beliefs. Second, as documented above, the professionals tend to roam around the support of the distributions of yield forecasts and therefore the identities of the mean/median forecasters change over time. Notwithstanding this lack of association of $\mathcal{CP}$ with a specific forecaster, features of individual forecasters' learning rules that are shared across these professionals may well be reflected in the pseudo-forecaster $\mathcal{CP}$'s implicit learning rule.

A natural question at this juncture is: what are the properties of $\mathcal{CP}$'s beliefs relative to those of $\mathcal{BE}$ or of the individual BCFF forecasters? The patterns of fitted forecast errors for $\mathcal{CP}$ in Figure 8, Panels A and B, are broadly consistent with those for $\mathcal{BE}$. The fitted error in forecasting $\mathcal{EH}^{1y}_{t+6}$ displays the same cyclical pattern as its counterpart for $\mathcal{BE}$ and, since the fitted error for $\mathcal{TP}^{1y}$ is near zero throughout the sample, we can infer that $(y^{1y}_{t+6} - E^{\mathcal{CP}}_t(y^{1y}_{t+6}|\mathcal{I}_t))$ follows a similar pattern as Figure 7, Panel A. Also notable is the finding that the partial correlation of disagreement H_t with $\mathcal{CP}$'s forecast errors is channeled though the forecast errors for $\mathcal{EH}^{\mathcal{CP}}$ (compare Figure 8, Panels C and D, as is the case for $\mathcal{BE}$. These results are quite striking given that the forecasts for $\mathcal{CP}$ and $\mathcal{BE}$ are constructed differently, the former being survey based and the latter conditioning only on $\mathcal{P}_t$. Evidently, in addition to fundamental risks, learning has a significant impact on $\mathcal{CP}$'s implied risk premiums for one-year bonds. As such, there is nothing about the cyclical patterns in these errors per se that indicates irrational or inefficient forecasting by the consensus professional.

Note, however, that these visual similarities mask several important differences between the forecasts of $\mathcal{CP}$ and $\mathcal{BE}$. Revisiting Figures 5 and 6, the sensitivities of $\mathcal{CP}$'s forecast errors to its subjective forecast revisions $\mathcal{FR}^{\mathcal{CP}}(f; 9, 3)$ are indicated by the dotted vertical lines. In the case of $PC1$, $\mathcal{CP}$ is systematically to the right of both $\mathcal{BE}$ and the median of the professional forecasters, especially for the 10-year yield, which implies a *relative* underreaction of $\mathcal{CP}$ to news about $PC1$. In contrast, for news about $PC2$, $\mathcal{CP}$'s responses shift further



Panel A: $Proj\ (\mathcal{EH}^{1y}_{t+6} - E^{\mathcal{CP}}_t(\mathcal{EH}^{1y}_{t+6})|\mathcal{P}_t, H_t, \mathcal{FR}^{\mathcal{CP}}_t(f))$ Panel B: $Proj\ (\mathcal{TP}^{1y}_{t+6} - E^{\mathcal{CP}}_t(\mathcal{TP}^{1y}_{t+6})|\mathcal{P}_t, H_t, \mathcal{FR}^{\mathcal{CP}}_t(f))$

Panel C: $Proj\ (\mathcal{EH}^{1y}_{t+6} - E^{\mathcal{CP}}_t(\mathcal{EH}^{1y}_{t+6})|\mathcal{P}_t, H_t)$ Panel D: $Proj\ (\mathcal{TP}^{1y}_{t+6} - E^{\mathcal{CP}}_t(\mathcal{TP}^{1y}_{t+6})|\mathcal{P}_t, H_t)$

Figure 8. Fitted projections of $\mathcal{CP}$'s forecast errors. This figure compares fitted projections of $\mathcal{CP}$'s forecast errors of $\mathcal{EH}^m$ (left panels) and $\mathcal{TP}^m$ (right panels) onto $\mathcal{P}_t$, $(\mathcal{P}_t, H_t)$, or $(\mathcal{P}_t, H_t, \mathcal{FR}_t(f))$ one-year bond yields. (Color figure can be viewed at wileyonlinelibrary.com)

leftward with increasing maturity. There is sizable *relative* overreaction at the 10-year maturity to news about $PC2$. These patterns are connected (see below) to GLS' findings that $\mathcal{BE}$'s forecasts of y^{10y} are substantially more accurate than those of $\mathcal{CP}$.

Table III provides a broader comparison of the sensitivities of $\mathcal{CP}$ and $\mathcal{BE}$ for the case of $h = 6m$ and $\mathcal{FR}(f; 6, 6)$ (the latter involving $12m$ forecasts).[24] The signs of the loadings on $PC2$ for all three forecast errors flip between $\mathcal{BE}$ and $\mathcal{CP}$. Additionally, the coefficients on $\mathcal{FR}_t(PC2)$ for $\mathcal{CP}$ tend to be much larger in absolute value than their counterparts for $\mathcal{BE}$, and they are often statistically significant. Even for $\mathcal{FR}_t(PC1)$, the absolute value of $\mathcal{CP}$'s coefficient is twice as large as that of $\mathcal{BE}$. For the one-year term premiums, $\mathcal{CP}$ appears to overreact to revisions in expectations about slope and level relative to $\mathcal{BE}$. Finally, for $\mathcal{CP}$ there is relatively more explained variation in the errors for $\mathcal{TP}$ compared to $\mathcal{EH}$, whereas these R^2s are comparable for $\mathcal{BE}$. Indeed, the R^2 in the projection of $\mathcal{TP}^{1y}_t - E^{\mathcal{CP}}_{t-6}(\mathcal{TP}^{1y}_t)$ onto $\mathcal{FR}^{\mathcal{CP}}_t(f)$ alone is 20%. That is, $(\mathcal{P}_t, H_t)$ adds little

[24] Crump, Eusepi, and Moench (2018) express the one-period bond yield in terms of output growth and inflation, and then use long-dated forecasts of these macro variables to compute survey forecasts of average future short rates. Prior research suggests that yield curve dynamics are not well described by a small number of macro factors. I therefore restrict attention to the one-year-ahead yield forecasts of the BCFF professionals.

Table III

Comparing Projections of Forecast Errors for One-Year Yields across $\mathcal{BE}$ and $\mathcal{CP}$

This table presents projections of six-month-ahead forecast errors of y_t^m, $\mathcal{EH}^m$, and $\mathcal{TP}^m$ onto $\mathcal{FR}_t(f;6,6)$ (forecast revisions), $\mathcal{P}$ (yield *PC*s), and H (disagreement) for $m=1$ year and for $\mathcal{BE}$ and $\mathcal{CP}$. Robust t-statistics are in brackets.

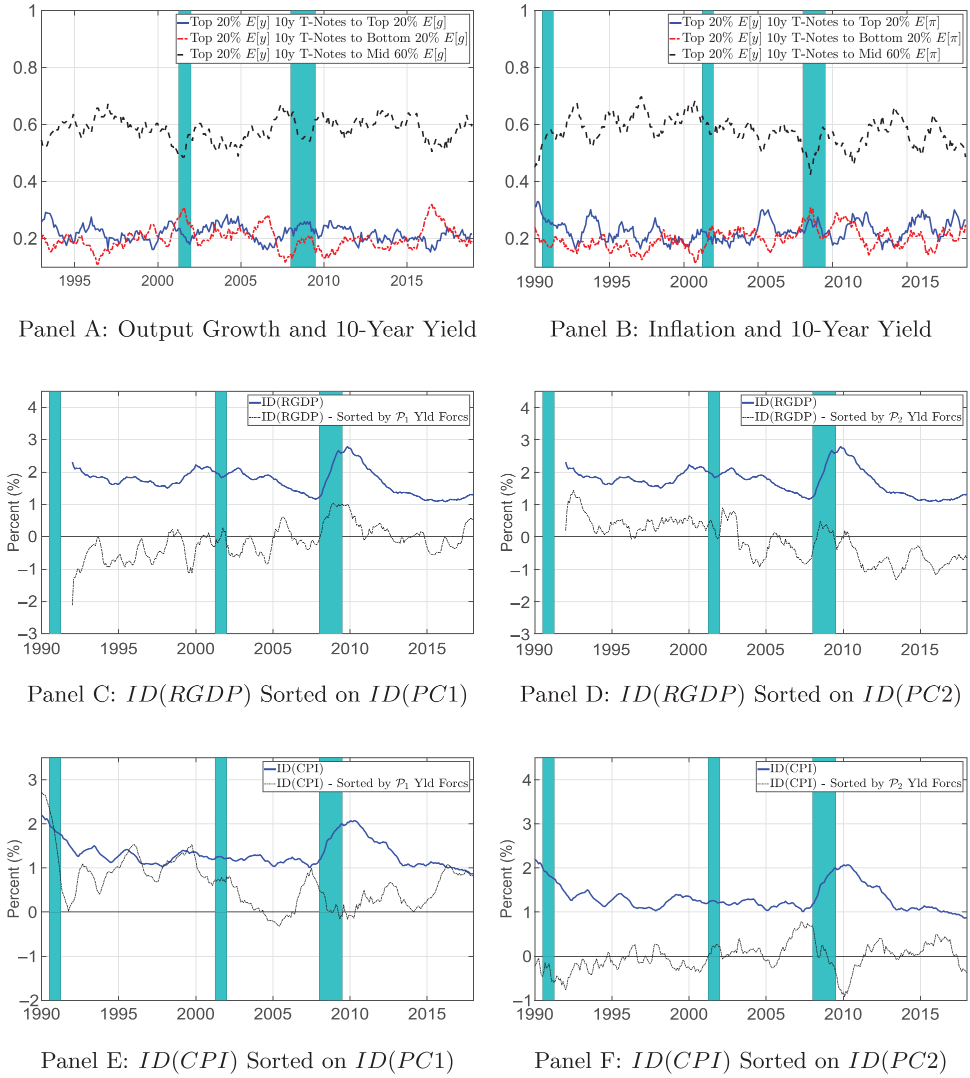
	$y_t^{1y}-E_{t-6}(y_t^{1y})$		$\mathcal{EH}_t^{1y}-E_{t-6}(\mathcal{EH}_t^{1y})$		$\mathcal{TP}_t^{1y}-E_{t-6}(\mathcal{TP}_t^{1y})$	
	$\mathcal{BE}$	$\mathcal{CP}$	$\mathcal{BE}$	$\mathcal{CP}$	$\mathcal{BE}$	$\mathcal{CP}$
$\mathcal{FR}_t(PC1)$	−0.019 [−0.345]	−0.004 [−0.070]	0.003 [0.067]	0.040 [0.742]	−0.022 [−2.320]	−0.042 [−3.278]
$\mathcal{FR}_t(PC2)$	0.165 [1.084]	0.171 [1.197]	0.173 [1.197]	0.253 [1.878]	−0.009 [−0.394]	−0.083 [−3.138]
$PC1$	−0.085 [−3.000]	−0.072 [−2.20]	−0.076 [−2.889]	−0.077 [−2.533]	−0.009 [−1.555]	0.005 [0.805]
$PC2$	−0.001 [−0.014]	0.018 [0.149]	−0.072 [−0.763]	0.057 [0.521]	0.071 [3.684]	−0.041 [−1.839]
$PC3$	0.859 [2.286]	0.570 [1.283]	0.865 [2.421]	0.620 [1.523]	−0.006 [−0.095]	−0.059 [−0.785]
$ID(2y)$	−0.079 [−0.314]	−0.530 [−1.865]	−0.078 [−0.314]	−0.437 [−1.612]	−0.002 [−0.030]	−0.095 [−1.830]
$ID(10y)$	0.677 [2.354]	0.811 [2.804]	0.615 [2.234]	0.768 [2.695]	0.061 [1.220]	0.047 [1.026]
Adj R^2	14.76%	11.58%	18.26%	14.81%	17.14%	23.61%

to the explained variation in $\mathcal{CP}$'s forecast error for $\mathcal{TP}$ after conditioning on $\mathcal{FR}_t^{\mathcal{CP}}(f)$.

These findings further document the inefficiencies of $\mathcal{CP}$'s forecasts relative to $\mathcal{BE}$'s constant-gain learning rule. Moreover, if one's goal is to better understand the median responses of professional forecasts to news in bond markets, focusing on the "representative consensus" $\mathcal{CP}$'s reactions will likely result in a distorted perspective.

V. Why Do Professional Forecasters Disagree?

Disagreement about bond yields in extant equilibrium models with belief dispersion is typically driven entirely by disagreement about core fundamentals (e.g., the long-run inflation rate or long-run rate of output growth). For the model of Xiong and Yan (2009), this is evident from (2). This tight connection arises because most models are essentially real business cycle models expanded to include inflation to price nominal bonds. This is true of the models discussed in Section I.A as well as of many representative agent models of the yield curve (e.g., Wachter (2006) and Bansal and Shaliastovich (2013)). Ehling et al. (2018) combine features of both frameworks to study the aggregate effects of inflation disagreement. This leads to direct mappings between disagreements about macroeconomic fundamentals and bond yields.



Panel A: Output Growth and 10-Year Yield

Panel B: Inflation and 10-Year Yield

Panel C: $ID(RGDP)$ Sorted on $ID(PC1)$

Panel D: $ID(RGDP)$ Sorted on $ID(PC2)$

Panel E: $ID(CPI)$ Sorted on $ID(PC1)$

Panel F: $ID(CPI)$ Sorted on $ID(PC2)$

Figure 9. This figure examines relationships between yield and macro disagreements. First row: fraction of BCFF professions holding the 20% highest one-year-ahead yield forecasts (for $m = 10$ years) who fall into the top 20%, mid 60%, and bottom 20% of forecasts for real GDP growth and inflation. Second and third rows: comparing the macro disagreements of those BCFF forecasters who determine the yield disagreements $ID(PC1)$ (left panels) or $ID(PC2)$ (right panels) with unrestricted macro disagreements. (Color figure can be viewed at wileyonlinelibrary.com)

A. Matching Yield and Macro ("Fundamental") Disagreements

To examine the relationship between forecasters' views on yields and macroeconomic fundamentals, I first report (first row of Figure 9) the fractions of BCFF professionals who hold the top 20% of one-year-ahead yield forecasts for $m = 10$ years who fall in the top 20%, mid 60%, and bottom 20% of one-year-ahead forecasts for real GDP growth and inflation. If each yield forecaster is randomly matched to macro forecasts, then we would expect 20% of these yield forecasters to fall within the highest macro forecasts, 60% to fall in the middle group, and 20% to fall in the bottom 20%. In fact, on average, the yield-fundamental matchings approximately satisfy this random assignment. The patterns for y^{2y} and y^{5y} are similar.

Taking another perspective, for each month I compute the cross-sectional disagreement about real GDP growth one year ahead, $ID(RGDP)$, using the interdecile range of BCFF forecasts and compare this to the disagreement about real GDP implied by those BCFF professionals who determine the cross-sectional disagreement about the one-year-ahead yield $PC1$ or $PC2$. Both measures of disagreement are smoothed by computing a rolling 12-month average. Figure 9, Panel C, for example, shows that those professionals who hold the 90% and 10% percentiles of $PC1$ forecasts—who determine $ID(PC1)$ today—have essentially no disagreement about future real economic growth (dashed line). However, there is substantial disagreement cross-sectionally among BCFF professionals about output growth (solid line). The same pattern emerges if yield disagreement is based on the slope of the yield curve (Figure 9, Panel D).

More nuanced results obtain for disagreement about consumer price index (CPI) inflation (last row of Figure 9). From roughly 1995 until 2000, those professionals whose views determine $ID(PC1)$ also disagree about future inflation, roughly to the extent of measured $ID(CPI)$. However, from about 2000 through 2016, there is no systematic and tight connection between actual and $ID(PC1)$-implied measures of inflation disagreement. Only in the last couple of years of my sample, when short rates were pegged near zero, does alignment return.

B. On the Term Structures of Disagreement

Thus far I focus on disagreement over a one-year forecast horizon. BCFF provides monthly data on forecasts at quarterly intervals out to one year, and it also surveys long-term forecasts at semiannual intervals. As Andrade et al. (2016) highlight, an interesting feature of the term structures of disagreement—measured disagreement for a given variable as the forecast horizon increases—is that they slope upward for bond yields and downward for inflation and output growth. This pattern, which is illustrated in Figure 10 for the yield on the 10-year Treasury bond and the real GDP growth rate, raises additional questions about the determinants of disagreements.

Focusing on inflation (π), real growth (g), and the federal funds rate (FFR), Andrade et al. (2016) rationalize these patterns within a reduced-form "sticky information" model that has each individual investor observing (π_t, g_t, FFR_t) with probabilities less than one. Their analysis is purely descriptive of the dynamics of disagreements about (π_t, g_t, FFR_t); they do not price assets. Similar to my concerns about the noisy information friction discussed in Section I.B, this informational structure does not seem natural within a model of bond

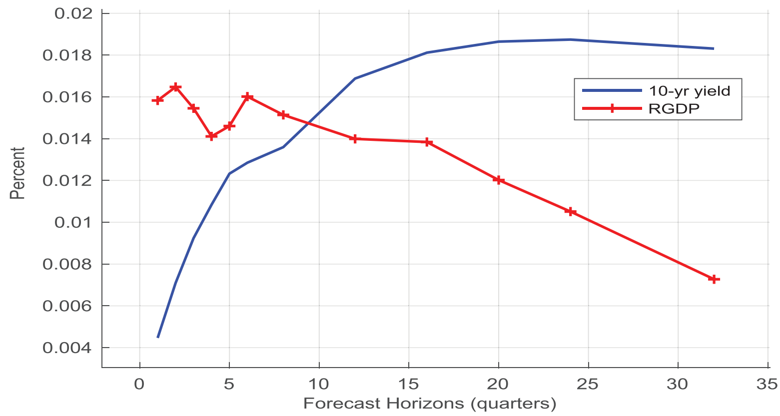


Figure 10. Term structures of disagreement. This figure shows the term structures of disagreement about the 10-year Treasury bond yield and the growth rate of real GDP. Sample: January 1984 to December 2018. (Color figure can be viewed at wileyonlinelibrary.com)

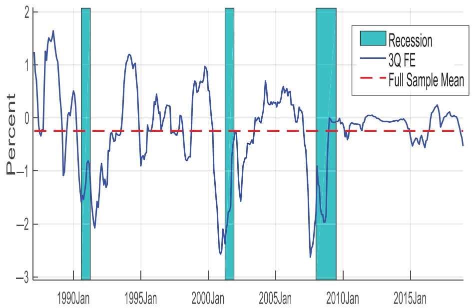
prices. The term structure of yield disagreement is upward-sloping for yields on bonds of all maturities (e.g., y^{10y} in Figure 10). Moreover, yields are easily observed by all market participants, which casts doubt on sticky information as the explanation for these patterns.

Of course, macroeconomic variables may well be observed with noise, especially upon their initial releases. Modifying term structure models to allow for imperfect measurement of macro variables, while having perfectly observed bond yields, therefore seems desirable. For example, entering (π_t, g_t) as risk factors measured without error in a low-dimensional factor model can be problematic because, as illustrated in an affine setting by Joslin, Le, and Singleton (2013b), the resulting models produce enormous pricing errors for bonds. Treating (π_t, g_t) as measured with error and using filtering in estimation solves this mispricing problem. However, the likelihood function reduces pricing errors by choosing filtered $(\hat{\pi}_t, \hat{g}_t)$ that mimic low-order yield *PC*s, for it is these yield *PC*s that accurately price bonds.

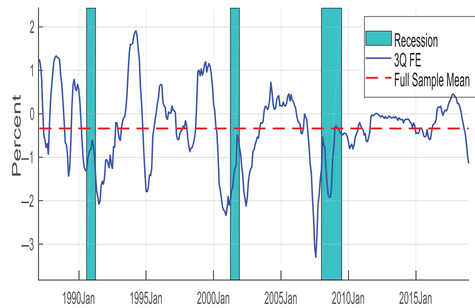
As the literature moves forward to develop models that rationalize the weak correlations between investor-matched yield and fundamental disagreements, the rich patterns in the term structures of disagreements are likely to provide useful discipline.

C. What Is Missing from Our Equilibrium Models?

While yield disagreement is only weakly connected to disagreement about real growth and inflation, there may be stronger connections with beliefs about other economic fundamentals.



Panel A: $FFR_{t+9} - E_t^{\mathcal{CP}}(FFR_{t+9})$



Panel B: $y_{t+9}^{1y} - E_t^{\mathcal{CP}}(y_{t+9}^{1y})$

Figure 11. Realized forecast errors from $\mathcal{CP}$. This figures shows $\mathcal{CP}$'s realized forecast errors for the *FFR* (Panel A) and one-year yield (Panel B) over a nine-month horizon. (Color figure can be viewed at wileyonlinelibrary.com)

C.1. Monetary Policy

Perhaps the most prominent candidate for a closer tie to yield disagreement is dispersion of beliefs about monetary policy. Since monetary policy is often closely tied to the short-term interest rate, one strategy for introducing a role for central bank interventions into dynamic term structure models is to view the core factor model for the short-term rate as representative of a Taylor-style policy rule (e.g., Ang et al. (2011)). This modeling strategy runs up against two conceptual issues, however, even before addressing investor heterogeneity. The first is the identification of the policy rule. Standard representations of Taylor rules have

$$r_t = \rho_0 + \ \rho_\pi \pi_t + \rho_g g_t + \rho_\ell \ell_t, \tag{26}$$

where g_t is interpreted as an output gap and ℓ_t as the policy shock. As discussed in Joslin, Le, and Singleton (2013a), without the imposition of additional economic structure, the parameters of (26) are not econometrically identified. Intuitively, this is because (26) can be rotated to observationally equivalent expressions within the family of arbitrage-free term structure models. The second is the large deterioration in fits to yields with (π_t, g_t) as factors.

These issues are circumvented entirely by treating the policy rate r_t as the monetary instrument and interpreting views about its future path as beliefs about monetary policy. Examples of this approach are the analyses of survey responses about the *FFR* (the policy rate) in Cieslak (2018) and Cao et al. (2020). Cieslak focuses on the errors of the cross-sectional median forecaster of *FFR* (her version of my consensus pseudo-agent $\mathcal{CP}$). Figure 11, Panel A, documents her central motivating evidence that *FFR* forecast errors are systematically negative during and shortly after NBER recessions, a pattern that she interprets as professionals underestimating the extent of monetary easing.[25]

It is instructive to parse interpretations of Figure 11, Panel A, into two questions: (i) Could optimal learning induce patterns in forecast errors for ***FFR*** that have been attributed to inefficient forecasting or omission of relevant information and (ii) how do the forecast errors of $\mathcal{CP}$ compare to those of an efficient learner in the Treasury market? Regarding (i), note that the pattern of the errors in Figure 11, Panel A, is not specific to ***FFR***; similar patterns arise for y^{1y} (Figure 11, Panel B), and for y^{6m} and y^{2y} as well. Shocks to ***FFR*** have impacts that extend far out the maturity spectrum (e.g., Fleming and Remolona (1999), Piazzesi (2005)). Therefore, as I document above, it would be natural for a Bayesian agent learning about the monetary policy rule and/or its impact on yield factor dynamics to have "optimally" persistent forecast errors with these documented cyclical patterns.

Yet, turning to (ii), notice that the sharp downward and upward spikes in $\mathcal{CP}$'s ***FFR*** forecast errors in Figure 11, Panel A, correspond quite closely in timing to the sharp moves in the fitted forecast errors of $\mathcal{CP}$ for $\mathcal{EH}^{1y}_{t+6}$ conditioned on $\mathcal{I}_t = (\mathcal{P}_t, H_t, \mathcal{FR}_t(f))$ (Figure 8, Panel A). As noted when discussing Table II, this fluctuation, which is not present when conditioning on $\mathcal{P}_t$ alone, is due to $\mathcal{CP}$'s heightened sensitivities to $\mathcal{FR}(f)$.[26] $\mathcal{CP}$'s strong response to $\mathcal{FR}(f)$ is not a characteristic of $\mathcal{BE}$'s errors for y^{1y} (Figure 7, Panel A). In other words, the cyclicality of $\mathcal{CP}$'s forecast errors per se is not anomalous relative to an efficient learning rule, but rather his exaggerated responses to *perceptions of news* about $(PC1, PC2)$ stand out as anomalous (relative to a candidate Bayesian rule).

A different, longer horizon, component of monetary policy is highlighted in Bianchi, Lettau, and Ludvigson (2021). They document long-term regime shifts in Federal Reserve policy that coincide with low-frequency shifts in the real ***FFR*** and equity market return premia. The representative agent in their model recognizes when a shift has occurred, but is learning about the duration of the regime they are currently in. $\mathcal{BE}$'s learning rule allows for shifts in the monetary policy rules through the unforeseen changes in the factor dynamics but, as framed here, her rule does not recognize the possibility of shifts across persistent regimes.

C.2. Fiscal Policy

Might disparate perspectives on fiscal policy's impact on bond markets drive dispersion in beliefs about future yields? Krishnamurthy and

[25] Note that to facilitate comparison with my earlier and subsequent analyses, I use a forecast horizon of nine months, with the patterns virtually identical for the 12-month horizon. Additionally, I report the median forecast error as of the date of the forecast, while Cieslak reports on the date of the realization of ***FFR***. That is, my figures shift the counterpart to her data forward by nine months.

[26] $\mathcal{CP}$'s fitted errors based on $\mathcal{I}_t = \mathcal{P}_t$ tend to adjust toward zero during recessions earlier and more smoothly.

Table IV

Full-Sample Projections of Realized Excess Returns

This table presents full-sample projections of the realized excess returns xr_{t+12} over a one-year horizon onto date t information $\mathcal{P}_t$ and either H_t or $E[PU_t^{fis}|H_t]$, the fitted projection of measured fiscal uncertainty onto disagreement H_t.

	xr^{2y}	xr^{2y}	xr^{10y}	xr^{10y}
$\mathcal{P}_1$	1.467 [3.051]	1.531 [3.315]	6.103 [1.902]	7.248 [2.153]
$\mathcal{P}_2$	−0.325 [−2.456]	−0.347 [−2.845]	−3.584 [−4.121]	−3.976 [−4.600]
$\mathcal{P}_3$	0.868 [1.178]	0.869 [1.175]	−6.575 [−1.401]	−6.562 [−1.341]
$ID(2y)$	0.579 [1.321]	– [–]	6.559 [2.362]	– [–]
$ID(10y)$	−1.666 [−3.771]	– [–]	−12.31 [−4.342]	– [–]
$E[PU^{fis}|H]$	– [–]	−0.601 [−4.232]	– [–]	−3.871 [−3.753]
Adj R^2	26.2%	26.1%	30.9%	29.5%

Vissing-Jorgensen (2011, 2012) show that the aggregate supply of bonds impacts risk premiums. Leippold and Matthys (2017) and Kucera, Kocenda, and Marsal (2019) document significant impulse responses of bond yields to shocks to news-based measures of policy uncertainty.

GLS project the text-based measure of fiscal uncertainty PU^{fis} from Backer, Bloom, and Davis (2016) onto $H' = (ID(y^{2y}), ID(y^{10y}))$ to get fitted values $E[PU^{fis}|H]$. When $E[PU^{fis}|H]$ was substituted for the conditioning information H in an expanded learning rule $\ell_{CG}(\mathcal{P}, H)$ for $\mathcal{BE}$, the out-of-sample root mean squared forecast errors (RMSEs) are no larger than, and sometimes smaller than, the RMSEs for $\ell_{CG}(\mathcal{P}, H)$. That is, the component $E[PU^{fis}|H]$ of fiscal uncertainty captures all of the relevant information in yield disagreement H for forecasting future excess returns on Treasury bonds.

The full-sample counterparts to their findings are displayed in Table IV. For excess returns over a one-year horizon on both the two- and 10-year bonds, substituting the fitted component of fiscal uncertainty for the disagreement measures H leads to nearly identical R^2s. The matched loadings on $\mathcal{P}_t$ are similar. Although $E[PU^{fis}|H]$ is not matched directly to the professionals' disagreement about fiscal policy, this evidence hints at intriguing interactions among fiscal activity, disagreement, and risk premiums in Treasury markets.

C.3. Funding Frictions and Capital Flows

More broadly, might disparate beliefs about asset demands or supplies, deployable/investable capital, or time-variation in the "marginal" investor underlie the strong predictive power of yield disagreement for bond market risk premiums? Financial institutions give considerable weight to the flows of funds in financial markets when forecasting future prices/yields, suggesting, at least implicitly, that dispersion of viewpoints about fund flows is relevant for understanding price dynamics.[27] Hanson et al. (2018) document a shift in the joint dynamics of short- and long-term bond yields around 2000, and they interpret this evidence within a model that builds on the arbitrage frictions in Vayanos and Villa (2009) and the presence of slow-moving capital as in Duffie (2010). Particular attention is given to the impact of past changes in short-term rates on term premiums.

In highlighting potential roles for fund flows, Hanson et al. (2018) abstract from learning. Their central empirical observation that the *PC*s of bond yields do not follow a (first-order) Markov process is an inherent feature of the learning friction faced by $\mathcal{BE}$.[28] As we have already seen, learning also gives rise to a significant role for conditioning information beyond $\mathcal{P}_t$ when forecasting excess returns. Most central to my focus, their analysis raises the question of whether the beliefs of the BCFF professionals or $\mathcal{BE}$ changed in a material way around 2000. I provide some evidence on this below.

Frictions related to fund flows are also often modeled in settings where investors have homogeneous beliefs and, as such, they do not address the "source of disagreement" puzzle. Heyerdahl-Larsen and Illeditsch (2020) focus specifically on this puzzle in their development of an equilibrium model in which different beliefs about future asset demands arise, even when economic fundamentals are known. Their framework formalizes some of these demand frictions in a way that gives rise to yield disagreement that shows weak correlation to output growth or inflation. Similar asset demand uncertainty arises in the model of Albuquerque et al. (2016), though without heterogeneity across agents.

VI. How Rational Are Professionals' Yield Forecasts?

The median across professionals' reactions to their own revisions in expectations about $PC1$ nine months ahead ($\mathcal{FR}^{\mathcal{L}}(PC1; 9, 3)$)—the cross-professional median $\beta^{10y,9}_{\mathcal{FR}_1}$—is similar to the Bayesian $\mathcal{BE}$'s reaction coefficient.[29] This finding raises the question as to whether the risk premiums of the median-$\beta^{10y,9}_{\mathcal{FR}1}$ forecaster and of $\mathcal{BE}$ are similar. This might reasonably be expected if (nearly) identical loadings on forecast revisions revealed that these two forecasters faced similar informational frictions. Figure 12 displays the time series of expected excess returns on the 10-year bond over a one-year horizon for $\mathcal{BE}$, $\mathcal{CP}$, and the pseudo-forecaster with median $\beta^{10y,9}_{\mathcal{FR}_1}$. Although the forecasts of $\mathcal{CP}$ and median-$\beta^{10y,9}_{\mathcal{FR}_1}$ are set in different ways, their implied risk premiums broadly track each other over time and are substantially lower than those of

[27] Complementary to the above-cited work on fiscal deficits, evidence that flows of funds in the mortgage markets have impacted risk premiums is provided in Hanson (2014) and Malkhozov et al. (2016).

[28] See also Joslin, Le, and Singleton (2013a) for dynamic term structure models with higher order lags.

[29] In contrast, the BCFF consensus forecaster $\mathcal{CP}$ not only placed different weight on $\mathcal{FR}(f)$, but also had very different loadings (from $\mathcal{BE}$) on the slope of the yield curve.

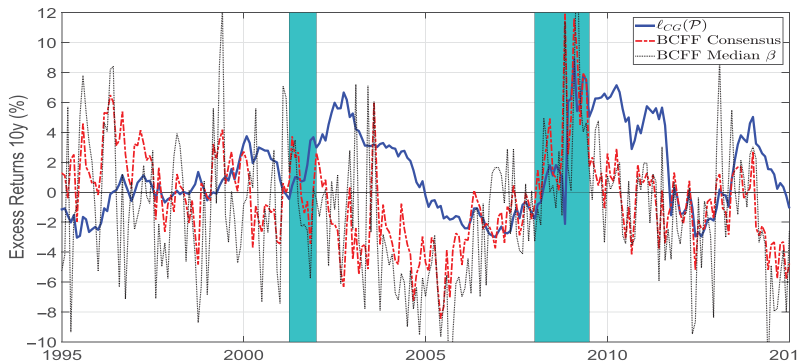


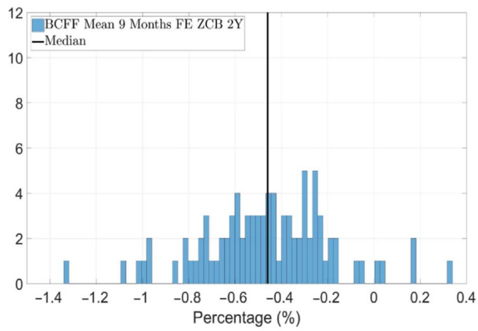
Figure 12. Expected excess returns on the 10-year bond. This figure shows the time series of expected excess returns on the 10-year bond over a one-year horizon for $\mathcal{BE}$ (GPS's rule $\ell_{CG}(\mathcal{P})$), $\mathcal{CP}$, and the forecaster who has the median reaction to revisions in expectations about $PC1$ nine months ahead ($\mathcal{FR}(PC1; 9, 3)$). (Color figure can be viewed at wileyonlinelibrary.com)

$\mathcal{BE}$ following recessions. Expressed in yield space, following NBER recessions $\mathcal{CP}$ and median-$\beta^{10y,9}_{\mathcal{FR}_1}$ expected the 10-year yield to rise more rapidly than $\mathcal{BE}$ expected. With hindsight, we know that $\mathcal{BE}$ was the more accurate forecaster (GLS), substantially so following recessions.

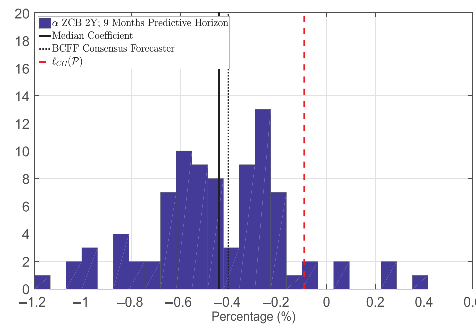
This finding deepens the puzzle regarding the outperformance of $\mathcal{BE}$ relative to median-$\beta^{10y,9}_{\mathcal{FR}_1}$. If it is not differences in the way $\mathcal{BE}$ and the individual BCFF forecasters react to their own revisions in expectation, then what is driving the outperformance of $\mathcal{BE}$ relative to the BCFF forecasters? The fitted forecast errors hint at a potential answer. The time-series mean of $\mathcal{BE}$'s error in forecasting y^{1y} is approximately zero (Figure 7, Panel A), while there was a small negative bias in her average forecast error for y^{10y}. In contrast, Figure 8, Panel A, shows a systematic and large negative bias in $\mathcal{CP}$'s fitted forecast error for y^{1y}. The latter is to be expected if most of the individual forecasters exhibit negatively biased errors. This is in fact the case as illustrated in Figure 13, Panel A, for y^{2y} over a nine-month horizon.[30] What aspects of learning rules might give rise to this shared systematic bias across these professionals?

When studying biases at the individual level, researchers face several practical limitations with the BCFF survey data. Computation of the mean of a professional's forecast errors is straightforward, because errors are sourced directly from the raw data. However, analyzing the sensitivity of these errors to various conditioning variables requires additional estimation. Also, survey participants enter and leave the survey, and some have short tenures as respondents. With these cautionary observations as background, I next examine how BCFF professionals adjust their beliefs in response to shocks to the level and slope of the yield curve. Toward this end, I restrict attention to respondents who participated in the BCFF survey for at least 36 months, and I focus on projections of their forecast errors onto the risk factors $\mathcal{P}_t$.

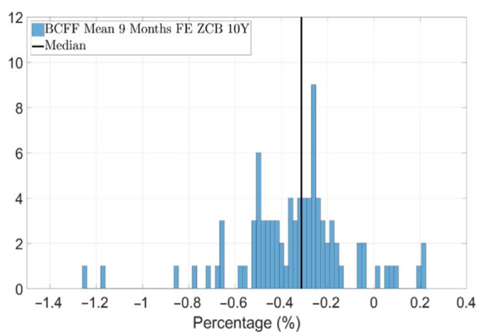
[30] This is equally true of other maturities, including the short-term rate y^{6m}.



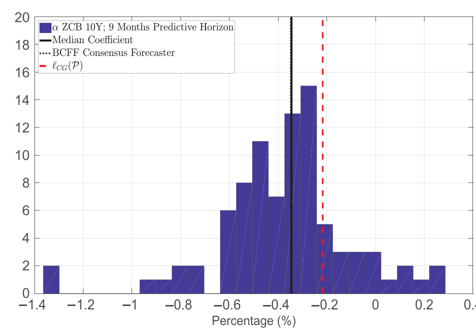
Panel A: y^{2y}: BCFF Forecast Error Means



Panel B: y^{2y}: BCFF $\hat{\alpha}_0^{2y,9}$



Panel C: y^{10y}: BCFF Forecast Error Means

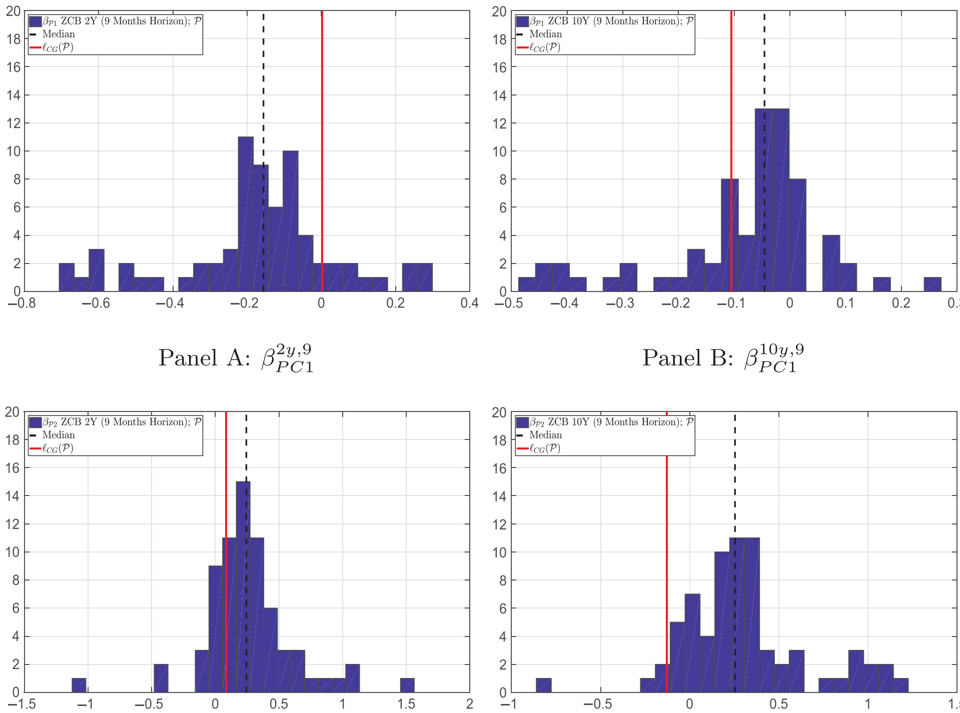


Panel D: y^{10y}: BCFF $\hat{\alpha}_0^{10y,9}$

Figure 13. Frequency distributions of mean forecast errors and $\hat{\alpha}^{m,9}$. For $m = 2y$ and $10y$, this figure shows the frequency distributions of mean forecast errors of the BCFF forecasters over a nine-month horizon (left panel) and fitted intercepts in the time-series projections of individual forecast errors onto a constant and $\mathcal{FR}(f; 9, 3)$ (right panel). (Color figure can be viewed at wileyonlinelibrary.com)

As a first step, it is instructive to revisit the projections (21) for individual professionals onto own forecast revisions $\mathcal{FR}_t^{\mathcal{L}}(f; 9, 3)$. If the means of these "news" variables are approximately zero, then the estimated intercepts α_0^{m9} in (21) should be equal to their biases. Comparison of the matching left and right panels of Figure 13 verifies that this is indeed the case. The medians in matched figures are about -0.45% for y^{2y} and -0.38% for y^{10y}, and all of these distributions are somewhat bimodal. Equally notable, the medians of the fitted intercepts $\hat{\alpha}_0^{m,9}$ for the BCFF professionals in the right-hand side figures are very close to the intercepts for $\mathcal{CP}$; the consensus BCFF forecaster inherits the median cross-forecaster bias.

To what extent are these biases associated with how the BCFF professionals process information about the level and slope of the yield curve? Figure 14 displays the frequency distributions of the loadings $\beta_{PC1}^{m,9}$ on $PC1$ from the forecast error projections for y^{2y} (left column) and y^{10y} (right column) onto $(1, \mathcal{P}_t)$. The median partial correlations with $PC1$ are negative, and there are



Panel A: $\beta^{2y,9}_{PC1}$ Panel B: $\beta^{10y,9}_{PC1}$

Panel C: $\beta^{2y,9}_{PC2}$ Panel D: $\beta^{10y,9}_{PC2}$

Figure 14. Frequency distributions of fitted $\boldsymbol{\beta^{m,9}_{f}}$. This figures shows the frequency distributions of fitted $\beta^{m,9}_{f}$ for $PC1$ and $PC2$ from projections of individual forecast errors over a nine-month horizon onto $(1, \mathcal{P}'_t)$ for the $m = 2y$ and $10y$ bond yields. (Color figure can be viewed at wileyonlinelibrary.com)

long left tails, especially for y^{2y}. As maturity m increases to 10 years, there is an increasingly rightward shift in the median partial correlation of the professionals' forecast errors with $PC1$ (toward zero), and the substantial right tail similarly shrinks toward zero. There is much weaker predictability of forecast errors by $PC1$ for most of the professionals for the five- and 10-year yields.

These patterns for the BCFF professionals stand in sharp contrast to $\mathcal{BE}$'s forecasts. Her errors in forecasting y^{2y} are relatively unpredictable by $PC1$ and overall they are unbiased. As maturity increases, the estimated $\beta^{m,9}_{PC1}$ shows a moderate leftward shift into the negative region. This is symptomatic of a small negative bias in $\mathcal{BE}$'s forecasts of y^{10y}_{t+6} as can be seen from the negative fitted errors in Figure 7, Panel B, and $\alpha^{10y,9}_{0} < 0$ in Figure 13, Panel D.

These findings align with professionals who were consistently surprised by the downward shifts in yields during most of the sample period; they were persistently predicting levels of yields that were too high (negative errors). Given that the corresponding figure for the six-month yield is very similar, one might reasonably wonder whether professionals were surprised by the dovish policies of the Federal Reserve. $\mathcal{BE}$, in contrast, perhaps owing to the accommodation of shifting long-run means in her learning rule as discussed in Section I.C, produced forecasts that were relatively unbiased and uncorrelated with the past level of bond yields. $\mathcal{BE}$'s rule more efficiently addressed *secular trends*.

The second row of Figure 14 shows the partial correlations of the professionals' forecast errors with *PC*2. Across the maturity spectrum, a steep yield curve predicts positive forecast errors. Moreover, the median correlations increase with maturity, implying larger forecast errors when the yield curve is steep as m increases. The coefficients β^{m9}_{PC2} for $\mathcal{BE}$, in contrast, are near zero across the maturity spectrum. Indeed, almost all of the support of the frequency distribution of the BCFF respondents lies to the right of $\mathcal{BE}$'s β^{m9}_{PC2}. I find this result striking given that $\mathcal{BE}$'s risk premiums track the slope of the Treasury curve quite closely (Figure 12). Here again, and now at the level of individual professional forecasters, there is evidence of anomalous reactions to news about *PC*2, relative to the Bayesian $\mathcal{BE}$. These patterns likely contribute to the *cyclical* outperformance of $\mathcal{BE}$ over the professionals, especially following recessions.

VII. Concluding Remarks

Financial theory, along with the practicalities arising from informational frictions, implies that rational agents learning about the dynamics of the risk factors in bond markets will have serially correlated yield forecast errors. Accordingly, in assessing whether professional forecasters hold rational beliefs, it is useful to compare the properties of individual forecasts to those of a benchmark rational forecaster. For this purpose, I use the learning rule of GLS's "outside observer-econometrician" $\mathcal{BE}$ who follows a (constrained) Bayesian rule for learning about the dynamics of the *PC*s of bond yields.

I find serial dependence in the forecast errors of the BCFF professionals at both the individual and the "consensus" levels. The vast majority of these forecasters have negatively biased forecast errors (forecasts persistently above the realized yields), both in an absolute sense and relative to $\mathcal{BE}$'s forecasts (especially for short-maturity bonds, for example, y^{2y} versus y^{10y}). Professionals' biases have a negative "level" component throughout much of the sample, as if survey respondents were systematically surprised by the extent to which the levels of yields were declining. The absolute magnitudes of their biases also fluctuate over the business cycle with the slope of the yield curve. When the yield curve was relatively steep following NBER recessions, the accuracy of the BCFF forecasts deteriorated substantially relative to $\mathcal{BE}$.

Exploration of the sources of these biases reveals that the common focus in the literature on revisions in forecasts—an agent's *subjective* news about the risk factors—provides an incomplete picture of informational frictions. Comparison of Tables I and II shows that the predictive power of $\mathcal{FR}(f;6,6)$ for $\mathcal{BE}$'s forecast errors is very weak (statistically zero) for yields on bonds out to two years to maturity. Rather, it is primarily the yield factors ($\mathcal{P}_t$) and disagreement (H_t) that drive persistence in $\mathcal{BE}$'s errors. Disagreement H and forecast revisions have roughly equal explanatory power for forecast errors on five- and 10-year bond yields. In contrast, the subjective revisions in expectations of the BCFF professionals are the dominant factors behind their forecast errors.

There are hints that perceptions about monetary policy affect these patterns in forecast errors, at least for the shorter end of the curve, and that uncertainty about fiscal policy may (at least partially) underlie the yield forecast errors for longer maturity bonds. These possibilities, and more broadly the potential roles for demand shocks, warrant further exploration using individual survey responses. Of equal interest, how should we interpret the explanatory power of disagreement for the forecast errors of the BCFF professionals? Must disagreement about future yields ultimately find its roots in disagreement about a set of fundamentals that we have heretofore not identified? Or might the role of disagreement have more subtle origins within the equilibrium price setting process?

More economic structure is needed to draw firmer conclusions about investors' learning challenges. Many of the characteristics of beliefs documented here show up, to varying degrees, along the entire maturity spectrum of yields. Cieslak (2018) and Gomez-Cram and Yaron (2020), using very different frameworks, argue that variation in interest rate risk was largely a real not an inflation phenomenon for most of my sample period. Kurmann and Ostrok (2013) argue that 50% or more of the unpredictable movements in the yield-curve slope are associated with news about future total factor productivity, with a key driver of this correlation being the *endogenous* response of monetary policy. I document a strong correlation between yield disagreements and future forecast errors for both $\mathcal{BE}$ and $\mathcal{CP}$. My reduced-form focus on learning about yields is silent about the contributions of endogenous versus exogenous shocks (be they fiscal, monetary, or of some other origin) to forecast errors.

Using the constant-gain learner $\mathcal{BE}$ as a reference, my evidence suggests that the majority of the professionals were submitting inefficient real-time forecasts based on widely available public information about yield curve dynamics. Yet. the evidence also points to $\mathcal{BE}$'s learning rule as having systematically failed to capture relevant shifts in the factor dynamics, particularly as they relate to the long end of the yield curve. GLS derive $\mathcal{BE}$'s Bayesian rule under some simplifying assumptions about the evolution of the unknown parameters and her priors. Among the features absent from her priors are long-term regime shifts.

More generally, how might ambiguity about which specifications of factor dynamics are most relevant for bond pricing have impacted learning by $\mathcal{BE}$ and the BCFF professionals? Hansen and Sargent (2021, 2020) explore the interplay between model ambiguity and potential misspecification that arises in the presence of uncertainty. Learning would further enrich decision making with ambiguity-averse preferences and the resultant model risk premiums (e.g., Leippold, Trojani, and Vanini (2008)). Exploration of these issues within this term structure setting represents an interesting avenue for future research.

Throughout this analysis I have strived to navigate an informative pathway between the properties of the beliefs of the individual BCFF professionals and the "consensus" BCFF forecaster $\mathcal{CP}$ who has been the focus of much of the extant empirical work using survey data. Depending on the dimension of beliefs of interest, I find that the median response across professionals to an informational signal may be very different than the response to this same signal by the pseudo-forecaster who holds the median forecast of bond yields. As such, some caution seems warranted when viewing $\mathcal{CP}$ as representative of the professional survey respondents. There is also tremendous heterogeneity among the professionals in terms of how their forecast errors correlate with their own perceptions of news about the risk factors. How then should we interpret equilibrium analyses of learning by a representative agent?

Finally, I have left two central issues virtually unaddressed. First, what explains the breadth of the reactions of professionals to economic "news" that is relevant for pricing bonds? Do plausible differences in priors or attitudes toward ambiguity/uncertainty induce the diverse response coefficients documented here? Or might incentives beyond the production of efficient forecasts in the classical econometric sense partially underlie this heterogeneity? Second, this analysis looks at survey data through the lens of an outside observer. What are the properties of equilibria in bond markets that arise when market participants follow *diverse* Bayesian learning rules or show *differing degrees of* ambiguity aversion in formulating their heterogeneous forecasts and investment strategies? My descriptive evidence challenges the plausibility of some recent framings of informational frictions underlying equilibrium dispersion of beliefs. Hopefully it also points to fruitful paths for enriching extant heterogeneous-agent bond pricing models.

REFERENCES

Acemoglu, Daron, Victor Chernozhukov, and Muhamet Yildiz, 2016, Faragility of asymptotic agreement under Bayesian learning, *Theoretical Economics* 11, 187–225.

Adam, K., and A. Marcet, 2011, Internal rationality, imperfect market knowledge and asset prices, *Journal of Economic Theory* 146, 1244–1252.

Albuquerque, Rui, Martin Eichenbaum, Victor Luo, and Sergio Rebelo, 2016, Valuation risk and asset pricing, *Journal of Finance* LXXI, 2861–2903.

Andrade, Philippe, Richard K. Crump, Stefano Eusepi, and Emanuel Moench, 2016, Fundamental disagreement, *Journal of Monetary Economics* 83, 106–128.

Ang, Andrew, Jean Boivin, Sen Dong, and Rudy Loo-Kung, 2011, Monetary policy shifts and the term structure, *The Review of Economic Studies* 78, 429–457.

Bacchetta, Philippe, and Eric Van Wincoop, 2006, Can information heterogeneity explain the exchange rate determination puzzle? *American Economic Review* 96, 552–576.

Backer, Scott R., Nicholas Bloom, and Steven J. Davis, 2016, Measuring economic policy uncertainty, *Quarterly Journal of Economics* 131, 1593–1636.

Bansal, R., and I. Shaliastovich, 2013, A long-run risks explanation of predictability puzzles in bond and currency markets, *Review of Financial Studies* 26, 1–33.

Bauer, Michael D., and Glenn D. Rudebusch, 2020, Interest rates and falling stars, *American Economic Review* 110, 1316–1354.

Bianchi, Francesco, Martin Lettau, and Sydney Ludvigson, 2021, Monetary policy and asset valuation, Technical report, Duke University.

Bianchi, Francesco, Sydney Ludvigson, and Sai Ma, 2021, Belief distortions and macroeconomic fluctuations, Technical report, NBER Working Paper 27406.

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

Buraschi, Andrea, Ilaria Piatti, and Paul Whelan, 2021, Subjective bond returns and belief aggregation, Technical report, Imperial College.

Buraschi, Andrea, and P. Whelan, 2016, Speculation, hedging, and interest rates, Technical report, Working paper, Copenhagen Business School.

Cao, Shuo, Richard Crump, Stefano Eusepi, and Emanuel Moench, 2020, Fundamental disagreement about monetary policy and the term structure of interest rates, Technical report, FRB New York.

Cieslak, A., 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.

Coibion, O., and Y. Gorodnichenko, 2012, What can survey forecasts tell us about informational rigidities? *Journal of Political Economy* 120, 116–159.

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

Collin-Dufresne, P., M. Johannes, and L. Lochstoer, 2016, Parameter learning in general equilibrium: The asset pricing implications, *American Economic Review* 106, 664–698.

Crump, Richard, Stefano Eusepi, and Emanuel Moench, 2018, The term structure of expectations and bond yields, Technical report, FRBNY Staff Report No. 775.

Cujean, Julien, and Michael Hasler, 2017, Why does return predictability concentrate in bad times? *Journal of Finance* LXXII, 2717–2754.

Dai, Qiang, and Kenneth J. Singleton, 2000, Specification analysis of affine term structure models, *Journal of Finance* 55, 1943–1978.

Dai, Qiang, and Kenneth J. Singleton, 2002, Expectations puzzles, time-varying risk premia, and affine models of the term structure, *Journal of Financial Economics* 63, 415–441.

d'Arienzo, Daniele, 2020, Maturity increasing overreaction and bond market puzzles, Technical report, Bocconi University.

David, Alexander, 2008, Heterogeneous beliefs, speculation, and the equity premium, *Journal of Finance* 63, 41–83.

Duffee, Gregory R., 2011, Forecasting with the term structure: The role of no-arbitrage, Technical report, Economics Working Paper No. 576, Johns Hopkins University.

Duffie, D., 2010, Asset price dynamics with slow-moving capital, *Journal of Finance* LXV, 1237–1267.

Duffie, D., and R. Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, 379–406.

Ehling, Paul, Michael Gallmeyer, Christian Heyerdahl-Larsen, and Philipp Illeditsch, 2018, Disagreement about inflation and the yield curve, *Journal of Financial Economics* 3, 459–484.

Fama, E., 2006, The behavior of interest rates, *Review of Financial Studies* 19, 359–379.

Farmer, Leland, Emi Nakamura, and Jon Steinsson, 2021, Learning about the long run, Technical report, UC Berkeley.

Fleming, M. J., and E. Remolona, 1999, The term structure of announcement effects, FRB New York Staff Report No. 76.

Giacoletti, M., K. Laursen, and K. Singleton, 2021, Learning from disagreement in the us treasury bond market, *Journal of Finance* 76, 395–441.

Gomez-Cram, Roberto, and Amir Yaron, 2020, How important are inflation expectations for the nominal yield curve? *Review of Financial Studies forthcoming*.

Hansen, Lars, and Thomas Sargent, 2020, Structured ambiguity and model misspecification, *Journal of Economic Theory* forthcoming.

Hansen, Lars, and Thomas Sargent, 2021, Macroeconomic uncertainty prices when beliefs are tenuous, *Journal of Econometrics* 223, 222–250.

Hanson, Samuel, 2014, Mortgage convexity, *Journal of Financial Economics* 113, 270–299.

Hanson, Samuel, David Lucca, and Jonathan Wright, 2018, The excess sensitivity of long-term rates: A tale of two frequencies, Technical report, FRBNY Staff Report No. 810.

Heyerdahl-Larsen, Christian, and Philipp Illeditsch, 2020, Demand disagreement, Technical report, Texas A&M University.
Joslin, Scott, Anh Le, and Kenneth J. Singleton, 2013a, Gaussian macro-finance term structure models with lags, *Journal of Financial Econometrics* 11, 581–609.
Joslin, Scott, Anh Le, and Kenneth J. Singleton, 2013b, Why Gaussian macro-finance term structure models are (nearly) unconstrained factor-VARs, *Journal of Financial Economics* 109, 604–622.
Joslin, Scott, Kenneth J. Singleton, and Haoxiang Zhu, 2011, A new perspective on Gaussian dynamic term structure models, *Review of Financial Studies* 24, 926–970.
Jouini, E., and C. Napp, 2007, Consensus consumer and intertemporal asset pricing with heterogeneous beliefs, *Review of Economic Studies* 74, 1149–1174.
Kozicki, Sharon, and Peter A. Tinsley, 2001, Shifting endpoints in the term structure of interest rates, *Journal of Monetary Economics* 47, 613–652.
Krishnamurthy, A., and A. Vissing-Jorgensen, 2011, The effects of quantitative easing on long-term interest rates, *Bookings Papers on Economic Activity* Fall 2011, 215–284.
Krishnamurthy, A., and A. Vissing-Jorgensen, 2012, The aggregate demand for treasury debt, *Journal of Political Economy* 120, 233–267.
Kucera, Adam, Evzen Kocenda, and Ales Marsal, 2019, Yield curve dynamics and fiscal policy, Working paper, National Bank of Slovakia.
Kurmann, Andre, and Christopher Ostrok, 2013, News shocks and the slope of the term structure of interest rates, *American Economic Review* 103, 2612–2632.
Lahiri, Kajal, and Xuguang Sheng, 2008, Evolution of forecast disagreement in a Bayesian learning model, *Journal of Econometrics* 144, 325–340.
Lamont, Owen, 2002, Macroeconomic forecasts and microeconomic forecasters, *Journal of Economic Behavior and Organization* 48, 265–280.
Laster, David, Paul Bennett, and In Sun Geoum, 1999, Rational bias in macroeconomic forecasts, *Quarterly Journal of Economics* 114, 293–318.
Le, A., K Singleton, and J. Dai, 2010, Discrete-time affine^q term structure models with generalized market prices of risk, *Review of Financial Studies* 23, 2184–2227.
Le, Anh, and Kenneth J. Singleton, 2012, A robust analysis of the risk structure of equilibrium term structures of bond yields, Working paper, University of North Carolina.
Leippold, Markus, and Felix Matthys, 2017, Government policy undercertainty and the yield curve, Working paper, University of Zurich.
Leippold, Markus, Fabio Trojani, and Paolo Vanini, 2008, Learning and asset prices under ambiguous information, *Review of Financial Studies* 21, 2565–2596.
Malkhozov, Aytek, Philippe Mueller, Andrea Vedolin, and Gyuri Venter, 2016, Mortgage risk and the yield curve, *Review of Financial Studies* 29, 1220–1253.
Van Nieuwerburgh, S., and L. Veldkamp, 2010, Learning asymmetries in real business cycles, *Journal of Monetary Economics* 53, 753–772.
Patton, Anderw J., and Allan Timmerman, 2010, Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion, *Journal of Monetary Economics* 57, 803–820.
Piazzesi, M., 2005, Bond yields and the federal reserve, *Journal of Political Economy* 113, 311–344.
Sargent, Thomas, and Noah Williams, 2005, Impacts of priors on convergence and escapes from Nash equilibrium, *Review of Economic Dynamics* 8, 360–391.
van Dijk, Dick, Siem J. Koopman, Michel van der Wel, and Jonathan H. Wright, 2014, Forecasting interest rates with shifting endpoints, *Journal of Applied Econometrics* 29, 693–712.
Vayanos, D., and J. Villa, 2009, A preferred habit model of the term structure of interest rates, Technical report, London School of Economics.
Wachter, J., 2006, A consumption-based model of the term structure of interest rates, *Journal of Financial Economics* 79, 365–399.
Wang, Chen, 2020, Under- and over-reaction in yield curve expectations, Technical report, Notre Dame University.
Xiong, Wei, and Hongjun Yan, 2009, Heterogeneous expectations and bond markets, *Review of Financial Studies* 23, 1433–1466.