High Frequency Evolution of Macro Expectation and Disagreement

Haotian Deng*
Shanghai University of Finance and Economics

Jiangyuan Li†
Shanghai University of Finance and Economics

Jinqiang Yang‡
Shanghai University of Finance and Economics

October 2024

*School of Finance, Shanghai University of Finance and Economics; Postal Address: Guoding Road 777, 200433, Shanghai, China.

†School of Finance, Shanghai University of Finance and Economics; Postal Address: Guoding Road 777, 200433, Shanghai, China; E-mail: lijiangyuan@mail.sufe.edu.cn.

‡School of Finance, Shanghai University of Finance and Economics; Postal Address: Guoding Road 777, 200433, Shanghai.

High Frequency Evolution of Macro Expectation and Disagreement

Abstract

This paper investigates the high-frequency dynamics of macroeconomic expectations and disagreement among professional forecasters. We propose a novel mixed-frequency estimation approach that integrates daily asset returns with quarterly expectation data from the Survey of Professional Forecasters. Our findings indicate that consensus forecasts are updated efficiently according to Bayes' rule, independent of prior forecasts. By employing "representative forecasters" as proxies for real-world agents, we derive a simple yet intuitive evolution equation for disagreement, revealing that changes in disagreement are primarily driven by different interpretations of new information. Furthermore, we reconstruct daily series of expectations and disagreement concerning macroeconomic growth, achieving impressive R^2 values of 93.3% and 84.5% against the true quarterly series.

JEL Classification: C53, D84, E17, E37

Keywords: Expectation Formation, Forecast Disagreement, Mixed-Frequency Method

1 Introduction

Expectation formation is fundamental to asset pricing, particularly in the context of disagreement among market participants. Traditional theories, such as the full information rational expectations (FIRE) framework, suggest that there should be no disagreement among agents (Muth, 1961; Lucas Jr, 1972; Sargent and Wallace, 1976). These models assume that all agents, learning from a homogeneous sequence of signals, should theoretically converge towards a uniform prior. Yet, empirical evidence shows persistent disagreement among agents (Jonung, 1981). Models of information rigidity offer compelling explanations for these observations. [1] Central to these models is the important role of information—or "news"—in shaping expectations and disagreement. However, the high-frequency nature of this news and the low-frequency survey data (typically collected quarterly or monthly) used in much of the empirical literature are misaligned, posing challenges for accurately capturing the dynamics of expectations and disagreement in real time.

Since the seminal work of Fama (1970), there has been extensive literature examining how economic agents adjust their expectations regarding future economic conditions in response to newly acquired information (Andreou, Ghysels, and Kourtellos, 2013; Coibion and Gorodnichenko, 2015; Coibion, Gorodnichenko, and Kumar, 2018; Hagenhoff and Lustenhouwer, 2023). This dynamic process of expectation adjustment occurs at a frequency that significantly exceeds that of conventional survey reports, which typically provide data only on a monthly or quarterly basis. For instance, while the Federal Reserve announces its interest rate decisions during its quarterly Federal Open Market Committee (FOMC) meetings, market participants do not passively await these announcements; instead, they continuously revise their interest rate expectations based on ongoing economic and financial data (Bernanke and Kuttner, 2005).

Thus, adopting a solely low-frequency perspective may present several critical limitations. Firstly, it may neglect the immediate effects of real-time information and therefore distort the true evolution of expectations. Researches by Nakamura and Steinsson (2018) and Chaudhry and Oh (2020) on the Fed Information Effect illustrate this point: studies utilizing monthly surveys, such as the Blue Chip survey for growth expectations, may inadequately capture the influence of monetary policy shocks due to the inherent low frequency of the data. Secondly, the low-frequency perspective may underestimate the presence of stickiness among professional forecasters. While previous researches by Sims (2003),

[1] For instance, the sticky information model posits that agents update their beliefs only intermittently in response to new data with a certain probability, while the noisy information model assumes that agents perceive signals with inherent inaccuracies (Mankiw and Reis, 2002; Sims, 2003).

Mankiw, Reis, and Wolfers (2003), and Lahiri and Sheng (2008) suggest that professional forecasters do not experience delays in gathering relevant information compared to others, such conclusions may be limited by their reliance on low-frequency survey data. In a quarterly framework, agents have sufficient time to collect information. This may mask underlying daily-frequency stickiness in their expectations. Despite these insights, there remains a notable absence of methodologies that simultaneously integrate high-frequency information with low-frequency expectation data, a gap that this research aims to address.

Our paper makes the following three contributions to the literature. First, we develop a novel framework to rigorously investigate the high-frequency dynamics of expectation. Utilizing this framework, we demonstrate that consensus forecasts are updated efficiently according to Bayes' rule at daily frequencies, such that $\mathbb{F}_{t+1}[x_{t+1}] = \mathbb{F}_t[x_{t+1}] + \boldsymbol{\beta}'\mathbf{r}_{t+1}$, where $\mathbf{r}$ represents news or asset returns. This finding indicates that forecast revisions are independent of prior forecasts and are solely responsive to new information. Second, by employing the concept of "representative forecasters", our analysis extends to understanding the high frequency evolution of disagreement in forecasts. We derive a similar pattern with consensus, showing that $\sigma\left(\mathbb{F}_{i,t+1}[x_{t+1}]\right) = \sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right) + (\Delta\boldsymbol{\beta})'\mathbf{r}_{t+1}$, where σ represents the standard deviation of agents' forecasts. This demonstrates that changes in disagreement are also independent to prior periods' disagreement and are primarily driven by different interpretations of new information among agents. Third, leveraging this new mix-frequency approach, we reconstruct the unobserved daily series of expectations and disagreement regarding macroeconomic growth that span the interval between two quarterly survey releases.

Our first contribution is a novel mixed-frequency estimation framework that enables the simultaneous analysis of high-frequency news and low-frequency expectations. This approach is grounded in a classical Bayesian framework and allows us to derive a simple yet intuitive evolution equation for consensus forecasts, given by $\mathbb{F}_{t+1}[x_{t+1}] = \alpha\mathbb{F}_t[x_{t+1}] + \boldsymbol{\beta}'\mathbf{r}_{t+1}$. This model delineates the sources of agents' forecasts into three fundamental components: (i) the initial prior beliefs of the agents, (ii) the weights assigned to these priors, and (iii) the interpretations of new public information. We then reformulate the evolution equation in a recursive form to establish the relationship between expectations data from two consecutive releases. Finally, we employ a random search method to estimate the high-order coefficients within the evolution equation.

It is quite common for time series observations to be sampled at different frequencies (Ghysels, 2016). In the literature addressing mixed-frequency issues, two primary approaches are commonly used: (i) a naive approach that aggregates high-frequency data to the lowest frequency (Deng, Wang, and Zhou, 2024), and (ii) statistical models that accommodate mixed-frequency data simultaneously. [2] However, simply aggregating high-frequency data may lead to the loss of valuable information, while those statistical mixed-frequency methods often lack clear economic interpretation and may be complex to implement. The method we propose effectively overcomes these challenges, ensuring that the estimation process retains high-frequency information while being easy to implement and aligned with economic intuition.

Specifically, we utilize daily-frequency asset returns as proxies for new information, which allows us to construct a high-frequency evolution equation for analysis. We select asset returns due to their public availability, frequent updates, and the valuable information they provide to agents for revising their expectations (Banerjee, 2011; Andreou, Ghysels, and Kourtellos, 2013; Chaudhry and Oh, 2020; Kerssenfischer and Schmeling, 2024). Additionally, we provide empirical evidence regarding the relationship between asset returns and forecast revisions. Our findings show that specific bivariate pairs of assets can explain a sizeable amount of variance in forecast revisions, with an R^2 of 25.8%.

Building upon our mixed-frequency method, we derive the coefficients within the evolution equation. The coefficient α, which corresponds to the influence of prior forecasts, equals one across all window periods and asset pairs. This finding indicates that consensus are updated efficiently according to Bayes' rule, exhibiting neither stickiness nor overreaction to new information. Given that the forecast data we utilize is sourced from the Survey of Professional Forecasters (SPF), these findings align with the observations made by Sims (2003) and Lahiri and Sheng (2008), which suggest that the sticky information model may not adequately characterize the consensus forecasts of professionals. However, it is essential to note that this efficiency in updates does not apply uniformly at the individual level. Since we find considerable heterogeneity among agents in their forecasting behaviour, as some appear to exhibit stickiness to previous forecasts while others tend to overreact to new information.

Our second contribution is to demonstrate that the evolution of forecast disagreement follows the equation $\sigma\left(\mathbb{F}_{i,t+1}\left[x_{t+1}\right]\right) = \sigma\left(\mathbb{F}_{i,t}\left[x_{t+1}\right]\right) + (\Delta\boldsymbol{\beta})'\mathbf{r}_{t+1}$, which shows a pattern similar to that of consensus forecasts. We begin with an explanation of the intuitive basis for cross-sectional variance of forecasts. The variance is formulated to capture the discrepancies between individual agents' forecasts and the consensus forecast. By integrating the evolution equations of both individual and consensus forecasts, we can derive a complex evolution equation for cross-sectional variance in forecasts. However, directly estimating this evolution

[2] See, for example, Ghysels, Sinko, and Valkanov (2007); Ghysels and Wright (2009); Andreou, Ghysels, and Kourtellos (2010); Ghysels (2016); Pettenuzzo, Timmermann, and Valkanov (2016); Chaudhry and Oh (2020).

equation for disagreement proves infeasible due to parameter constraints. [3] If we relax these constraints and apply the previous mixed-frequency method, the estimated parameters may be biased and fail to represent the true variance.

To address this issue, we utilize two "representative forecasters" as proxies for real-world agents: one making optimistic predictions (above the consensus) and another making pessimistic predictions (below the consensus). Both representative forecasters are positioned symmetrically around the mean expectation, with their deviations from this mean representing the cross-sectional variance of real agents. This setup allows us to apply the previous mixed-frequency method to separately estimate the evolution equations for both representative forecasters. Then, we can directly calculate their variances as an estimation of the cross-sectional variance of the real agents.

Additionally, using "representative forecasts" as proxies enables us to demonstrate that disagreement follows an evolution equation similar to that of the cross-sectional mean, i.e. $\sigma\left(\mathbb{F}_{i,t+1}[x_{t+1}]\right) = \alpha\sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right) + (\Delta\boldsymbol{\beta})'\mathbf{r}_{t+1}$, where the coefficient α also equals one. This result aligns with our earlier finding regarding the evolution of consensus forecasts, where a random walk outperforms other parameter values in capturing the dynamic of expectations. Despite a formal difference, both our evolution equation and the model proposed by Lahiri and Sheng (2008), imply the same three components of disagreement: (i) prior-mean heterogeneity, $\sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right)$, (ii) the weights attached to these priors, α, and (iii) diverse interpretations of new information, $\Delta\boldsymbol{\beta}$. [4]

Our third contribution is the construction of a daily time series of forecasters' mean expectations and their disagreement regarding macroeconomic growth. We employ our mixed-frequency method within a recursive out-of-sample estimation framework, utilizing data from SPF. While existing literature has recovered daily series of expectations using quarterly-frequency panel of forecasts, we are the first to extract daily-frequency series of disagreement (Andreou, Ghysels, and Kourtellos, 2013; Chaudhry and Oh, 2020). Our method generates daily estimates of both the cross-sectional mean and standard deviation that closely approximate actual quarterly values, achieving high R-squared values against the true quarterly series—specifically, an R^2 of 93.3% for the mean and 84.5% for the standard deviation.

[3] The parameter constraints will be discussed in detail in Section 4.

[4] We broadly interpret "diverse interpretations of new information" to include scenarios where agents utilize different models to interpret new public information, possess private information, exhibit rational inattention or asymmetric attention, or rely on past experiences (Patton and Timmermann, 2010; Andrade and Le Bihan, 2013; Giacomini, Skreta, and Turen, 2020; Kohlhas and Walther, 2021; Bordalo, Conlon, Gennaioli, Kwon, and Shleifer, 2023). We consider all these factors as contributing to distinct coefficient $\boldsymbol{\beta}$.

We then compare our method with other mixed-frequency approaches, such as MIDAS, the Kalman filter, and reinforcement learning (Ghysels and Wright, 2009; Chaudhry and Oh, 2020). Our method has been shown to have significant advantages in terms of interpretability, efficiency, and empirical performance. Additionally, we offer an economic rationale for why specific setups—such as fixing the coefficient α at one and jointly estimating the cross-sectional mean and variance—can enhance performance in reinforcement learning algorithms. This helps us better understand and interpret the predictions generated by these black-box models.

This paper connects two strands of literature: the evolution of disagreement and mixed-frequency models. Numerous models discuss how agents' disagreement evolve at low frequencies (Lahiri and Sheng, 2008; Patton and Timmermann, 2010; Giacomini, Skreta, and Turen, 2020). However, there is still a lack of a systematic method to capture the high-frequency dynamics of this belief dispersion. Technically speaking, there also exist various mixed-frequency methods that analyze the high-frequency dynamics of financial and economic variables. [5] Our work provides a novel framework that rigorously investigates the high-frequency evolution of expectation and disagreement among agents, thereby bridging the gap between these two strands of literature.

The remainder of this paper is organized as follows: Section 2 summarizes the data and variables; Section 3 develops the framework for analyzing the high-frequency dynamics of expectations; Section 4 derives and estimates the evolution equation of forecasters' disagreement; Section 5 presents the results of the empirical estimation of daily expectations and disagreement series; and finally, Section 6 concludes the study.

2 Data

As with some related previous work, we focus on the expectations of professional forecasters. Compared to other agents, professional forecasters obviously invest more resources in predicting future macroeconomic variables and, as a result, cannot fully represent the perspectives of all agents (Carroll, 2003). Nevertheless, their views are disseminated and significantly influence the expectations and decisions of other agents. Therefore, the level of attention that professional forecasters pay to new information can be viewed as an upper limit on how much other agents focus on overall economic conditions (Andrade and Le Bihan, 2013). We utilize data from the Survey of Professional Forecasters

[5] See, for example, Ghysels, Sinko, and Valkanov (2007); Andreou, Ghysels, and Kourtellos (2010, 2013); Ghysels (2016); Pettenuzzo, Timmermann, and Valkanov (2016); Mariano and Ozmucur (2020).

(SPF), a quarterly survey conducted by the Federal Reserve Bank of Philadelphia since 1990. The SPF gathers forecasts on a range of macroeconomic variables from approximately 40 anonymous professional forecasters, providing data at both the consensus and individual levels. Our analysis centers on the one-quarter-ahead forecasts for real GDP growth and covers the period from the first quarter of 1991 to the fourth quarter of 2023. [6] For instance, a forecaster's expectation of real GDP growth in 2023Q3 (t) is from survey conducted in mid 2023Q2($t-1$). [7] Additionally, we employ the specific survey release dates to align with daily-frequency asset price data. Overall, the SPF provides a detailed perspective on macroeconomic forecasts over an extended period, which allows us to analyze the evolution of expectations and disagreement among professionals.

We adopt a broad interpretation to asset prices, encompassing rate changes, spreads, returns, and other value-related metrics of financial assets. To enable the construction of a daily time series analysis, we restrict our focus to assets with available liquid daily returns. For equities, we consider returns on the market index and Fama-French factors—namely market, size, and value. For fixed income, we consider returns on treasuries and changes in credit spreads. For exchange rates, we consider changes in the weighted average of the U.S. dollar's foreign exchange. For derivatives, we consider changes in the VIX index. Table 1 provides a comprehensive summary of the data sources used in our analysis.

Table 2 presents the summary statistics for the SPF forecasts. On average, approximately 39 participants engaged in the survey each quarter during the sample period. Interestingly, we observe that the degree of divergence in nowcasts is, on average, greater than that in

[6] Although the SPF survey was established in the fourth quarter of 1968, our sample begins in the first quarter of 1991 for several reasons. Originally managed by the American Statistical Association and the National Bureau of Economic Research, the Federal Reserve Bank of Philadelphia took over administration of the survey in 1990. Therefore, the precise deadline and release dates for surveys conducted before the second quarter of 1990 are not available. Additionally, the sample sizes in 1990 were small and varied considerably (with 14 respondents in Q1, 9 in Q2, 13 in Q3, and 30 in Q4). To ensure data consistency and reliability, we select the first quarter of 1991 as the starting point for our analysis, as it marks a period of stabilization under the management of the Federal Reserve Bank of Philadelphia.

[7] We define the mean growth expectations by first computing the mean forecast for the level of real GDP (X_t) and then computing the rate of growth (x_t), using the formulas:

$$\mathbb{F}_t\left[x_{t+1}\right] = 100 \times \left[\left(\frac{\mathbb{F}_t\left[X_{t+1}\right]}{\mathbb{F}_t\left[X_t\right]}\right)^4 - 1 \right]$$

$$\mathbb{F}_{t+1}\left[x_{t+1}\right] = 100 \times \left[\left(\frac{\mathbb{F}_{t+1}\left[X_{t+1}\right]}{X_t}\right)^4 - 1 \right]$$

where $\mathbb{F}_t\left[x_{t+1}\right]$ represents the forecast for quarter-over-quarter growth in period $t+1$ made on the basis of observations known through period $t-1$. It is worth noting that when calculating the nowcast $\mathbb{F}_{t+1}\left[x_{t+1}\right]$, we use X_t instead of $\mathbb{F}_{t+1}\left[X_t\right]$, since X_t reflects the real-time quarterly historical value for the previous quarter, i.e., the quarter before the quarter when the SPF is conducted.

forecasts, which may stem from heterogeneous interpretations of new information among agents.

3 High-Frequency Evolution of Growth Expectations

3.1 Relation between News and Growth Expectations

To establish a foundation for our empirical analysis, we consider a straightforward framework for forecast revisions. We begin with a Bayesian agent who updates expectations according to the classical Bayes' rule. Assume that the data-generating process is defined by $x_t = \rho x_{t-1} + u_t$, where $u_t \sim \mathcal{N}\left(0, \sigma_u^2\right)$ is i.i.d. over time and $\rho > 0$. Agent i observes the noise signal $s_t^i = x_t + \epsilon_t^i$, where $\epsilon_t^i \sim \mathcal{N}\left(0, \sigma_\epsilon^2\right)$ represents forecaster-specific i.i.d. noise. According to the Kalman filter, beliefs should be updated as follows

$$\underbrace{\mathbb{F}_{i,t+1}\left[x_{t+1}\right]}_{\text{nowcast}} = \underbrace{\mathbb{F}_{i,t}\left[x_{t+1}\right]}_{\text{forecast}} + \frac{\Sigma}{\Sigma + \sigma_\epsilon^2} \underbrace{\left(s_t^i - \mathbb{F}_{i,t}\left[x_{t+1}\right]\right)}_{\text{new information}} \tag{1}$$

where $\mathbb{F}_{i,t}\left[x_{t+1}\right]$ represents the forecast by agent i for the target variable x_{t+1} at time t. Σ is the steady state variance of the prior $f\left(x_{t+1} \mid s_t^i, s_{t-1}^i, \cdots\right)$. To help with interpretation of this equation, we transform it into a simpler form:

$$\mathbb{F}_{i,t+1}\left[x_{t+1}\right] = \mathbb{F}_{i,t}\left[x_{t+1}\right] + News_{t+1} \tag{2}$$

The $News_{t+1}$ term subsumes the coefficients on the variables that represent new information, which would reflect the information content of those variables for predicting x_{t+1}. According to Nordhaus (1987) and Fuhrer (2018), we interpret equation (2) as representing an efficient Bayesian forecast, since the forecast revision $FR_{i,t+1}$ from period $t-1$ to period t is influenced exclusively by new information:

$$FR_{i,t+1} \equiv \mathbb{F}_{i,t+1}\left[x_{t+1}\right] - \mathbb{F}_{i,t}\left[x_{t+1}\right] = News_{t+1} \tag{3}$$

However, if the coefficient α_i on $\mathbb{F}_{i,t}\left[x_{t+1}\right]$ differs significantly from one (say $\alpha_i < 1$), it suggests that the revision responds inefficiently to news. We assume that the time-invariant constant α_i and the time-varying function $News_t$ are uniform across all agents. [8] Next, we modify the $News_{t+1}$ term to represent observable signals $\mathbf{r}_{t+1}$ and reintroduce the coefficient vector $\boldsymbol{\beta}_i$:

$$\begin{aligned} \mathbb{F}_{i,t+1}[x_{t+1}] &= \alpha_i \mathbb{F}_{i,t}[x_{t+1}] + \boldsymbol{\beta}_i' \mathbf{r}_{t+1} \\ FR_{i,t+1} \equiv \mathbb{F}_{i,t+1}[x_{t+1}] - \mathbb{F}_{i,t}[x_{t+1}] &= (\alpha_i - 1)\mathbb{F}_{i,t}[x_{t+1}] + \boldsymbol{\beta}_i' \mathbf{r}_{t+1} \end{aligned} \tag{4}$$

where α_i is a scalar and $\boldsymbol{\beta}_i$ is a vector of scalars. Currently, our model does not impose any specific assumptions regarding the formation process of agents' beliefs. Agents could be Bayesians, efficiently updating their beliefs in response to new information, or they might hold degenerate priors, ignoring any new information that becomes available (Anderson, Ghysels, and Juergens, 2005). If $\alpha_i = 1$, then $\alpha_i - 1 = 0$. According to equation (4), revisions $FR_{i,t+1}$ are independent to the prior forecasts $\mathbb{F}_{i,t}[x_{t+1}]$, and agents update their beliefs following Bayes' rule. If $\alpha_i < 1$, then $\alpha_i - 1 < 0$, indicating an inefficiency that leads to a muted or smoothed response to news and agents are sticky to their prior beliefs. If $\alpha_i > 1$, then $\alpha_i - 1 > 0$, indicating that agents overreact to news. Taking the average across all agents, we derive the following expression for the cross-sectional mean of growth expectations at period t:

$$\mathbb{F}_{t+1}[x_{t+1}] = \frac{1}{N}\sum_{i=1}^{N}\mathbb{F}_{i,t+1}[x_{t+1}] = \frac{1}{N}\sum_{i=1}^{N}\alpha_i \mathbb{F}_{i,t}[x_{t+1}] + (\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\beta}_i)'\mathbf{r}_{t+1} \tag{5}$$

Motivated by this expression, we employ the following approximating moment:

$$\mathbb{F}_{t+1}[x_{t+1}] = \alpha \mathbb{F}_t[x_{t+1}] + \boldsymbol{\beta}'\mathbf{r}_{t+1} \tag{6}$$

Equation (6) implies the same relationship established by Chaudhry and Oh (2020). In their setup, $\mathbf{r}_t$ represents a vector of asset returns in period t, since they contain informative content for forecast revisions. We adopt this setting by conducting several time series regressions. As described earlier, we define the forecast revision for period t as the difference between the mean forecast of period t-growth reported in period t (nowcast) and the mean forecast of period t-growth reported in period $t-1$ (past forecast), i.e. $FR_t = \mathbb{F}_t[x_{t+1}] - \mathbb{F}_{t-1}[x_{t+1}]$. Two immediate issues arise: First, there is a frequency mismatch.

[8] This is similar to the setup of agents' heterogeneous beliefs about the conditional mean of aggregate consumption growth $\log \mu_{g,i,t+1|t}$ proposed by Anderson, Ghysels, and Juergens (2005):

$$\log \mu_{g,i,t+1|t} = \alpha \log \mu_{m,i,t+1|t} + q_t$$

where α is a constant, q_t is an unknown function of current information, and $\mu_{m,i,t+1|t}$ is agent i's forecast of the market return.

Forecast revisions are at a quarterly frequency, whereas asset returns are daily. Commonly, the average of daily asset returns between two survey release dates is calculated to align with the quarterly revisions. However, we choose to compute the sum rather than the average for two reasons: (i) if the number of asset return observations remains consistent between each pair of release dates, then calculating the mean and sum will yield equivalent results in terms of coefficients, differing only in magnitude; (ii) this can be unified with the method we introduce later. The second issue to address is the selection of assets. Consistent with Chaudhry and Oh (2020), we restrict our analysis to bivariate pairs of assets, since we assume that a linear combination of two asset returns may approximate others.

Table 3 reports the results of these regressions. In alignment with Chaudhry and Oh (2020), we observe that certain pairs of assets yield sizable R^2. For instance, returns on 5-year treasury constant maturity and change in BAA corporate bond yield relative to yield on 10-year treasury constant maturity spread explain the greatest amount of variance in forecast revision, with an R^2 of 25.8%. This empirical evidence reinforces our use of asset returns as proxies for news, as they indeed contain valuable information about forecast revisions.

3.2 Estimation Method

In this subsection, we focus on analyzing the evolution of mean growth expectations. Specifically, our objective is to estimate the coefficients in equation (6), with a particular emphasis on the coefficient α associated with the past forecast $\mathbb{F}_t[x_{t+1}]$. We still encounter the issue with mixed frequencies: forecasts are made quarterly, while the asset returns that represent news are recorded daily. We refrain from simply aggregating daily asset returns into quarterly periods, as shifts in expectations seem to occur more frequently than what the quarterly SPF data releases would imply. To effectively address this issue, we propose the following method.

The intuition behind our method is to find the relationship between two quarterly-released forecasts. We start by revisiting equation (6) and reformulating it in a recursive form:

$$
\begin{aligned}
\mathbb{F}_1[x] &= \alpha\mathbb{F}_0[x] + \boldsymbol{\beta}'\mathbf{r}_1 \\
\mathbb{F}_2[x] &= \alpha\mathbb{F}_1[x] + \boldsymbol{\beta}'\mathbf{r}_2 \\
&\cdots \\
\mathbb{F}_T[x] &= \alpha\mathbb{F}_{T-1}[x] + \boldsymbol{\beta}'\mathbf{r}_T
\end{aligned}
\tag{7}
$$

where $\mathbb{F}_0[x], \cdots, \mathbb{F}_T[x]$ denote daily-frequency forecasts between two release dates. [9] We omit the subscript t from the target variable x, as the predicted target remains unchanged for the quarter, irrespective of whether it is day 0 or day T. This recursive approach allows us to clearly establish the relationship between $\mathbb{F}_0[x]$ and $\mathbb{F}_T[x]$ for two consecutive release dates:

$$\mathbb{F}_T[x^p] = \alpha^T \mathbb{F}_0[x^p] + \sum_{k=0}^{T-1} \alpha^k \boldsymbol{\beta}' \mathbf{r}_{T-k}^p \tag{8}$$

Equation (8) has the same form as the model given by Barberis (2018) for return extrapolation. [10] It is worth noting that we assign a superscript p to the target variable x and the asset return $\mathbf{r}$ to signify the quarter. This can help us distinguish it from the subscript t that denotes daily-frequency forecasts. However, estimating the coefficients α and $\boldsymbol{\beta}$ remains challenging due to the presence of higher-order terms in the equation. By setting the coefficient on the one-day lag cross-sectional mean α to a fixed value α_0, we simplify the equation and therefore allow it to estimate using the ordinary least squares (OLS) method:

$$\mathbb{F}_T[x^p] - \alpha_0^T \mathbb{F}_0[x^p] = \boldsymbol{\beta}' (\sum_{k=0}^{T-1} \alpha_0^k \mathbf{r}_{T-k}^p) \tag{9}$$

Let $y = \mathbb{F}_T[x^p] - \alpha_0^T \mathbb{F}_0[x^p]$ and $\mathbf{X} = \sum_{k=0}^{T-1} \alpha_0^k \mathbf{r}_{T-k}^p$. Then, according to OLS, the estimator $\hat{\boldsymbol{\beta}}$ is given by $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y$. Additionally, we can calculate the sum of squared residuals $\mathrm{SSR} = (y - \mathbf{X}\boldsymbol{\beta})'(y - \mathbf{X}\boldsymbol{\beta})$ and the R-squared value for a given α_0. Given that our aim also

[9]For instance, if the target variable x is the real GDP growth for the second quarter of 2024, then $\mathbb{F}_0[x]$ represents the mean growth expectation (past forecast) on February 9, 2024 (the release date of SPF in 2024Q1), $\mathbb{F}_T[x]$ represents the mean growth expectation (nowcast) on May 10, 2024 (the release date of SPF in 2024Q2), and $\mathbb{F}_1[x], \cdots, \mathbb{F}_{T-1}[x]$ represent the mean growth expectations between these two dates.

[10]The setup of return extrapolators in Barberis (2018) follows

$$X_t = \theta X_{t-1} + (1-\theta)(P_{t-1} - P_{t-2})$$

where X_t denotes the extrapolators' belief at time t, and $(P_{t-1} - P_{t-2})$ represents recent price change. Thus, their belief about the future price change is

$$X_T = \theta^{T-1} X_1 + (1-\theta) \sum_{k=1}^{T-1} \theta^{k-1} (P_{t-k} - P_{t-k-1})$$

Despite the subtle differences, it has a recursive form similar to Equation (8), reflecting both the decay of initial belief over time and the accumulation of new information.

includes estimating the coefficient α, the current task is to identify the value of α_0 that minimizes the SSR: [11]

$$
\begin{aligned}
\frac{\partial \mathrm{SSR}}{\partial \alpha_0} = {} & 2T\alpha_0^{T-1}\mathbb{F}_0[x^p]'\mathbf{X}\hat{\boldsymbol{\beta}} - 2y'\left(\sum_{k=0}^{T-1} k\alpha_0^{k-1}\mathbf{r}_{T-k}^p\right)\hat{\boldsymbol{\beta}} \\
& + \hat{\boldsymbol{\beta}}'\left[\left(\sum_{k=0}^{T-1} k\alpha_0^{k-1}\mathbf{r}_{T-k}^p\right)'\mathbf{X} + \mathbf{X}'\left(\sum_{k=0}^{T-1} k\alpha_0^{k-1}\mathbf{r}_{T-k}^p\right)\right]\hat{\boldsymbol{\beta}}
\end{aligned}
\tag{10}
$$

Equation (10) is nonlinear with respect to α_0 and may resist straightforward analytical solution. [12] Motivated by the grid search technique commonly used for hyper-parameter optimization in machine learning, we develop a random search method to identify the optimal coefficient α on past forecast $\mathbb{F}_t[x_{t+1}]$. The intuition is to search the optimal α that corresponds to the minimum SSR within a specified range.

The initial step in our method involves establishing the range and granularity for the parameter α. We systematically construct a grid over the feasible domain of α and divide it into discrete intervals. Each grid point corresponds to a potential value of α for testing. The domain for α is defined based on theoretical considerations. According to the definition of efficient forecasts by Fuhrer (2018), an α less than 1 suggests a muted response to news, while an α greater than 1 indicates an overreaction. Consequently, it is logical to hypothesize that the optimal α should be around 1. In the formal estimation process, we begin by setting a broad sampling range for α and gradually narrowing the interval to enhance accuracy. This iterative refinement approach helps avoid missing the optimal solution while reducing computational costs. The next step is to determine the appropriate sample period for parameter estimation. We employ a rolling window of 40 quarters for parameter estimation rather than using the full sample, as it is obviously that the parameters cannot remain constant over the entire sample period. For example, with our sample beginning in 1991Q1, the first rolling window spans from 1991Q1 to 2000Q4. The length of the rolling window does not necessarily have to be fixed at 40 quarters. However, it should be carefully chosen to balance statistical power and parameter stability. A window that is too short may yield insufficient samples for robust OLS estimation while one that is too long may violate the assumption that the parameters do not remain constant over a long period. To reduce estimation variance, we can train six models using overlapping lookback windows of T = 40, 44, 48, 52, 56, and 60 quarters, and then average their outputs. These results are proved to be robust, but for simplicity, our benchmark estimation

[11] Detailed derivation can be found in Appendix A.

[12] We can indeed use some nonlinear optimization algorithms to solve this.

process employ a 40-quarter rolling window. The third step, as previously mentioned, is to select asset returns that represent news. We continue to consider bivariate asset pairs, specifically choosing those that yield the highest R^2.

3.3 Estimation Results

We apply our random search method to estimate the optimal coefficients. According to our predefined model, we calculate the corresponding $\boldsymbol{\beta}$ and SSR for each α in the specified grid. We then identify the optimal α that minimizes SSR within the grid. With this estimation framework established, we now proceed to formally estimate the coefficients.

Figure 1 illustrates the relationship between SSR and the coefficient α over the first rolling window period (1991Q1 to 2000Q4). Unreported results show that this relationship remains consistent across other window periods and is robust to different asset returns. Near the optimal coefficient, SSR exhibits a monotonic decrease followed by an increase as α rises. This suggests an optimal α value of 1. Figure 2 presents the estimated optimal α coefficients and their standard deviations across different estimation windows. As evident from the first horizontal error bars, the mean values, represented by dot, fall within 0.002 of 1, and all estimated coefficients cluster tightly around 1, with minimal deviations. This indicates a consistency that α equals one across all rolling windows. Figure 3 compares SSR across window periods for various α. We find that SSR is minimized when $\alpha = 1$ across the entire sample periods, which further supports our results.

To be rigorous, we also examine the case that includes an intercept term in Equation (6), i.e. $\mathbb{F}_{t+1}[x_{t+1}] = \alpha\mathbb{F}_t[x_{t+1}] + \boldsymbol{\beta}'\mathbf{r}_{t+1} + c$. Through iteration, we can derive the relationship between forecast revision, asset returns, and the intercept term as follows:

$$\mathbb{F}_T[x^p] - \alpha_0^T\mathbb{F}_0[x^p] = \boldsymbol{\beta}'(\sum_{k=0}^{T-1}\alpha_0^k\mathbf{r}_{T-k}^p) + (\sum_{k=0}^{T-1}\alpha_0^k)c \tag{11}$$

By substituting $\alpha = 1$ into this equation and adding the term of prior belief, we have:

$$FR^p \equiv \mathbb{F}_T[x^p] - \mathbb{F}_0[x^p] = \boldsymbol{\beta}'(\sum_{k=0}^{T-1}\mathbf{r}_{T-k}^p) + (k\mathbb{F}_0[x^p] + b) \tag{12}$$

Assuming $\alpha = 1$ and $c = 0$, the coefficients k and $b = Tc$ in Equation (12) are expected to be insignificantly different from zero. [13] Table 4 reports the estimated results from the first window period. As expected, the coefficients $\boldsymbol{\beta}$ corresponding to asset returns are significantly different from zero since they contain useful information in explaining forecast revisions. Conversely, the intercept term b and the coefficients k, corresponding to prior belief $\mathbb{F}_0[x^p]$, do not show a significant difference from zero. This further supports our findings that $\alpha = 1$.

Numerous evidences suggest that the coefficient α, corresponding to past forecasts, equals one. This raises two immediate questions: What are the estimated results for individuals, and what distinguishes individual from the consensus? To address these questions, we conduct less rigorous tests at the individual level. [14] Figure 4 displays the frequency distribution of the optimal coefficients at the individual level. While these coefficients predominantly cluster around 1.0, there is no evidence indicating that conclusions drawn at the individual level align with those at the consensus level. Using the same survey data but a different modeling strategy, we confirm the conclusion of Manzan (2021), where the author argues that non-Bayesian behavior is common among most professional forecasters in the sample.

In this section, we demonstrate that the evolution of mean growth expectations is efficient; that is, forecast revisions only reflect new information and are uncorrelated with prior beliefs. This dynamic process can be expressed by a simple yet intuitive equation, i.e., $\mathbb{F}_{t+1}[x_{t+1}] = \mathbb{F}_t[x_{t+1}] + \boldsymbol{\beta}'\mathbf{r}_{t+1}$. The fact that the coefficient α on past forecasts equals one is crucial for our subsequent analysis of the evolution of disagreement. Therefore, we adopt this as a foundation in the following section. We now turn to exploring the intuition behind the evolution of the cross-sectional variance in growth expectations.

[13] Chaudhry and Oh (2020) also provide evidence for an intercept term of 0 through parameter calibration, i.e. $\mathbb{F}_{t+1}[x_{t+1}] = \alpha\mathbb{F}_t[x_{t+1}] + \boldsymbol{\beta}'\mathbf{r}_{t+1} + c$.

[14] The term “less rigorous” is applied to these tests due to limitations in the individual-level data. First, the data comprises an unbalanced panel, with forecasters participating incompletely throughout the survey period. Second, some forecasts are not continuously present; for instance, the forecast with ID 446 entered the survey in the second quarter of 1993 and reappeared in the fourth quarter of the same year. Third, unlike our approach for estimating mean growth expectations, it is impractical to conduct rolling window estimations or select optimal asset pairs for each individual’s data.

4 High-Frequency Evolution of Forecast Disagreement

4.1 Intuition

In this subsection, we begin by discussing the intuition behind cross-sectional variance. The variance is defined as $\mathrm{Var}(\mathbb{F}_{i,t}[x_{t+1}]) = \frac{1}{N}\sum_{i=1}^{N}(\mathbb{F}_{i,t}[x_{t+1}] - \mathbb{F}_t[x_{t+1}])^2$. By incorporating the approximations from equation (4) and equation (6), we prove in Appendix B that:

$$\mathrm{Var}(\mathbb{F}_{i,t+1}[x_{t+1}]) = \gamma \mathrm{Var}(\mathbb{F}_{i,t}[x_{t+1}]) + \boldsymbol{\zeta}'\mathbf{r}_{t+1}\mathbf{r}'_{t+1}\boldsymbol{\zeta} + \boldsymbol{\delta}'\mathbf{r}_{t+1}\mathbb{F}_t[x_{t+1}] \tag{13}$$

where γ is a scalar related to α and α_i, $\boldsymbol{\zeta}$ is a vector of scalars related to $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_i$, and $\boldsymbol{\delta}$ is a vector of scalars related to α, α_i, $\boldsymbol{\beta}$, and $\boldsymbol{\beta}_i$.

Equation (13) suggests a pattern similar to the evolution of mean growth expectations. However, direct estimation of these parameters using previous method is infeasible due to inherent constraints, as they are all tied to the coefficients in equation (4) and (6). While we have already estimated the coefficients α and $\boldsymbol{\beta}$ for consensus forecasts in Section 3, we cannot similarly estimate the coefficients α_i and $\boldsymbol{\beta}_i$ for each individual in equation (4) due to the limitation of individual-level survey data. Consequently, we are unable to derive the explicit constraint conditions required for parameter estimation in equation (13). If we relax these constraints and apply the previous mixed-frequency method, the results, although computable, do not reflect the true "variance" parameters as needed. In an unreported set of results, we try to freely estimate coefficients in equation (13). However, the predicted variance occasionally becomes negative, indicating that what is being estimated is not the true cross-sectional variance. To address this issue, a more direct estimation method is required.

4.2 Representative Forecasters

The intuition of our method is to approximate the entire distribution by linking higher-order moments of agents' beliefs to the true distribution. To achieve this, we utilize "representative forecasters" as proxies for real-world agents that are impractical to directly estimate. We posit the existence of two such forecasters: one, denoted as $\mathbb{F}_t^H[x_{t+1}]$, whose predictions exceed the average growth expectations $\mathbb{F}_t[x_{t+1}]$, and another, denoted as

$\mathbb{F}_t^L[x_{t+1}]$, whose predictions fall below these expectations. [15] These forecasters are termed "representative" because they capture the mean and variance of individual forecasters:

$$\begin{aligned}
\mathbb{F}_t[x_{t+1}] &= \frac{1}{N}\sum_{i=1}^{N}\mathbb{F}_{i,t}[x_{t+1}] = \frac{1}{2}\left(\mathbb{F}_t^H[x_{t+1}]+\mathbb{F}_t^L[x_{t+1}]\right)\\
\mathrm{Var}(\mathbb{F}_{i,t}[x_{t+1}]) &= \frac{1}{N}\sum_{i=1}^{N}(\mathbb{F}_{i,t}[x_{t+1}]-\mathbb{F}_t[x_{t+1}])^2\\
&= \frac{1}{2}\left[(\mathbb{F}_t^H[x_{t+1}]-\mathbb{F}_t[x_{t+1}])^2+(\mathbb{F}_t[x_{t+1}]-\mathbb{F}_t^L[x_{t+1}])^2\right]
\end{aligned} \tag{14}$$

The advantage of this setup is that $\mathbb{F}_t^H[x_{t+1}]$ and $\mathbb{F}_t^L[x_{t+1}]$ are symmetrically positioned around the mean growth expectations, with their deviations from the mean constituting one standard deviation:

$$\begin{aligned}
\mathbb{F}_t^H[x_{t+1}] &= \mathbb{F}_t[x_{t+1}]+\sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right)\\
\mathbb{F}_t^L[x_{t+1}] &= \mathbb{F}_t[x_{t+1}]-\sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right)
\end{aligned} \tag{15}$$

where $\sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right) = \sqrt{\mathrm{Var}(\mathbb{F}_{i,t}[x_{t+1}])}$ denotes the cross-sectional standard deviation at time t for the forecast period $t+1$. A relevant question arises regarding the stability of the "representative forecaster": if agents are considered representative forecasters at time t, do they maintain this status after updating their beliefs at time $t+1$? Specifically, if equation (15) holds at time t, does it continue to be valid at time $t+1$? In our framework, we assume that the position of representative forecasts in relation to the mean growth expectations remains constant over time. [16] This assumption ensures that, regardless of changes at time t, the representative forecasters can accurately reflect the consensus forecast and the disagreement among all real agents.

Two immediate questions arise: how does the representative forecaster update their expectations, and what is their relationship with consensus forecasts? To answer these questions, we begin with the relationship between our two representative forecasters and the mean growth expectations:

$$\begin{aligned}
2\mathbb{F}_{t+1}[x_{t+1}] &= \mathbb{F}_{t+1}^H[x_{t+1}]+\mathbb{F}_{t+1}^L[x_{t+1}]\\
2\alpha\mathbb{F}_t[x_{t+1}]+2\boldsymbol{\beta}'\mathbf{r}_{t+1} &= (\alpha^H+\alpha^L)\mathbb{F}_t[x_{t+1}]+\left[(\boldsymbol{\beta}^H)'+(\boldsymbol{\beta}^L)'\right]\mathbf{r}_{t+1}\\
&\quad+(\alpha^H-\alpha^L)\sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right)
\end{aligned} \tag{16}$$

[15] The situation involving N representative forecasters will be discussed further in Subsection 4.4.

[16] The "position" of individuals relative to consensus forecasts will be discussed in Subsection 4.4.

This equation derives from the evolutionary relationship we propose for both consensus forecasts and representative forecasts. [17] Notice that equation (16) contains three terms on the right side and only two on the left. If the equation holds, we may, with less rigor, assume that the coefficients of each corresponding term must be equal. [18] Thus we have

$$\begin{cases} 2\alpha = \alpha^H + \alpha^L \\ 0 = \alpha^H - \alpha^L \\ 2\boldsymbol{\beta} = \boldsymbol{\beta}^H + \boldsymbol{\beta}^L \end{cases} \tag{17}$$

The implications of the equations are straightforward and intuitive. First, we establish that $\alpha^H = \alpha^L = \alpha$. This finding suggests that the representative forecasters adhere to the same information processing pattern as the mean growth expectations. Since we have already proven that α is equal to one, both representative and consensus forecasts are efficiently updated according to Bayes' rule. Second, equation (17) illustrates the differences in how the representative forecaster and the consensus forecaster react to the same news. Let $\Delta\boldsymbol{\beta} = \boldsymbol{\beta}^H - \boldsymbol{\beta} = \boldsymbol{\beta} - \boldsymbol{\beta}^L$ represent the differential interpretation of news between two representative forecasts and the consensus forecast, where the formulation of $\Delta\boldsymbol{\beta}$ ensures the symmetry in these interpretations. Since the forecast revision only reflect news, change in disagreement also stems exclusively from differential interpretations.

The advantage of our setup lies in its ability to utilize the disagreement between two representative forecasters as a proxy for the disagreement among all real forecasters. Therefore, two points are worth noting: First, the position of the representative forecaster relative to the consensus should remain unchanged over time. This stability ensures that both the mean and the disagreement of the representative forecaster consistently align with real-world agents. Second, one might be concerned that this conclusion only holds in a situation involving two forecasters; we will address this concern in Subsection 4.4 and demonstrate that the conclusion also holds for N individuals.

[17] Since we have

$$\mathbb{F}^H_{t+1}[x_{t+1}] = \alpha^H \mathbb{F}^H_t[x_{t+1}] + (\boldsymbol{\beta}^H)'\mathbf{r}_{t+1}$$
$$\mathbb{F}^L_{t+1}[x_{t+1}] = \alpha^L \mathbb{F}^L_t[x_{t+1}] + (\boldsymbol{\beta}^L)'\mathbf{r}_{t+1}$$

substituting in the approximations

$$\begin{aligned} &\mathbb{F}^H_{t+1}[x_{t+1}] + \mathbb{F}^L_{t+1}[x_{t+1}] \\ &= \left[\alpha^H \mathbb{F}^H_t[x_{t+1}] + (\boldsymbol{\beta}^H)'\mathbf{r}_{t+1}\right] + \left[\alpha^L \mathbb{F}^L_t[x_{t+1}] + (\boldsymbol{\beta}^L)'\mathbf{r}_{t+1}\right] \\ &= \alpha^H\left[\mathbb{F}_t[x_{t+1}] + \sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right)\right] + \alpha^L\left[\mathbb{F}_t[x_{t+1}] - \sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right)\right] + \left[(\boldsymbol{\beta}^H)' + (\boldsymbol{\beta}^L)'\right]\mathbf{r}_{t+1} \\ &= (\alpha^H + \alpha^L)\mathbb{F}_t[x_{t+1}] + \left[(\boldsymbol{\beta}^H)' + (\boldsymbol{\beta}^L)'\right]\mathbf{r}_{t+1} + (\alpha^H - \alpha^L)\sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right) \end{aligned}$$

[18] The reasonableness of this assumption will be discussed more rigorously in Subsection 4.4.

4.3 Learning the Cross-sectional Variance of Expectation Growth

In this subsection, we utilize representative forecasters to analyze changes in the cross-sectional variance of expectation growth. Referring to equation (15) and drawing from the conclusion that $\alpha^H = \alpha^L = \alpha = 1$, we derive the following

$$
\begin{aligned}
\left[\mathbb{F}_{t+1}[x_{t+1}] + \sigma\left(\mathbb{F}_{i,t+1}[x_{t+1}]\right)\right] &= \left[\mathbb{F}_t[x_{t+1}] + \sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right)\right] + (\boldsymbol{\beta}^H)'\mathbf{r}_{t+1} \\
\mathbb{F}_{t+1}[x_{t+1}] &= \mathbb{F}_t[x_{t+1}] + \boldsymbol{\beta}'\mathbf{r}_{t+1} \\
\left[\mathbb{F}_{t+1}[x_{t+1}] - \sigma\left(\mathbb{F}_{i,t+1}[x_{t+1}]\right)\right] &= \left[\mathbb{F}_t[x_{t+1}] - \sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right)\right] + (\boldsymbol{\beta}^L)'\mathbf{r}_{t+1}
\end{aligned}
\tag{18}
$$

Through these equations, it becomes apparent that the cross-sectional standard deviation follows a straightforward evolution equation, similar to that of the cross-sectional mean:

$$
\sigma\left(\mathbb{F}_{i,t+1}[x_{t+1}]\right) = \alpha\sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right) + (\Delta\boldsymbol{\beta})'\mathbf{r}_{t+1} \tag{19}
$$

where the coefficient α corresponding to the prior disagreement (cross-sectional standard deviation of expectation growth) also equals one. This result implies that disagreement revision is also solely influenced by news.

With equation (19), we implement the iterative method introduced previously:

$$
\sigma\left(\mathbb{F}_{i,T}[x^p]\right) - \alpha_0^T\sigma\left(\mathbb{F}_{i,0}[x^p]\right) = (\Delta\boldsymbol{\beta})'\left(\sum_{k=0}^{T-1}\alpha_0^k\mathbf{r}_{T-k}^p\right) \tag{20}
$$

Similar to equation (9), let $y = \sigma\left(\mathbb{F}_{i,T}[x^p]\right) - \alpha_0^T\sigma\left(\mathbb{F}_{i,0}[x^p]\right)$ and $\mathbf{X} = \sum_{k=0}^{T-1}\alpha_0^k\mathbf{r}_{T-k}^p$. This formulation allows us to directly apply the OLS estimation to determine the coefficients. Different values of α_0 yield different SSR, with the optimal α_0 being the one that minimizes the SSR.

Next, we provide empirical evidence to support this analysis. Figure 5 illustrates the relationship between the sum of squared residuals (SSR) and the coefficient α on prior disagreement during the first rolling window period from 1991Q1 to 2000Q4. The variation of SSR as a function of α aligns closely with the patterns observed in Figure 1. This suggests that the optimal coefficient α for the evolution of disagreement is also one. Additionally, Figure 2 depicts the mean and standard deviations of optimal coefficients across different estimation windows for Expectation, Expectation High, Expectation Low, and Disagreement. The results corroborate our previous analysis. Moreover, Table 5 confirms the symmetric relationship among the $\boldsymbol{\beta}$ coefficients, specifically that $\Delta\boldsymbol{\beta} = \boldsymbol{\beta}^H - \boldsymbol{\beta} = \boldsymbol{\beta} - \boldsymbol{\beta}^L$.

Another interesting finding is that the news (asset returns) capable of explaining forecast revisions may not necessarily account for revisions in disagreement. Table 5 exemplifies this outcome: the bivariate pair of assets (5-year fixed-term index and change in AAA-10Y spread), which provides a considerable R-squared value for the revisions of representative forecasters, explains only a minimal degree of revisions in disagreement ($R^2 = 0.014$). Conversely, another bivariate asset pair (change in AAA-10Y spread and factor HML) that does not substantially explain the forecast revisions can explain a larger degree of revisions in disagreement ($R^2 = 0.152$). This result is intuitively reasonable. During certain periods, market participants may uniformly update their expectations based on pivotal news, yet this does not imply significant variances in their perceptions of this information. For instance, should the Federal Reserve signal a rate hike in response to strong economic indicators, market participants might uniformly adjust their forecasts concerning future economic conditions. Although interpretations of this information may vary slightly, the general consensus about the central bank's strategy to manage inflation typically leads to closely aligned market expectations. This example illustrates that while certain information may universally influence expectations, it does not necessarily provoke marked disagreement among market participants.

In this section, we establish the evolution equation of forecasts disagreement. Similar to the consensus forecasts, the update of disagreement is independent from the previous disagreement. These conclusions are drawn by utilizing representative forecasters, since this setup helps us avoid estimating coefficients for each real-world agents. However, further discussions concerning representative forecasters are warranted.

4.4 Discussion

In the previous setup, we assume that the two representative forecasters adhere to equation (15). Now, we extend it to a more general case:

$$\begin{aligned} \mathbb{F}_{i,t}\left[x_{t+1}\right] &= \mathbb{F}_{t}\left[x_{t+1}\right] + k_i \sigma\left(\mathbb{F}_{i,t}\left[x_{t+1}\right]\right) \\ \mathbb{F}_{i,t+1}\left[x_{t+1}\right] &= \mathbb{F}_{t+1}\left[x_{t+1}\right] + k_i \sigma\left(\mathbb{F}_{i,t+1}\left[x_{t+1}\right]\right) \end{aligned} \tag{21}$$

where $k_i \neq 0$ denotes the position of a representative forecaster i relative to the consensus forecaster, and remains constant over time. This constancy is crucial as it allows the representative forecasters to represent the consensus and disagreement among all real agents in any period. In the previous section, we introduce $\mathbb{F}_t^H\left[x_{t+1}\right]$ and $\mathbb{F}_t^L\left[x_{t+1}\right]$ as instances where k_i equals 1 and -1, respectively. It is straightforward to demonstrate that our conclusions are valid for other values of k_i (say, ± 2).

We now turn to the situation of N representative forecasters. By substituting equation (25) into the evolution equations (4) and (6) regarding expectations for individuals and consensus, we can obtain the evolution equation for disagreement directly: [19]

$$\sigma\left(\mathbb{F}_{i,t+1}[x_{t+1}]\right) = \alpha_i \sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right) + \frac{\alpha_i - \alpha}{k_i}\mathbb{F}_t[x_{t+1}] + \left(\frac{\boldsymbol{\beta}_i - \boldsymbol{\beta}}{k_i}\right)'\mathbf{r}_{t+1} \tag{22}$$

This equation holds for any representative forecaster i. Assume there exists a representative forecaster i such that $\alpha_i = \alpha$. Substituting it into equation (22), we have

$$\sigma\left(\mathbb{F}_{i,t+1}[x_{t+1}]\right) = \alpha \sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right) + \left(\frac{\boldsymbol{\beta}_i - \boldsymbol{\beta}}{k_i}\right)'\mathbf{r}_{t+1} \tag{23}$$

This is equivalent to the evolution equation (19) for the cross-sectional standard deviation, where $\Delta\boldsymbol{\beta} = (\boldsymbol{\beta}_i - \boldsymbol{\beta})/k_i$.

We then apply the methodology outlined in Equation (16) across N representative forecasters: [20]

$$\begin{aligned}
\mathbb{F}_{t+1}[x_{t+1}] &= \frac{1}{N}\sum_{i=1}^{N}\mathbb{F}_{i,t+1}[x_{t+1}] \\
&= \left(\frac{1}{N}\sum_{i=1}^{N}\alpha_i\right)\mathbb{F}_t[x_{t+1}] + \left(\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\beta}_i'\right)\mathbf{r}_{t+1} \\
&\quad + \left(\frac{1}{N}\sum_{i=1}^{N}\alpha_i k_i\right)\sigma(\mathbb{F}_{i,t}[x_{t+1}]) \\
&= \alpha\mathbb{F}_t[x_{t+1}] + \boldsymbol{\beta}'\mathbf{r}_{t+1}
\end{aligned} \tag{24}$$

[19] We begin with the evolution equation for agent i:

$$\begin{aligned}
\mathbb{F}_{i,t+1}[x_{t+1}] &= \alpha_i\mathbb{F}_{i,t}[x_{t+1}] + \boldsymbol{\beta}_i'\mathbf{r}_{t+1} \\
\mathbb{F}_{t+1}[x_{t+1}] + k_i\sigma\left(\mathbb{F}_{i,t+1}[x_{t+1}]\right) &= \alpha_i\mathbb{F}_t[x_{t+1}] + \alpha_i k_i\sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right) + \boldsymbol{\beta}_i'\mathbf{r}_{t+1} \\
\alpha\mathbb{F}_t[x_{t+1}] + \boldsymbol{\beta}'\mathbf{r}_{t+1} + k_i\sigma\left(\mathbb{F}_{i,t+1}[x_{t+1}]\right) &= \alpha_i k_i\sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right) + \alpha_i\mathbb{F}_t[x_{t+1}] + \boldsymbol{\beta}_i'\mathbf{r}_{t+1} \\
k_i\sigma\left(\mathbb{F}_{i,t+1}[x_{t+1}]\right) &= \alpha_i k_i\sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right) + (\alpha_i - \alpha)\mathbb{F}_t[x_{t+1}] + (\boldsymbol{\beta}_i - \boldsymbol{\beta})'\mathbf{r}_{t+1} \\
\sigma\left(\mathbb{F}_{i,t+1}[x_{t+1}]\right) &= \alpha_i\sigma\left(\mathbb{F}_{i,t}[x_{t+1}]\right) + \frac{\alpha_i - \alpha}{k_i}\mathbb{F}_t[x_{t+1}] + \left(\frac{\boldsymbol{\beta}_i - \boldsymbol{\beta}}{k_i}\right)'\mathbf{r}_{t+1}
\end{aligned}$$

[20] Since we have

$$\begin{aligned}
\mathbb{F}_{i,t+1}[x_{t+1}] &= \alpha_i\mathbb{F}_{i,t}[x_{t+1}] + \boldsymbol{\beta}_i'\mathbf{r}_{t+1} \\
&= \alpha_i\left[\mathbb{F}_t[x_{t+1}] + k_i\sigma(\mathbb{F}_{i,t}[x_{t+1}])\right] + \boldsymbol{\beta}_i'\mathbf{r}_{t+1} \\
&= \alpha_i\mathbb{F}_t[x_{t+1}] + \boldsymbol{\beta}_i'\mathbf{r}_{t+1} + \alpha_i k_i\sigma(\mathbb{F}_{i,t}[x_{t+1}])
\end{aligned}$$

Given that k_i represents the position of the exogenously determined representative forecaster, it should remain independent of α_i. Consequently, their covariance, $\text{Cov}(\alpha_i, k_i)$, should be zero. [21] And the coefficient $\frac{1}{N}\sum_{i=1}^{N}\alpha_i k_i$ corresponding to the third term $\sigma(\mathbb{F}_{i,t}[x_{t+1}])$ on the right side of equation (24), equals zero, too. [22] Finally, we establish that $\alpha = \frac{1}{N}\sum_{i=1}^{N}\alpha_i = 1$, $\boldsymbol{\beta} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\beta}_i$ and $\boldsymbol{\beta}_i = \boldsymbol{\beta} + k_i\Delta\boldsymbol{\beta}$. However, it is critical to note that these results apply exclusively to representative forecasters and not to real agents, given that the position k_i relative to the consensus forecaster remains constant and independent of α_i.

5 High-Frequency Expectations and Disagreement

In this section, we apply our mixed-frequency method to construct a daily time series of agents' mean expectations and disagreement regarding macroeconomic growth. Our method is related to the following literature. Ghysels and Wright (2009) introduce the MIDAS (mixed frequency data sampling) regression and Kalman filter to utilize asset price data for constructing daily forecasts of upcoming survey releases; Andreou, Ghysels, and Kourtellos (2013) extract a small set of daily financial factors from a large panel of about 1000 daily financial assets and rely on MIDAS to provide daily forecasts; Chaudhry and Oh (2020) propose a framework based on reinforcement learning (RL) which has shown superior empirical performance. Next, we first detail our method and subsequently compare it with these established methods.

5.1 Extract Daily Time Series of the Cross-sectional Mean and Variance

In our previous discussion, we introduce a mixed-frequency method to analyze forecast revisions on a daily frequency and illustrate that changes in expectations occur more frequently than the quarterly survey intervals suggest. This method also enables us to reconstruct the unobserved daily series of expectations that span the interval between two quarterly survey releases by estimating coefficients within the evolution equations. Let

[21] It is easy to know that $\bar{k} = \frac{1}{N}\sum_{i=1}^{N} k_i = 0$. Thus we have

$$\text{Cov}(\alpha_i, k_i) = \sum_{i=1}^{N}(\alpha_i - \bar{\alpha})(k_i - \bar{k}) = \sum_{i=1}^{N}(\alpha_i k_i - \bar{\alpha}k_i - \bar{k}\alpha_i + \bar{\alpha}\bar{k}) = \sum_{i=1}^{N}\alpha_i k_i$$

[22] This is also the reason why we assumed this term to be 0 in the previous discussion.

$\mathbb{F}_0[x^p]$ and $\sigma\left(\mathbb{F}_{i,0}\left[x^p\right]\right)$ represent the true cross-sectional mean and standard deviation for period p as reported in the previous quarter's survey, respectively. Applying the recursive form in equation (7), we derive the estimated daily-frequency time series:

$$\begin{bmatrix} \mathbb{F}_1\left[x^p\right] \\ \mathbb{F}_2\left[x^p\right] \\ \vdots \\ \mathbb{F}_T\left[x^p\right] \end{bmatrix} = \mathbf{D} \begin{bmatrix} \mathbb{F}_0\left[x^p\right] \\ \mathbb{F}_0\left[x^p\right] \\ \vdots \\ \mathbb{F}_0\left[x^p\right] \end{bmatrix} + \mathbf{T} \begin{bmatrix} \mathbf{r}_1^p \\ \mathbf{r}_2^p \\ \vdots \\ \mathbf{r}_T^p \end{bmatrix} \tag{25}$$

where $\mathbf{D} = \operatorname{diag}\left(\hat{\alpha}^1, \hat{\alpha}^2, \cdots, \hat{\alpha}^T\right)$ is a scaling matrix that represents the contribution of the initial state $\mathbb{F}_0\left[x^p\right]$ at each point in time. The input vector $\mathbf{r}_t^p$ denotes the asset returns at time t. Its impact is expressed through a Toeplitz matrix $\mathbf{T}$, [23] which captures how each input (asset returns) affects changes in the state over time. As previously discussed, the estimated coefficient $\hat{\alpha} = 1$. Consequently, the estimated forecast for day t, starting from the release date of the previous quarter, is computed as $\mathbb{F}_t\left[x^p\right] = \mathbb{F}_0\left[x^p\right] + \hat{\boldsymbol{\beta}}'\left(\sum_{k=0}^{t-1} \mathbf{r}_{t-k}^p\right)$. The estimation of daily-frequency disagreement follows a similar process with expectations, requiring only a substitution of $\mathbb{F}_t\left[x^p\right]$ with $\sigma\left(\mathbb{F}_{i,t}\left[x^p\right]\right)$, specifically, $\sigma\left(\mathbb{F}_{i,t}\left[x^p\right]\right) = \sigma\left(\mathbb{F}_{i,0}\left[x^p\right]\right) + \Delta\hat{\boldsymbol{\beta}}'\left(\sum_{k=0}^{t-1} \mathbf{r}_{t-k}^p\right)$.

We implement a recursive out-of-sample estimation procedure to generate the daily series for the cross-sectional mean and standard deviation of GDP growth forecasts. For each quarter from 2000 to 2023, the daily-frequency series is estimated using our mixed-frequency method over a rolling lookback window of $T = 40$ quarters. These windows span from the SPF release date in quarter $t - T$ to the release in quarter t. Each window consists of an estimating period and a predicting period. During the estimating period, the coefficients in equations (9) and (20) are estimated using historical data from $t - T$ to $t - 1$. We still select a bivariate pair of assets that yield the highest R-squared for each period, as discussed previously, to enhance the accuracy of model fitting. In the subsequent

[23] A Toeplitz matrix is defined by constant diagonals, with each row shifted right from the one above. It's commonly used in signal processing for representing linear systems. In our analysis, the Toeplitz matrix is defined as follows:

$$\mathbf{T} = \begin{bmatrix} \hat{\boldsymbol{\beta}}' & 0 & \cdots & 0 \\ \hat{\alpha}\hat{\boldsymbol{\beta}}' & \hat{\boldsymbol{\beta}}' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\alpha}^{T-1}\hat{\boldsymbol{\beta}}' & \hat{\alpha}^{T-2}\hat{\boldsymbol{\beta}}' & \cdots & \hat{\boldsymbol{\beta}}' \end{bmatrix}$$

This matrix shares the structure used by Chahrour and Jurado (2021) in computing the foresight quantity for the income process within the information structure. The intuition is that the asset returns we employ and their associated foresight serve a comparable function: providing agents with information that extends beyond the current and historical data of a stochastic process.

predicting period, these estimated coefficients are integrated into the diagonal matrices $\mathbf{D}$ and the Toeplitz matrix $\mathbf{T}$. The daily time series is then derived according to equation (25).

Figures 6 and 7 illustrate the out-of-sample estimation results. Our method reliably generates daily estimates of the cross-sectional mean and standard deviation that closely approximate the true quarterly values. To assess the accuracy of our daily series, we compare the estimated cross-sectional means and standard deviations derived from our approach on SPF release dates with their actual counterparts. For example, on November 11th, 2000, the true cross-sectional mean reported by the SPF is 3.221%, whereas our daily estimate was 3.215%, yielding an absolute error of 0.005%. Similarly, the true cross-sectional standard deviation is 0.867%, compared to our estimate of 0.972%, yielding an absolute error of 0.105%. Over the entire out-of-sample period, the estimated daily cross-sectional mean realizes an R^2 of 93.3% relative to the true quarterly series, and the estimated daily cross-sectional standard deviation realizes an R^2 of 84.5%.

5.2 Comparison with RL Approach: Machine Learning Interpretability

In this subsection, we undertake a detailed comparison of our cross-sectional estimation method with other established methods documented in the literature. Chaudhry and Oh (2020) provide a comprehensive review and comparative analysis of MIDAS, the Kalman filter, and reinforcement learning (RL) for estimating latent expectations at a daily frequency, Their findings indicate that the RL method significantly improves empirical performance by reducing the number of estimated parameters. Therefore, our discussion will primarily focus on illustrating both the similarities and differences between our mixed-frequency method and the RL-based method. We begin with a concise overview of the RL-based method, followed by a comparative analysis with our method. This comparison is crucial as it allows us to better understand and interpret the predictions generated by black-box models, particularly those based on reinforcement learning.

In general, reinforcement learning (RL) algorithms allow an agent to learn the optimal policy for action decisions based on the current state. Within the RL framework, the state vector for period t consists of the cross-sectional mean and variance from period $t-1$, along with the asset returns of period t:

$$\varphi(s_t) = \begin{pmatrix} \mathbb{F}_{t-1}[x^p] \\ \mathrm{Var}(\mathbb{F}_{i,t-1}[x^p]) \\ \mathbf{r}_t \end{pmatrix} \in \mathbb{R}^{m+2} \tag{26}$$

where m denotes the number of assets. The initial state is

$$\varphi(s_1) = \begin{pmatrix} \mathbb{F}_0[x^p] \\ \mathrm{Var}(\mathbb{F}_{i,0}[x^p]) \\ \mathbf{r}_1 \end{pmatrix} \tag{27}$$

where $\mathbb{F}_0[x^p]$ and $\mathrm{Var}(\mathbb{F}_{i,0}[x^p])$ represent the true cross-sectional mean and variance from the previous quarter's survey release. According to equation (6) and equation (13), the RL algorithm employs the following policy function

$$g_{\boldsymbol{\lambda}}(s_t) \equiv \begin{pmatrix} \mathbb{F}_t[x^p] \\ \mathrm{Var}(\mathbb{F}_{i,t}[x^p]) \end{pmatrix} = \begin{pmatrix} \alpha\mathbb{F}_{t-1}[x^p] + \boldsymbol{\beta}'\mathbf{r}_t \\ \gamma\mathrm{Var}(\mathbb{F}_{i,t-1}[x^p]) + \boldsymbol{\zeta}'\mathbf{r}_t\mathbf{r}_t'\boldsymbol{\zeta} + \boldsymbol{\delta}'\mathbf{r}_t\mathbb{F}_{t-1}[x^p] \end{pmatrix} \in \mathbb{R}^2 \tag{28}$$

where

$$\boldsymbol{\lambda} = \begin{pmatrix} \alpha \\ \boldsymbol{\beta} \\ \gamma \\ \boldsymbol{\zeta} \\ \boldsymbol{\delta} \end{pmatrix} \in \mathbb{R}^{3m+2} \tag{29}$$

According to Chaudhry and Oh (2020), the superior performance of reinforcement learning (RL) is attributed to its efficiency in parameter estimation. The RL method requires only $3m+2$ parameters, which is significantly fewer than those required by MIDAS and Kalman filtering methods. The objective is to minimize the Euclidean distance between the predicted values and the true values. Consequently, the rewards within the RL framework are defined as follows:

$$r_t(s^t) = \begin{cases} 0 & \text{if } t < T \\ -\left\| \begin{pmatrix} \hat{\mathbb{F}}_T[x^p] \\ \hat{\sigma}(\mathbb{F}_{i,T}[x^p]) \end{pmatrix} - \begin{pmatrix} \mathbb{F}_T[x^p] \\ \sigma(\mathbb{F}_{i,T}[x^p]) \end{pmatrix} \right\| & \text{if } t = T \end{cases} \tag{30}$$

where $\mathbb{F}_T[x^p]$ and $\sigma(\mathbb{F}_{i,T}[x^p])$ are the observed moments, computed from the cross-section of forecasts released quarterly. This setup allows for the direct application of the RL algorithm to learn the optimal parameters $\boldsymbol{\lambda}$. The learned function $g_{\boldsymbol{\lambda}}$ is subsequently used to estimate the daily cross-sectional mean and variance.

Both the RL method and our mixed-frequency method demonstrate strong empirical performance, realizing R^2 values of 82.3% and 93.3% for the cross-sectional mean, respectively. However, several important details warrant discussion. First, the RL method estimates the cross-sectional mean and variance jointly, which yields better estimates of the cross-sectional mean. This outcome may seem counterintuitive, as the empirical performance improves despite an actual increase in the number of parameters (from $m+1$

to $3m+2$). Our research provides an economic rationale for why this setup improves the performance of such black-box RL algorithms. As shown in equation (28), incorporating higher-order terms introduces additional dimensional information beyond simple mean estimation. As discussed in Section 3, these parameters are actually determined by the coefficients in the individual-level evolution equations. While directly estimating these parameters without constraints may lead to a better fit, it is at the expense of introducing bias, and the estimates cannot be considered as true variances. However, by utilizing our proposed representative forecasters, we can avoid the direct estimation of constrained parameters, therefore transforming the problem into a simple and unconstrained linear regression.

Second, in their estimation process, they fix the coefficient on the one-day lag cross-sectional mean (α) at one. The results show that this constraint also yields better performance than freely estimating α. The decision to fix α at one can also be explained within our analytical framework. Firstly, it reduces the number of parameters required, thereby potentially increasing the efficiency of the RL algorithms. Secondly, our substantial theoretical and empirical analysis supports the setting of α to one. This prior knowledge simplifies the model and consequently improves its empirical performance.

Third, if only the cross-sectional mean is estimated separately and the coefficient α is not fixed at one, the parameters that the RL method needs to estimate —specifically, the $\boldsymbol{\beta}$ coefficients associated with asset returns —are identical to those we consider. This raises an immediate question: why do the parameters estimated by two methods appear the same, yet yield different estimation results and empirical performances? The primary reason lies in the differing choices of estimation techniques. Ordinary Least Squares (OLS) is employed in our analysis to optimize linear regression models by minimizing the sum of squared residuals. It yields a Best Linear Unbiased Estimator (BLUE) under classical assumptions. In contrast, RL focuses on maximizing long-term cumulative rewards rather than directly minimizing prediction errors. This makes RL more suitable for dynamic, sequential decision-making environments, as it relies on environmental interactions and requires sufficient historical data for effective learning. However, when model is clear and linear, RL may introduce extra bias and variance, particularly when the hyper-parameters are improperly set or the data is not sufficient enough. Another reason contributing to the difference in empirical performance between the two methods is asset selection. Due to the complexity of implementing and computing RL algorithms, it is difficult to select asset returns that yield the highest R^2 within each window period. Model convergence already requires many iterations, and hyper-parameter tuning further increases these computational costs.

However, with OLS, we can efficiently explore all binary regressions to identify the optimal assets for each window period, due to its analytical solution for parameter estimation.

6 Conclusion

This paper investigates the high-frequency dynamics of macroeconomic expectations and the corresponding evolution of disagreement among forecasters. Recognizing the critical role that expectations play in financial markets and economic policy, our study addresses the limitations of existing low-frequency frameworks that often overlook the rapid adjustments made by agents in response to new information.

Our primary contributions are threefold. First, we introduce a mixed-frequency estimation method that effectively integrates high-frequency asset returns with low-frequency macroeconomic survey data. Through rigorous analysis, we demonstrate that consensus forecasts are updated efficiently in accordance with the Bayesian framework. The forecast revision for consensus exhibit independence from prior forecasts and a strong responsiveness to new information. Second, we propose a new concept of "representative forecasters" as proxies for real agents to derive a simple yet intuitive evolution equation for forecast disagreement. This setup helps us to show that changes in disagreement stem primarily from heterogeneous interpretations of incoming information rather than from prior disagreement. Third, we recover the unobserved daily series of expectations and disagreement between two quarterly survey releases dates. While there are some mixed-frequency methods in the literature for estimating the first moment (expectation) at a daily frequency, we are the first to recover the second moment (disagreement) to a daily frequency.

The implications of our findings extend to both economic theory and practical applications. By highlighting the critical relationship between high-frequency information and expectation dynamics, we emphasize the necessity of incorporating high-frequency data into forecasting models. Our framework is straightforward and user-friendly, making it suitable for economists and policymakers. Additionally, our framework offers valuable insights and economic explanations for the settings of certain black-box models, such as reinforcement learning. Finally, our mixed-frequency framework can be applied to other macroeconomic variables to enhance the robustness of forecasting models in various economic contexts.

References

Anderson, E. W., E. Ghysels, and J. L. Juergens. 2005. Do heterogeneous beliefs matter for asset pricing? *Review of Financial Studies* 18:875–924.

Andrade, P., and H. Le Bihan. 2013. Inattentive professional forecasters. *Journal of Monetary Economics* 60:967–982.

Andreou, E., E. Ghysels, and A. Kourtellos. 2010. Regression models with mixed sampling frequencies. *Journal of Econometrics* 158:246–261.

Andreou, E., E. Ghysels, and A. Kourtellos. 2013. Should macroeconomic forecasters use daily financial data and how? *Journal of Business and Economic Statistics* 31:240–251.

Banerjee, S. 2011. Learning from prices and the dispersion in beliefs. *Review of Financial Studies* 24:3025–3068.

Barberis, N. 2018. Psychology-based models of asset prices and trading volume. In *Handbook of Behavioral Economics: Applications and Foundations 1*, vol. 1, pp. 79–175. Elsevier.

Bernanke, B. S., and K. N. Kuttner. 2005. What explains the stock market's reaction to Federal Reserve policy? *Journal of Finance* 60:1221–1257.

Bordalo, P., J. J. Conlon, N. Gennaioli, S. Y. Kwon, and A. Shleifer. 2023. Memory and probability. *Quarterly Journal of Economics* 138:265–311.

Carroll, C. D. 2003. Macroeconomic expectations of households and professional forecasters. *Quarterly Journal of Economics* 118:269–298.

Chahrour, R., and K. Jurado. 2021. Optimal foresight. *Journal of Monetary Economics* 118:245–259.

Chaudhry, A., and S. Oh. 2020. High-Frequency Expectations from Asset Prices: A Machine Learning Approach. *Available at SSRN 3694019* .

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.

Coibion, O., Y. Gorodnichenko, and S. Kumar. 2018. How do firms form their expectations? New survey evidence. *American Economic Review* 108:2671–2713.

Deng, Y., Y. Wang, and T. Zhou. 2024. Macroeconomic Expectations and Expected Returns. *Journal of Financial and Quantitative Analysis* pp. 1–70.

Fama, E. F. 1970. Efficient capital markets. *Journal of Finance* 25:383–417.

Fuhrer, J. C. 2018. Intrinsic expectations persistence: evidence from professional and household survey expectations. *Federal Reserve Bank of Boston* .

Ghysels, E. 2016. Macroeconomics and the reality of mixed frequency data. *Journal of Econometrics* 193:294–314.

Ghysels, E., A. Sinko, and R. Valkanov. 2007. MIDAS regressions: Further results and new directions. *Econometric Reviews* 26:53–90.

Ghysels, E., and J. H. Wright. 2009. Forecasting professional forecasters. *Journal of Business and Economic Statistics* 27:504–516.

Giacomini, R., V. Skreta, and J. Turen. 2020. Heterogeneity, inattention, and Bayesian updates. *American Economic Journal: Macroeconomics* 12:282–309.

Hagenhoff, T., and J. Lustenhouwer. 2023. The role of stickiness, extrapolation and past consensus forecasts in macroeconomic expectations. *Journal of Economic Dynamics and Control* 149:104638.

Jonung, L. 1981. Perceived and expected rates of inflation in Sweden. *American Economic Review* 71:961–968.

Kerssenfischer, M., and M. Schmeling. 2024. What moves markets? *Journal of Monetary Economics* p. 103560.

Kohlhas, A. N., and A. Walther. 2021. Asymmetric attention. *American Economic Review* 111:2879–2925.

Lahiri, K., and X. Sheng. 2008. Evolution of forecast disagreement in a Bayesian learning model. *Journal of Econometrics* 144:325–340.

Lucas Jr, R. E. 1972. Expectations and the Neutrality of Money. *Journal of Economic Theory* 4:103–124.

Mankiw, N. G., and R. Reis. 2002. Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve. *Quarterly Journal of Economics* 117:1295–1328.

Mankiw, N. G., R. Reis, and J. Wolfers. 2003. Disagreement about inflation expectations. *NBER Macroeconomics Annual* 18:209–248.

Manzan, S. 2021. Are professional forecasters Bayesian? *Journal of Economic Dynamics and Control* 123:104045.

Mariano, R. S., and S. Ozmucur. 2020. High-mixed frequency forecasting methods in R—With applications to Philippine GDP and inflation. *Handbook of Statistics* 42:185–227.

Muth, J. F. 1961. Rational expectations and the theory of price movements. *Econometrica: Fournal of the Econometric Society* pp. 315–335.

Nakamura, E., and J. Steinsson. 2018. High-frequency identification of monetary non-neutrality: the information effect. *Quarterly Journal of Economics* 133:1283–1330.

Newey, W. K., and K. D. West. 1994. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61:631–653.

Nordhaus, W. D. 1987. Forecasting Efficiency: Concepts and Applications. *The Review of Economics and Statistics* 69:667–674.

Patton, A. J., and A. Timmermann. 2010. Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion. *Journal of Monetary Economics* 57:803–820.

Pettenuzzo, D., A. Timmermann, and R. Valkanov. 2016. A MIDAS approach to modeling first and second moment dynamics. *Journal of Econometrics* 193:315–334.

Sargent, T. J., and N. Wallace. 1976. Rational expectations and the theory of economic policy. *Journal of Monetary Economics* 2:169–183.

Sims, C. A. 2003. Implications of rational inattention. *Journal of Monetary Economics* 50:665–690.

A Appendix: The Optimization of SSR

Let $y = \mathbb{F}_T[x^p] - \alpha_0^T \mathbb{F}_0[x^p]$ and $\mathbf{X} = \sum_{k=0}^{T-1} \alpha_0^k \mathbf{r}_{T-k}^p$. According to the definition of the sum of squared residuals:

$$\begin{aligned}
\mathrm{SSR} &= (y - \mathbf{X}\boldsymbol{\beta})'(y - \mathbf{X}\boldsymbol{\beta}) \\
&= y'y - y'\mathbf{X}\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}'\mathbf{X}'y + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}
\end{aligned}$$

First Term: Since $y'y$ does not explicitly depend on α_0, the first term $\frac{\partial(y'y)}{\partial \alpha_0}$ is 0.

Derivative of y: Since y linearly depends on α_0^T, the derivative of y with respect to α_0 is calculated as $\frac{\partial y}{\partial \alpha_0} = -T\alpha_0^{T-1}\mathbb{F}_0[x^p]$

Derivative of X: The derivative of $\mathbf{X}$ with respect to α_0 is $\frac{\partial \mathbf{X}}{\partial \alpha_0} = \sum_{k=0}^{T-1} k\alpha_0^{k-1}\mathbf{r}_{T-k}^p$

Second and Third Term:

$$\begin{aligned}
\frac{\partial\left(y'\mathbf{X}\hat{\boldsymbol{\beta}}\right)}{\partial \alpha_0} &= \frac{\partial\left(\hat{\boldsymbol{\beta}}'\mathbf{X}'y\right)}{\partial \alpha_0} \\
&= \frac{\partial y'}{\partial \alpha_0}\mathbf{X}\hat{\boldsymbol{\beta}} + y'\frac{\partial \mathbf{X}}{\partial \alpha_0}\hat{\boldsymbol{\beta}} \\
&= -T\alpha_0^{T-1}\mathbb{F}_0[x^p]'\mathbf{X}\hat{\boldsymbol{\beta}} + y'\left(\sum_{k=0}^{T-1} k\alpha_0^{k-1}\mathbf{r}_{T-k}^p\right)\hat{\boldsymbol{\beta}}
\end{aligned}$$

Last Term:

$$\begin{aligned}
\frac{\partial\left(\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}\right)}{\partial \alpha_0} &= \hat{\boldsymbol{\beta}}'\left(\frac{\partial \mathbf{X}'}{\partial \alpha_0}\mathbf{X} + \mathbf{X}'\frac{\partial \mathbf{X}}{\partial \alpha_0}\right)\hat{\boldsymbol{\beta}} \\
&= \hat{\boldsymbol{\beta}}'\left[\left(\sum_{k=0}^{T-1} k\alpha_0^{k-1}\mathbf{r}_{T-k}^p\right)'\mathbf{X} + \mathbf{X}'\left(\sum_{k=0}^{T-1} k\alpha_0^{k-1}\mathbf{r}_{T-k}^p\right)\right]\hat{\boldsymbol{\beta}}
\end{aligned}$$

Combining these derivatives, we derive the overall gradient of SSR with respect to α_0:

$$\begin{aligned}
\frac{\partial\,\mathrm{SSR}}{\partial \alpha_0} = 2T\alpha_0^{T-1}\mathbb{F}_0[x^p]'\mathbf{X}\hat{\boldsymbol{\beta}} - 2y'\left(\sum_{k=0}^{T-1} k\alpha_0^{k-1}\mathbf{r}_{T-k}^p\right)\hat{\boldsymbol{\beta}} \\
+ \hat{\boldsymbol{\beta}}'\left[\left(\sum_{k=0}^{T-1} k\alpha_0^{k-1}\mathbf{r}_{T-k}^p\right)'\mathbf{X} + \mathbf{X}'\left(\sum_{k=0}^{T-1} k\alpha_0^{k-1}\mathbf{r}_{T-k}^p\right)\right]\hat{\boldsymbol{\beta}}
\end{aligned}$$

B Appendix: The Evolution of the Cross-sectional Variance

Since we have the following expression for the cross-sectional variance:

$$\text{Var}(\mathbb{F}_{i,t}[x_{t+1}]) = \frac{1}{N}\sum_{i=1}^{N}(\mathbb{F}_{i,t}[x_{t+1}] - \mathbb{F}_t[x_{t+1}])^2$$

substituting the evolution equations at both the individual and consensus level

$$\begin{aligned}
\mathbb{F}_{i,t+1}[x_{t+1}] &= \alpha_i \mathbb{F}_{i,t}[x_{t+1}] + \boldsymbol{\beta}_i' \mathbf{r}_{t+1} \\
\mathbb{F}_{t+1}[x_{t+1}] &= \alpha \mathbb{F}_t[x_{t+1}] + \boldsymbol{\beta}' \mathbf{r}_{t+1}
\end{aligned}$$

we have

$$\begin{aligned}
\text{Var}(\mathbb{F}_{i,t+1}[x_{t+1}]) &= \frac{1}{N}\sum_{i=1}^{N}\left[(\alpha_i \mathbb{F}_{i,t}[x_{t+1}] + \boldsymbol{\beta}_i' \mathbf{r}_{t+1}) - (\alpha \mathbb{F}_t[x_{t+1}] + \boldsymbol{\beta}' \mathbf{r}_{t+1})\right]^2 \\
=& \frac{1}{N}\sum_{i=1}^{N}\left[\left(\alpha_i \mathbb{F}_{i,t}[x_{t+1}] - \alpha \mathbb{F}_t[x_{t+1}]\right)^2 + \left[\left(\boldsymbol{\beta}_i' - \boldsymbol{\beta}'\right)\mathbf{r}_{t+1}\right]^2 + 2\left(\alpha_i \mathbb{F}_{i,t}[x_{t+1}] - \alpha \mathbb{F}_t[x_{t+1}]\right)\left(\boldsymbol{\beta}_i' - \boldsymbol{\beta}'\right)\mathbf{r}_{t+1}\right] \\
=& \gamma \text{Var}(\mathbb{F}_{i,t}[x_{t+1}]) + \boldsymbol{\zeta}' \mathbf{r}_{t+1} \mathbf{r}_{t+1}' \boldsymbol{\zeta} + \boldsymbol{\delta}' \mathbf{r}_{t+1} \mathbb{F}_t[x_{t+1}]
\end{aligned}$$

where γ is a scalar related to α and α_i, $\boldsymbol{\zeta}$ is a vector of scalars related to $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_i$, and $\boldsymbol{\delta}$ is a vector of scalars related to α, α_i, $\boldsymbol{\beta}$, and $\boldsymbol{\beta}_i$. The detailed expression is as follows:

$$\begin{cases}
\gamma = \dfrac{\frac{1}{N}\sum_{i=1}^{N}\left(\alpha_i \mathbb{F}_{i,t}[x_{t+1}] - \alpha \mathbb{F}_t[x_{t+1}]\right)^2}{\frac{1}{N}\sum_{i=1}^{N}\left(\mathbb{F}_{i,t}[x_{t+1}] - \mathbb{F}_t[x_{t+1}]\right)^2} \\
\zeta = \dfrac{1}{N}\sum_{i=1}^{N}(\boldsymbol{\beta}_i - \boldsymbol{\beta})
\end{cases}$$

The expression for the cross-sectional variance decomposes into three terms: (i) the disagreement of the individual forecasts from the previous time period; (ii) the difference in response coefficient $\boldsymbol{\beta}_i$ from the consensus $\boldsymbol{\beta}$; and (iii) the covariance between the news and consensus prior.

C Tables

Table 1. Summary of Data Sources

Type	Data	Source
Survey		
Forecasts	Survey of Professional Forecasters	Philadelphia Fed
Equity		
Stock Market Indexes	Value-weighted Return	CRSP
Fama-French Factors	Factor Returns	Fama-French Library
Fixed Income		
Treasuries	Return on Fixed-term Indexes	CRSP
Credit	Change in AAA- and BAA-10Y Spread	FRED
Exchange Rates		
USD	Change in U.S. Dollar Index	FRED
Derivatives		
Options	Change in VIX Index	CRSP

This table summarizes our data sources: For real GDP growth forecasts, we employ quarterly data from the Survey of Professional Forecasters (SPF). For Equities, we consider daily returns from the CRSP value-weighted portfolio and Fama-French 3 factors (market, size and value). For treasuries, we consider daily returns from the CRSP fixed-term indexes, which cover periods of 1, 5, 10, 20, and 30 years, recognizing that yields from short-term and long-term government bonds convey distinct information. For credit, we consider changes in the spread between the yields on corporate bonds with AAA and BAA Ratings. For exchange rates, we consider changes in the weighted average of the U.S. dollar's foreign exchange value (following the discontinuation of variable "DTWEXM" from FRED, we have integrated it with variable "DETOXIFIES"). And for derivatives we use the change in VIX index.

Table 2. Summary Statistics of GDP Growth Forecasts

	GDP Growth Expectation		GDP Growth Disagreement	
	Nowcast	Forecast	Nowcast	Forecast
Count	38.634	38.634	38.634	38.634
Mean	2.262	2.680	1.031	0.994
Std Dev	3.623	1.269	0.820	1.062
Min	-31.478	-1.517	0.420	0.436
P25	1.826	2.316	0.656	0.609
Median	2.520	2.591	0.846	0.760
P75	3.102	3.076	1.106	1.079
Max	19.405	10.422	7.168	11.736

The table reports the summary statistics of the cross-sectional moments of GDP growth forecasts from the Survey of Professional Forecasters (SPF). We construct time series for both the cross-sectional mean, called "Expectation", and the cross-sectional standard deviation, called "Disagreement". The "Nowcast" represents the forecast of GDP growth for the current period, while the "Forecast" represents the prior belief (one-quarter-ahead forecast) regarding GDP growth for the same target period. Except for the "Count" row, the units in the other rows are expressed as percentages (%), as we focus on the forecasts of the GDP growth rate.

Table 3. Regressions of Forecast Revisions on Asset Returns

Asset 1	Asset 2	Coeff. on Asset 1	Coeff. on Asset 2	R^2
5YR Fixed-term Index Return	Change in BAA-10Y Spread	-0.096**	-0.022**	0.258
CRSP Value-weighted Return	5YR Fixed-term Index Return	0.021**	-0.141***	0.225
Market Return	5YR Fixed-term Index	0.021**	-0.137***	0.224
5YR Fixed-term Index	Change in AAA-10Y Spread	-0.102**	-0.011	0.221
Change in AAA-10Y Spread	Change in BAA-10Y Spread	-0.005	-0.031***	0.220
Change in BAA-10Y Spread	Change in the foreign exchange value	-0.037***	0.003	0.218
5YR Fixed-term Index	Change in the foreign exchange value	-0.143***	-0.028	0.210
5YR Fixed-term Index	Change in VIX index	-0.126***	-0.003	0.201
Change in BAA-10Y Spread	Change in VIX index	-0.032***	-0.002	0.197
Market Return	Change in BAA-10Y Spread	0.001	-0.034***	0.191
CRSP Value-weighted Return	Change in BAA-10Y Spread	-0.001	-0.035***	0.191
Change in AAA-10Y Spread	Change in the foreign exchange value	-0.023***	-0.014	0.177
Change in AAA-10Y Spread	Change in Change in VIX index index	-0.020***	-0.003	0.170
Market Return	Change in AAA-10Y Spread	0.011	-0.020***	0.164

The table reports the results from time series regressions of forecast revisions on asset returns from 1991Q1 to 2023Q4 (Due to space limitations, only the top-ranked asset portfolios of R-squared are reported). Forecast revisions defined as the difference between the mean forecast of quarter t made at quarter t (nowcast) and quarter $t-1$ (past forecast) and the average is computed after winsorizing the data at 5% level. We use the sum of asset returns between the consecutive release dates for SPF forecasts corresponding to the revisions. Standard errors are calculated using the Newey-West method with an automatic bandwidth selection procedure as described by Newey and West (1994). * p<0.10, ** p<0.05, *** p<0.01.

Table 4. Regressions of equation (12) using data from the first window period

	(1)	(2)
β_1 (5YR Fixed-term Index)	-0.235***	-0.362***
	(0.074)	(0.128)
β_2 (Change in AAA-10Y Spread)	0.024*	0.035*
	(0.013)	(0.019)
b (constant)		0.609
		(0.636)
k (past forecast $\mathbb{F}_0[x^p]$)		-0.086
		(0.253)
R^2	0.224	0.280

This table reports the estimated results of equation (12) using data from the first window period (1991Q1-2000Q4). We choose the combination of 5YR Fixed-term Index and Change in AAA-10Y Spread because they provide the highest R-squared (or the lowest SSR) in the first window period. In unreported set of results, we use data from other windows and other asset returns for testing, and the results are basically consistent with those reported in this table. Changing the length of the rolling window period from 40 quarters to other reasonable lengths will not affect the robustness of the results. The R-squared of the regression depends on the sample period and asset returns we choose, but in most cases, an appropriate pair can provide a sizeable R-squared. Standard errors are calculated using the Newey-West method with an automatic bandwidth selection procedure as described by Newey and West (1994). * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 5. Regressions of Different Types of Revisions on Asset Returns

	Expectation	Expectation High	Expectation Low	Disagreement
Panel A				
β_1 (5YR Fixed-term Index)	-0.235	-0.233***	-0.238***	0.003
	(0.074)	(0.073)	(0.079)	(0.019)
β_2 (Change in AAA-10Y Spread)	0.024	0.022	0.026*	-0.002
	(0.013)	(0.013)	(0.014)	(0.003)
R^2	0.224	0.226	0.203	0.014
Panel B				
β_1 (Change in AAA-10Y Spread)	-0.007	-0.011	-0.003	-0.004**
	(0.008)	(0.009)	(0.007)	(0.002)
β_2 (HML)	-0.010	0.002	-0.023	0.013***
	(0.023)	(0.023)	(0.024)	(0.004)
R^2	0.022	0.025	0.031	0.152

This table presents the regression results for consensus forecast, representative forecasters and disagreement using data from the first window period (1991Q1-2000Q4 Panel A provides the estimated results for the first bivariate asset pair (5-year fixed-term index and change in AAA-10Y spread) that generates a significant R^2 for the revisions of consensus and representative forecasts. Panel B provides the estimated results for another bivariate asset pair (change in AAA-10Y spread and factor HML) that explains a large degree of revisions in disagreement. Standard errors are calculated using the Newey-West method with an automatic bandwidth selection procedure as described by Newey and West (1994). * p<0.10, ** p<0.05, *** p<0.01.

D Figures

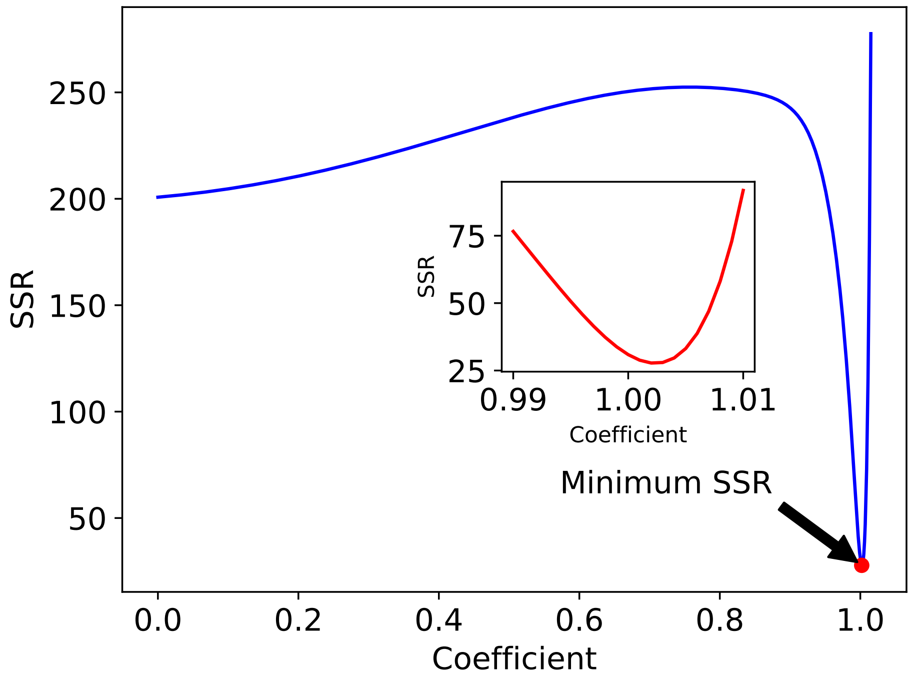

Figure 1. Relationship between the sum of squared residuals (SSR) and the coefficient α on past forecasts

This figure illustrates the relationship between the sum of squared residuals (SSR) and the coefficient α on past forecasts across different values of the coefficient using data from the first window period. As the coefficient increases from 0 to approximately 0.8, the SSR displays a mild upward trend, indicating a gradual increase in the discrepancy between the predicted and observed values. At approximately 0.8, the SSR begins to sharply decrease, reaching its minimum at a coefficient value of 1.0. This point represents the optimal coefficient value where the fit of the model is maximized and the error is minimized. As the coefficient exceeds 1.0, the SSR escalates dramatically, as depicted by the steep upward trajectory on the right side of the figure. This surge in SSR suggests that coefficients greater than 1.0 lead to a significant overfitting of the model, where the model predictions deviate increasingly from the actual data. The inset graph provides a magnified view of the SSR near the optimal coefficient, showing a clear convex shape and pinpointing the minimum SSR, thus suggesting the most efficient point for the model fit within the specific range.

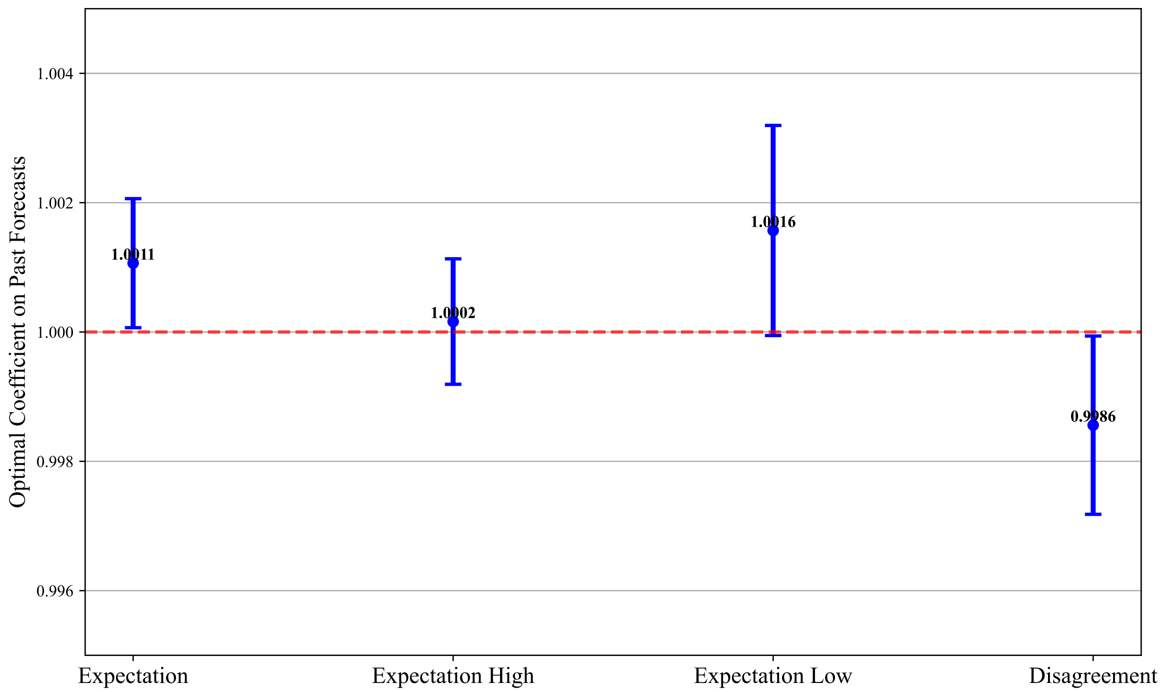

Figure 2. Optimal coefficient α on past forecasts for each rolling windows

This figure presents the estimated optimal α coefficients and their standard deviations across different estimation windows. The dots represent the mean of the optimal coefficients calculated using data from different window periods. Despite minor differences, the optimal coefficient α values for consensus forecast, representative forecasts, and disagreement are all very close to 1.0 throughout the entire sample period.

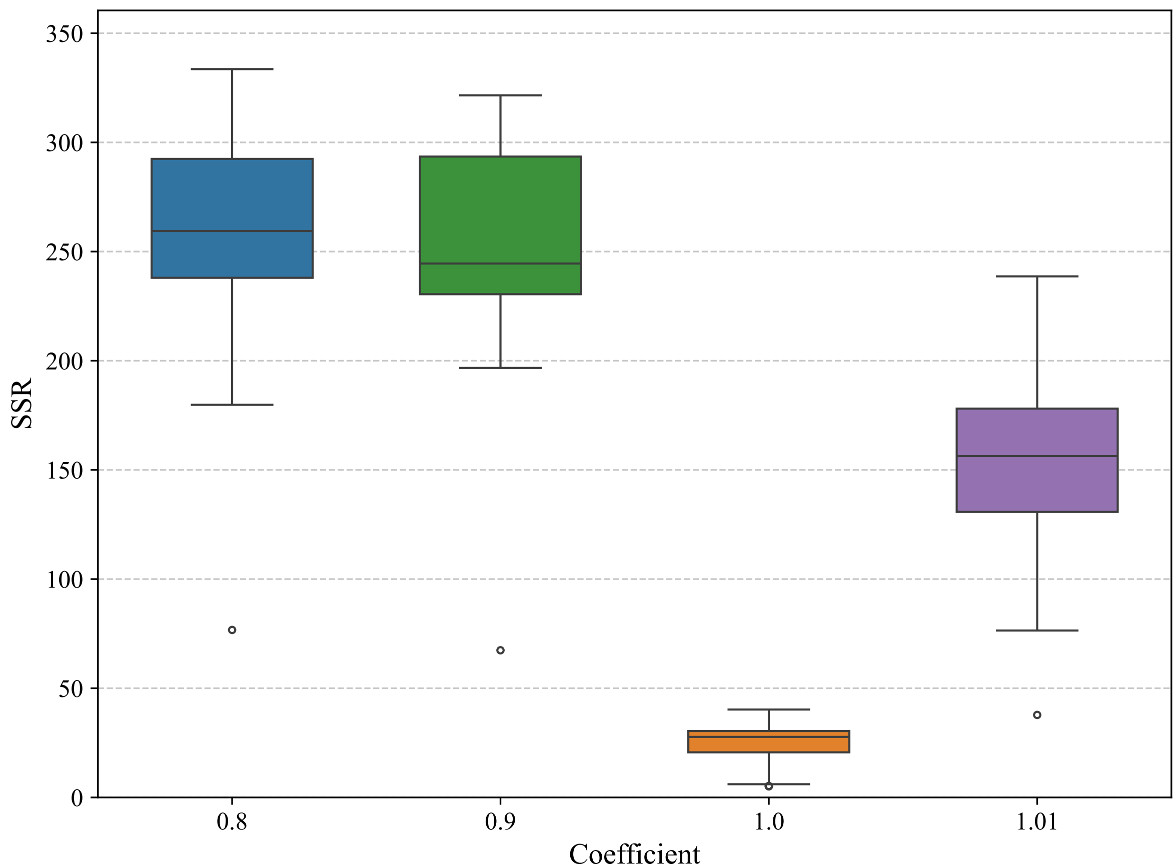

Figure 3. Sum of squared residuals in each window period under different given alphas

This figure illustrates the relationship between the coefficient α on past forecasts and the sum of squared residuals (SSR) across various window periods. The box plots represents the distribution of SSR given different values of α and clearly demonstrate that the SSR reaches its minimum when coefficient α equals 1.0 across the entire sample periods. Since the COVID-19 pandemic has significantly increased prediction errors, we only use data before 2020. However, this does not change the fact that the optimal coefficient still equals one.

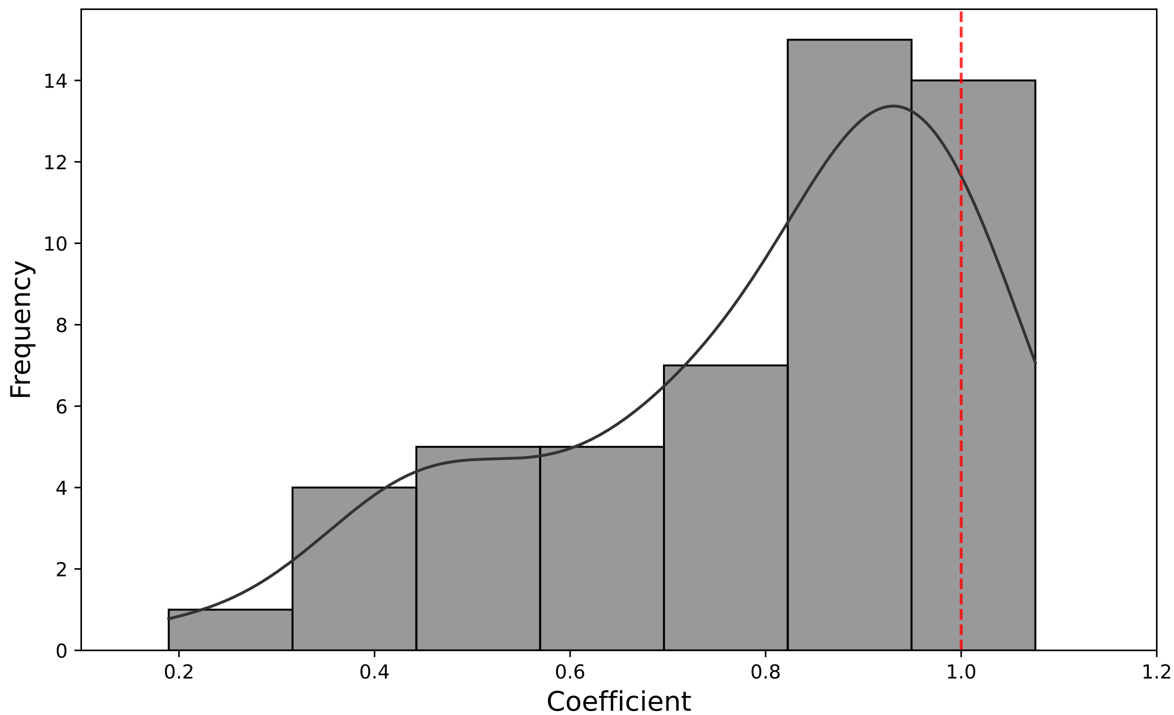

Figure 4. Distribution of Optimal Coefficients α on Past Forecasts at the Individual Level

This figure displays the frequency distribution of the optimal coefficients at the individual level. To ensure that the optimal coefficients can be correctly estimated, we select individuals who participated in the SPF more than 30 times from the first quarter of 1991 to the fourth quarter of 2023 (a total of 51 individuals).

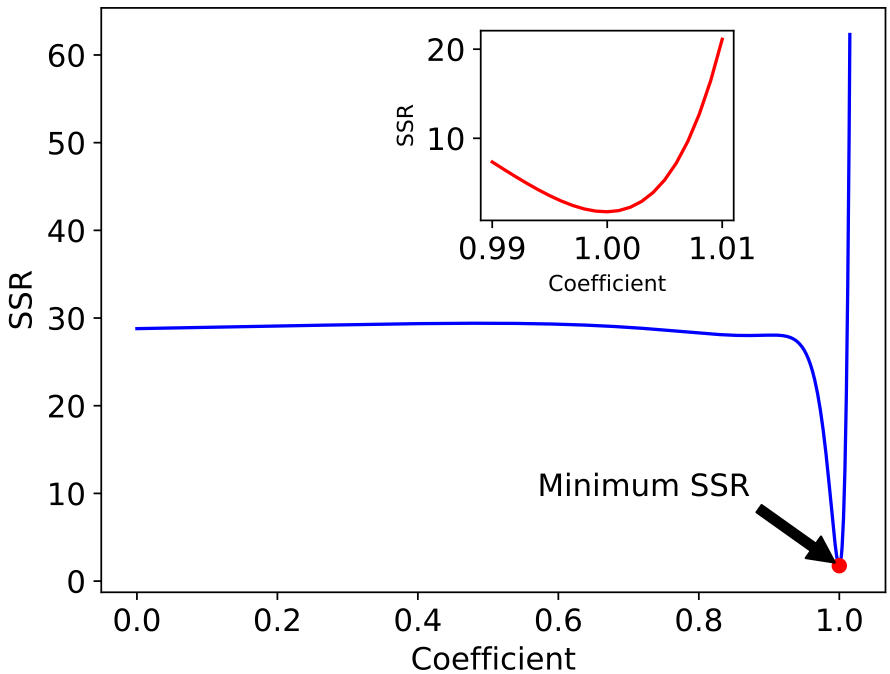

Figure 5. Relationship between the sum of squared residuals (SSR) and the coefficient α on past disagreement

This figure plots the relationship between the sum of squared residuals (SSR) and the coefficient α on past disagreement for the first window period. The variation shows a similar pattern observed in Figure 1, suggesting the optimal coefficient for the evolution equation of disagreement is also one.

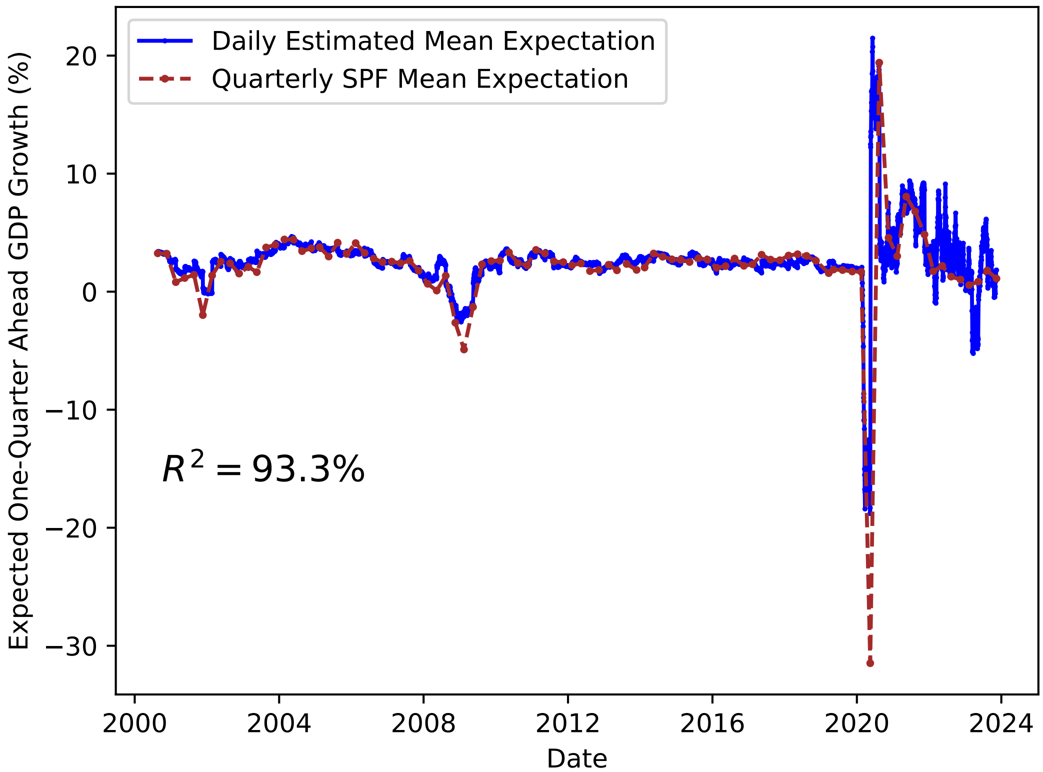

Figure 6. Estimated Daily Series and True Quarterly Series (Expectations)

The figure plots the daily cross-sectional mean series estimated using our mixed-frequency approach and the true quarterly SPF cross-sectional mean series. The daily series is constructed from out-of-sample estimates.

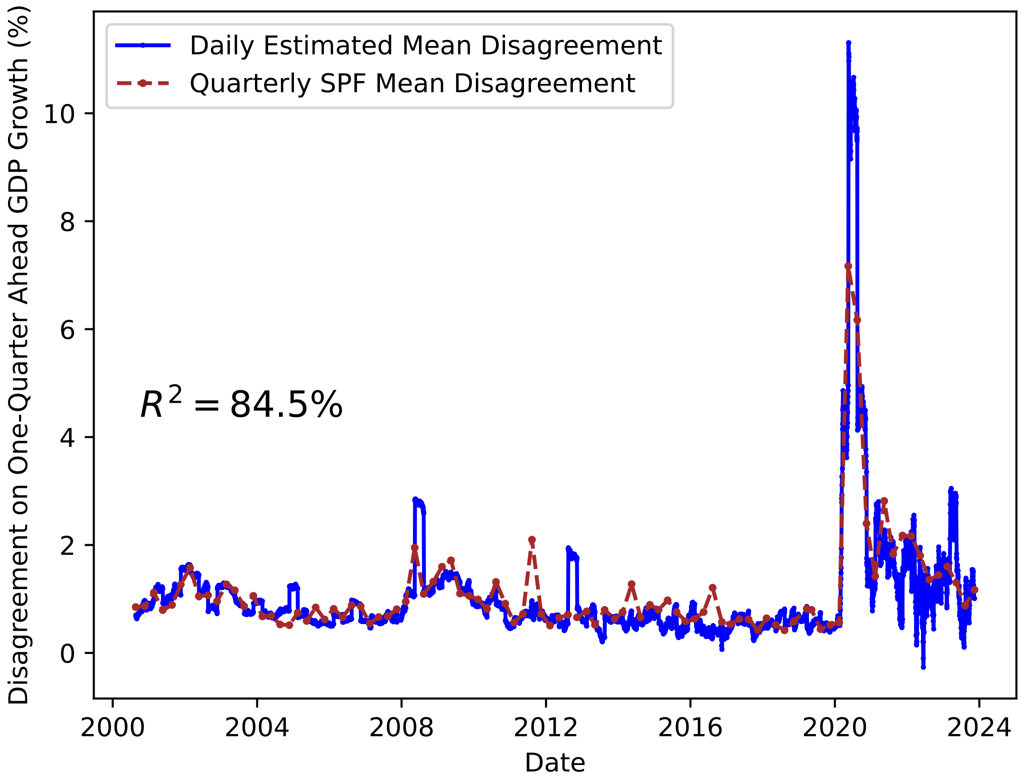

Figure 7. Estimated Daily Series and True Quarterly Series (Disagreement)

The figure plots the daily cross-sectional standard deviation series estimated using our mixed-frequency approach and the true quarterly SPF cross-sectional standard deviation series. The daily series is constructed from out-of-sample estimates.