STAFF REPORTS

NO. 1150
APRIL 2025

How Do We Learn About the Long Run?

Richard K. Crump | Stefano Eusepi | Emanuel Moench
Bruce Preston

FEDERAL RESERVE BANK *of* NEW YORK

How Do We Learn About the Long Run?

Richard K. Crump, Stefano Eusepi, Emanuel Moench, and Bruce Preston

Federal Reserve Bank of New York Staff Reports, no. 1150
April 2025
https://doi.org/10.59576/sr.1150

Abstract

Using a novel and unique panel dataset of individual-level professional forecasts at short, medium, and very-long horizons, we provide new stylized facts about survey forecasts. We present direct evidence that forecasters use multivariate models in an environment with imperfect information about the current state, leading to heterogenous non-stationary expectations about the long run. We show forecast revisions are consistent with the predictions of a multivariate unobserved trend and cycle model. Our results suggest models of expectations formation which are either univariate, stationary, or both, are inherently misspecified and that macroeconomic modelling should reconsider the conventional assumption that agents operate in a well-understood stationary environment.

JEL classification: D83, D84
Key words: expectations formation, shifting endpoint models, imperfect information, survey forecasts

---

Crump: Federal Reserve Bank of New York (email: richard.crump@ny.frb.org). Eusepi: Brown University (email: stefano_eusepi@brown.edu). Moench: Frankfurt School of Finance and Management, CEPR (email: e.moench@fs.de). Preston: University of New South Wales (email: bruce.preston@unsw.edu.au). The authors would like to thank Katerina Petrova for helpful comments and discussions, and participants at the 9th Continuing Education in Macroeconometrics conference. Oliver Kim, Nick Ritter, and Ignacio Lopez Gaffney provided excellent research assistance. Eusepi and Preston thank the Australian Research Council for financial support under the grant DP210103427.

1 Introduction

Muth (1961) argued for a more sophisticated treatment of expectations than naive or adaptive expectations. He advanced the idea that firm expectations should be understood as being determined by the relevant economic theory characterizing the environment inhabited by decision makers. In the rational expectations revolution of the 1970s, this idea came to be interpreted as model-consistent expectations. But Muth warned that his proposal should not be interpreted as a theory of what firms should actually do and acknowledged that measured industry expectations were clearly inconsistent with the predictions of theory, displaying under reaction to contemporary economic developments. Model-consistent expectations needed to be twisted.

This paper adduces direct evidence that the formation of individual-level expectations are indeed sophisticated, with long-term expectations about various macroeconomic time series depending on short-run developments in the variable itself, and of real output, unemployment, inflation and nominal interest rates. But despite the sophistication of multivariate forecasting models, long-term expectations are unlikely to be consistent with full-information about a stationary economic environment, features typical of contemporary macroeconomic models. Long-term expectations exhibit frequent revision and variation at the individual level. We show that these properties are intrinsic to multivariate unobserved trend and cycle models and argue that such models of imperfect information represent a promising twist to full-information model-consistent expectations.

To arrive at these conclusions, we leverage a new and unique panel dataset from Blue Chip Economic Indicators. The dataset comprises *individual* professional forecasts spanning short-, medium- and very-long horizons. Covering the period from 1998 to 2016, the dataset includes 16 US macroeconomic variables with forecast horizons ranging from one quarter to six to eleven years ahead. Tracking the evolution of forecasts across a wide range of macro-variables for each forecaster permits us to investigate the formation of expectations under imperfect knowledge about the long-run behavior of the economy. The richness of this dataset provides new insights into the dynamics of expectation formation and its implications for economic outcomes.

We establish three new stylized facts:

F1. The cross-sectional rankings of forecasters based on their short- and long-term forecasts are only weakly correlated. For instance, the most optimistic forecaster for next quarter's GDP growth may not necessarily be the most optimistic about long-run growth. In contrast, the ranking of forecasters based on longer-horizon forecasts is strongly correlated, with their ranking remaining largely unchanged.

F2. Long-term and short-term forecast revisions systematically co-vary, with long-run forecasts exhibiting substantial time-variability and frequent revision.

F3. Agents forecast macroeconomic variables as *joint stochastic processes*. Revisions in long-term forecasts of most variables depend on revisions to short-run expectations of that variable, and also measures of real output, inflation, nominal interest rates and unemployment.

We show that multivariate unobserved component forecasting models can explain these stylized facts. Forecasters assume that the data are determined by two components, a permanent "trend" and a transitory "cycle". There are two sources of forecaster-specific heterogeneity: i) different model parameters; and ii) different noisy signals of the true underlying data. These minimal assumptions on the nature of forecaster heterogeneity and imperfect information, which nest existing models such as Patton and Timmermann (2010) and Andrade, Crump, Eusepi, and Moench (2016), are sufficient to generate heterogenous time-varying long-run expectations. Importantly, we make no assumption on the degree of rationality of the forecasters nor do we assume that forecasters have the correct law of motion for the economy. This is an important accommodation since the imposition of rational expectations is most natural in stationary economic environments (Lucas, 1986).

The model provides testable predictions, which we exploit to establish the stylized facts.

i. Because forecasts are determined by a stationary and permanent component, the stationary component becomes increasingly negligible as the forecast horizon lengthens. As a result, the rank correlation of forecasters across forecast horizons, from most optimistic to most pessimistic, is stronger for progressively longer-horizon forecast pairs, which are almost entirely determined by the persistent component. This explains Fact 1;

ii. Because forecasters have imperfect information about the current state, short-term forecast errors lead to revisions in the estimates of the permanent and transitory components by amounts which are determined by the Kalman gain. This underscores individual long-term forecasts respond to new information and are not pinned down to some exogenous long-run fundamental. These revisions lead to updates in the entire term structure of expectations, so that revisions in short-run and long-run expectations are correlated. We estimate a significant correlation. This explains Fact 2 and;

iii. Because forecasters use multivariate models with imperfect information, revisions to long-run forecasts must be correlated with revisions to short-run forecasts of all variables. The data reject the null hypothesis that long-term forecast revisions do not co-move with revisions to short-run forecasts of GDP, inflation, unemployment and short-term interest rate forecasts for most macroeconomic variables. This is strong evidence against the hypothesis that forecasters use univariate models, because revisions in long-term forecasts for a given variable would only co-move with short-run forecast revisions of that individual variable. This explains Fact 3.

These findings are consistent with other survey evidence demonstrating that long-term expectations of economic and financial variables vary over time. For example, the Survey of Professional Forecasters annually surveys participants on their estimates of the non-accelerating inflation rate of unemployment, while the Federal Reserve Bank of New York's Survey of Primary Dealers includes questions on "longer-run" values of economic variables such as output, inflation, and the target interest rate. Additionally, policy-makers' views on the long-term evolution of the economy can shift over time, as evidenced by the FOMC Survey of Economic Projections. Such variations in long-term expectations highlight the dynamic nature of economic decision-making and the need for macroeconomic models that can explain their behavior. However, these surveys provide only summary statistics about the distribution of individual long-term expectations in contrast to our rich dataset of individual expectations.

The results have important implications for how we should model long-term expectations formation and raise questions about what is the "relevant economic theory". Our evidence strongly suggests that models of expectations formation which are either stationary, univariate or both are fundamentally misspecified. In particular, univariate stationary models are missing key mechanisms regarding the behavior of expectations. This signals caution is required when interpreting regression analyses that endeavor to identify particular types of information frictions (Coibion and Gorodnichenko (2015) and Bordalo, Gennaioli, Ma, and Shleifer (2020)). That forecasting models are multivariate in nature suggests that more structural approaches are required to identify dynamic properties of expectations. Angeletos, Huo, and Sastry (2020) provides an example of the required approach in the case of the recent literature on under- and over-reaction of forecasts to new information.

Modern rational expectations economic theory generally assumes households and firms make decisions in well-understood stationary environments, with full information about the state of the economy and, when forming expectations, process their information optimally. Motivated by empirical evidence challenging the rational expectations assumption, a growing literature introduces various forms of information frictions and bounded rationality.[1] But regardless of the specific departure from full-information rational expectations, these analyses ultimately give focus to short-run dynamic implications, as they too assume people make decisions in a stable environment, with long-run expectations anchored to a well-understood fundamental equilibrium. Our results call for models in which information frictions matter for both short- and long-run expectations.

As a practical matter, households, firms and policymakers are likely to have imperfect knowledge about the long run when making decisions. Uncommonly large and persistent shocks, non-recurring shifts in policy regimes, prolonged swings in productivity growth, financial innovation, and other forms of structural change all cloud the long-term outlook of the economy. Incorporating such sources of uncertainty, not only account for the stylized facts of this paper, but can also have first-order implications for our understanding of the macroeconomy. Relaxing the cross-equation restrictions of rational expectations equilibrium analysis and allowing long-run beliefs to depart from steady state, permits a class of model with considerably richer dynamics than rational expectations. Not only does this promise progress on quantitative questions that are considered puzzles through the lens of a rational expectations equilibrium analysis, but it facilitates addressing certain questions that such models either cannot address, or are poorly suited to address.

For example, learning about the long run reduces the need for ad hoc sources of persistence—such as habit formation and price indexation—and provide a source of amplification of shocks (e.g.,

[1] This includes papers on level-k thinking, myopia and cognitive discounting, noisy information and rational inattention.

Milani (2007), Eusepi and Preston (2011)). In asset pricing models, learning re-weights income and substitution effects to resolve a range of puzzles, including return predictability and excess volatility (e.g., Sinha (2016), Adam, Marcet, and Beutel (2017), Eusepi and Preston (2018), Farmer, Nakamura, and Steinsson (2024), and Crump, Eusepi, and Moench (2024)). In analyses of monetary policy, departures from rational expectations permit developing coherent models of central bank communication and the anchoring of inflation expectations (e.g., Eusepi and Preston (2010) and Carvalho, Eusepi, Moench, and Preston (2023)). Finally, instability in long-run expectations can make inflation control more difficult in both normal times and crisis periods, when the policy is constrained by the zero lower bound on nominal interest rates, and when public debt is high (e.g., Eusepi, Giannoni, and Preston (2024), Eusepi, Gibbs, and Preston (2025) and Eusepi and Preston (2018)).

2 Relation to the Literature

The vast majority of earlier empirical work on survey measures of expectations focuses on *short-term* economic developments—see, for example, Angeletos, Huo, and Sastry (2020). This choice is partly driven by what data are available, as there is substantially less information on long-horizon forecasts. But it also reflects the common assumption in macroeconomic models that economic agents operate in a stationary environment and, consequently, that they can quickly and efficiently come to understand the long-run behavior of the economy. Any information frictions that might be relevant to the expectation formation process, are only relevant to short-run economic dynamics.

Our paper contributes to a growing literature emphasizing that professional forecasters act as if data have short-run and long-run components that must be disentangled (Kozicki and Tinsley (2001), Andrade, Crump, Eusepi, and Moench (2016), Crump, Eusepi, Moench, and Preston (2023), Crump, Eusepi, and Moench (2022) and Farmer, Nakamura, and Steinsson (2024)). Trend-cycle decompositions have a long tradition in theoretical and empirical macroeconomic research. For instance, the seminal real-business cycle model in Kydland and Prescott (1982) assumes agents cannot perfectly observe the short- and long-term components of technical progress. Stock and Watson (1989) and Stock and Watson (2007) model inflation as having a trend and a transitory component. This approach has also been incorporated in countless structural models of inflation dynamics of which Cogley, Primiceri, and Sargent (2010) is a prominent example. Various studies apply trend-cycle decompositions to other macroeconomics variables, showing that models which embed slow-moving time-varying drifts capture well the dynamics properties of real GDP growth (Stock and Watson (1989), Cogley and Sargent (2005) and Laubach and Williams (2003)) and the federal funds rate (Kozicki and Tinsley (2001) and Gürkaynak, Sack, and Swanson (2005)). Consensus expectations from professional forecasters have been used to explore the link between forecast revisions and forecast errors using pass-through regressions of either macroeconomic news or movements in short-term expectations to long-term expectations, see for example Gürkaynak, Levin, and Swanson (2010) and Beechey, Johannsen, and Levin (2011). In addition, Bems, Caselli, Grigoli, and Gruss (2021) provide time-series evidence of this link for a large set of countries. To our knowledge, we are the first to use detailed *individual-level* data including on long-horizon forecasts and all of our stylized facts are new to the literature. This permits adducing direct supportive evidence on unobserved component models.[2]

Although disagreement among forecasters has been studied by some researchers (e.g., Lahiri and Sheng (2008), Patton and Timmermann (2010), Andrade, Crump, Eusepi, and Moench (2016), Cao, Crump, Eusepi, and Moench (2021)), there has been little work on disagreement at long-horizon forecasts. And the work that has been done, such as Andrade, Crump, Eusepi, and Moench (2016), relies on summary statistics as individual-level forecasts were not previously available.[3] However, summary statistics could mask the degree of disagreement (e.g., if forecasts were bimodal).

We argue that a multivariate trend-cycle decomposition model should be the benchmark model of expectations formation. In the recent literature there are relatively few examples of this approach. In models of professional forecast behavior they have been used by Andrade, Crump, Eusepi, and Moench (2016), Crump, Eusepi, and Moench (2022), and Crump, Eusepi, Moench, and Preston (2023)). In structural models used for policy evaluation, Sargent, Williams, and Zha (2006) emphasize multivariate forecasts, and Eusepi, Giannoni, and Preston (2024) multivariate long-horizon forecasts.

3 A Model of Expectations Formation

We assume that forecasters use an unobserved component model to make projections of future macroeconomic variables. The model incorporates three features that are essential to capture the behavior of observed expectations. First, agents have imperfect knowledge about the long-run behavior of the economy: one of the unobserved components captures slow-moving long-run trends. Second, forecasters have heterogenous sources of information and therefore update their beliefs responding to both public and private sources of information. While we consider a restricted model environment where agents share the same perceived law of motion for the economy, we discuss the implications of some degree of heterogeneity along this dimension as well. Third, agents employ multivariate forecasting models to capture economic inter-dependencies amongst macroeconomic variables. The realization that market participants have complex forecasting models goes back to the seminal work of Muth:

> Averages of expectations in an industry are more accurate than naive models and as accurate as elaborate equation systems, although there are considerable cross-sectional differences of opinion. (Muth, 1961)

[2]Notable exceptions focus only on individual long-run inflation expectations—see, for example, Weber, D'Acunto, Gorodnichenko, and Coibion (2022).

[3]For example, in the case of the BCEI and the Blue Chip Financial Forecasts (BCFF) only the average of the top-10 and bottom-10 forecasts are reported. We should note there are two exceptions. First, the Survey of Professional Forecasters, since the third quarter of 2005, provides individual forecasts for the five-year horizon beginning in five years (e.g., Binder, Janson, and Verbrugge (2019)). Second, Cao, Crump, Eusepi, and Moench (2021) also focused on a single variable (3-month T-Bill) but used the same data set as we use in this paper.

Because no assumption is made about the true data-generating process, our modelling makes no assumption about the rationality of forecasters.

To fix ideas, we start with a univariate forecasting model and then move to the more realistic case in which forecasters employ a multivariate model to make predictions.

3.1 A Simple Univariate Model

Consider a set of N forecasters predicting the future values of the variable y_t. We make no assumptions on the true data-generating process of y_t. Forecasters do not fully observe y_t, with each agent $i = 1, \ldots, N$ receiving a noisy signal

$$s_t^{(i)} = y_t + \eta_t^{(i)}, \tag{1}$$

where the forecaster-specific innovations $\eta_t^{(i)}$ are Gaussian processes with mean zero and variance $\sigma^2_{i,\eta}$. These shocks are $i.i.d$ across both time and forecasters. Imperfect observability of the variable y_t can be justified on the ground that agents do not have real-time measures of aggregate variables such as GDP and therefore form estimates using a wealth of information, including monthly indicators of real economic activity. In addition, published macroeconomic data can in many cases be viewed as noisy indicators of the underlying fundamental economic variables. For example, various volatile measures of price inflation are often combined to provide better forecasts of inflation.

Individual agents face an additional source of uncertainty. Their forecasting model for the evolution of y_t is

$$y_t = \omega_t + x_t + \sigma_\nu \varepsilon_t^\nu \tag{2}$$

where

$$\omega_t = \omega_{t-1} + \sigma_\omega \varepsilon_t^\omega \tag{3}$$

$$x_t = \phi x_{t-1} + \sigma_x \varepsilon_t^x, \tag{4}$$

and $\varepsilon_t^\nu, \varepsilon_t^\omega$ and ε_t^x are mutually independent, zero mean, and unit variance $i.i.d$ Gaussian processes. Importantly agents cannot observe either the transitory noise, ε_t^ν, the cyclical component, x_t, or the persistent component, ω_t, which jointly determine the evolution of y_t. The persistent component captures "structural change" in the economy, possibly from shifts in long-run productivity tied to the pace of technological innovation; shifts in the implicit or explicit inflation target of the central bank; or shifts in the natural rate of interest, because of either demographics or changes in the global preference for borrowing and saving. As such the persistent component is slower-moving with less volatile innovations than the cycle component so that $0 < \sigma_\omega << \sigma_x$.

We assume that agents share the same forecasting model. However, we allow individual-specific priors about the parameter ϕ regulating the persistence of the cycle component, as well as the volatility of innovations of the three unobserved components. Each forecaster is then associated to a set of parameters: $\theta^{(i)} \equiv (\phi_i, \sigma_{i,\nu}, \sigma_{i,\omega}, \sigma_{i,x}, \sigma_{i,\eta})$. For further discussion of these properties see Lahiri and Sheng (2008), Patton and Timmermann (2010), and Andrade, Crump, Eusepi, and Moench (2016). The model can be extended in two dimensions. First, by allowing agents to receive additional signals about the unobserved components ω_t and x_t. These signals capture information from other indicators of the business cycle, such as survey measures of economic activity and the slope of the yield curve, policy announcements affecting expectations about the economic outlook at different horizons, such as forward guidance about monetary policy, and announced fiscal reforms. Second, following Andrade, Crump, Eusepi, and Moench (2016) we could incorporate some degree of information stickiness so that not all forecasts are updated each period to reflect all currently available information. Such extensions do not affect the model implications discussed below. We therefore omit them without loss of generality.

Given their parameters $\theta^{(i)}$, agents estimate the trend and cycle components, denoted by $\omega^i_{t|t}$ and $x^i_{t|t}$, using the observed $s^{(i)}_t$ and the Kalman filter.[4] This gives rise to the updating equations

$$\omega^i_{t|t} = \omega^i_{t|t-1} + \kappa^i_\omega \left(y_t - \omega^i_{t|t-1} - x^i_{t|t-1} + \eta^{(i)}_t \right) \tag{5}$$

$$x^i_{t|t} = x^i_{t|t-1} + \kappa^i_x \left(y_t - \omega^i_{t|t-1} - x^i_{t|t-1} + \eta^{(i)}_t \right). \tag{6}$$

Forecasters revise their estimates of trend and cycle components in response to their forecast error. In turn the forecast error is a function of realized innovations in both the trend and cycle components of y_t and the agent-specific noise $\eta^{(i)}_t$. The size of the revision depends on the Kalman gains, κ^i_ω and κ^i_x, which reflect priors about the volatility and persistence of both unobserved processes and the observation noise. For example, higher perceived volatility of the noise, $\eta^{(i)}_t$, leads to a lower gain. And higher persistence and lower volatility of the drift ω_t implies a relatively lower gain κ^i_ω compared to κ^i_x.

Two properties of the forecasting model deserve comment. First, the Kalman gains are forecaster-specific given their different priors about model parameters. Second, the Kalman gains are not necessarily optimal, given that we make no assumptions about the true data-generating process of y_t. The estimates of the trend and cycle components might display 'over-reaction' or 'under-reaction' relative to the rational expectations update. This includes the possibility that agents over-extrapolate from recent observations, consistent with various behavioral theories of expectations formation (Angeletos, Huo, and Sastry (2020), Bordalo, Gennaioli, Ma, and Shleifer (2020) and Afrouzi, Kwon, Landier, Ma, and Thesmar (2023)). Furthermore, the forecast errors in (5) and (6) need not be *i.i.d.* across time and can be autocorrelated, reflecting different degrees of model mis-specification (Eusepi and Preston (2011)).

Given the current estimate of trend and cycle components, the entire term structure of forecasts

[4] Agents also obtain current estimates of the noise components but, since forecasts of these variables are zero, we can ignore the role of these objects in this discussion.

for the target variable y_t follow. Forecasts for any horizon h are determined as

$$\mathbb{E}_t^{(i)}\left[y_{t+h}\right] = \omega_{t|t}^i + \phi_i^h x_{t|t}^i. \tag{7}$$

We can now use this simple model to derive testable predictions related to Fact #1 and Fact #2.

(MI #1*a*.) Individual long-term forecasts vary over time. Let $h^* > \bar{h}_\phi$ be sufficiently large such that $\phi_i^{\bar{h}_\phi} \approx 0$. From equation (7) we then have $\mathbb{E}_t^{(i)}\left[y_{t+h^*}\right] \approx \omega_{t|t}^i$ so that long-horizon forecasts are dominated by the random-walk component of the forecasting model. A direct implication is that *individual* long-horizon forecasts will be revised frequently in response to new information. Furthermore, because forecasters have different priors and received signals, their long-term forecasts will, in general, be different.

(MI #1*b*.) The ranking of individual forecasts is approximately invariant at long forecast horizons. Consider any two forecasters, i and j, at time t. Compare their forecasts at two different long-horizons $h^{**} > h^* > \bar{h}$. Given their forecasts, their *relative ranking* will not change in the sense that if

$$\mathbb{E}_t^{(i)}\left[y_{t+h^*}\right] \approx \omega_{t|t}^i > \mathbb{E}_t^{(j)}\left[y_{t+h^*}\right] \approx \omega_{t|t}^j \tag{8}$$

then it must also be that $\mathbb{E}_t^{(i)}\left[y_{t+h^{**}}\right] > \mathbb{E}_t^{(j)}\left[y_{t+h^{**}}\right]$.

Conversely, if we consider the ranking of the same forecasters at a shorter horizon $\tilde{h} < \bar{h}_\phi$, where the business cycle component x_t remains a relevant determinant of forecasts, then the ranking among forecasters can be arbitrary. Formally,

$$\mathbb{E}_t^{(i)}\left[y_{t+\tilde{h}}\right] = \omega_{t|t}^i + \phi_i^{\tilde{h}} x_{t|t}^i \lesseqqgtr \omega_{t|t}^j + \phi_j^{\tilde{h}} x_{t|t}^j = \mathbb{E}_t^{(j)}\left[y_{t+\tilde{h}}\right]. \tag{9}$$

This delivers the testable prediction that the correlation between the ranking of forecasters at short- and long-term horizons is lower than the correlation between the rankings at two longer forecasting horizons. This prediction is consistent with a perceived law of motion for the data which includes two components, one persistent and one transitory. We emphasize that this model implication continues to be true even if forecasters must estimate the parameters of the model as part of their updating.

(MI #2*a*.) Forecast revisions across horizons are correlated. We have assumed that the two factors driving the evolution of the forecast variable are unobserved to the agents and that their estimates are updated according to equations (5) and (6). These equations are difficult to test directly because, as econometricians, we do not observe forecaster-specific signals s_t. We can, however, use the updating equation to generate the prediction that long-horizon and short-horizon forecast revisions will be related according to

$$\mathbb{C}_t\left[\omega^i_{t+1|t+1}-\omega^i_{t+1|t}, y^i_{t+1|t+1}-y^i_{t+1|t}\right]=\kappa^i_\omega\left(\kappa^i_\omega+\kappa^i_x\right)\cdot\mathbb{V}_t\left[y_{t+1}-\omega^i_{t+1|t}-x^i_{t+1|t}+\eta^{(i)}_{t+1}\right]>0, \tag{10}$$

almost surely, where $\mathbb{C}_t(\cdot)$ measures the covariance and $\mathbb{V}_t(\cdot)$ the variance which is nonzero and, in this simple case, strictly positive.

Note that here no assumptions are needed about forecast errors. Under rational expectations the Kalman gain reflects the true data-generating process and the forecast errors are mean zero and serially uncorrelated. Here we assume none of that. The expressions above are consistent with both rational and boundedly rational beliefs updating. What is tested is whether the term structure of forecasts of individual forecasters is consistent with the proposed model of statistical updating: forecasters behave as if they are "econometricians."

3.2 A General Model

The univariate model just introduced is useful but has important counterfactual implications (see Andrade, Crump, Eusepi, and Moench (2016), Crump, Eusepi, and Moench (2022) and the discussion in the sequel). We therefore derive a more general set of predictions in a situation in which agents must forecast more than one variable.[5] To that end, consider a collection of N agents forecasting the future values of a vector of variables, y_t. The perceived law of motion is

$$y^{(i)}_t = y_t + \eta^{(i)}_t, \qquad t=1,\ldots,T \tag{11}$$

$$y_t = H_i' z^{(i)}_t, \tag{12}$$

$$z^{(i)}_t = F_i z^{(i)}_{t-1} + V_i \varepsilon^{(i)}_t, \tag{13}$$

where $\varepsilon^{(i)}_t$ and $\eta^{(i)}_t$ are mutually independent multivariate Gaussian white noise for $i=1,\ldots,N$ where the variance of $\varepsilon^{(i)}_t$ is the identity matrix and the variance of $\eta^{(i)}_t$ is $\Sigma_{\eta,i}$.

Here $z^{(i)}_t=\left(x^{(i)\prime}_t, \tilde{x}^{(i)\prime}_t, \omega^{(i)\prime}_t\right)'$ with $x^{(i)}_t, \omega^{(i)}_t \subseteq \mathbb{R}^q$ and $\tilde{x}^{(i)}_t \in \mathbb{R}^{m_i}$ with F_i of the form

$$F_i = \begin{bmatrix} F_{i,11} & 0_{(m_i+q)\times q} \\ 0_{q\times(m_i+q)} & F_{i,22}\end{bmatrix}, \tag{14}$$

where $\text{rank}(F_{i,22})>1$.[6] The addition of the states $\tilde{x}_t$, whose dimension can vary across forecasters, ensures that the model can accommodate popular multivariate stationary processes such as a VAR(p) or a VARMA$(1,1)$. In words, we allow agents to have different time-series models for the

[5] We could make the model above even more general but at the expense of more notation. In particular, we could allow for forecasters to observe signals about either of the two components as in Crump, Eusepi, and Moench (2024). In such a case, all model implications highlighted above continue to hold.

[6] This choice of F is without loss of generality because if it was not block diagonal then all components would be non-stationary, in general.

cyclical (e.g., VAR(p) vs. VARMA$(1,1)$) and trend components (e.g., imposing the Fisher equation in the long-run). In sum, this structure allows for different, forecaster-specific factors to determine dynamics at both high and low frequencies. Finally, just as we had in the univariate model, each forecaster is associated with the set of parameters $\Theta^{(i)} \equiv (H_i, F_i, V_i, \Sigma_{i,\eta})$.

The key assumption we make is $\lim_{h\to\infty} F^h_{i,11} = 0_{(m_i+q)\times(m_i+q)}$ so that the perceived law of motion features *additively separable* stationary and non-stationary components. Agents receive a noisy signal, $y^{(i)}_t$, about y_t which they must further disentangle between permanent and transitory sub-components.[7]

We can now present the model implications for Fact #3:

(MI #3*a*.) Forecasts are updated jointly across multiple variables. Estimates of the current states are determined recursively using the updating equation

$$z^{(i)}_{t|t} = z^{(i)}_{t|t-1} + \mathsf{K}_i(y_t - H_i' z^{(i)}_{t|t-1} + \eta^{(i)}_t), \tag{15}$$

which implies the term structure of expectations at every horizon h

$$z^{(i)}_{t+h|t} = z^{(i)}_{t+h|t-1} + F^h_i \mathsf{K}_i(y_t - H_i' z^{(i)}_{t|t-1} + \eta^{(i)}_t), \tag{16}$$

where K_i is the steady state Kalman matrix.[8] More generally the forecast for any individual variable r at horizon $h > 0$ is computed as

$$y^{(i)}_{t+h,r|t} = y^{(i)}_{t+h,r|t-1} + \iota_r' H_i' F^h_i \mathsf{K}_i(y_t - H_i' z^{(i)}_{t|t-1} + \eta^{(i)}_t) \tag{17}$$

where ι_r is a $(2q+m)\times 1$ vector with rth element equal to one and zeros elsewhere. Importantly, equations (16) and (17) imply that revisions to the estimated states and forecasts for each variable are *linked to forecast errors of all other variables.*

Next, consider a sufficiently long forecast horizon $h^* > \bar{h}_F$, such that $F^h \approx \bar{F}$, where $\bar{F}$ is a $(2q+m)\times(2q+m)$ matrix with lower right elements equal to $F^h_{i,22}$ and zeros elsewhere. Equation (16) implies

$$\lim_{h\to\infty}\left(y^{(i)}_{t+h,r|t} - y^{(i)}_{t+h,r|t-1}\right) = \iota_r' H_i' \bar{F}_i \mathsf{K}_i(y_t - H_i' z^{(i)}_{t|t-1} + \eta^{(i)}_t). \tag{18}$$

Long-run forecast revisions for the rth variable are therefore linked not only to forecast errors and idiosyncratic noise of the rth variable, but also to all other variables in the vector of observables y_t. The multivariate model delivers the additional testable prediction that forecast revisions in the long-term forecast of each variable should be related to multiple forecast errors. In contrast to the

[7] This general framework nests a number of popular specifications in the literature such as the multivariate VAR(1) plus random walk which was utilized in, for example, Andrade, Crump, Eusepi, and Moench (2016), Crump, Eusepi, and Moench (2022), Del Negro, Giannone, Giannoni, and Tambalotti (2018), and Crump, Eusepi, Moench, and Preston (2023)

[8] The Kalman matrix corresponds to: $\mathsf{K} = PH\left(H'PH + \Sigma_\eta\right)^{-1}$, with $P = FPF' - F\mathsf{K}H'PF' + VV'$

univariate case in (10), there is no restriction on the sign of the correlation other than it should be nonzero for some variables considered. Model predictions about the ranking of forecasters remain unchanged in the more general setting.

MI (#3*b*.) Forecasters always disagree, even about variables they can perfectly observe. In a standard univariate model, disagreement is generated by imperfect information about the current state of the economy (e.g., measures of inflation or economic activity). However, financial variables such as the yield of the 3-month Treasury bill are clearly observable in contrast to, for example, the current state of the economy. A key feature of the multivariate model above is that it allows for a subset of variables to be observed without error, but will still *always* produce disagreement about future outcomes, regardless of each forecasters' model parameters (e.g., even if they are identical across forecasters).

4 Testing Model Implications

4.1 Data

The model presented above provides a general framework to think about the evolution of agent forecasts at different horizons. With extant data sources we can not test the model implications summarized in the previous section. In particular, because of data limitations there is scant evidence on *individuals*' term structure of expectations, especially in the context of longer-horizon forecasts. Surveys that do ask respondents about their beliefs far in the future, aggregate the individual survey responses and provide only limited summary statistics.[9]

Here we exploit a unique data set of individual medium- and long-horizon forecasts from the Blue Chip Economic Indicators (BCEI) survey, which provides the individual responses to the "Long-Range Consensus U.S. Economic Projections" summarized in the March and October issues. Our data cover the sample period from 1998 to 2016 and all of the variables queried by the BCEI. In addition, we also construct forecasts of the real 3-month Treasury bill by subtracting individual CPI inflation forecasts from the corresponding nominal 3-month Treasury bill forecasts. From October 1998 to March 2006, we can link all individual forecasts to the associated short-run forecasts. However, outside this data range, we can only link forecasters across variables and horizon, not across survey dates. Nonetheless, the richness of these data allow us to introduce a number of new stylized facts about individual long-horizon forecasts. To our knowledge, ours is the first empirical study using the cross-section of individual professional forecasts across all horizons including long horizons.

The BCEI Survey is a survey of professional forecasters that has been in publication since 1976.

[9] For example, both the Blue Chip Financial Forecasts and the BCEI present the cross-sectional average along with the average of the top-10 forecasters and bottom-10 forecasters at each forecast horizon. As another example, the Survey of Primary Dealers (SPD) provides the cross-sectional median along with the 25th and 75th percentiles. To our knowledge, the sole exception is that the Survey of Professional Forecasters provides individual five-year/five-year inflation forecasts but only since 2005 and for only this variable.

The survey is typically released on the 10th of each month, and is based on approximately 50 responses that have been collected during the first week of the same month. The survey focuses primarily on economic variables such as those in the NIPA tables, but also includes forecasts for the unemployment rate, total industrial production, housing starts, and vehicle sales along with forecasts for the 3-month Treasury bill and the 10-year note. The participants of the survey range from large commercial banks, broker dealers, insurance companies, large manufacturers, economic consulting firms, GSEs and others. Beginning in March 1979, BCEI began querying respondents on their forecasts for a selection of variables over the following five years. Later that year, these special questions addressed longer horizons including 6-to-11 years ahead. These biannual questions have generally been conducted in the March and October surveys thereafter.

4.2 Properties of the Cross-Section and Time Series of Survey Forecasts

Model Implication #1(a): Most existing evidence about time-variation in longer-horizon forecasts has focused on movements in consensus expectations. However, consensus expectations can mask important features of the underlying data. For example, consensus expectations may move because only a small number of forecasters meaningfully change their forecasts. Similarly, consensus forecasts may change only modestly over time but mask substantial time variation in individual forecasts that offset each other. We can directly investigate the time variation in individual forecasts.

Figure 1 presents our panel data of long-horizon forecasts for six key variables from the BCEI data.[10] We see that disagreement in individual long-horizon forecasts is ubiquitous across all variables and does not reflect bi-modality or other simple cross-sectional patterns. Since existing papers only used summary statistics it is possible that the degree of disagreement was overstated but with our unique data we are able to show this is not the case. Instead, disagreement is always and everywhere a feature of longer-horizon forecasts. On its face, this directly contradicts the assumption of full information about the long-run behavior of the economy.

Table 2 presents summary statistics across forecasters measuring the degree of time variation in their long-horizon forecasts. For each variable, we calculate the relative standard deviation, that is, the standard deviation over the absolute value of the average forecast as a standardized measure of variability over time. We then report the maximum, minimum, 75th, 50th, and 25th percentile of this distribution for each variable. Table 2 shows that almost all forecasters change their longer-horizon forecasts over time. For real GDP growth, for instance, the standard deviation of the median long-horizon forecast will be approximately 10% of its unconditional mean. This shows that the data are consistent with the model implication, MI #1a.

[10]To conserve space, we restrict our graphs in the main text to six primary variables: real GDP growth, the unemployment rate, the real 3-month T-bill, CPI inflation, real personal consumption expenditures, and real nonresidential fixed investment. However, in the Supplemental Appendix we report results for the remaining 10 variables. All results in the paper hold for all 16 variables.

Model Implication #1(b): Because of the property of additive separability between the non-stationary and stationary components in the law of motion, as the forecast horizon grows, the influence of the stationary component subsides. If we order the forecasts across the N forecasters for horizon h, and fixed t, then the ordering for forecast horizon $h+1$ will be increasingly similar as h grows. Said differently, the cross-sectional correlation goes to one.

A key advantage of our panel data set is that we are able to link individual forecasts across multiple horizons at each point in time. For the full-sample period we can link respondents' forecasts of 2-year, 3-year, 4-year, 5-year, 6-year, and 7-11 year ahead horizons. For the subsample of 1998–2006 we can additionally link the current (0-year ahead) and 1-year ahead forecasts. Figure 2 shows the cross-sectional relationship between forecasts at different horizons and long-horizon forecasts. Specifically, we calculate the cross-sectional rank correlation between forecasts across individual forecasters at each point in time. We then take the time series average over the sample to produce a term structure of rank correlations. Across all variables, there is a clear upward shape of the term structure with rank orderings becoming positive and then moving toward one at the longest horizon. Figure A.2 in the Supplemental Appendix shows that this property also holds for the other 10 variables.

The monotonic relation between cross-sectional correlations and horizon pairs holds in the time series as well. Figure 2 shows the time series of the cross-sectional correlation for the three horizon pairs: Y0/YLR, Y2/YLR, and Y6/YLR. In general, the strongest correlation holds for the Y6/YLR horizon pair. These cross-sectional correlations are themselves time varying. This is also consistent with the prediction of an unobserved components model, as the degree of cross-sectional correlation (even in the limit with a continuum of forecasters) will be state dependent. The results for the other target variables are presented in Figures A.2 and A.3 in the Supplemental Appendix and follow a similar pattern.

4.3 The Relation Between Short-Run and Longer-Run Forecasts

Model Implications #2(a), #3(a): The unobserved components model implies that revisions to long-horizon forecasts should systematically co-vary with revisions to short-horizon forecasts (#3*a*.). To investigate this relation in the data, for the period between 1998 and 2006 we can construct an unbalanced panel across forecasters in the BCEI survey. We will run regressions of the form:

$$\begin{aligned}\mathrm{LH\ F{\scriptstyle CST}R{\scriptstyle EV}}^j_{i,t} = \alpha^j_i + \beta^j \cdot \mathrm{SH\ F{\scriptstyle CST}R{\scriptstyle EV}}^j_{i,t} + \\ + \delta^j_{\mathsf{Unemp}} \cdot \mathrm{SH\ F{\scriptstyle CST}R{\scriptstyle EV}}^{\mathsf{Unemp}}_{i,t} + \delta^j_{\mathsf{RGDP}} \cdot \mathrm{SH\ F{\scriptstyle CST}R{\scriptstyle EV}}^{\mathsf{RGDP}}_{i,t} \\ + \delta^j_{\mathsf{CPI}} \cdot \mathrm{SH\ F{\scriptstyle CST}R{\scriptstyle EV}}^{\mathsf{CPI}}_{i,t} + \delta^j_{\mathsf{TBill}} \cdot \mathrm{SH\ F{\scriptstyle CST}R{\scriptstyle EV}}^{\mathsf{TBill}}_{i,t} + \varepsilon^j_{i,t}\end{aligned} \tag{19}$$

where $\mathrm{LH\ F{\scriptstyle CST}R{\scriptstyle EV}}^j_{i,t}$ represents the year-over-year revision in forecasts for a far in the future forecast for respondent i at time t and variable j. $\mathrm{SH\ F{\scriptstyle CST}R{\scriptstyle EV}}^j_{i,t}$ represents the counterpart for year-over-year revisions in short-horizon forecasts. For each target variable we also include four additional variables, the short-run forecast revisions for real GDP growth, the unemployment rate, CPI inflation, and the 3-month Treasury bill. These are natural choices for our multivariate specification as they capture the aggregate state of real activity, inflation and monetary policy. The regression includes forecaster-specific fixed effects because, even if it were true that an agent had the correct model, in any finite sample the time series average of forecast errors in equations (17) and (18) would not be identically zero.

To estimate equation (19) we have to calculate forecast revisions using our panel data. To do so we calculate the change in the one-year-ahead forecast from the previous year to the current-year forecast. Thus, we use March to March observations and October to October observations, reflecting forecast timing and horizons in the survey. In particular, SH $\text{FcstRev}^j_{i,t}$ represents the difference between the current year forecast (Y0) and the one-year ahead forecast (Y1) taken one year earlier. This ensures that the target variable remains the same across the two survey periods. Similarly, for LH $\text{FcstRev}^j_{i,t}$, we use the one-year change in the longest-horizon forecast (Y7-11) since a Y8-12 forecast is unavailable. However, given how far in the future this forecast is there should be little difference between this variable and one based on a Y8-12 forecast. To construct appropriately clustered standard errors that are robust to the overlapping nature of the data, we use the variance estimator introduced in Chiang, Hansen, and Sasaki (2024).

Table 1: **Relation Between Short-Run and Longer-Run Forecast Revisions.** This table presents the results of panel-data regressions of the form given in equation (19). The last two columns report p-values corresponding to the null hypotheses $\dot{\mathbb{H}}_0$ and $\ddot{\mathbb{H}}_0$, respectively. Bolded p-values denote a significance level smaller than 10%.

Target Variable	Num. Obs.	$\hat{\beta}$	$\hat{\delta}_{\text{Unemp}}$	$\hat{\delta}_{\text{RGDP}}$	$\hat{\delta}_{\text{CPI}}$	$\hat{\delta}_{\text{TBill}}$	F-Test ($\dot{\mathbb{H}}_0$)	F-Test ($\ddot{\mathbb{H}}_0$)
Auto Sales	207	0.61	0.26	-0.19	0.31	-0.22	**0.00**	**0.00**
Corporate Profits	221	-0.02	-0.35	0.27	0.04	-0.15	**0.00**	**0.00**
Disposable Pers. Income	246	-0.00	-0.14	0.06	-0.15	0.02	**0.00**	**0.00**
GDP Deflator	255	0.09	-0.15	-0.02	0.03	-0.07	**0.00**	**0.00**
Housing Starts	247	0.20	0.02	0.02	-0.02	0.01	**0.00**	**0.01**
Industrial Production	245	0.01	0.16	0.16	-0.14	0.09	**0.00**	**0.04**
Nominal GDP Growth	253	0.07	-0.07	-0.00	0.09	-0.02	**0.04**	0.59
Net Exports	248	-0.01	-57.24	8.04	-2.62	-35.81	**0.03**	0.15
Nonres. Fixed Inv.	248	-0.08	0.04	0.53	0.38	0.18	**0.00**	**0.10**
Pers. Consumption Exp.	250	0.14	-0.05	-0.06	-0.03	0.02	**0.00**	0.85
Real 3-Mo. T-Bill	253	0.40	-0.42	-0.02	-0.08	-0.23	**0.00**	**0.00**
10-Year Treasury Yield	255	0.05	-0.10	0.01	0.03	-0.07	0.13	0.18
Unemployment Rate	255	0.09	–	-0.03	-0.05	-0.09	**0.00**	0.11
Real GDP Growth	255	0.08	0.05	–	-0.01	0.08	**0.00**	0.31
Consumer Price Index	255	0.15	0.15	0.01	–	-0.01	0.16	**0.09**
Nominal 3-Mo. T-Bill	255	0.12	0.54	-0.31	-0.08	–	**0.00**	**0.00**

Table 1 presents the results of our panel data regressions. In the third column we present the point estimate of β^j across the 16 target variables. For most variables, the point estimate is positive suggesting that, all else equal, short-run forecast revisions are positively associated with long-horizon forecast revisions. However, as discussed in Section 3, in a multivariate model the β^j

may feature negative coefficients; indeed, this is the case for four of the variables, corporate profits, disposable personal income, net exports, and nonresidential fixed investment. Importantly, for these four variables the coefficients associated with short-run forecast revisions for other variables (columns 4–7) feature much larger magnitudes. For example, revisions to long-horizon corporate profit forecasts load much more heavily on revisions to short-run forecasts of the state of the economy (as measured by the unemployment rate and real GDP growth) along with revisions to the forecast for the short-term interest rate. This is a useful example to illustrate the key mechanism of the multivariate model.

In terms of statistical significance, we can use an F-test for the null hypothesis that the five main coefficients are equal to zero, i.e.,

$$\dot{\mathbb{H}}_0 : \beta^j = \delta^j_{\mathsf{Unemp}} = \delta^j_{\mathsf{RGDP}} = \delta^j_{\mathsf{CPI}} = \delta^j_{\mathsf{TBill}} = 0. \tag{20}$$

The penultimate column reports p-values for this test. In general, the null hypothesis is strongly rejected for most variables despite the relatively modest sample size we have available. For the remaining two variables, the p-values are only modestly higher than 10%. These results provide strong evidence in favor of the unobserved components model. Through the lens of the model, short-term forecast errors propagate in a way that shifts forecasts at all horizons, including very-long horizons.

To further refine our empirical results, in the last column of the table we report an F-test of the joint null hypothesis that

$$\ddot{\mathbb{H}}_0 : \delta^j_{\mathsf{Unemp}} = \delta^j_{\mathsf{RGDP}} = \delta^j_{\mathsf{CPI}} = \delta^j_{\mathsf{TBill}} = 0. \tag{21}$$

In words, we are testing for the joint significance of the additional variables in each specification.[11] For 10 out of the 16 of our variables we can clearly reject this null hypothesis. Said differently, for the majority of our variables there is sufficient evidence in these data in favor of a *multivariate* model of expectations formation. These results provide strong evidence in favor of the multivariate unobserved components model. Importantly, for variables such as inflation (CPI inflation or the GDP deflator) or the short-term interest rate, we clearly reject the null suggesting that univariate models of expectations formation are incomplete (see, e.g., Coibion and Gorodnichenko (2015), Farmer, Nakamura, and Steinsson (2024) for univariate approaches to model expectations formation for these variables).

One possible concern with the panel data regressions shown in Table 1 is that the individual slopes change with each forecaster and are cross-sectionally correlated with the right-hand side variables, the short-run forecast revisions. This could compromise the consistency of the pooled OLS estimator. Standard approaches to circumvent this issue are challenging with the limited number of time series observations we have available. However, in unreported results we find that

[11] When the target variable is one of these four additional variables, then the null hypothesis is modified as a joint test that the coefficients associated with the remaining three variables are zero.

this bias is negligible for empirically-relevant choices of the true parameters in the model, giving confidence in the pooled OLS estimates presented in Table 1.

5 Implications for Models of Expectations Formation

The stylised facts of this paper have implications for how we conduct statistical inference on expectations and how we build behavioral theories of expectations formation. First, we discuss the consequences of multivariate forecasting models. Second, the consequences of movements in long-run beliefs. While this separation is somewhat artificial given the main finding that forecasters have multivariate models of the long run, it provides a useful way to structure the discussion.

5.1 Consequences of Multivariate Forecasting Models

That expectations are given by a multivariate model potentially reflects that forecasters have some understanding of economic theory and a degree of sophistication in expectations formation that much concerned Muth (1961). Of course, this is a known, if not mostly overlooked insight, of the early imperfect information literature. Using the Livingston Survey data, Mullineaux (1980) and Caskey (1985) provide evidence that short-run expectations of inflation are influenced by observed money growth rates, consistent with the quantity theory of money. Evidence is also adduced that macroeconomic variables thought important for inflation, such as the fiscal surplus, also matter for expectations formation. Observing a positive correlation, Caskey (1985) speculates that movements in unemployment might matter for inflation expectations through the anticipated effects from future policy adjustments. Lewis (1995) extends these insights on the importance of money growth rates to exchange rate expectations. More recently these ideas have been explored in the context of contemporary procedures for monetary policy. For example, Carvalho and Nechio (2014) demonstrate that professional forecasters and higher-educated and higher-income households tend to hold expectations that are consistent with a Taylor rule, commonly used to describe interest rate decisions of the Federal Reserve.[12]

A corollary of these findings is that univariate analyses of expectations formation are likely misspecified. Building on Coibion and Gorodnichenko (2015), a large literature explores whether household, firm and policymaker expectations display departures from full-information rational expectations. A specific interest is whether expectations display "over-reaction" or "under-reaction" to new information. Using consensus short-horizon forecast data from the Survey of Professional Forecasters, they estimate regressions of the form

$$\textsc{FcstErr}_{t+h} \quad = \quad \alpha + \beta \cdot \textsc{FcstRev}_t^{(h)} + \varepsilon_t$$

[12] See also closely related work by Dräger, Lamla, and Pfajfar (2016) and Bauer, Pflueger, and Sunderam (2024).

where

$$
\begin{aligned}
\text{FcstErr}_{t+h} &\equiv X_{t+h} - \mathbb{E}_t[X_{t+h}] \\
\text{FcstRev}_t^{(h)} &\equiv \mathbb{E}_t[X_{t+h}] - \mathbb{E}_{t-1}[X_{t+h}]
\end{aligned}
$$

for variable X_t, $\mathbb{E}_t(\cdot)$ the expectations operator conditioned on period t information, and $h = 3$ quarters. The sign of the coefficient β indicates whether expectations tend to under or overshoot actual future values in response to new information, and, under certain assumptions, is informative about the nature of information frictions relevant to decision making. Bordalo, Gennaioli, Ma, and Shleifer (2020) extend this analysis to panel regressions of individual-level data.

But if beliefs are formed using multivariate econometric models, it is unclear what such single equation methods reveal about the nature of expectations formation. For example, if the true beliefs are given by a vector-autoregression model, the dynamics of any single variable will be given by an ARMA process, that depends on all structural shocks—and so too the revisions in expectations at any point in time. To infer expectations have the property of over-reaction or under-reaction requires a structural approach. Angeletos, Huo, and Sastry (2020) provides an example of what is required. They argue that the impulse response function of forecast errors to an identified shock provides a complete summary of the dynamic properties of expectations. Importantly, they demonstrate that various results in the literature—notably Coibion and Gorodnichenko (2015), Bordalo, Gennaioli, Ma, and Shleifer (2020) and Kohlhas and Walther (2021)—arise because they are each a particular but distinct weighted combination of the moments that define the impulse response function. Not only does this integrate apparently disparate and conflicting findings, but it also suggests that expectations about key macroeconomic variables may display under-reaction in the short run and over-reaction in the long run.

These insights serve as warning against single-equation methods. The degree to which it matters for inference is clearly demonstrated by a number of recent papers on the term structure of interest rates. Andrade, Crump, Eusepi, and Moench (2016) show using forecast data from Blue Chip Financial Forecasts, that, to explain the term structure of disagreement—proxied by the difference of the average top and bottom ten forecasters—about interest rates from the near term to the very long term, requires modeling the joint dynamics of interest rates, inflation, and the output gap. Without expectations data on output and inflation, unobserved components models—which otherwise provide an excellent fit of the data—are unable to characterize rising disagreement along the term structure of interest rate expectations. In related work that focuses on consensus expectations, Crump, Eusepi, and Moench (2024) use the universe of professional forecast data in the United States to model the term structure of expectations of interest rates, inflation and the output gap. Again they show that univariate models provide a poor fit of the term structure of interest rate expectations, particularly in the medium- to longer-run—revealing expectations about inflation and the output gap to be important, presumably because of market understanding of the Federal Reserve's reaction function.

Use of univariate expectations formation models is very common in the literature. For example, Farmer, Nakamura, and Steinsson (2024) use a single equation unobserved component model of the term structure of interest rates. It is also common in structural models where expectations are not assumed to be rational. In the adaptive learning literature, Slobodyan and Wouters (2012) study an estimated new Keynesian model with learning. Households and firms use univariate forecasting models to construct short-term forecasts relevant to their decisions. Forecast errors for a particular macroeconomic time series lead to a revision of expectations about that variable alone. Hommes, Mavromatis, Özden, and Zhu (2023) also study a new Keynesian model, introducing "behavioral learning equilibria". This concept of bounded rationality asserts that for each variable, agents use simple AR(1) forecasting rules, whose parameters are consistent with the observed sample mean and autocorrelation of past data for that variable. Both papers appeal to simplicity as a criterion for agents to select univariate forecasting models. But our results, along with various other papers discussed above, suggest that multivariate beliefs may be critical to a proper understanding of the joint dynamics of inflation, interest rates and output and the constraints that such beliefs place on policy design.

5.2 Consequences of Learning about the Long Run

The central result of this paper is not just that forecasters use multivariate models, but that these multivariate models explain individual-specific long-horizon forecasts. How then should we build economic models consistent with this evidence? To proceed we must acknowledge that our results do not take a stand on rationality. But whether forecasts are optimal or not, the results force us to confront certain features of standard economic analyses. Most models assume a stationary economic environment that has a well-defined, unique, steady state. By assumption, then, long-run expectations must be largely determined by this steady state. Emphasis is given to questions relating to how short-to-medium-run expectations mediate the impulse and propagation mechanisms of the model and the effects of policy. But what are the consequences of shifting long-run expectations for such questions?

A broad literature in macroeconomics and finance has recognized the first-order implications of modeling imperfect knowledge about the long-run behavior of the economy. One approach has relied on the assumption of rational expectations. In such models long-run beliefs are tied to a specific source of uncertainty. Erceg and Levin (2003) evaluate the impact of disinflationary policies when the central bank's inflation target (the long-run mean of inflation) is not fully credible and market participants infer it from observed outcomes. Fuhrer (1996) and Kozicki and Tinsley (2005) appeal to uncertainty about the prevailing monetary regime to reconcile the term structure of interest rate expectations (including the very long-run) with the observed behavior of yields. Results suggest that imperfect information about the long run can potentially revive the expectations hypothesis of the term structure of interest rates. Edge, Laubach, and Williams (2007) focuses on how long-run expectations about labor productivity can shape the macroeconomic response to persistent shocks to total factor productivity. Further, imperfect information about the trend of productivity can solve the co-movement problem in an otherwise standard real business cycles model (Barro and King (1984)). Finally, Timmermann (1993) shows the learning about long-run dividend growth can lead to excess volatility in asset prices and predictability of returns in a conventional asset pricing model (see also, Nagel and Xu (2021)).

The assumption of rational expectations is often justified as the outcome of a learning process in a stationary environment but is less amenable to an environment with continuous transformation (Sargent, 1993). Another strand of research that places strong emphasis on long-run expectations and multivariate forecasting models is the adaptive learning literature. These papers depart from those just discussed, assuming both information frictions and bounded rationality. Under this approach market participants do not necessarily know the underlying factors driving the long run behavior of the economy. Additionally, agents' learning process can induce long-run volatility even in absence of changes in fundamentals. Preston (2005) showed how to solve the new Keynesian model under arbitrary beliefs. The paper focused on the stability properties of the model under recursive least squares learning, exploring what types of monetary policy rules ensured learnability of rational expectations equilibrium. A key insight from this work is that stability properties were generally determined by the interaction of monetary policy with long-run beliefs.

But the approach offers much more. Relaxing the cross-equation restrictions of rational expectations equilibrium analysis and allowing long-run beliefs to depart from steady state, permits a class of model with considerably richer dynamics than rational expectations. Not only does this promise progress on quantitative questions that are considered puzzles through the lens of a rational expectations equilibrium analysis, but it facilitates addressing certain questions that such models either cannot address, or are poorly suited to address.

For example, learning dynamics reduce the need for ad hoc sources of persistence—such as habit formation and price indexation—and provide a source of amplification of shocks (Milani (2007), Eusepi and Preston (2011)). Learning as an information friction can also generate both under-reaction and over-reaction (Eusepi and Preston (2011), Carvalho, Eusepi, Moench, and Preston (2023), Eusepi, Giannoni, and Preston (2024), and Hajdini (2023)). In asset pricing models, departures from full-information rational expectations fundamentally re-weight income and substitution effects to resolve a range of puzzles, including return predictability and excess volatility (Adam, Marcet, and Beutel (2017), Sinha (2016), Eusepi and Preston (2018), Farmer, Nakamura, and Steinsson (2024), and Crump, Eusepi, and Moench (2024)).

Evaluating central bank credibility and communication is difficult in the full-information rational expectations environment. By assumption, expectations are consistent with monetary policy strategy. Households and firms understand the systematic conduct of policy and hold expectations that are consistent with the maintained inflation target. Adaptive learning models permit studying the conditions under which expectations about future interest rates, output and inflation should be consistent with policy, and therefore "anchored" (Eusepi and Preston (2010) and Carvalho, Eusepi, Moench, and Preston (2023)). Furthermore, instability in long-run expectations can make inflation control more difficult in both normal times and crisis periods, when the policy is constrained by the zero lower bound on nominal interest rates, and when public debt is high (see Eusepi, Giannoni, and Preston (2024), Eusepi, Gibbs, and Preston (2025) and Eusepi and Preston (2018)).

6 Conclusion

This paper provides novel empirical and theoretical insights into how professional forecasters form and revise long-run expectations. Muth (1961) challenged the economics profession to articulate a more sophisticated model of decision maker expectations than naive or adaptive expectations. Despite expressing reservations, he argued that firms should construct forecasts using the "relevant economic theory", an idea that became synonymous with model-consistent expectations during the rational expectations revolution of the 1970s. We provide direct evidence that professional forecasters are indeed sophisticated, but show that these expectations raise important questions about the relevant economic theory that is consistent with professional forecaster beliefs.

Using a unique, novel, and rich panel dataset on forecasts across multiple variables and horizons, we show that long-horizon expectations display considerable heterogeneity across forecasters, evolve dynamically in response to short-run developments, and are shaped by multivariate considerations. These findings challenge the conventional assumption of anchored or rational long-run expectations and motivate models in which information frictions are relevant at short- and long-run horizons. Our results suggest that incorporating multivariate unobserved-component learning mechanisms into macroeconomic models is essential for better understanding expectations formation, belief dynamics, and their implications for policy design and macroeconomic stability.

References

Adam, K., A. Marcet, and J. Beutel (2017). Stock price booms and expected capital gains. *American Economic Review 107*(8), 2352–2408.

Afrouzi, H., S. Y. Kwon, A. Landier, Y. Ma, and D. Thesmar (2023). Overreaction in expectations: Evidence and theory. *Quarterly Journal of Economics 138*(3), 1713–1764.

Andrade, P., R. K. Crump, S. Eusepi, and E. Moench (2016). Fundamental disagreement. *Journal of Monetary Economics 83*(1), 106–128.

Angeletos, G.-M., Z. Huo, and K. A. Sastry (2020). Imperfect macroeconomic expectations: Evidence and theory. In *NBER Macroeconomics Annual 2020*, Volume 35, pp. 1–86. National Bureau of Economic Research.

Barro, R. J. and R. G. King (1984). Time-separable preferences and intertemporal-substitution models of business cycles. *Quarterly Journal of Economics 99*(4), 817–839.

Bauer, M. D., C. E. Pflueger, and A. Sunderam (2024). Perceptions about monetary policy. *Quarterly Journal of Economics 139*(4), 2227–2278.

Beechey, M. J., B. K. Johannsen, and A. T. Levin (2011). Are long-run inflation expectations anchored more firmly in the Euro area than in the United States? *American Economic Journal: Macroeconomics 3*(2), 104–29.

Bems, R., F. Caselli, F. Grigoli, and B. Gruss (2021). Expectations' anchoring and inflation persistence. *Journal of International Economics 132*.

Binder, C. C., W. Janson, and R. J. Verbrugge (2019). Thinking outside the box: Do SPF respondents have anchored inflation expectations? Working Paper 19-15, Federal Reserve Bank of Cleveland.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020). Overreaction in macroeconomic expectations. *American Economic Review 110*(9), 2748–2782.

Cao, S., R. K. Crump, S. Eusepi, and E. Moench (2021). Fundamental disagreement about monetary policy and the term structure of interest rates. Staff Report 934, Federal Reserve Bank of New York.

Carvalho, C., S. Eusepi, E. Moench, and B. Preston (2023). Anchored inflation expectations. *American Economic Journal: Macroeconomics 15*(1), 1–47.

Carvalho, C. and F. Nechio (2014). Do people understand monetary policy? *Journal of Monetary Economics 66*(C), 108–123.

Caskey, J. (1985). Modeling the formation of price expectations: A Bayesian approach. *American Economic Review 75*(4), 768–776.

Chiang, H. D., B. E. Hansen, and Y. Sasaki (2024). Standard errors for two-way clustering with serially correlated time effects. *Review of Economics and Statistics*. (forthcoming).

Cogley, T., G. E. Primiceri, and T. J. Sargent (2010). Inflation-gap persistence in the US. *American Economic Journal: Macroeconomics 2*(1), 43–69.

Cogley, T. and T. J. Sargent (2005). Drift and volatilities: Monetary policies and outcomes in the post WWII U.S. *Review of Economic Dynamics 8*(2), 262–302.

Coibion, O. and Y. Gorodnichenko (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review 105*(8), 2644–2678.

Crump, R. K., S. Eusepi, and E. Moench (2022). The term structure of expectations and bond yields. Staff Report 775, Federal Reserve Bank of New York.

Crump, R. K., S. Eusepi, and E. Moench (2024). Is there hope for the expectations hypothesis? Staff Report 1098, Federal Reserve Bank of New York.

Crump, R. K., S. Eusepi, E. Moench, and B. Preston (2023, November). The term structure of expectations. In R. Bachmann, G. Topa, and W. van der Klaauw (Eds.), *Handbook of Economic Expectations* (1 ed.), Volume 1, Chapter 17, pp. 507–540. Elsevier.

Del Negro, M., D. Giannone, M. P. Giannoni, and A. Tambalotti (2018). Safety, liquidity, and the natural rate of interest. *Brookings Papers on Economic Activity 49*(1), 235–94.

Dräger, L., M. J. Lamla, and D. Pfajfar (2016). Are survey expectations theory-consistent? The role of central bank communication and news. *European Economic Review 85*(C), 84–111.

Edge, R. M., T. Laubach, and J. C. Williams (2007). Learning and shifts in long-run productivity growth. *Journal of Monetary Economics 54*(8), 2421–2438.

Erceg, C. J. and A. T. Levin (2003). Imperfect credibility and inflation persistence. *Journal of Monetary Economics 50*(4), 915–944.

Eusepi, S., M. Giannoni, and B. Preston (2024). The short-run policy constraints of long-run expectations. *Journal of Political Economy*. (forthcoming).

Eusepi, S., C. G. Gibbs, and B. Preston (2025). Monetary policy as insurance. manuscript, UNSW Business School.

Eusepi, S. and B. Preston (2010). Central bank communication and macroeconomic stabilization. *American Economic Journal: Macroeconomics 2*, 235–271.

Eusepi, S. and B. Preston (2011). Expectations, learning and business cycle fluctuations. *American Economic Review 101*(6), 2844–2872.

Eusepi, S. and B. Preston (2018). Fiscal foundations of inflation: Imperfect knowledge. *American Economic Review 108*(9), 2551–2589.

Farmer, L. E., E. Nakamura, and J. Steinsson (2024). Learning about the long run. *Journal of Political Economy 132*(10), 3334–3377.

Fuhrer, J. C. (1996). Monetary policy shifts and long-term interest rates. *Quarterly Journal of Economics 111*(4), 1183–1209.

Gürkaynak, R., A. Levin, and E. Swanson (2010). Does inflation targeting anchor long-run inflation expectations? Evidence from the U.S., UK, and Sweden. *Journal of the European Economic Association 8*(6), 1208–1242.

Gürkaynak, R., B. Sack, and E. T. Swanson (2005). The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models. *American Economic Review 95*(1), 425–436.

Hajdini, I. (2023). Mis-specified forecasts and myopia in an estimated new Keynesian model. *American Economic Journal: Macroeconomics*. (forthcoming).

Hommes, C., K. Mavromatis, T. Özden, and M. Zhu (2023, November). Behavioral learning equilibria in new Keynesian models. *Quantitative Economics 14*(4), 1401–1445.

Kohlhas, A. N. and A. Walther (2021). Asymmetric attention. *American Economic Review 111*(9), 2879–2925.

Kozicki, S. and P. A. Tinsley (2001). Shifting endpoints in the term structure of interest rates. *Journal of Monetary Economics 47*(3), 613–652.

Kozicki, S. and P. A. Tinsley (2005). What do you expect? Imperfect policy credibility and tests of the expectations hypothesis. *Journal of Monetary Economics 52*(2), 421–447.

Kydland, F. E. and E. C. Prescott (1982). Time to build and aggregate fluctuations. *Econometrica 50*(6), 1345–70.

Lahiri, K. and X. Sheng (2008). Evolution of forecast disagreement in a Bayesian learning model. *Journal of Econometrics 144*(2), 325–240.

Laubach, T. and J. C. Williams (2003). Measuring the natural rate of interest. *Review of Economics and Statistics 85*(4), 1063–1070.

Lewis, K. (1995). Puzzles in international financial markets. In G. Grossman and K. Rogoff (Eds.), *Handbook of International Economics*, pp. 1913–1971. Elsevier, Amsterdam.

Lucas, R. E. (1986). Adaptive behavior and economic theory. *Journal of Business 59*(4), S401–S426.

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

Mullineaux, D. J. (1980). Inflation expectations and money growth in the United States. *The American Economic Review 80*(1), 149–161.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica 29*(3), 315–335.

Nagel, S. and Z. Xu (2021). Asset pricing with fading memory. *Review of Financial Studies 35*(5), 2190–2245.

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*(7), 803–820.

Preston, B. (2005). Learning about monetary policy rules when long-horizon expectations matter. *International Journal of Central Banking 1*(2), 81–126.

Sargent, T., N. Williams, and T. Zha (2006). Shocks and government beliefs: The rise and fall of American inflation. *American Economic Review 96*(4), 1193–1224.

Sargent, T. J. (1993). *Bounded Rationality in Macroeconomics.* Oxford, UK: Oxford Economic Press.

Sinha, A. (2016). Learning and the yield curve. *Journal of Money, Credit and Banking 48*(2-3), 513–547.

Slobodyan, S. and R. Wouters (2012). Learning in a medium-scale DSGE model with expectations based on small forecasting models. *American Economic Journal: Macroeconomics 4*(2), 65–101.

Stock, J. H. and M. W. Watson (1989). New indexes of coincident and leading economic indicators. In *NBER Macroeconomics Annual 1989*, Volume 4, pp. 351–409. National Bureau of Economic Research.

Stock, J. H. and M. W. Watson (2007). Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking 39*(s1), 3–33.

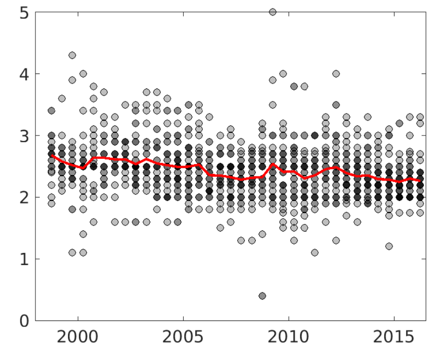
Timmermann, A. G. (1993). How learning in financial markets generates excess volatility and predictability in stock prices. *Quarterly Journal of Economics 108*(4), 1135–1145.

Weber, M., F. D'Acunto, Y. Gorodnichenko, and O. Coibion (2022). The subjective inflation expectations of households and firms: Measurement, determinants, and implications. *Journal of Economic Perspectives 36*(3), 157–184.

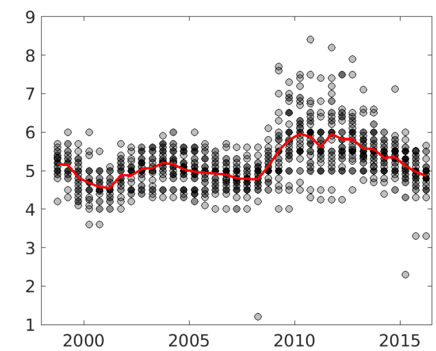
Figure 1: **Cross-Section of Long-Horizon Forecasts.** This figure shows the individual time series for the longest-horizon forecast available (seven-eleven years ahead) from the Blue Chip Economic Indicators (BCEI) for a selection of economic indicators. The red line represents the consensus (mean) forecast. The sample period is October 1998–March 2016.

(a) *Real GNP/GDP Growth*

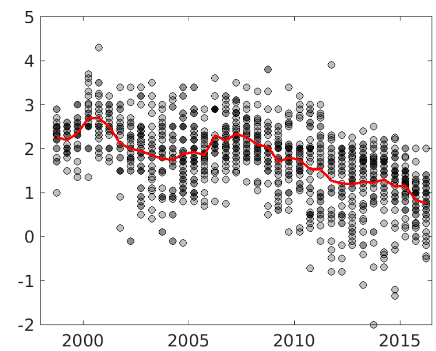
(b) *CPI Inflation*



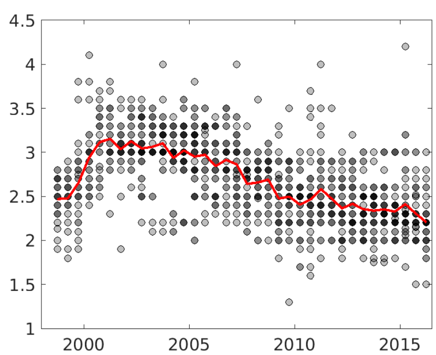
(c) *Unemployment Rate*



(d) *Real 3-Month T-Bill*



(e) *Personal Cons. Exp.*



(f) *Nonres. Business Fix. Inv.*

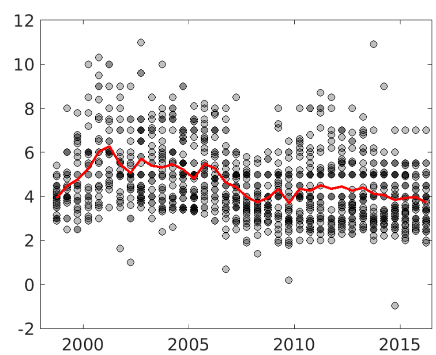
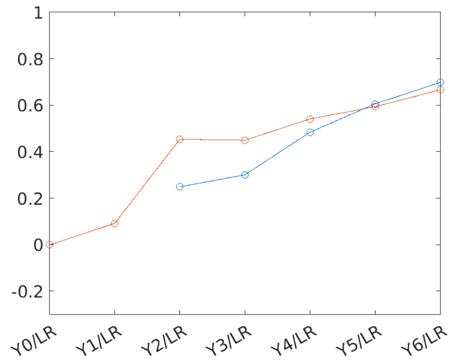


Table 2: **Time Variability of Long-Horizon Forecasts.** This table reports summary statistics of the individual long-horizon BCEI forecasts. We report the maximum, minimum, 75th, 50th, and 25th percentile of the standard deviation relative to the absolute value of the mean. We exclude forecasters with fewer than three long-horizon forecasts in our sample. The sample period is October 1998–March 2006.

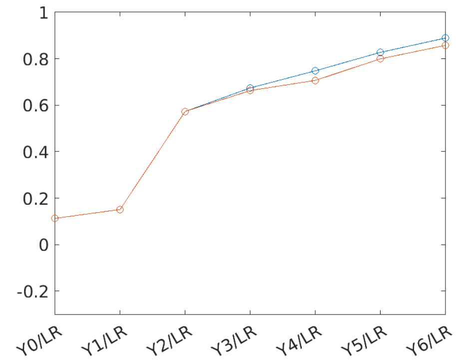
Variable	Max	75th pct.	Median	25th pct.	Min	Consensus
RGDP	0.18	0.12	0.1	0.08	0.02	0.127
GDPPI	0.45	0.21	0.17	0.13	0	0.036
NGDP	0.16	0.1	0.09	0.07	0.02	0.080
CPI	0.35	0.19	0.16	0.12	0	0.052
IndProd	0.31	0.17	0.14	0.12	0.02	0.129
DispPerInc	0.5	0.15	0.12	0.09	0.03	0.113
PerConsExp	0.21	0.15	0.12	0.08	0	0.111
NonResFixInv	0.47	0.33	0.26	0.2	0	0.151
CorProfits	0.59	0.3	0.25	0.16	0	0.120
TBill3M	0.28	0.17	0.13	0.11	0.02	0.137
TNote10Y	0.22	0.12	0.09	0.07	0.02	0.100
Unempl	0.13	0.1	0.08	0.06	0.03	0.079
HouseStMil	0.16	0.1	0.08	0.06	0.01	0.066
AutoSaleMil	0.11	0.08	0.06	0.04	0.01	0.048
NetX	0.68	0.48	0.34	0.26	0.04	0.249
RShortRate	0.74	0.42	0.31	0.2	0.09	0.281

Figure 2: **Term Structure of Cross-Sectional Correlation.** This figure shows the time series average of cross-sectional rank correlations of forecasts across different horizon pairs. The orange line presents results for the sample period of October 1998–March 2006 whereas the blue line presents results for the full sample period of October 1998–March 2016.

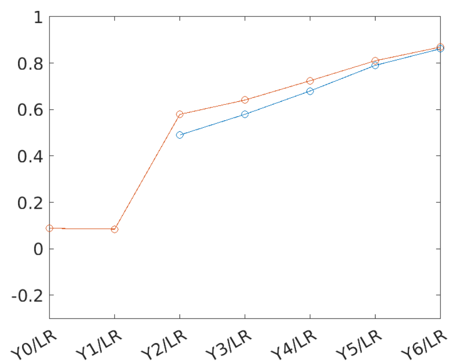
(a) *Real GNP/GDP Growth*



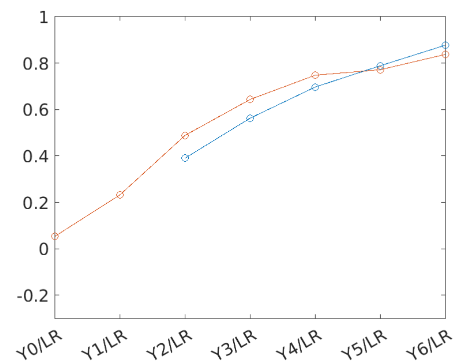
(b) *CPI Inflation*



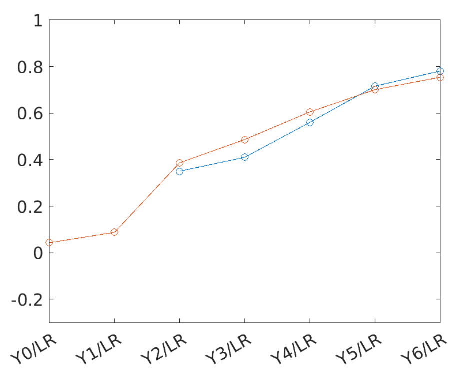
(c) *Unemployment Rate*



(d) *Real 3-Month T-Bill*



(e) *Personal Cons. Exp.*



(f) *Nonres. Business Fix. Inv.*

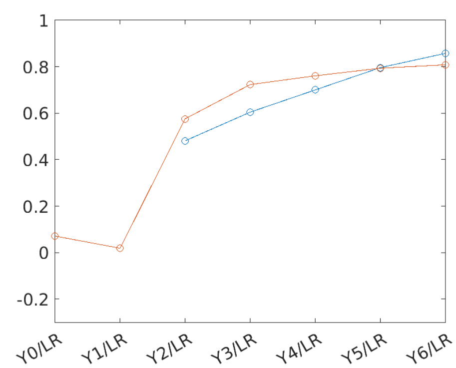
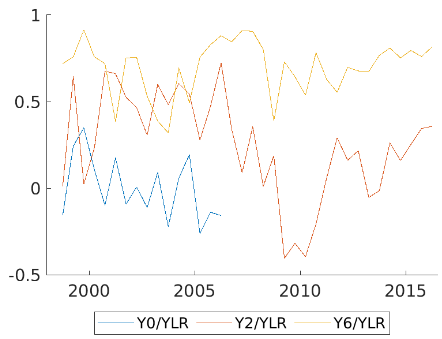
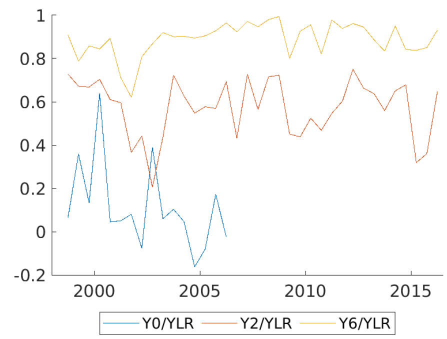


Figure 3: **Time Series of Cross-Sectional Correlation.** This figure shows the time series of cross-sectional rank correlations of forecasts across selected horizon pairs. The Y0/YLR pair (blue line) is presented for the sample period October 1998–March 2006 whereas the other horizon pairs span the full sample period October 1998–March 2016.

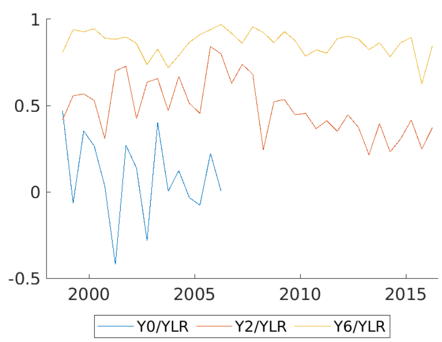
(a) *Real GNP/GDP Growth*



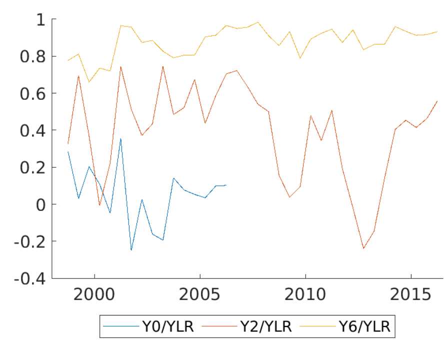
(b) *CPI Inflation*



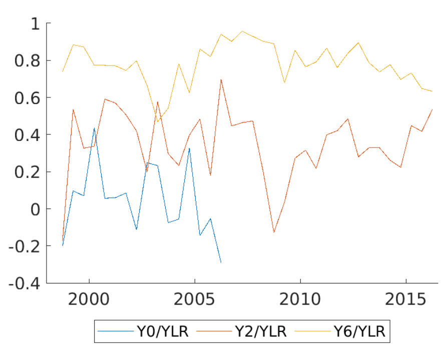
(c) *Unemployment Rate*



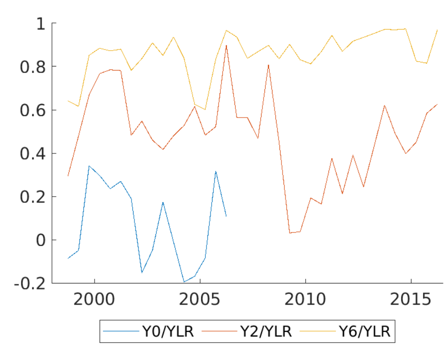
(d) *Real 3-Month T-Bill*



(e) *Personal Cons. Exp.*



(f) *Nonres. Business Fix. Inv.*



Supplement to "How Do We Learn About the Long Run?"

Richard K. Crump
Stefano Eusepi
Bruce Preston
Emanuel Moench

April 13, 2025

S-A Additional Results

This Supplemental Appendix provides additional empirical results based on the Blue Chip Economic Indicators (BCEI) Survey data introduced in the main text.

We define the six "main" variables as: (1) real GNP/GDP growth; (2) CPI inflation; (3) the unemployment rate; (4) the real 3-month T-bill; (5) real personal consumption growth; (6) nonresidential fixed investment growth.

We define the ten "supplement" variables as: (1) auto sales; (2) corporate profits growth; (3) disposable personal income growth; (4) GNP/GDP deflator; (5) housing starts; (6) industrial production; (7) net exports; (8) nominal GNP/GDP growth; (9) 3-month T-bill; (10) 10-year note.

Figure A.1. Cross-Section of Long-Horizon Forecasts. The charts below replicate those in Figure 1 for the additional ten variables. This figure shows the individual time series for the longest-horizon forecast available (seven-eleven years ahead) from the Blue Chip Economic Indicators (BCEI) for a selection of economic indicators. The red line represents the consensus (mean) forecast. The sample period is October 1998–March 2016.

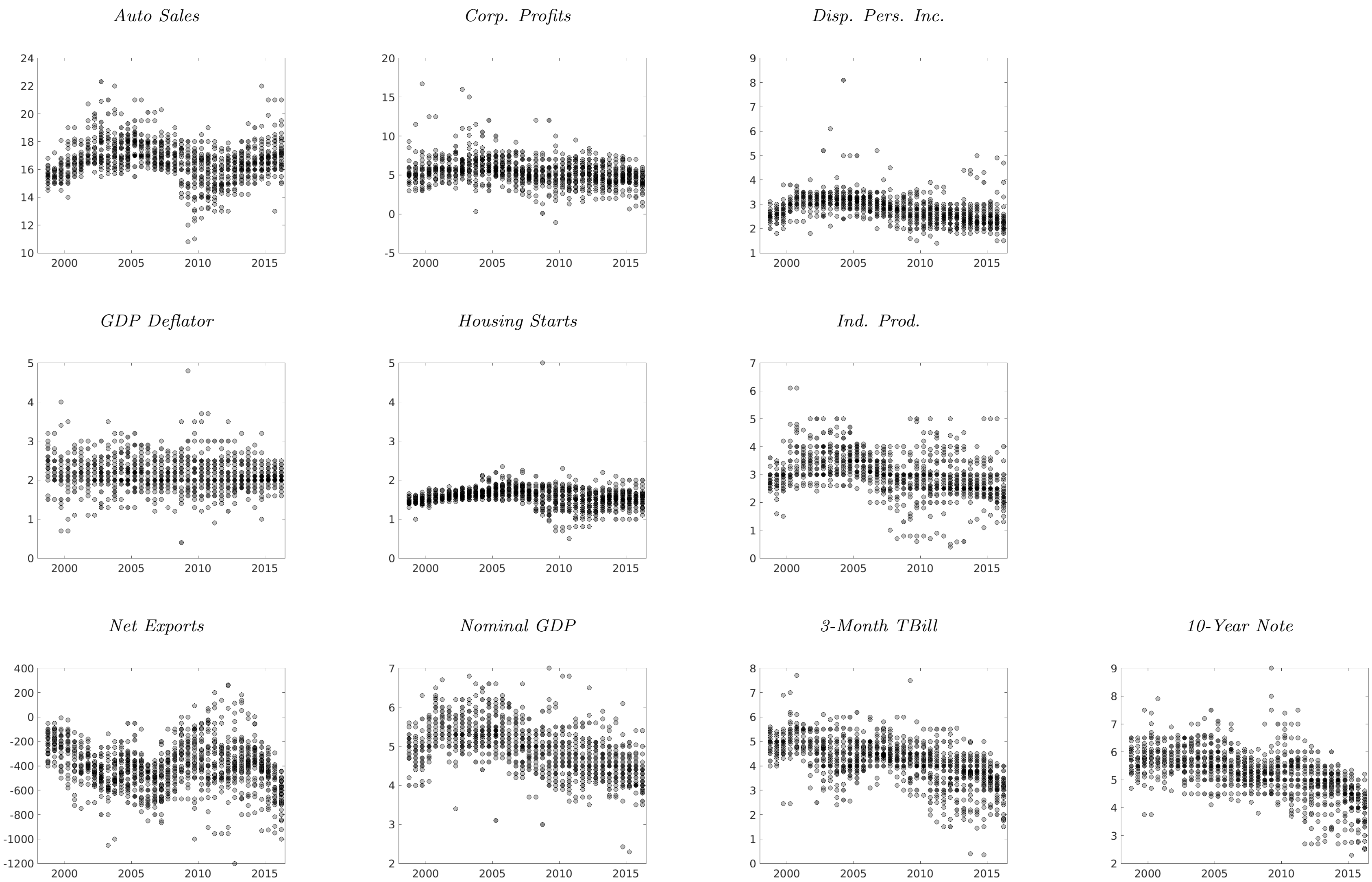


Figure A.2. Term Structure of Cross-Sectional Correlation. The charts below replicate those in Figure 2 for the additional ten variables. This figure shows the time series average of cross-sectional rank correlations of forecasts across different horizon pairs. The blue line presents results for the sample period of October 1998–March 2006 whereas the orange line presents results for the full sample period of October 1998–March 2016.

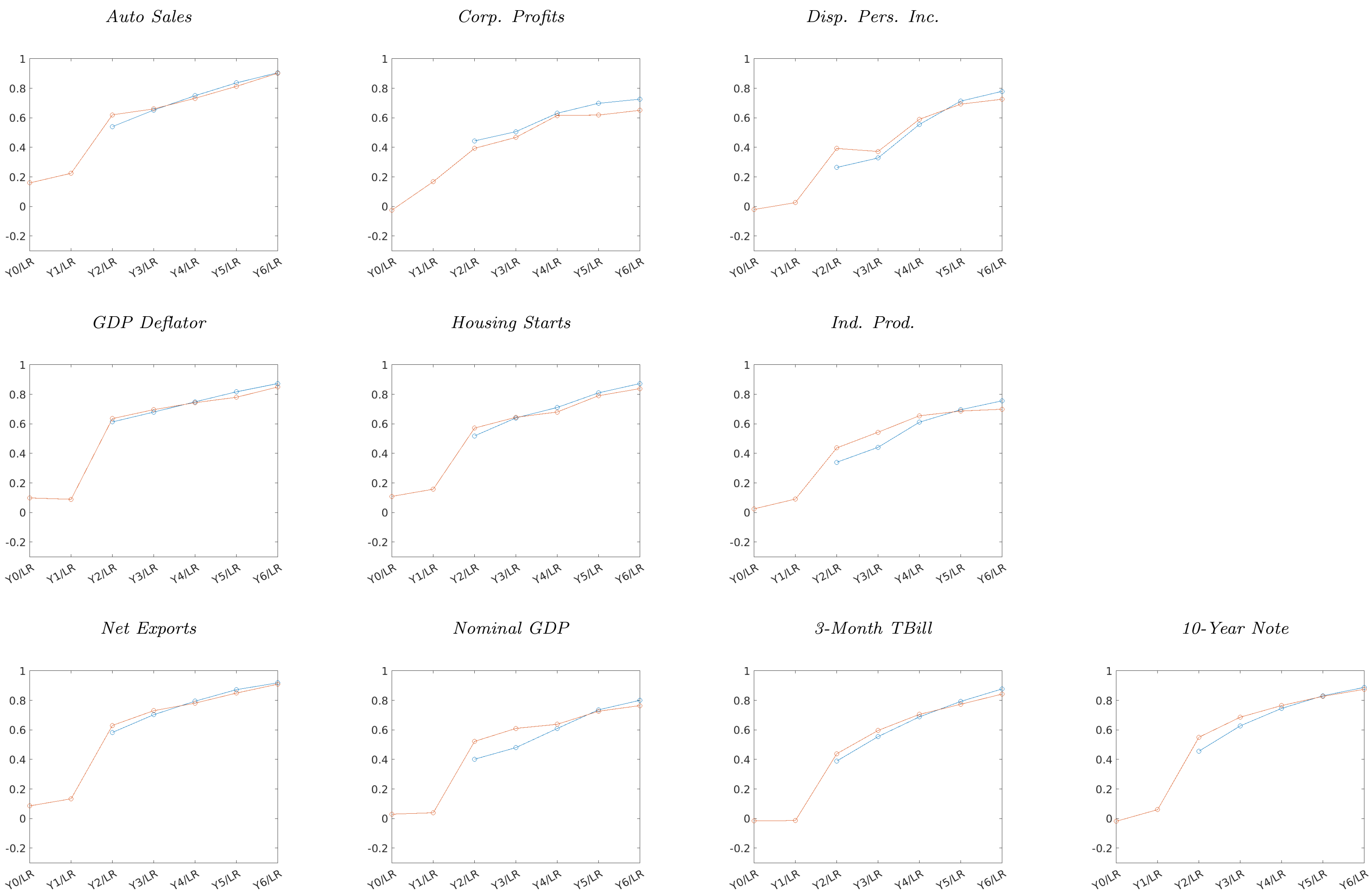


Figure A.3. The Time Series of Cross-Sectional Correlation. This figure shows selected horizon pairs of the underlying time series of the cross-sectional correlation of forecasts which are used to construct the term structures of cross-sectional correlation shown in Figure A.2. The correlation measure used is the Spearman correlation coefficient.