What Hundreds of Economic News Events Say About Belief Overreaction in the Stock Market

Francesco Bianchi* — Johns Hopkins, CEPR, and NBER

Sydney C. Ludvigson† — NYU, CEPR, and NBER

Sai Ma‡ — Federal Reserve Board

September 12, 2024

Preliminary

Abstract

We measure the nature and severity of a variety of belief distortions in market reactions to hundreds of economic news events using a new methodology that synthesizes estimation of a structural asset pricing model with algorithmic machine learning to quantify bias. We estimate that investors systematically overreact to perceptions about multiple fundamental shocks in a macro-dynamic system, generating asymmetric compositional effects when several counteracting shocks occur simultaneously in real-world events. We show that belief overreaction to all shocks can lead the market to over- *or* underreact to events, amplifying or dampening volatility.

JEL: E7, E27, E32, E17, G4. Keywords: AI, News, Overreaction, Beliefs

*Department of Economics, Wyman Park Building, 3100 Wyman Park Drive, Baltimore, MD 21211. (francesco.bianchi@jhu.edu)

†Department of Economics, NYU, 19 W.4th St, 6th Floor, New York, NY 10012. (sydney.ludvigson@nyu.edu)

‡Federal Reserve Board of Governors, C Ave & 20th Street NW, Washington, DC 20551. (sai.ma@frb.gov)

We are grateful to Nicolò Ceneri, Doruk Gökalp, Do Lee, and Steven Zheng for excellent research assistance, and to Nicholas Barberis, Aditya Chaudhry, Nicola Gennaioli, Samuel Hartzmark, Stefan Nagel, Andrei Shliefer, and seminar participants at the NBER Spring 2024 Behavioral Finance meeting, UNC Chapel Hill, LSE, Queen Mary University London, Wharton, and University of Texas Austin for helpful comments. Bianchi and Ludvigson received financial support from the National Science Foundation under Grant 2116641. The views expressed are those of the authors and do not necessarily reflect those of the Federal Reserve Board or the Federal Reserve System.

1 Introduction

The pronounced volatility of world equity markets is difficult to reconcile with textbook models in which the price of a stock is the rational expectation of future cash flow fundamentals, discounted at a constant rate. These theories imply that stock markets should be far more stable than observed, leading a vast literature to explain "excess" stock market volatility with discount rate variation.[1] But recent advancements in the field of behavioral finance point toward a different explanation, namely that investors may exhibit systematic expectational errors ("belief distortions") that lead them to overreact to news relevant for cash flow growth. A standard result is that overreaction amplifies market volatility, offering an explanation for observed equity markets that does not rely on variable discount rates.

Documenting evidence of overreaction (or belief distortions more generally) requires both a measure of what investors subjectively expect, and a benchmark for gauging any distortion in subjective growth expectations. The traditional approach to this problem is to use surveys of analysts or investors to measure subjective expectations, and to use in-sample regressions of survey forecast errors on lagged forecast revisions to measure overreaction. Despite valuable insights, the very simplicity and convenience of the traditional approach necessarily leaves several pertinent questions unanswered.

For one, the precise news events to which investor beliefs purportedly overreact are left unspecified in the forecast-error-on-forecast-revision regression approach. This lack of specificity opens the door to challenges from antipodal behavioral models proposing that investors underreact to at least some forms of news, and/or that revisions in subjectively expected future cash flow growth respond to the stock market rather than drive it. If markets overreact to news, which events have historically been responsible for such reactions and why?

The traditional approach is also ill-suited to the task of developing a deeper understanding of the perceived sources of primitive economic risk that must jointly drive beliefs and markets. This is relevant when considering evidence that price pressure from unexplained shifts in demand–"flows" in and out of the stock market evidently unrelated to cash flow news–are responsible for substantial stock market volatility. But if news about cash flow growth is not the main driver of stock market volatility, what are the perceived sources of economic risk that are? Such questions are difficult or impossible to answer without a structural model that stipulates the types of shocks that investors may be contemplating.

Finally, contrary to the traditional approach, new measures of belief distortion generated from dynamic, real-time machine learning algorithms find little evidence that survey forecast

[1] For textbook treatments of this issue, see Chapters 7 and 8 of Campbell, Lo and MacKinlay (1997), and Chapter 20 of Cochrane (2005).

errors are related to lagged forecast revisions. This raises immediate questions about the traditional methodology, since it means that the standard regression approach to measuring over- or underreaction may not always provide a reliable means of quantifying systematic expectational error.

In this paper we study how the stock market reacts to news using a new methodology developed here that (i) measures the stock market's response to specific news events, (ii) estimates revisions in the representative investor's subjective expectations and perceived sources of risk as a result of those events, and (iii) gauges the quantitative importance (if any) of a range of belief distortions in driving the market's reactions to news. The approach, which we refer to as "structural AI synthesis" for brevity, combines insights from artificial intelligence (AI) and mixed-frequency Bayesian econometrics to provide a benchmark for measuring the nature and severity of investors' distorted beliefs. Its definitive feature is the *synthesis* of algorithmic machine learning and a structural asset pricing model in the analysis of data. A general premise of this synthesis is that a wide variety of observable data constitute important signals of what real-world market participants actually believe and expect. These include not only direct measures of subjective asset market expectations from surveys of equity analysts and investors (as in the traditional approach), but also fluctuations in spot prices, futures markets, and professional forecasts of the broader economy, both at high frequency around news events and over longer periods of time. At its most general level, the procedure is a methodology for using machine learning algorithms, survey forecasts, and other forward-looking data produced from a complicated real-world setting to inform a stylized belief formation framework that is an explicit approximation of reality.

This approach–as applied to the present investigation–has four central ingredients. First, it considers high frequency market reactions to hundreds of specific news events spanning macroeconomic data releases, corporate earnings announcements, and central bank communications from the Federal Reserve (the Fed). Second, we specify and estimate a structural asset pricing model to empirically decompose the representative investor's high frequency reactions to these events into perceived shocks driving primitive macroeconomic risks that together span cash flow and discount rate news. Third, investor beliefs in the structural model are allowed to potentially depart from rationality in range of ways by magnitudes that are freely estimated. The framework nests general forms of over- and underreaction that arise from distortions in the perceived laws of motion driving the aggregate economy, as well as specific belief formation frameworks such as inattention, which typically delivers underreaction (Sims (2003), Gabaix (2019)), or diagnostic expectations (DE), which delivers overreaction (Bordalo, Gennaioli and Shleifer (2018), Bordalo, Gennaioli, LaPorta and

Shleifer (2019), and Bordalo, Gennaioli, LaPorta and Shleifer (2022)), while retaining rational expectations (RE) as a special case. Fourth, we use the AI approach of Bianchi, Ludvigson and Ma (2022a) (BLM1) to construct a real-time RE benchmark designed to quantify subjective biases in human beliefs present in markets and surveys.

In the structural model studied here, multiple primitive macroeconomic risks are relevant for the subjective growth expectations that underpin shareholder value. News causes investors to revise their understanding of which primitive shocks moving the economy, recognizing that the fundamentals they drive interact dynamically. In this context, any distortions in perception about the laws of motion driving fundamentals operate, ipso facto, on a system of macro dynamics rather than on a univariate earnings or payout process and translate directly into distorted perceptions about the shocks. This implies that investors may misattribute one primitive impulse to a mixture others, with consequences for asset prices.

The multivariate generalization also gives rise to a key distinction between the present framework and typical univariate DE models where overreaction only amplifies market volatility (e.g., Bordalo et al. (2019), Bordalo et al. (2022)). Here, depending on the composition of perceived shocks that a news event elicits, DE-driven overreaction to each shock individually can imply that the stock market as a whole *under*reacts to real-world events, thereby *dampening* volatility. This happens because many real-world news events are associated with changes in more than one perceived shock with counteracting effects. As an example, suppose that the same degree of diagnosticity is applied to all perceived shocks while a news event elicits two that push the market in opposite directions (e.g., bad cash flow news paired with good discount rate news). Under these circumstances, diagnostic expectations can dampen the market's response because the resulting shock-specific overreactions–even if generated by a single, common DE parameter–can still be asymmetric across shocks that differ by their volatility and propagation properties. In general, non-standard predictions arise whenever overreaction to a positive (negative) force for the market partially offsets a lesser *relative* overreaction to a predominating negative (positive) force.

We begin the presentation of our findings with results from a preliminary analysis demonstrating the systematically superior forecasting performance of the machine algorithm compared to investor, analyst, and professional forecaster surveys. This motivates our use of the algorithm as a measure of non-distorted expectation formation. We then present our main findings, which can be summarized as follows.

First, while the structural estimation treats as equally likely the opposing belief formation frameworks of inattention (which would imply underreaction to news) and DE (which would imply overreaction), our parameter estimates imply that the representative investor exhibits

belief overreaction in a manner consistent with DE. The estimated baseline model with DE-type overreaction fits the post-war behavior of the stock market with little to no error.

Second, market reactions to big real-world news events sometimes exhibit overreaction as well, creating "excess" volatility in response to such events. This additional volatility is driven by the DE distortion and occurs when overreaction to each shock individually amplifies the effects of all shocks combined, in keeping with standard predictions of univariate DE models. For these events, markets often overreact strongly to movements in payout-relevant factors, intensifying the price response compared to the RE counterfactual. However, when several counteracting shocks happen simultaneously, we often find that DE leads instead to market underreaction, once the effects of overreaction to all shocks are netted out.

This brings us to our third result: The market's reaction to news is often characterized by strong DE-driven overreaction to good (bad) news about transitory payout-relevant factors that partially offsets a lesser relative overreaction to predominating bad (good) news stemming from other factors, while under RE any offsetting force would be small or absent entirely. This result–which dampens the market reaction relative to the fully rational response–is not attributable to inattention and occurs even though a single free parameter controls the magnitude of the DE distorted reactions to all shocks. Our estimates indicate that this type of asymmetric compositional effect describes stock market behavior in several episodes of post-millennial history, most notably the global financial crisis, in which behavioral overreaction was a force for stability rather than volatility.

Taken together, these findings imply that, when confronted with the sequence of estimated shocks that actually occurred over the post-millennial period, an empirically grounded rational expectations benchmark that allows discount rates to vary implies that the stock market should be highly volatile–more so than in data–thereby creating a puzzle of "excess stability" rather than excess volatility. Yet a model with DE-style belief overreaction can perfectly explain the data, not because it creates more volatility but because it creates less.

Relation to the Literature A large and growing body of literature studies overreaction in subjective expectations and its relation to stock market behavior (Barberis, Shleifer and Vishny (1998), Chen, Da and Zhao (2013), Bordalo et al. (2018), Bordalo, Gennaioli, Ma and Shleifer (2020), Bordalo et al. (2019), Nagel and Xu (2022), Bordalo et al. (2022), De La O and Meyers (2021, 2023) Hillenbrand and McCarthy (2021).) At the same time, other researchers have argued that at least some types of news are not tended to and thus met with underreaction (e.g., Mankiw and Reis (2002), Woodford (2002), Sims (2003), Gabaix (2019), Kohlhas and Walther (2021)). We extend these literatures by undertaking a structural estimation to freely estimate the direction and severity of a range of biases (if any) in

the stock market's reaction to hundreds of real-world news events, delineating the role of perceptions about multiple fundamental macro shocks in driving these reactions.

Other studies hypothesize that any link between subjectively expected future cash flow growth and stock price variation occurs because the former responds to the latter rather than drives it (Bastianello and Fontanier (2022), Chaudhry (2023), Jin and Li (2023)) or, relatedly, that unexplained flows in and out of the stock market–evidently disconnected from genuine cash flow news–are responsible for substantial stock market volatility (e.g., Gabaix and Koijen (2021), Hartzmark and Solomon (2022)). These studies often link price movements to innovations in instruments that would reasonably seem unrelated to any evident news, without taking a stand on what may have caused the price movement or flow to change in the first place. We take the complimentary and converse approach of explicitly modeling equilibrium price movements and then studying reactions to specific events in which relevant information not known previously to market participants was revealed. Since actual news causes adjustments in forward-looking asset prices only when investors' subjective expectations are revised, such reactions should be highly informative about investor beliefs.

We follow the tradition of many papers in using equity analysts' survey forecasts of earnings growth as one observable indicator of subjective cash flow expectations in our analysis. As emphasized by Adam and Nagel (2023), however, the extent to which equity analysts' forecasts are representative of broader market expectations remains an open question. The methodology adopted here takes a step toward addressing this limitation by employing a structural estimation that substantially broadens the set of observable indicators relevant for understanding investors' underlying beliefs.

The methodology of this paper builds off of the structural mixed-frequency approach of Bianchi, Ludvigson and Ma (2022b) (BLM2) for inferring what markets learn from news. Viewed through the lens of a multivariate structural model, news is a multifaceted signal about which primitive shocks are moving the economy. For example, a higher-than-expected inflation reading might cause investors to revise upward their perception of cost-push supply shocks (directly affecting inflation and thus real payout growth), of aggregate demand shocks (indirectly affecting inflation through GDP growth), and of mark-up/pricing power shocks (which may merely correlate with inflation but directly affect the earnings share of output and ultimately payout growth). However, unlike the present study, BLM2 investigates market reactions to news without addressing whether those reactions may be nonrational and if so why, a gap this paper fills.

Our approach, which integrates machine-learning insights from Bybee, Kelly, Manela and Xiu (2021), Gu, Kelly and Xiu (2020), and Cong, Tang, Wang and Zhang (2021), follows the

dynamic, real-time AI methodology of BLM1 for measuring systematic expectational errors. A foundational principle of this methodology recognizes that market participants have access to thousands of pieces of potentially relevant information in real time, the importance of which may vary dramatically over the course of a sample, while the canonical standard for rational expectation formation is predicated on the efficient use of all of it. It is in this real-world context that the dynamic machine algorithm of BLM1 is trained to measure ex ante expectational errors embedded in human judgments by efficiently coping with the problems of overfitting and structural change in a procedure designed to be free from look-ahead bias, i.e., without relying on in-sample estimation, data that would not have been known in real time, or specification choices that may improve fit but could only have been implemented with hindsight. Adherence to these principles is important to avoid overstating biases in real-world expectation formation implied by the structural model and is a central motivating precept of the structural-AI synthesis.

The next section presents a simplified framework to explain the key elements of our model and structural-AI synthesis approach. There we clarify that our methodology derives results that are conditional on a stylized structural model, but one that we explicitly treat as an approximation of reality. We show how the estimation step accommodates specification error present in the parametric representation of aggregate dynamics, in investor beliefs, and in machine beliefs. Sections 3 and 4 describe our machine learning and full structural model frameworks, respectively. Section 5 presents our estimation procedure and mapping to data for the full structural model. Section 6 presents our findings. Section 7 concludes. Throughout the paper we use lowercase letters to denote log variables, i.e., $d_t = \ln(D_t)$, and "$\sim$" to denote features of the model under the subjective beliefs of the investor that may depart from full rationality.

2 Simplified Model

Let real stock market payout, D_t, be a time-varying share K_t of real output Y_t, i.e., $D_t = K_t Y_t$. With arbitrary time-variation in K_t, the specification $D_t = K_t Y_t$ is a tautology. Empirically, however, log growth Δd_t is better described by the specification $d_t = k_t + y_t$ than by a univariate process for Δd_t, due to distinct trend/cycle components in k_t and y_t.[2]

[2]This is because k_t is stationary while y_t has a trend. The highest frequency variation in Δd_t is entirely driven by a transitory component of k_t, while its lowest frequency variation is entirely driven by trend growth in y_t. This is especially visible for earnings growth and the earnings share. In our estimation, we use data on S&P 500 earnings, corporate sector payout, and S&P 500 dividend shares of GDP, each of which are treated as noisy signals on K_t.

Consider a theoretical setting in which a representative investor forms subjective beliefs about log real stock market payouts, d, which follow the law of motion:

$$\begin{aligned} \Delta d_t &= \Delta y_t + k_t - k_{t-1} \\ k_t &= (1-\rho_k)k + \rho_k k_{t-1} + \varepsilon_{k,t} \\ \Delta y_t &= (1-\rho_{\Delta y})\Delta y + \rho_{\Delta y}\Delta y_{t-1} + \varepsilon_{\Delta y,t}. \end{aligned}$$

As explained below, our approach explicitly treats the above "objective" law of motion is an approximate description of the true data generating process, and takes this into account at the estimation stage. Write the above as a system in deviations from steady-state using "hats," i.e., $\hat{k}_t \equiv k_t - k$:

$$\underbrace{\begin{bmatrix} \hat{\Delta d}_{t+1} \\ \hat{k}_{t+1} \\ \hat{\Delta y}_{t+1} \end{bmatrix}}_{\hat{S}^M_{t+1}} = \underbrace{\begin{bmatrix} 0 & \rho_k - 1 & \rho_{\Delta y} \\ 0 & \rho_k & 0 \\ 0 & 0 & \rho_{\Delta y} \end{bmatrix}}_{T^M(\theta^M)} \underbrace{\begin{bmatrix} \hat{\Delta d}_t \\ \hat{k}_t \\ \hat{\Delta y}_t \end{bmatrix}}_{\hat{S}^M_t} + \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}}_{R^M} \underbrace{\begin{bmatrix} \varepsilon_{k,t+1} \\ \varepsilon_{\Delta y,t+1} \end{bmatrix}}_{\varepsilon^M_{t+1}}, \tag{1}$$

or, letting $\theta^M \equiv (\rho_k, \rho_{\Delta y})'$, in matrix notation as

$$\hat{S}^M_{t+1} = T^M\left(\theta^M\right)\hat{S}^M_t + R^M\varepsilon^M_{t+1}. \tag{2}$$

We assume that both k_t and Δy_t are stationary with $0 \leq \rho_k, \rho_{\Delta y} < 1$. The system (1) therefore implies that $\hat{\Delta d}_{t+1} = (\rho_k - 1)\hat{k}_t + \rho_{\Delta y}\hat{\Delta y}_t$ has both a negatively autocorrelated component originating from the payout share k_t, and a positively autocorrelated component originating from output growth Δy_t. It follows that a *negative* impulse to $\varepsilon_{k,t}$ implies $\hat{\Delta d}_{t+1} > 0$, i.e., *positive* catch-up growth next period, while a *positive* impulse to $\varepsilon_{k,t}$ implies $\hat{\Delta d}_{t+1} < 0$, i.e., *negative* fall-back growth next period.

We consider two types of distortion in investor beliefs about stock market fundamentals S^M_t. First, the perceived process for fundamentals growth may differ from (2) and instead evolve according to

$$\hat{S}^M_{t+1} = T^M\left(\widetilde{\theta}^M\right)\hat{S}^M_t + R^M\widetilde{\varepsilon}^M_{t+1}. \tag{3}$$

In (3), $T^M\left(\widetilde{\theta}^M\right)$ is the same as $T^M\left(\theta^M\right)$ except the investor's subjective value $\widetilde{\theta}^M \equiv$

$(\widetilde{\rho}_k, \widetilde{\rho}_{\Delta y})'$ of the persistence of fundamentals, θ^M, differs from its objective value, and the perceived shock vector $\widetilde{\varepsilon}_t^M$ could differ from the objective innovation ε_t^M.

Second, investor expectations may be subject to a time-varying distortion η_t, analogous to univariate models with diagnostic expectations (Bordalo et al. (2018), Bordalo et al. (2019), Bordalo et al. (2022)), here generalized to the multivariate system (3). As in those models, investors are unaware that they have a distortion but behave *as if* their subjective expectations $\widetilde{\mathbb{E}}_t[\cdot]$ were conditional on additional news contained here in the 3×1 vector η_t:

$$\widetilde{\mathbb{E}}_t\left[\hat{S}_{t+1}^M\right] = T^M\left(\widetilde{\theta}^M\right)\left(\hat{S}_t^M + \zeta\eta_t\right), \tag{4}$$

where $\eta_t \equiv \left(\eta_{\Delta d,t}, \eta_{k,t}, \eta_{\Delta y,t}\right)'$ has an innovation equal to perceived cash-flow news $\widetilde{\varepsilon}_t^M$. The scalar parameter ζ controls the magnitude and nature of the distortion and nests different models. If $\zeta > 0$, investor expectations overreact their perceived news as in models with DE or earlier models of belief overreaction (e.g., Barberis et al. (1998)). If $\zeta < 0$, investors underreact to perceived news, as in models with inattention (Sims (2003), Gabaix (2019)). In the remainder of this paper, we refer to η_t simply as the "DE distortion" for brevity, even though strictly speaking the reference to diagnostic expectations only applies when $\zeta > 0$. The empirical relevance of either type of distortion–captured by the sign and magnitude of ζ–will be subject to estimation in the full structural model.

Further generalizing univariate approaches to DE, we can allow excess optimism/pessimism to gradually revert over time with η_t following a VAR(1) (rather than AR(1)) process $\eta_t = \rho_\eta \widetilde{T}^M \eta_{t-1} + R^M \widetilde{\varepsilon}_t^M$, where $\widetilde{T}^M \equiv T^M\left(\widetilde{\theta}^M\right)$ and $0 \leq \rho_\eta < 1$. Note that when $\rho_\eta = 0$, $\eta_{\Delta d} = \widetilde{\varepsilon}_{k,t} + \widetilde{\varepsilon}_{\Delta y,t}$.

Equation (3) implies that the investor may misperceive the law of motion for cash flow growth. It is important to clarify that this does not mean the investor misperceives Δd_{t+1} itself, once observed. That is, investors do not suffer from delusions about the facts of cash flow growth once they learn those facts. What the distinction between (2) and (3) *does* imply is that investors may disagree with a fully rational agent about how they got to those facts.

Suppose for this simplified model that investors price in a constant risk-premium and risk-free rate r_f under their subjective beliefs. Let P_t^D denote the stock price and apply a Campbell and Shiller (1989) approximate present value identify based on a Taylor expansion of $r_{t+1}^D \equiv \ln(P_{t+1}^D + D_{t+1}) - \ln(P_t^D)$ around a point $P_t^D/D_t = PD$:

$$r_{t+1}^D = \kappa_{pd,0} + \beta pd_{t+1} - pd_t + \Delta d_{t+1}, \tag{5}$$

where r_t^D is the stock market return, $pd_t \equiv p_t^D - d_t$, $\beta \equiv \frac{PD}{1+PD}$, $\kappa_{pd,0} \equiv \ln(1+\exp(pd)) - \beta(pd)$, and $pd = \ln(PD)$. With the constant risk-premium and risk-free rate and imposing $\lim_{j\to\infty} \beta^j pd_{t+j} = 0$, the price-payout ratio is

$$pd_t = pd + \widetilde{\mathbb{E}}_t \sum_{h=0}^{\infty} \beta^h \hat{\Delta} d_{t+1+h} \tag{6}$$

$$= pd + \left(\frac{\widetilde{\rho}_{\Delta y}}{1-\widetilde{\rho}_{\Delta y}\beta}\right)\left(\hat{\Delta y}_t + \zeta\eta_{\Delta y,t}\right) + \left(\frac{\widetilde{\rho}_k - 1}{1-\widetilde{\rho}_k\beta}\right)\left(\hat{k}_t + \zeta\eta_{k,t}\right), \tag{7}$$

where $pd \equiv \left(\kappa_{pd,0} - r^D + \Delta d\right)/(1-\beta)$, with r^D equal to the constant subjectively expected return, and $\Delta d = \Delta y$ equal to steady-state payout growth. The price-payout ratio observed in the data thus reflects the investor's perceived law of motion and the time-varying distortion η_t. Combining (7) and (5) and noting that $\widetilde{\mathbb{E}}_t\left[\eta_{t+1}\right] \equiv 0$ because the agent is unaware of the distortion, we verify $\widetilde{\mathbb{E}}_t\left[r_{t+1}^D\right] = r^D$.

Now consider a benchmark for objective beliefs based on a machine learning algorithm that is by construction free from human cognitive biases and that "knows" (from a large sample of data) the approximating objective law of motion (2) and the asset pricing dynamics (7). (We discuss the sense in which the algorithm can "know" further below.) Taking expectations of (5) under these objective beliefs $\mathbb{E}_t[\cdot]$ yields

$$\mathbb{E}_t\left[r_{t+1}^D\right] = r^D + \left[\frac{\rho_{\Delta y} - \widetilde{\rho}_{\Delta y}}{1-\widetilde{\rho}_{\Delta y}\beta}\right]\hat{\Delta y}_t + \left[\frac{\beta\rho_{\Delta y}\rho_\eta - 1}{1-\widetilde{\rho}_{\Delta y}\beta}\right]\widetilde{\rho}_{\Delta y}\zeta\eta_{\Delta y,t}$$
$$+ \left[\frac{(\rho_k - \widetilde{\rho}_k)(1-\beta)}{1-\widetilde{\rho}_k\beta}\right]\hat{k}_t + \left[\frac{\beta\rho_k\rho_\eta - 1}{1-\widetilde{\rho}_k\beta}\right](\widetilde{\rho}_k - 1)\zeta\eta_{k,t} \tag{8}$$

While the subjective expected return of the investor is always r^D, the terms in square brackets give the impact on future returns that the machine can anticipate given knowledge of the investor's subjective beliefs and time-varying distortion η_t. The special case of rational expectations occurs in this simplified model when (i) $\zeta = 0$, and (ii) $\widetilde{T}^M = T^M$, in which case objective expected returns in (8) are always r^D, and investors rationally price in a constant risk-free rate and risk premium.

More generally, the machine will find returns predictable and the algorithm can be used to identify different systematic biases. To illustrate, consider the roles of different economic shocks in shaping subjective and objective expected returns. (The shocks themselves can be identified using the empirical methodology described below.) For ease of exposition in what follows, we set $\rho_\eta = 0$ in (4), implying $\eta_{k,t} = \widetilde{\varepsilon}_{k,t}$, $\eta_{\Delta y,t} = \widetilde{\varepsilon}_{\Delta y,t}$, and consider several shock

and parameterization cases below.

To start, suppose there is a one-unit negative impulse to $\widetilde{\varepsilon}_{k,t}$ from steady state–i.e., a negative shock the payout share–while $\widetilde{\varepsilon}_{\Delta y,t} = 0$. Case 1.1: let $\zeta > 0$ as in models with DE, while $\widetilde{\rho}_k = \rho_k = 0$. With $\widetilde{\varepsilon}_{k,t} < 0$ and $\zeta > 0$, the model implies belief overreaction to the decline in k_t, which generates excessive optimism about catch-up growth:

$$\left(\widetilde{\mathbb{E}}_t - \mathbb{E}_t\right)\left[\hat{\Delta d}_{t+1}\right] = -\zeta\widetilde{\varepsilon}_{k,t} > 0.$$

This inflates the price impact as investors behave as if they expect to rebound to a higher *level* of k_t. By contrast, under rational expectations the ex-dividend price would not respond at all to the purely i.i.d. shock. The inevitable investor disappointment in future growth causes a price reversal and lower future returns that the machine can predict, therefore (8) shows that $\mathbb{E}_t\left[r^D_{t+1}\right] = r^D + \zeta\widetilde{\varepsilon}_{k,t} < r^D$. Case 1.2: with $\widetilde{\varepsilon}_{k,t} < 0$ and $\zeta = 0$ while $\widetilde{\rho}_k > \rho_k$, there is no overreaction to the shock but the investor over-extrapolates today's lower payout share to the future, generating excessive pessimism about catch-up growth:

$$\left(\widetilde{\mathbb{E}}_t - \mathbb{E}_t\right)\left[\hat{\Delta d}_{t+1}\right] = \left(\widetilde{\rho}_k - \rho_k\right)\widetilde{\varepsilon}_{k,t} < 0.$$

In this case prices fall too much, and the machine can anticipate that investors will be positively surprised subsequently, leading to a price reversal and predictably higher future returns: $\mathbb{E}_t\left[r^D_{t+1}\right] = r^D + \left[\left(\rho_k - \widetilde{\rho}_k\right)(1-\beta)/(1-\widetilde{\rho}_k\beta)\right]\widetilde{\varepsilon}_{k,t} > r^D$.

We can also contemplate a one-unit negative impulse to $\widetilde{\varepsilon}_{\Delta y,t}$ from steady state, while $\widetilde{\varepsilon}_{k,t} = 0$. Case 2.1: with $\widetilde{\varepsilon}_{\Delta y,t} < 0$ and $\zeta > 0$ while $\widetilde{\rho}_{\Delta y} = \rho_{\Delta y} > 0$, the model implies belief overreaction to the decline in today's growth, resulting in excessive pessimism about subsequent growth:[3] $\left(\widetilde{\mathbb{E}}_t - \mathbb{E}_t\right)\left[\hat{\Delta d}_{t+1}\right] = \widetilde{\rho}_{\Delta y}\zeta\widetilde{\varepsilon}_{\Delta y,t} < 0$. Prices fall too much, and the machine can predict the price reversal as (8) shows it would forecast higher future returns. Case 2.2: if $\zeta = 0$, while $\widetilde{\rho}_{\Delta y} > \rho_{\Delta y}$, the investor over-extrapolates today's lower growth to the future, which likewise generates excessive pessimism about subsequent growth: $\left(\widetilde{\mathbb{E}}_t - \mathbb{E}_t\right)\left[\hat{\Delta d}_{t+1}\right] = \left(\widetilde{\rho}_{\Delta y} - \rho_{\Delta y}\right)\widetilde{\varepsilon}_{\Delta y,t} < 0$. Once more the machine would forecast higher future returns, as (8) shows.

We show below that most real-world news events elicit revisions in more than one perceived shock. Using the cases just described, we observe that if a news or economic event is associated with simultaneous negative impulses in $\widetilde{\varepsilon}_{k,t}$ and $\widetilde{\varepsilon}_{\Delta y,t}$, computing its overall

[3]Here we need $\widetilde{\rho}_{\Delta y} \neq 0$ because, as (4) shows, DE operates on news, which requires that the innovation have some predictability for future growth.

market impact requires that the above effects be combined. This gives rise to an important distinction with single-shock DE models that apply to a univariate earnings or payout process such as Bordalo et al. (2019), Bordalo et al. (2022). In particular, depending on how the composition of perceived shocks is revised in response to news, overreaction to each shock individually, i.e., $\zeta > 0$, can imply that the market as a whole *under*reacts to real-world news. This happens because the *relative* overreaction across shocks is not the same, owing to their distinct volatility and propagation properties. For example, consider the situation where cases 1.1 and 2.1 apply simultaneously. In case 1.1 DE causes p_t^D to rise too much, whereas there would be no price response to the purely i.i.d. impulse under RE. In case 2.1 DE causes prices to fall to much, whereas with $\widetilde{\rho}_{\Delta y} > 0$ there would be some decline under RE, albeit smaller. All combined, with DE the upside overreaction in 1.1 can under some parameter values partially offset the downside overreaction in 2.1, dampening the overall market response relative to RE. This empirically relevant example illustrates how asymmetric compositional effects can imply that DE stabilizes the market during events or crises with overreactions to multiple counteracting shocks.

How can we use these theoretical insights in a realistic empirical setting with observable data? A challenge with doing so is that the structural model is a stylized representation of reality subject to error, while the machine beliefs, survey forecasts, and other data are the product of much more complicated real-world phenomena. To this address this challenge, we exploit a state-space representation that treats the structural model as an explicit (reasonable) approximation, and the machine beliefs, survey forecasts, and other data as noisy signals that map onto the structural model equations.

To see how this works, consider the model solution for this simplified framework, which implies that the state vector $S_t = \left[\Delta d_t, k_t, \Delta y_t, pd_t, pd_{t-1}, r_t^D, \eta_t, S_t^*\right]'$ evolves according to a vector autoregression (VAR) state equation

$$S_t = C(\Theta) + T(\Theta)S_{t-1} + R(\Theta)Q\varepsilon_t^M,$$

where S_t^* is discussed below, C, T, R are matrices comprised of the model's primitive parameters $\Theta = \left(\rho_k, \rho_{\Delta y}, \widetilde{\rho}_k, \widetilde{\rho}_{\Delta y}, \zeta, r^D, \beta, \rho_\eta, k, \Delta y\right)'$, Q is a matrix of shock volatilites, and η_t is the latent DE distortion to be estimated. The relation between the variables of the model and data counterparts X_t can be written as a vector observation equation taking the form

$$X_t = D + ZS_t + Uv_t,$$

where D and Z are parameters, and v_t is a vector of observation errors with standard deviation U. The observation error v_t is important for modeling noise due to various sources–including as arises from the approximating structural model itself–and is discussed below. By combing $X_t = D + ZS_t + Uv_t$ with $S_t = C + TS_{t-1} + RQ\varepsilon_t^M$, the model parameters and theoretical states S_t can be estimated using state-space methods.

Investor expectations in the model are conditioned on the state vector S_t. However, some of its elements will be observable from historical data, and a subset of those are "noisy" because they undergo subsequent revision, so their real-time values will be of lesser quality than their final estimates. For example, asset price data p_t^D are not subject to revision, but d_t is *real* payout and must computed using data on inflation that is subject to revision. To better match the conditions of real-world decision making, we assume that investors operating in real time have access only to a noisy measure of any indicators subject to subsequent revision, and price assets on that basis. To illustrate the general approach, suppose for this simplified model we include data on k_t, Δy_t, pd_t, and returns r_t^D, and empirical forecasts of Δd_t and r_t^D from surveys and/or prediction markets that serve as noisy signals of investor expectations. In this approach, the vector of econometric observables X_t must also include machine forecasts to measure non-distorted beliefs, computed in a first-step as described below. Let $S_t^* = \left[\Delta d_t^*, k_t^*, \Delta y_t^*, pd_t^*, pd_{t-1}^*, r_t^{D*}\right]$ denote elements of S_t observed in real time.

Let $\mathbb{F}_t[y_{t+h}]$ generically denote a vector of observed subjective forecast measures made at time t of variable y at time $t+h$ measured from surveys, futures markets, etc., and let $\mathbb{E}_t^{ML}[y_{t+h}]$ denote an observed objective machine forecast produced in an outer estimation (discussed below). Let matrices with a subscript, e.g., Z_x, denote the parameter subvector of Z that when multiplied by S_t or S_t^* and added to $D_x + U_x v_{x,t}$ picks out the appropriate model variable to map back into empirical observations on x_t, e.g., $Z_k S_t$ simply picks out the element of S_t corresponding to k_t, $Z_{s^M} S_t$ picks out the subvector $[\Delta d_t, k_t, \Delta y_t]'$, $Z_{\mathbb{E}(r),\Delta y}$ is the first term in brackets in (8) showing the effect of Δy_t on $\mathbb{E}_t\left[r_{t+1}^D\right]$, and so on. The

observation equation $X_t = D + ZS_t + Uv_t$ takes the form

$$\begin{bmatrix} [\Delta d_t, k_t, \Delta y_t]' \\ pd_t \\ r_t^D \\ \mathbb{F}_t\left[\hat{\Delta} d_{t+1}\right] \\ \mathbb{E}_t^{ML}\left[\hat{\Delta} d_{t+1}\right] \\ \mathbb{F}_t\left[r_{t+1}^D\right] \\ \mathbb{E}_t^{ML}\left[r_{t+1}^D\right] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ r^D \\ r^D \end{bmatrix} + \begin{bmatrix} Z_{s^M} S_t \\ Z_{pd} S_t \\ Z_r S_t \\ \left((\widetilde{\rho}_k - 1)Z_k + \widetilde{\rho}_{\Delta y} Z_{\Delta y} + \zeta Z_\eta\right) S_t^* \\ \left((\rho_k - 1)Z_k + \rho_{\Delta y} Z_{\Delta y}\right) S_t^* \\ 0 \\ \left(Z_{\mathbb{E}(r),k} + Z_{\mathbb{E}(r),\Delta y} + \zeta Z_{\mathbb{E}(r),\eta}\right) S_t^* \end{bmatrix} + Uv_t. \tag{9}$$

The above illustrates key steps in our structural-AI synthesis approach:

1. **Historical data** $[\Delta d_t, k_t, \Delta y_t]'$ are mapped onto the model's approximating objective laws of motion $Z_{s^M} S_t$ to obtain best-fitting descriptions of data dynamics. Observation errorsv_t account for both estimation and specification error arising from the assumption that the stylized model is an approximation of more complicated (true) data generating processes.

2. **Real time data**. Investors and machine base their expectations on data available in *real time*, i.e., on S_t^*, as in real-world forecasting.

3. **Multiple signals identify** $\widetilde{\mathbb{E}}_t[\cdot]$. Subjective beliefs are identified from a range of empirical indicators including financial market variables (e.g., r_t^D and pd_t), survey and futures markets forecasts in $\mathbb{F}_t[\cdot]$, other forward-looking data, all treated as noisy signals of the underlying subjective expectations process $\widetilde{\mathbb{E}}_t[\cdot]$ of the investor. This can be seen from mappings r_t^D, pd_t, $\mathbb{F}_t\left[r_{t+1}^D\right]$, and $\mathbb{F}_t\left[\Delta d_{t+1}\right]$ above, all of which inform estimates of the parameters driving perceived dynamics and the DE distortion.[4] For example, multiple subjective expectations measures $\mathbb{F}_t\left[\hat{\Delta} d_{t+1}\right]$ map into $\left((\widetilde{\rho}_k - 1)Z_k + \widetilde{\rho}_{\Delta y} Z_{\Delta y} + \zeta Z_\eta\right) S_t^*$, informing estimates of $\widetilde{\rho}_k$, $\widetilde{\rho}_{\Delta y}$, and ζ.

4. **Machine forecasts identify** $\mathbb{E}_t[\cdot]$. Iterative machine forecasts $\mathbb{E}^{ML}[\cdot]$ designed to be free from look-ahead bias load on S_t^* but map into the objective law of motion to identify a notion of *real time* rational expectations, e.g., $\mathbb{E}_t^{ML}\left[\hat{\Delta} d_{t+1}\right]$ is mapped

[4]Short samples for survey expectations or other data are not technically a problem for this methodology since missing values are filled in with estimates using a filter and structural model.

into $\mathbb{E}_t[\hat{\Delta} d_{t+1}] = \left((\rho_k - 1)Z_k + \rho_{\Delta y} Z_{\Delta y} S_t^*\right)$. This ensures that estimates of $\mathbb{E}_t[\cdot]$ are consistent with an expectations process that does not assume knowledge investors could not possibly possess.

The above steps merit further discussion. The core of our structural AI synthesis is a strategy for using information from a high dimensional, nonparametric machine-based representation of objective beliefs to identify systematic expectational errors as represented in stylized parametric frameworks of human behavior. There are three links in this chain: (i) the machine representation of objective beliefs, (ii) the systematic expectational errors (if any) embedded in market behavior, and (iii) the stylized parametric framework of human behavior. We discuss each of them in turn focusing on their roles in providing a measure of objective expectation formation, a measure of distorted beliefs, and a conceptual framework.

(i) As explained in the next section, our machine benchmark for objective expectation formation $\mathbb{E}_t^{ML}[\cdot]$ is not based on any parametric model, but instead on a (nonparametric) multi-layer neural network used to estimated iteratively from a very large sample of real-time data inputs. Two aspects of this are central to our approach. First, the real time nature of this process is important because estimates of rational expectations that fail to take into account the conditions of real-world expectation formation are subject to over-fitting and hindsight bias, often leading the very behavioral distortions we are interested in measuring to be overstated (BLM1). Second, it is well known that a multi-layer neural network can approximate virtually any unknown function arbitrarily well given a large enough set of data inputs.[5] This clarifies the sense in which the machine can "know" the approximating objective law of motion of the stylized model: with access to large datasets to repeatedly estimate $\mathbb{E}_t^{ML}[\cdot]$ while using these as signals to map onto equations of stylized model parameters, an appropriate machine algorithm can effectively figure out unbiased estimates of the approximating objective laws of motion, *up to* an observation residual that accounts for estimation, specification, and data revision errors. This suggests that known big-data information-processing algorithms–using only data that agents verifiably could have possessed in real time–provide an appealing real time RE benchmark.

(ii) The representative investor that drives markets may have distorted beliefs, as captured by parameters driving wedges with the RE benchmark. Forward-looking data from multiple sources, including markets themselves, are used as signals on the representative investor's beliefs. However, both machine and investors must cope with estimation, specification, and data-revision errors, as well as with structural change in an evolving environment.

[5] See Hecht-Nielsen (1987) for the well-known Kolmogorov universal representation theorem that applies to arbitrary continuous functions and Ismailov (2023) for the theorem extending to discontinuous functions.

All these aspects represent noise in the model estimation that is subsumed in the observations errors of equations defining machine and investor beliefs.

(iii) As in any economic model, the parametric structural models above and in the next section serve the purpose of providing a conceptual framework for interpreting the data. Such frameworks are necessary to relate output to existing theoretical literatures and to make descriptive concepts precise.

Putting this all together, investor beliefs in the model can be thought of as the sum of the machine forecast, bias, and net noise. Conditional on a stylized parametric model, the procedure uses surveys and other forward-looking data to inform subjective parameters and distortions, machine forecasts inform objective parameters, and observation errors capture noise.

The final step in the empirical analysis is to measure market reactions to news, which we do by employing the mixed-frequency filtering algorithm developed in BLM2 to estimate revisions in investor perceptions in tight windows surrounding news events. The nature and severity of any behavioral biases in market reactions to news is estimated by comparing jumps in investor beliefs with those informed by objective machine learning expectations. This leads us to discuss machine beliefs, which are compiled from algorithmic output and produced in a first-stage for use in X_t.

3 Machine Learning

The objective of dynamic machine learning in our context is to provide a benchmark for unbiased, information-efficient expectation formation under the conditions of real-world decision making. Following BLM1, the algorithm has two important features. First, it must account for the data-rich environment in which investors operate. Second, it must form forecasts in an entirely ex ante way, i.e., without relying on ex post historical outcome data, in sample estimation, or specification choices that could only have been known or implemented with hindsight. This section provides an abbreviated description of the machine estimation. We refer the reader to BLM1 and the Online Appendix for further details.

We are interested in forming a machine expectation of a time series $y_{j,t+h}$ indexed by j whose value in period $h \geq 1$ the machine is asked to predict. The following machine specification is estimated over rolling samples:

$$y_{j,t+h} = G^e\left(\mathcal{X}_t, \boldsymbol{\beta}_{j,h,t}\right) + \epsilon_{jt+h}. \tag{10}$$

where $\mathcal{X}_t$ is a large input dataset available in real time including an intercept, and $G^e(\cdot)$ is a machine learning estimator that can be represented by a high dimensional set of finite-valued parameters $\boldsymbol{\beta}_{j,h,t}$.[6] With this estimator, we follow the six step algorithmic approach of BLM1: 1. Sample partitioning,[7] 2. In-sample estimation, 3. Training and cross-validation, 4. Grid and sample partition re-optimization, 5. Out-of-sample prediction, 6. Roll forward and repeat. Step 2 includes variable selection, shrinkage, and hyper-parameter tuning. The end product of this procedure is a time-series of objective time t machine "beliefs" about $y_{j,t+h}$, denoted $\mathbb{E}_t^{(i)}[y_{j,t+h}]$.

Two points about the algorithm bear emphasizing. First, the machine expectations are based on only that information at t that we can verify would have been available to investors in real time. Second, human *bias* in this methodology is a genuine ex ante expectational error, not an ex post forecast errors. In particular, bias in expectations is measured relative to the machine forecast, not relative to the ex post outcome. We show below that the machine achieves sizable reductions in the mean-square-forecast-errors relative to survey forecasts over an extended evaluation period. Additional analysis of the machine forecasts can be found in a companion paper.

4 Structural Model

We now apply the ideas presented for the simplified model to the full structural model. We work with a risk-adjusted log-linear approximation to the model, in which all random variables are conditionally log-normally distributed.

Macro Dynamics As above, let aggregate stock market payout, D_t, be a time-varying share K_t of real output Y_t, i.e., $D_t = K_t Y_t$. Macro dynamics are described by a series of equations for the nominal short rate i_t, general price inflation π_t, output growth Δy_t, and the log payout share of output $k_t \equiv d_t - y_t$. For each of these, we specify "trend" or "long-run"

[6] We use the Long Short-Term Memory (LSTM) deep sequence recurrent neural network estimator with N hidden layers $h_t^n \in \mathbb{R}^{D_{h^n}}$

$$G^{LSTM}(\mathcal{X}_t, \theta_{jh}) = \sum_{n=1}^{N} \underbrace{W^{(yh^n)}}_{1\times D_{h^n}} \underbrace{h_t^n(\mathcal{X}_t, \theta_{jh})}_{D_{h^n}\times 1} + \underbrace{b_y}_{1\times 1}.$$

[7] At time t, a prior training sample of size $\dot{T}$ is partitioned into two subsample windows: an "estimation" subsample consisting of the first T_E observations, and a hold-out "validation" sample of T_V subsequent observations so that $\dot{T} = T_E + T_V$.

components denoted with "bars" that evolve according to

$$\overline{x}_t = (1-\phi_x)\overline{x}_{t-1} + \phi_x x_t + \sigma_{\overline{x},\xi_t}\varepsilon_{\overline{x},t}, \quad \forall x = \{i, \pi, \Delta y, k\}, \tag{11}$$

where $\varepsilon_{\overline{x},t} \sim N(0,1)$ is an i.i.d. shock to the trend component of x with a time-varying volatility $\sigma_{\overline{x},\xi_t}$ discussed below, and ϕ_x is a parameter governing its persistence. We assume that i_t, π_t, Δy_t, and k_t vary around these components as follows.

We assume that the nominal short rate is set by the central bank and follows the process

$$i_t - i = (1-\psi_i)\left[\psi_\pi\left(\overline{\pi}_t - \pi\right) + \psi_{\Delta y}\left(\overline{\Delta y}_t - g\right)\right] + \psi_i\left(\overline{i}_{t-1} - i\right) + \sigma_{i,\xi_t}\varepsilon_{i,t}, \tag{12}$$

where $\varepsilon_{i,t} \sim N(0,1)$ is an i.i.d. monetary policy shock, and i, π, and g are parameters. The dynamics of inflation and output growth follow similar primitive processes:

$$\pi_t - \pi = \beta_{\pi,\pi}\left(\overline{\pi}_{t-1} - \pi\right) + \beta_{\pi,\Delta y}\left(\overline{\Delta y}_t - g\right) + \beta_{\pi,i}\left(\overline{i}_{t-1} - i\right) + \sigma_{\pi,\xi_t}\varepsilon_{\pi,t} \tag{13}$$

$$\Delta y_t - g = \beta_{\Delta y,\pi}\left(\overline{\pi}_{t-1} - \pi\right) + \beta_{\Delta y,\Delta y}\left(\overline{\Delta y}_{t-1} - g\right) + \beta_{\Delta y,i}\left(\overline{i}_{t-1} - i\right) + \sigma_{\Delta y,\xi_t}\varepsilon_{\Delta y,t}, \tag{14}$$

where $\beta_{i,j}$ are parameters and $\varepsilon_{\pi,t} \sim N(0,1)$ and $\varepsilon_{\Delta y,t} \sim N(0,1)$ are i.i.d. shocks that represents short-run variation in these variables. The log payout share, k_t, is modeled as a primitive process following:

$$k_t - k = \rho_{k,k}\left(\overline{k}_{t-1} - k\right) + \beta_{k,\Delta y}\left(\overline{\Delta y}_t - g\right) + \sigma_{k,\xi_t}\varepsilon_{k,t}, \tag{15}$$

where $\varepsilon_{k,t} \sim N(0,1)$ is an i.i.d. shock.[8]

We refer to i.i.d. innovations without the bars–which capture the more cyclical components in the series–by their component names (e.g., $\varepsilon_{k,t}$ is the "payout share shock"), and those with the bars as the trend component shocks (e.g., $\varepsilon_{\overline{k},t}$is the "trend payout share shock"). It should be kept in mind, however, that the "trend" components are latent random variables that are hybrids of i.i.d. and persistent processes and are contemporaneously correlated with multiple economic variables in the simultaneous system above. We use these hybrid specifications to introduce parsimoniously parameterized persistence in the variables in a manner similar to a vector autoregression, but with fewer estimable parameters.

[8]This specification for macro dynamics is consistent with a triangular identification strategy for monetary policy shocks.

The shock volatilities in all of primitive processes above vary with the discrete valued random variable ξ_t, which evolves according to a N-state Markov-switching process with transition matrix $\mathbf{H}$. Collect the parameters $\psi_i, \phi_\pi, ...$ etc., of the above equations including $\mathbf{H}$ into a vector θ^M. Equations (12)-(15), along with the expression for payout growth, $\Delta d_t = \Delta k_t + \Delta y_t$, represent a macro-dynamic system that can be expressed as a Markov-switching vector autoregression (MS-VAR) law of motion (LOM) taking the form:

$$S_t^M = C^M\left(\theta^M\right) + T^M(\theta^M)S_{t-1}^M + R^M(\theta^M)Q_{\xi_t}^M\varepsilon_t^M, \tag{16}$$

where $S_t^M \equiv \left[\Delta y_t, \overline{\Delta y}_t, \Delta d_t, \pi_t, \overline{\pi}_t, i_t, \overline{i}_t, k_t, \overline{k}_t\right]'$, $C^M(\cdot)$, $T^M(\cdot)$, $R^M(\cdot)$ are matrices of primitive parameters θ^M, $\varepsilon_t^M = \left[\varepsilon_{\Delta y,t}, \varepsilon_{\overline{\Delta y},t}, \varepsilon_{\pi,t}, \varepsilon_{\overline{\pi},t}, \varepsilon_{i,t}\varepsilon_{\overline{i},t}, \varepsilon_{k,t}, \varepsilon_{\overline{k},t}\right]'$ is a vector of primitive macro shocks, and $Q_{\xi_t}^M(\cdot)$ is a diagonal matrix of shock volatilities that varies stochastically with ξ_t. Due to the endogeneity of macro dynamics, $R^M(\cdot)$ has non-zero off diagonal elements, implying that multiple fundamental shocks affect a single state variable.

Perceived Macro Dynamics Investors have subjective beliefs $\widetilde{\theta}^M$ about the parameters governing macro dynamics in (12)-(15) that could differ from the objective θ^M. Let these differences be captured by a wedge vector w_θ: $\widetilde{\theta}^M = \theta^M + w_\theta$. We assume that investors apply these perceived dynamics to a noisy measure of S_t^M that they observe in real time, denoted S_t^{M*}. The two are related by $A_o S_t^{M*} = A_o S_t^M + Q^v\varepsilon_{v,t}$, where $\varepsilon_{v,t} \sim N(0,1)$ is an i.i.d. "vintage" error attributable to data revisions.[9] Elements of S_t^M for which there is no post-publication revision are assumed to have no such vintage errors. Investors take S_t^{M*} as given and price assets accordingly.[10] Taken together, these assumptions imply that investors' perceived counterpart to (16) takes the form

$$S_t^{M*} = C^M\left(\widetilde{\theta}^M\right) + T^M\left(\widetilde{\theta}^M\right)S_{t-1}^{M*} + R^M\left(\widetilde{\theta}^M\right)\widetilde{Q}_{\xi_t}^M\widetilde{\varepsilon}_t^M \tag{17}$$

$$S_t^{M*} \equiv \left[\Delta y_t^*, \overline{\Delta y}_t^*, \Delta d_t^*, \pi_t^*, \overline{\pi}_t^*, i_t^*, \overline{i}_t^*, k_t^*, \overline{k}_t^*\right]' \tag{18}$$

$$\widetilde{\varepsilon}_t^M \equiv \left[\widetilde{\varepsilon}_{\Delta y,t}, \widetilde{\varepsilon}_{\overline{\Delta y},t}, \widetilde{\varepsilon}_{\pi,t}, \widetilde{\varepsilon}_{\overline{\pi},t}, \widetilde{\varepsilon}_{i,t}\widetilde{\varepsilon}_{\overline{i},t}, \widetilde{\varepsilon}_{k,t}, \widetilde{\varepsilon}_{\overline{k},t}\right]', \tag{19}$$

[9] The A_o matrix emphasizes that vintage errors can be on a linear combination of elements of S_t^{M*} and/or that they apply only to specific elements.

[10] This treats S_t^{M*} as an unbiased signal of the underlying "true" state vector S_t^M that is precise enough to reasonably ignore any uncertainty about the signal when pricing assets.

where $\widetilde{\varepsilon}_t^M$ is a vector of perceived primitive macroeconomic shocks. The perceived volatilities $\widetilde{Q}_{\xi_t}^M$ of these shocks vary with the same discrete valued random variable ξ_t but have a perceived transition matrix $\widetilde{\mathbf{H}}$ that may differ from $\mathbf{H}$. In this model, the primitive shocks perceived by the investor–all of which have some relevance for shareholder cash flow growth– are no longer a scalar but instead a vector $\widetilde{\varepsilon}_t^M$ that can differ from ε_t^M. Moreover, as in the simplified model, $\widetilde{R}^M$is neither square nor diagonal, so distorted beliefs translate directly into distorted perceptions about the shocks, leading investors to misattribute a change in one primitive shock to a mixture of others.

Let $\widetilde{T}^M \equiv T^M\left(\widetilde{\theta}^M\right)$ and analogously for $R^M\left(\widetilde{\theta}^M\right)$ and $C^M\left(\widetilde{\theta}^M\right)$. As in the simplified model, investors may exhibit a time-varying DE distortion η_t such that subjective expectations follow:

$$\widetilde{\mathbb{E}}_t\left[S_{t+h}^{M*}\right] = C_h^M\left(\widetilde{\theta}^M\right) + \left[T^M\left(\widetilde{\theta}^M\right)\right]^h S_t^{M*} + \left[T^M\left(\widetilde{\theta}^M\right)\right]^h \zeta\eta_t \tag{20}$$

where $C_h^M\left(\widetilde{\theta}^M\right) \equiv \widetilde{C}^M + \widetilde{T}^M\widetilde{C}^M + \left[\widetilde{T}^M\right]^2\widetilde{C}^M + ... + \left[\widetilde{T}^M\right]^{h-1}\widetilde{C}^M$. The scalar parameter ζ governs the strength of the over- or underreaction to all shocks, with $\zeta > 0$, implying overreaction, and $\zeta < 0$ implying underreaction. As above, the distortion η_t follows a VAR(1) process, with an innovation that is proportional to the vector of perceived news $\widetilde{\varepsilon}_t^M$:

$$\eta_t = \rho_\eta\widetilde{T}^M\eta_{t-1} + \widetilde{R}^M\widetilde{Q}_{\xi_t}^M\widetilde{\varepsilon}_t^M, \qquad \rho_\eta \in [0,1]\,. \tag{21}$$

Thus, η_t is a vector with elements comprised of unique decaying sums of multiple past perceived innovations $\left\{\widetilde{\varepsilon}_t^M, \widetilde{\varepsilon}_{t-1}^M, \widetilde{\varepsilon}_{t-2}^M, ...\right\}$.

Asset Pricing Dynamics The economy is populated by a continuum of identical investors who earn all income from trade in a stock market and a one-period nominal risk-free bond in zero net supply. Assets are priced by a representative investor who consumes per-capita aggregate shareholder payout, $D_t = K_tY_t$.

The representative investor's intertemporal marginal rate of substitution in consumption is the stochastic discount factor (SDF) with logarithm:

$$m_{t+1} = \ln\left(\beta_p\right) + \vartheta_{pt} - \gamma_{ra}\left(\Delta d_{t+1}\right). \tag{22}$$

where γ_{ra} is a curvature parameter and where the time discount factor is subject to an

aggregate externality in the form of a patience shifter ϑ_{pt} that individual investors take as given.[11] A time-varying specification for the subjective time-discount factor is essential for ensuring that investors are willing to hold the nominal bond at the interest rate set by the central bank's policy rule.

The first-order-condition for optimal holdings of the one-period nominal risk-free bond with a face value equal to one nominal unit is

$$LP_t^{-1} Q_t = \widetilde{\mathbb{E}}_t \left[M_{t+1} \Pi_{t+1}^{-1} \right], \tag{23}$$

where Q_t is the nominal bond price, $\widetilde{\mathbb{E}}_t$ denotes the subjective expectations of the investor, and $\Pi_{t+1} = P_{t+1}/P_t$ is the gross rate of general price inflation. We assume that investors have a time-varying preference for nominal risk-free assets over equity, accounted for by the term $LP_t > 1$ in (23), implying that the bond price Q_t is higher than it would be absent these benefits, i.e., when $LP_t = 1$. Taking logs of (23) and using the properties of conditional log-normality delivers an expression for the real interest rate as perceived by the investor:

$$i_t - \widetilde{\mathbb{E}}_t \left[\pi_{t+1} \right] = -\widetilde{\mathbb{E}}_t \left[m_{t+1} \right] - .5 \widetilde{\mathbb{V}}_t \left[m_{t+1} - \pi_{t+1} \right] - lp_t \tag{24}$$

where the nominal interest rate $i_t = -\ln(Q_t)$, $\pi_{t+1} \equiv \ln(\Pi_{t+1})$ is net inflation, $\widetilde{\mathbb{V}}[\cdot]$ is the conditional variance under the subjective beliefs of the investor, and $lp_t \equiv \ln(LP_t) > 0$. Variation in lp_t follows an AR(1) process

$$lp_t - \overline{lp} = \rho_{lp} \left(lp_{t-1} - \overline{lp} \right) + \sigma_{lp,\xi_t} \varepsilon_{lp,t} \tag{25}$$

subject to an i.i.d. shock $\varepsilon_{lp,t} \sim N(0,1)$. Since lp_t is a component of preferences, distorted perceptions play no role in (25).

Let P_t^D denote total value of market equity. Using (5), $pd_t \equiv \ln\left(P_t^D / D_t\right)$ obeys the

[11] This specification for ϑ_{pt} is a generalization of those considered in previous work (e.g.,Ang and Piazzesi (2003); Campbell and Cochrane (1999); Lettau and Wachter (2007)). Combining (24) and (22), we see that $\vartheta_{p,t}$ is implicitly defined as

$$\vartheta_t^p = -\left[i_t - \widetilde{\mathbb{E}}_t \left[\pi_{t+1} \right] \right] + \widetilde{\mathbb{E}}_t \left[\gamma_{ra} \Delta d_{t+1} \right] - .5 \widetilde{\mathbb{V}}_t \left[-\gamma_{ra} \Delta d_{p,t+1} - \pi_{t+1} \right] - lp_t - \ln\left(\beta_p\right).$$

following approximate log Euler equation:

$$\begin{aligned} pd_t &= \kappa_{pd,0} + \widetilde{\mathbb{E}}_t \left[m_{t+1} + \Delta d_{t+1} + \beta pd_{t+1} \right] + \\ &\quad + .5\widetilde{\mathbb{V}}_t \left[m_{t+1} + \Delta d_{t+1} + \beta pd_{t+1} \right]. \end{aligned} \tag{26}$$

Rewriting as a function of r^D_{t+1} and subtracting off (24), the log equity premium as perceived by the investor is:

$$\underbrace{\widetilde{\mathbb{E}}_t \left[r^D_{t+1} \right] - \left(i_t - \widetilde{\mathbb{E}}_t \left[\pi_{t+1} \right] \right)}_{\text{subj. equity premium}} = \underbrace{\begin{bmatrix} -.5\widetilde{\mathbb{V}}_t \left[r^D_{t+1} \right] - \widetilde{\mathbb{COV}}_t \left[m_{t+1}, r^D_{t+1} \right] \\ +.5\widetilde{\mathbb{V}}_t \left[\pi_{t+1} \right] - \widetilde{\mathbb{COV}}_t \left[m_{t+1}, \pi_{t+1} \right] \end{bmatrix}}_{\text{subj. risk premium}} + \underbrace{lp_t}_{\text{liquidity Premium}}, \tag{27}$$

where $\widetilde{\mathbb{COV}}_t \left[\cdot \right]$ is the conditional covariance under the subjective beliefs of the investor. The subjective equity premium has two components. The component labeled "subj. risk premium" is attributable to the agent's subjective perception of the quantity of risk, which varies in the model with fluctuations in the stochastic volatilities of the macro shocks, driven by ξ_t. The term labeled "liquidity premium" comes from the time-varying preference for risk-free nominal debt over equity. We treat this component as a latent random variable to be estimated that captures fluctuations in the pricing of risk due to factors not explicitly modeled, such as time variation in sentiment or implied risk aversion (e.g., from leverage constraints), flights to quality, or changes in the perceived liquidity and safety attributes of nominal risk-free assets (e.g., Krishnamurthy and Vissing-Jorgensen (2012)).

Equilibrium An equilibrium is defined as a set of prices (bond prices, stock prices), macro quantities (interest rates, inflation, output growth), laws of motion, and investor beliefs such that macro dynamics in (11)-(15) and thus (16) are satisfied, asset pricing dynamics in (22)-(26) are satisfied, with investor beliefs characterized by (17), (20) and (21).

Model Solution We solve the system of structural model equations that must hold in equilibrium, where agents form expectations taking into account the probability of regime change ξ_t, using standard algorithms that preserve log-normality of the entire system. Details are provided in the Online Appendix.

Let $S^A_t \equiv \left[m_t, pd_t, lp_t, \widetilde{\mathbb{E}}_t \left(m_{t+1} \right), \widetilde{\mathbb{E}}_t \left(pd_{t+1} \right) \right]$ be a set of asset pricing state variables describing the dynamics of (22)-(26), and let $S_t \equiv \left[S^M_t, S^{M*}_t, S^A_t, \widetilde{\varepsilon}^M_t, \eta_t \right]'$. The solution to

the complete structural model can be expressed as a MS-VAR in S_t:

$$S_t = \bar{C}\left(\theta_{\xi_t}, \widetilde{\theta}_{\xi_t}\right) + \bar{T}\left(\theta_{\xi_t}, \widetilde{\theta}_{\xi_t}\right) S_{t-1} + \bar{R}\left(\theta_{\xi_t}, \widetilde{\theta}_{\xi_t}\right) Q_{\xi_t}\varepsilon_t, \tag{28}$$

where $\bar{C}(\cdot)$, $\bar{T}(\cdot)$, $\bar{R}(\cdot)$ are matrices of primitive parameters involving elements of θ_{ξ_t} and $\widetilde{\theta}_{\xi_t}$, some of which vary with the Markov-switching variable ξ_t, and $Q_{\xi_t}(\cdot)$ is a matrix of shock volatilities that vary stochastically with ξ_t. The structural shocks are contained in $\varepsilon_t = \left(\varepsilon_t^M, \varepsilon_{lp,t}, \varepsilon_{v,t}\right)'$, which stacks the primitive macro shocks ε_t^M, the liquidity premium shock $\varepsilon_{lp,t}$ (a feature of preferences), and the vintage errors $\varepsilon_{v,t}$.[12]

5 Estimation and Mapping to Data

State-Space Estimation and Filter The system of estimable equations is placed in state-space form by combining the state equation (28) with an observation equation taking the form

$$\begin{aligned} X_t &= D_{\xi_t,t} + Z_{\xi_t,t} S_t' + U_t v_t \\ v_t &\sim N(0, I), \end{aligned} \tag{29}$$

where X_t denotes a vector of observable data and machine forecasts at time t, v_t is a vector of observation errors, U_t is a diagonal matrix with the standard deviations of v_t on the main diagonal, and $D_{\xi_t,t}$, and $Z_{\xi_t,t}$ are parameters that mapX_t into corresponding theoretical counterparts, which are functions of S_t. The parameters $Z_{\xi_t,t}$, U_t, and $D_{\xi_t,t}$ depend on t independently of ξ_t because some series in X_t are not available at all frequencies and/or over the full sample. As a result, the state-space estimation uses different measurement equations to include these series when the data are available, and exclude them when they are missing.

We estimate the state-space representation using Bayesian methods based on a modified version of Kim's (Kim (1994)) basic filter and approximation to the likelihood for Markov-switching state space models, specifying two volatility regimes (high/low). A random-walk metropolis Hastings MCMC algorithm is used to characterize uncertainty. The model parameters can be estimated on mixed-frequency monthly, quarterly, and biannual data and, following BLM2, used subsequently along with high-frequency forward-looking data to con-

[12]Neither $\widetilde{\varepsilon}_t^M$ or η_t appear separately in ε_t because $\widetilde{\varepsilon}_t^M = \left(\widetilde{R}^M \widetilde{Q}^M\right)^{-1}\left(S_t^{M*} - \widetilde{C}^M - \widetilde{T}^M S_{t-1}^{M*}\right)$ is entirely pinned down S_t^{M*} (and thus by ε_t^M and $\varepsilon_{v,t}$), while η_t has an innovation that is proportional to $\widetilde{\varepsilon}_t^M$.

duct an event study in tight windows around news events to characterize market reactions to news. We outline this procedure below.

Priors A complete description of the priors is provided in the Online Appendix.[13] Here we discuss priors on parameters governing investor beliefs. For the wedge vector $w_\theta \equiv \widetilde{\theta}^M - \theta^M$, we use a prior that is Normal, centered on zero, with standard deviation $\pm 5\%$ deviation from the objective parameter, i.e., $\widetilde{\theta} = \theta(1 + w_\theta)$ where $w_\theta \sim N(0, .05^2)$. For the parameter ζ governing the extent to which investors over- or underreact to perceived shocks, we use a prior that is Normal, centered on zero, with informative but loose tightness set to unit standard deviation to achieve modest shrinkage. Importantly, the priors for all of these parameters are *symmetric*, i.e., centered on zero, and are therefore without bias regarding the nature of the distortion. This is essential for our investigation because whether $\widetilde{\theta} \gtrless \theta$ could have quite different consequences for asset pricing dynamics. Likewise, $\zeta > 0$ would imply that investors overreact to perceived shocks, while $\zeta < 0$ would imply underreaction. In both cases, our estimation treats these polar parametric possibilities as equally likely and accordingly ensures that that both their sign and magnitude are approached as open empirical questions to be investigated.

Machine Expectations The structural-AI synthesis requires that machine forecasts of different variables be included in X_t.[14] As discussed above, machine forecasts are mapped into theoretical counterparts that load on a subvector of S_t, denoted $S_t^{MF} \equiv \left[S_t^{M*}, S_t^A, \widetilde{\varepsilon}_t^M, \eta_t\right]'$, that excludes S_t^M, thereby forcing the measure of non-distorted expectation formation to rely solely on real-time data S_t^{M*}. But, unlike investor expectations, mappings for machine expectations reflect the assumption that the machine has learned (from access to a large sample of data) the objective laws of motion for macro dynamics and the implied asset pricing dynamics (22)-(27). Thus machine forecasts of macro fundamentals are based on forward iterations of:

$$S_t^{M*} = C^M\left(\theta^M\right) + T^M(\theta^M) S_{t-1}^{M*} + R^M(\theta^M) Q_{\xi_t}^M \varepsilon_t^{M*}, \tag{30}$$

which twists estimates of θ^M in the objective LOM (16) toward values consistent with the machine forecasts. The resulting estimator of θ^M therefore strikes a balance between providing a good description of historical data and ensuring that the parameters describing the

[13] Priors for most parameters are standard and specified to be loosely informative except where stronger restrictions are dictated by theory, e.g., risk aversion must be non-negative.

[14] We use machine forecasts of excess stock market returns, S&P 500 earnings growth, GDP growth, and inflation in our estimation–see below.

objective data evolution are free overfitting and look-ahead bias characteristic of a purely ex post estimation.

Inferring Belief Reactions to News To infer how investor beliefs are affected by news, we apply the high frequency filtering algorithm developed in BLM2 for inferring jumps in investor perceptions about the current economic state, in tight windows around news events. To do this, we assume that investors can observe monthly values for S_t^{M*} and the volatility regime ξ_t at the *end* of each month even though they price assets continuously. It follows that news arriving *within* the month must lead to *updated beliefs* about the end-of-month values of S_t^{M*} and ξ_t investors expect to prevail.[15] We refer to these intramonth updates in beliefs as revisions in *nowcasts*, which are are equivalent to revisions in perceived shocks and the objects the filtering algorithm is designed to estimate. We discuss the procedure briefly below, leaving detailed coverage of the general approach to BLM2.

Data and Measurement The meta data-set used for this project consists of thousands of economic time series at mixed sampling intervals and spans the period January 1961 through December 2021. For the structural estimation, the observation vector often uses multiple noisy signals of the objective underlying theoretical concept. The measurement equation allows for observation errors in order to soak up variation in the signal that does not move identically with the theoretical concept. In what follows, we provide a brief summary of the data and how it is used. A complete description of the data, sources, and mapping to the model is provided in the Online Appendix.

Data used in structural estimation We estimate the model structural parameters on data at monthly or lower frequency sampling intervals (as available) from 2001:01-2021:12. Many series are used because they have obvious model counterparts, e.g., GDP growth, CPI inflation, the federal funds rate (FFR), stock market returns, the S&P 500 market capitalization. We use real-time versus historical versions of these, as appropriate for mappings involving S_t^{M*} versus S_t^M. The ratios of U.S. corporate sector payout-to-GDP, S&P 500 earnings-to-GDP and S&P 500 dividends-to-GDP are all used as noisy signals on the payout share K_t. Investor expectations over multiple horizons are informed by (i) surveys of expectations on future stock returns from UBS/Gallup, the Michigan Survey of Consumers (SOC), the Conference Board (CB), the CFO Survey from the Richmond Federal Reserve

[15] Investors can observe the objective volatility regime sequence $\{\xi_t, \xi_{t-1}, ...\}$ at the end of each t, but their perceived volatilities $\widetilde{Q}_{M,\xi_t}$ may still differ from the objective Q_{M,ξ_t}.

Bank, and the Consensus Forecasts of the S&P 500 index from Bloomberg (BBG), (ii) equity analysts' S&P 500 one-year-ahead earnings growth expectations from IBES and their long-term-growth expectations using the LTG expectation variable (see below), (iii) dividend growth expectations using S&P dividend futures data following the procedure of Gormsen and Koijen (2020), (iv) expectations of future inflation and GDP growth from the Survey of Professional Forecasters (SPF), BBG, the Livingston (LIV) Survey (inflation only), and the Blue Chip (BC) Survey, (v) interest rate expectations using Federal Funds Futures (FFF), Eurodollar (ED) futures, both at multiple contract horizons, and the Blue Chip (BC) survey expectations of the FFR 12 months ahead.[16] Data on the spread between the Baa corporate bond return and the 20-year Treasury bond return ("Baa spread" hereafter) are used as a noisy signal on the liquidity premium, lp_t.

Data used for news events and high-frequency filtering To estimate news-driven revisions in perceptions of the economic state S_t^{M*}, we use pre- and post-news event observations on a subset of the above series available at high frequency. These include tick level data on stock returns, the S&P 500 market capitalization, FFF and ED contract rates with different expiries, daily BBG survey expectations on multiple variables, and daily data on the Baa spread.[17] In our analysis, the pre-event value is always either 10 minutes before or the day before the news event, depending on data availability (daily versus minutely/tick level), and the post-event value is either 20 minutes after or the day after. Our sample of news events contain (i) 1,482 macroeconomic data releases for GDP, the Consumer Price Index (CPI), U.S. unemployment, and U.S. payroll data spanning the period 1980:01-2021:12, (ii) 16 corporate earnings announcement days spanning 1999:03-2020:05, and (iii) 219 Federal Open Market Committee (FOMC) press releases from the Fed spanning 1994:02-2021:12. The corporate earnings news events are from Baker, Bloom, Davis and Sammon (2019) who conduct textual analyses of *Wall Street Journal* articles to identify days in which there were large jumps in the aggregate stock market that could be attributed primarily to corporate earnings news with high confidence.[18] We run the filter to obtain estimates of S_t at high and low frequency from 2008:01-2021:12.

[16] In principle, fed funds futures market rates may contain a risk premium that varies over time. If such variation exists, it is absorbed in the estimation by the observation error for these equations (Piazzesi and Swanson (2008)).

[17] For events that occur when the market is closed we use minutely data on the S&P 500 E-mini futures market.

[18] Baker et al. (2019) (BBDS) examine next-day newspaper accounts of big daily moves ("jumps") in the stock market. Trained human readers classify the proximate cause of each jump into distinct categories and code the confidence with which the journalist advances an explanation for the jump. We are grateful to the authors of Baker et al. (2019) for providing us with their data for these event days.

Data inputs for machine learning algorithm The machine learning algorithm used to produce dynamic machine expectations uses thousands of initial data inputs, many of which are converted to diffusion index factors before being passed to the machine estimator following BLM1. The initial data inputs include a real-time macro dataset on 92 indicators, a panel dataset of 147 monthly financial indicators, and daily "up-to-the-forecast" financial market information from five broad classes of financial assets: (i) commodities prices (ii) corporate risk variables including credit spreads (iii) equities (iv) foreign exchange, and (v) government securities. A number of other inputs are used, including consensus forecast surprises, market jumps around past news events, and daily text-based factors estimated by Latent Dirichlet Allocation (LDA) analysis from around one million articles published in the *Wall Street Journal* between January 1984 to June 2022.[19]

6 Results

This section presents our estimation results. We begin with preliminary analysis of machine forecasting performance as compared to surveys and other benchmarks.

Preliminary Analysis Table 1 summarizes the forecasting performance of the machine relative to a survey or other benchmark for several different variables. Relative forecasting performance is reported as the ratio of the mean-square-forecast-error (MSE) of the machine, "$\text{MSE}_{\mathbb{E}}$" to that of the benchmark "$\text{MSE}_{\mathbb{F}}$". The sample for these comparisons varies depending on the availability of the survey data.

Panel (a) of Table 1 considers forecasts of the stock market. Surveys vary according to whether they ask about returns or price growth. To compare survey with machine forecasts, we compute machine forecasts for either the annual Center for Research in Securities Prices value-weighted return series (CRSP-VW, includes dividends) or S&P 500 price index growth, depending on which variable most closely aligns with the concept that survey respondents are asked to predict (see the Online Appendix for discussion of survey-by-survey conceptual alignment). Comparisons between machine and survey forecasts are made for four surveys of the stock market: the University of Michigan's Survey of Consumers (SOC), the Gallup/UBS poll of corporate executives, the CFO survey of corporate executives, and the Livingston Survey of forecast professionals.[20] We also compare the machine performance to a benchmark

[19]The results here are based a randomly selected sub-sample of 200,000 articles over the same period. This procedure follows Bybee et al. (2021), and estimates topic weights for individual articles to construct a time series of news attention by topic.

[20]The SOC surveys are qualitative up/down forecasts that we convert to point forecasts using the method-

Table 1: Forecasting Performance: Machine v.s. Benchmarks

ML: $y_{j,t+h} = G_{jh\mathbb{M}}(\mathbf{Z}_t) + \epsilon_{jt+h}$				
(a) CRSP Stock Excess Returns				
Horizon h quarters	4	4	4	4
Benchmark	SOC	Gallup/UBS	CFO	Livingston
$\text{MSE}_{\mathbb{E}}/\text{MSE}_{\mathbb{F}}$ (LSTM)	0.830	0.723	0.808	0.774
OOS R^2	0.170	0.277	0.193	0.226
Horizon h quarters	4	8	12	
Benchmark	Mean	Mean	Mean	
$\text{MSE}_{\mathbb{E}}/\text{MSE}_{\mathbb{F}}$ (LSTM)	0.680	0.535	0.532	
OOS R^2	0.320	0.465	0.468	
(b) S&P 500 Earnings Growth				
Horizon h quarters	4	8	12	LTG
Benchmark	IBES	IBES	IBES	IBES
$\text{MSE}_{\mathbb{E}}/\text{MSE}_{\mathbb{F}}$ (LSTM)	0.534	0.507	0.450	0.846
OOS R^2	0.466	0.493	0.550	0.154
(c) 1-year ahead Inflation and GDP Growth				
	Inflation	GDP		
Benchmark	SPF	SPF		
$\text{MSE}_{\mathbb{E}}/\text{MSE}_{\mathbb{F}}$ (LSTM)	0.771	0.876		
OOS R^2	0.229	0.124		

$\text{MSE}_{\mathbb{E}}$ and $\text{MSE}_{\mathbb{F}}$ are the machine and benchmark mean-squared-forecast-errors, respectively. OOS R^2 is defined as 1-$\text{MSE}_{\mathbb{E}}/\text{MSE}_{\mathbb{F}}$. The out-of-sample evaluation period is 2005:Q1 to 2021:Q4 for stock excess returns when the benchmark is the SOC and CFO survey, 2005:Q1 to 2007:Q3 when the benchmark is the Gallup/UBS survey, and semi-annual over 2005Q1 to 2021Q4 for Q2 and Q4 of each calendar year when the benchmark is the Livingston survey. For earnings growth, the out-of-sample evaluation period is 2005:Q1 to 2021:Q4, with h-quarter ahead forecasts constructed from the middle month of each quarter. For inflation and GDP growth the evaluation period is 1995:Q1 to 2021:Q4. Forecasts the historical mean ("Mean") are based on a first estimate of the mean from 1959:Q1 to 2004Q4 and then recursively updating the estimate, adding one observation at a time, over period 2005Q1 to 2021Q4.

based on a recursive estimate of the historical mean, given findings elsewhere that it is difficult for statistical models to beat such a measure when predicting excess stock market returns out-of-sample (Goyal and Welch (2008)). Panel (a) shows that the machine algorithm produces forecasts of excess returns that are consistently and substantially more accurate than all of these benchmarks. The ratio of the machine MSE to survey MSE is 0.62, 0.71, 0.72, 0.78 compared to the Gallup/UBS, CB, SOC, and CFO surveys, respectively. This corresponds

ology described in the Online Appendix. To obtain subjective expected excess returns, we subtract the one-year Treasury bill rate measured at the beginning of the survey month from stock market return expectations within each quarter.

to out-of-sample R^2 statistics of 0.38, 0.29, 0.28, and 0.16, respectively, implying that the machine forecasts are 38%, 29%, 28%, and 16% more accurate than the respective survey forecasts. The machine is also substantially more accurate than the historical mean, with the ratio of the machine MSE to historical mean MSE equal to 0.80, 0.81, and 0.81 for one-year, two-year, and three-year excess return forecasts.

Table 1, panel (b), considers forecasts of earnings growth on the S&P 500. We compare the machine expectations to the median IBES equity analyst forecasts of one-, two-, and three-year-ahead earnings growth on the S&P 500, following the procedure of De La O and Myers (2021) to convert firm-level earnings per share forecasts to S&P 500 forecasts by aggregating over the value-weighted firm-level forecasts and converting to growth forecasts. When using the IBES long-term growth forecasts (LTG), we follow Bordalo et al. (2019) in aggregating the value-weighted firm-level long-term growth forecasts of the median analyst to obtain LTG at the S&P 500 level. We treat LTG as measuring annual five-year forward growth expectations, i.e., annual earnings growth from four to five years ahead.[21] The machine substantially improves on the predictive accuracy of the IBES analysts, where the ratio of the machine MSE to IBES MSE is 0.68, 0.55, and 0.54 for one-, two-, and three-year-ahead earnings growth, respectively, implying the machine is 32%, 45%, and 46% more accurate than the median analyst for these predictions. The machine is about 69% more accurate than LTG when forecasting annual earnings growth five years forward.

Finally, Figure 1, panel (c) compares machine forecasts of inflation and Gross Domestic Product (GDP) growth with the median SPF forecast of these variables and shows that the ratio of the machine MSE to MSE of the median SPF forecast is 0.77 for inflation and 0.88 for GDP growth.

Overall, these results are consistent with the premise that a relatively unbiased, information-efficient machine using only real-time information is able to detect patterns in widely available data that notably improve predictive accuracy over human forecasts. This systematically superior performance motivates our use of the machine benchmark for measuring non-distorted expectation formation in the structural estimation.

Parameter Estimates Table 2 reports the posterior mode values for model parameters. Where applicable, separate values are reported for estimates of parameters governing the

[21] IBES states that "The long term growth rate represents an expected annual increase in operating earnings over the company's next full business cycle. These forecasts refer to a period of between three and five years." Although there is ambiguity in the question framing, interpreting LTG as an expected annual n-year forward growth rate (rather than the expected annualized n-year growth rate), is consistent with the reference to the *next* full business cycle and moreover makes the stable median LTG forecast easier to reconcile with the volatile median one-year growth forecast.

Table 2: Parameters

	Objective	Perceived		Objective	Perceived
ψ_i	0.0100	0.0100	$\beta_{\pi,i}$	0.0025	0.0025
ψ_π	0.0019	0.0019	$\beta_{\pi,\pi}$	-0.0001	-0.0001
$\psi_{\Delta y}$	0.0001	0.0001	$\beta_{\pi,\Delta y}$	0.0012	0.0012
ϕ_i	0.0484	0.0484	$\beta_{\Delta y,i}$	-0.0072	-0.0072
ϕ_π	0.0016	0.0016	$\beta_{\Delta y,\pi}$	0.0002	0.0002
$\phi_{\Delta y}$	0.0022	0.0022	$\beta_{\Delta y,\Delta y}$	0.0001	0.0002
ϕ_k	0.0383	0.0382	$\beta_{k,\Delta y}$	-9.0540	-8.9530
γ_{ra}	2.2975	–	$\rho_{k,k}$	0.9929	1.0353
ζ	3.8720	–	ρ_{lp}	0.9382	–
			ρ_η	0.9977	–

Notes: Posterior mode values of the parameters. The estimation sample spans 1961:M1-2021:M12.

objective macro LOM ("Actual") and the perceived macro LOM ("Perceived"). We highlight several results.

First, the scalar parameter ζ is estimated to be a positive value equal 3.87, implying a DE distortion indicative of overreaction to each perceived macro innovation in $\widetilde{\varepsilon}_t^M$. Second, for many parameters governing perceived macro dynamics, there is little deviation from the corresponding objective parameter value. However, there are appreciable differences in a few. We find $\widetilde{\rho}_{k,k} > \rho_{k,k}$, implying that today's payout share news is over-extrapolated to future payout share movements. There are also differences in the perceived and actual values of $\beta_{k,\Delta y}$, which in both cases is negative. This implies that increases in trend growth $\overline{\Delta y}_t$ drive *down* the payout share k_t. Because k_t affects $\overline{k}_t$ and ultimately future k_t through (11) and (15), the estimates say that increases in $\overline{\Delta y}_t$ cause a long-lasting decline in k_t. Yet because $0 > \widetilde{\beta}_{k,\Delta y} > \beta_{k,\Delta y}$, we find that investors *under*estimate the absolute impact of $\overline{\Delta y}_t$ on k_t, so that observed declines in k_t originating from increases in $\varepsilon_{\overline{\Delta y},t}$ will be partly misattributed to another perceived shock that moves k_t in the same direction. We return to this below. At the same time, $\beta_{\Delta y,\Delta y}$ is positive but small, indicating that $\overline{\Delta y}_t$ has only modest predictive power for future output growth, consistent with the fact that economic growth less autocorrelated than the other variables in our system. Putting this all together, this shows that increases in $\varepsilon_{\overline{\Delta y},t}$ are tantamount to bad cash flow news: the positive effects on economic growth are outweighed by the persistent negative effects on growth in the payout share.[22]

Table 3 shows basic asset pricing moments for stock returns and the real interest rate

[22]This result echoes findings in Greenwald, Lettau and Ludvigson (2019), which shows that the U.S. stock market grew far faster during decades with sluggish economic growth but rapid growth in the earnings share, than in decades with rapid economic growth but a relatively stable earnings share.

Table 3: Asset Pricing Moments

Moments	Model		Data	
	Mean	StD	Mean	StD
Log Stock Return	8.77	12.35	8.96	12.29
Log Excess Return	7.29	14.81	7.42	14.85
Real Interest Rate	1.48	2.92	1.54	2.53

Notes: Model moments based on modal parameter and latent state estimates. Annualized monthly statistics (means multiplied by 12, standard deviations by $\sqrt{12}$) and reported in units of percent. The log return (data) is the log difference in the S&P 500 market cap; excess returns subtract off FFR. The real interest rate is FFR minus the average one-year ahead forecasts of inflation from the BC, SPF, SOC, and Livingston surveys. The sample is 2001:M1 - 2019:M12.

implied by these estimates. The model based moments for these series are based on the modal parameter and latent state estimates and match their data counterparts closely.

Market Reactions to News: High-Frequency Structural Event Study To make progress on how markets react to this news, we use the BLM2 filtering algorithm to infer revisions in investor perceptions about the economic state, at high frequency around news events.

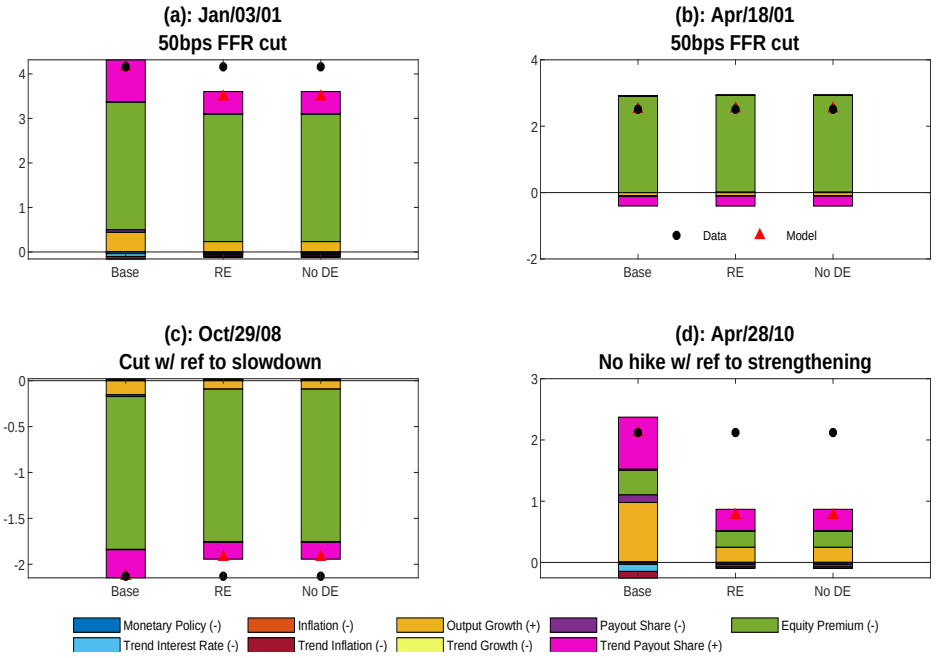
This procedure can be summarized as follows. Consider news events that occur within a given month t of our sample. Let $\delta_h \in (0,1)$ represent the number of time units that have passed during month t up to and including some particular point $t-1+\delta_h$. Let $S^{M*(i)}_{t|t-1+\delta_h}$ denote a filtered estimate of investor beliefs at time $t-1+\delta_h$ about the time t economic state they expect to prevail when it is observed at the end of the month, conditional on the volatility regime $\xi_t = i$. This is an estimate of the investor's nowcast of $S^{M*(i)}_t$. Let the associated filtered volatility regime probabilities be denoted $\pi^i_{t|t-1+\delta_h} \equiv \Pr\left(\xi_t = i | X_{t-1+\delta_h}, X^{t-1}\right)$, where X^{t-1}denotes the history $\{X_{t-1}, X_{t-2}, ...\}$. Finally, let δ_h assume distinct values δ_{pre} and δ_{post} that denote the times right before and right after the news event. Announcement-driven revisions in $S^{M*(i)}$ and in π^i are computed using high-frequency, forward-looking data by taking the difference between the estimated values for these variables pre- and post-news event. These differences can be linked back to jumps in specific variables in X_t (e.g., the stock market) using the mapping (29) and further decomposed into contributions coming from revisions in *perceived shocks* and volatility regimes using (28). We refer to these as "shock decompositions" and report them below. For recording the contribution of movements in subjective return premia, we report the combined contributions of lp_t and the volatility regimes to fluctuations in $\widetilde{\mathbb{E}}_t\left[r^D_{t+1}\right] - \left(i_t - \widetilde{\mathbb{E}}_t\left[\pi_{t+1}\right]\right)$ in (27), labeled "equity premium" in the figures below.

For the macroeconomic data releases and Fed news events we have an exact time stamp indicating when the information was released to the public. This allows us to construct precise 30 minute windows for these events (δ_{pre} =10 minutes before to δ_{post} =20 minutes after). We then run the filter at these times pre- and post-news using minutely or tick-level financial market data. We also use daily data on the day before and the day after these events for those series that are available daily but not at higher frequency. For the corporate earnings news–where events span an entire day–we run the filter using information on all high-frequency series from the close of the market on the day before to the opening of the market on the day after.

Our structural event study findings are divided by news category and displayed in a series of bar charts. For each shock decomposition figure, we report the market jump in the data (as measured by the S&P 500) with a black dot and the jump implied by the estimated model with a red triangle. Decompositions for the estimated baseline model ("Base", shown in the first bar from the left), always have black dots that lie on top of the red triangles because the baseline model is able to match both the direction and magnitude of the actual market jump with negligible error. We then compare these decompositions to two counterfactuals to reveal how the market would have behaved in the absence of distortions: (i) rational expectations, i.e., $\zeta = 0$ and $w_\theta = 0\ \forall\theta$, "RE" in the figures, and (ii) No DE, i.e., only $\zeta = 0$. The news event itself is described briefly in the subpanel titles, e.g., "Cut w/ ref to slowdown" refers to the Fed's decision to cut rates citing concerns about an economy-wide slowdown. As we have over 1700 separate news events, for the purposes of the plots below we focus on the news by category associated with the biggest stock market jumps in absolute value. (The results for all events are summarized in a subsequent table.)

Figure 1 reports the results for Fed news events. Panel (a) depicts the market response to the most quantitatively important Fed announcement in our sample, which occurred on January 3, 2001 when the central bank announced it would decrease the target federal funds rate by 50 basis points and the S&P 500 surged 4.2% in the 30 minute window surrounding the news. What did markets take away from this announcement? The estimates for the baseline model (left most bar) show that the biggest contributors to the jump were upward revisions in the perceived shocks to the trend payout share and output growth, and a downward revision in the subjective equity premium. The next two bars show that, under RE, the market would have jumped up 3.5% rather than 4.2%, an overreaction driven almost entirely by the DE distortion. (This can be observed by noting that the "No DE" counterfactual results in virtually the same jump and decomposition as the RE counterfactual.) DE causes investors to react with excessive optimism to the trend payout share and output growth shocks, inflating

Figure 1: Biggest Market Events: FOMC

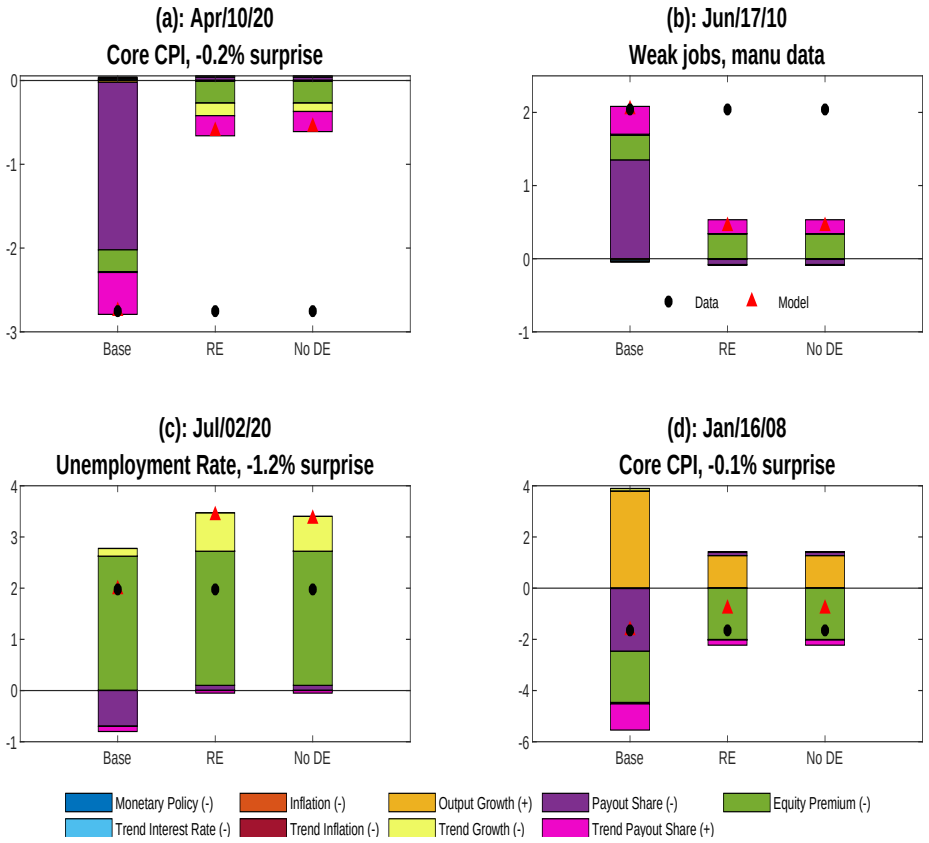


Notes: The figure reports shock decompositions of the pre-/post- FOMC announcement change in S&P 500 attributable to revisions in the perceived macro shocks and the subjective equity premium (the combined effect of shocks to lp_t and stochastic volatility). The specific FOMC events reported on are those coinciding with the four largest jumps in the S&P 500 in the high-frequency event window. The modifiers (+) or (-) refer to the sign of the baseline response to a positive increment in the fundamental shock labeled in the legend. The sample is 2001:M1-2021:M12.

the price response. This same pattern leads to even greater overreaction for the FOMC event depicted in panel (d), when the market jumped up by 2.10%, while it would have jumped only .77% under RE. The flavor of these results are broadly consistent with ideas put forth in recent work by Jin and Li (2023) who hypothesize that equity analysts overreact in their long-term earnings growth forecasts because they wrongly infer stock price movements as cash flow driven when they are in fact discount rate driven. Unlike Jin and Li, in our analysis equity analyst forecasts are but one of several signals on investor beliefs, and price movements are endogenous outcomes determined in equilibrium. However, the larger finding here that news is overreacted to because investors systematically overstate its implications for payout-relevant factors (while understating in relative terms its implications for discount rates) confirms their basic hypothesis, and is supported by our estimation for other events.

Figure 2 turns to market reactions to news about the macro economy. Panel (a) shows that, on April 10, 2020, early in the Covid-19 outbreak, the market fell 2.8% in the 30 minutes surrounding the Bureau of Labor Statistics (BLS) release of the CPI report, which came in

Figure 2: Biggest Events: Macro News



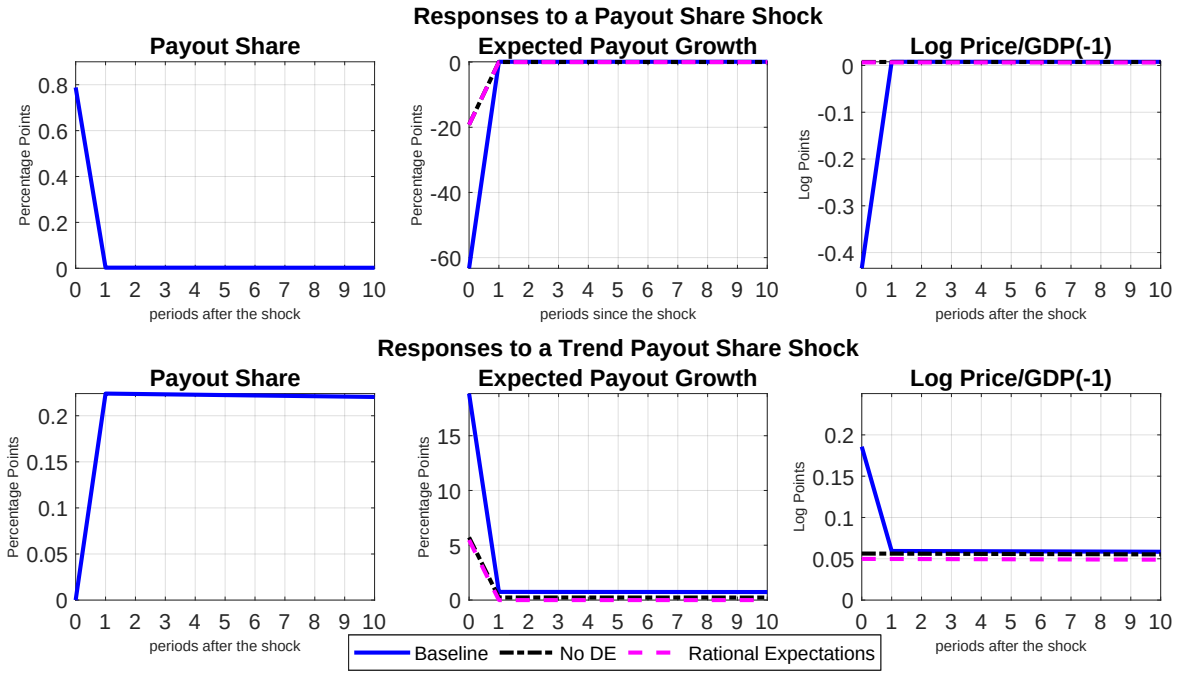
Notes: See Figure 1. The figure reports shock decompositions for the four biggest macro news events based on absolute jumps in the S&P 500 in the high-frequency event window.

.2% below consensus forecasts.[23] In this case the main driver of the 2.8% decline was an outsize reaction to a higher perceived payout share shock (this time the cyclical component, shown in purple). Yet because this shock plays virtually no role in the rational response, under the RE counterfactual the market would have declined just 0.55%. The same pattern but in the opposite direction plays out for the event in panel (b) surrounding the release of several pieces of data showing surprising economic weakness, which the estimates imply led to a downward revision in the perceived cyclical payout share shock, which contributed positively to the 2% market jump, while under RE it would have increased 0.37%.

To understand this result, we must first explain why a positive increment to the payout share shock causes a sharp decline in the stock market in the baseline case, but not under RE. For this we refer to Figure 3, which plots estimated impulse responses, in deviations

[23]The BLS releases typically occur at 8:30 am. We use S&P E-mini futures data to gauge market reactions to these events.

Figure 3: Responses to Payout Share Shocks



Notes: This figure plots estimated impulse responses at the posterior model parameter values, in deviations from steady-state, to positive payout share shock (top panels) and a positive trend payout share shock (bottom panels), respectively.

from steady state, to 2 standard deviation increases in the payout share shock, $\varepsilon_{k,t}$, (top row) and trend payout share shock, $\varepsilon_{\overline{k},t}$, (bottom row). From the top row, left panel, we see that a positive payout share shock leads to a highly transitory increase in payout relative to GDP that quickly mean reverts. Under both RE and in the baseline model, this mean reversion in the payout *share* creates the expectation of negative fall-back *growth* in payout next period (top row, middle panel), consistent with the common understanding that the payout share is stationary. However, in the RE case, expected fall-back growth is just negative enough to (almost) restore the k_t to its steady state level within one period, so that expected growth from period 1 onward is approximately zero.[24] Because the shock is rationally perceived to be a highly transitory deviation from the steady-state payout-output share, it has a negligible effect on the level of the stock market (top row, right panel). By contrast, in the baseline model with DE, the investor strongly overreacts to the increase in $\varepsilon_{k,t}$, giving rise to excessive pessimism about fall-back growth next period. This effect (demonstrated above

[24]It would be exactly zero if $\varepsilon_{k,t}$ were exactly i.i.d, but since k_t affects $\overline{k}_t$, which enters into (15), the shock causes expected payout growth in period 1 to rise imperceptibly above zero before declining slowly back to zero over time.

in the simplified model) is equivalent to temporarily believing that the *level* of the payout share will revert to below steady state. This erroneous belief causes the market to crash before recovering next period when actual growth is observed and investors learn that they had been excessively pessimistic (top row, right panel).

These results can be contrasted with the responses to a trend payout share shock in the bottom row of Figure 3. Unlike an increase in $\varepsilon_{k,t}$, an increase in $\varepsilon_{\overline{k},t}$ has highly persistent effects on k_t, implying that mean reversion takes decades.[25] When the shock hits, we get a one-time jump up in expected payout growth for next period that falls back toward zero in period 1. In the baseline model, however, overreaction to the initial good news from $\varepsilon_{\overline{k},t}$ is overreacted to, creating excessive optimism and inflating the price response. The same phenomenon leads to disappointment the following period when actual growth is observed and investors learn that they had been excessively optimistic, causing a price reversal back toward the RE response.[26] Note that while expected growth overreacts to both $\varepsilon_{k,t}$ and $\varepsilon_{\overline{k},t}$, the estimated magnitudes differ–overreaction to $\varepsilon_{k,t}$ is greater– leading to commensurately differing magnitudes of overreaction in the stock price response.

Panel (c) of Figure 2 shows how the specific news content of an event can deliver starkly different results depending on how the composition of perceived shocks is revised in response. In reaction to the event depicted–the July 2, 2020 BLS release of the unemployment rate–the stock market rose 1.97%, an increase that coincides with the baseline model response. Yet the market's response this news would have been to jump 3.35% had expectations been formed rationally, implying that distorted beliefs and DE in particular led the market to *under*react to the BLS announcement. This happens because, while the news causes investors to revise down their perception of the equity premium, which pushes the market up, it also causes a partially offsetting upward revision in the payout share shock, which pushes the market down. As in panel (a), this contributes negatively to the market change because it causes excessive pessimism about expected fall-back growth next period. Under RE there is no excessive pessimism to the payout share shock and thus no erroneous partially offsetting contribution that dampens the market response, as shown in Figure 3. In addition, under RE we see that a downward revision in the perceived shock to trend growth$\overline{\Delta y}_t$ (yellow bar) contributes positively to the market change because, as the parameter estimates in Table 2 indicate, this generates the expectation of *higher* future payout growth. By contrast, this positive force for the market makes a smaller positive contribution to the market change in

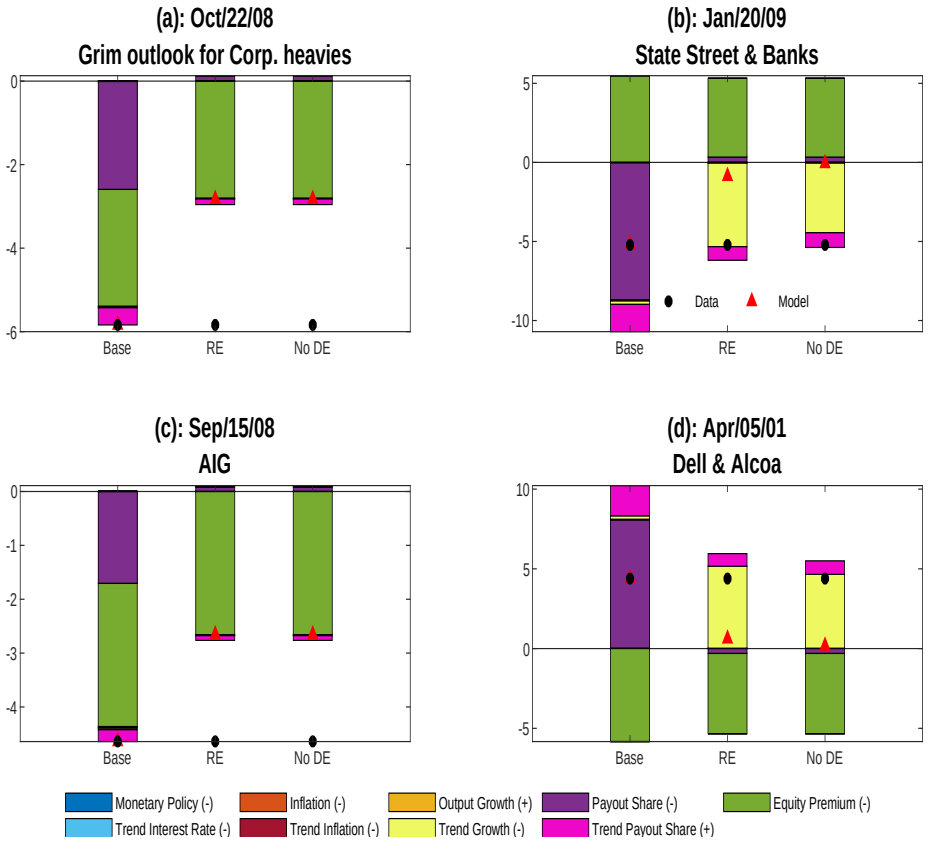
[25]Payout rises in period 1 because $\varepsilon_{\overline{k},t}$ affects k_t with lag–see (15).

[26]The baseline price remains slightly above the RE level for some time before finally converging. This happens because excessive optimism or pessimism generated by DE is modulated by the shock's perceived persistence, which in this case is estimated to be high–see (21).

the no DE case and only a very small positive contribution in the baseline model with DE. As we explain below, the reason is that part of the negative impulse to $\varepsilon_{\bar{\Delta}y,t}$ is misperceived as a positive (cyclical) payout share shock, which has a less persistent effect on payout growth than $\varepsilon_{\bar{\Delta}y,t}$ thereby dampening its effect on the market, a misperception that is necessarily amplified by DE. This event shows how DE, especially when combined with distortions in the perceived law of motion, can result in market underreaction to events because the degree of overreaction to counteracting perceived shocks can be asymmetric across shocks.

It is important to emphasize that this underreaction is not due to inattention, since a single parameter with an estimated value indicative of behavioral overreaction controls the distorted reactions to all shocks. The event in panel (c) underscores the capacity of DE to generate asymmetric compositional effects capable of either amplifying or dampening market fluctuations.

Figure 4: Biggest Events: Earnings News

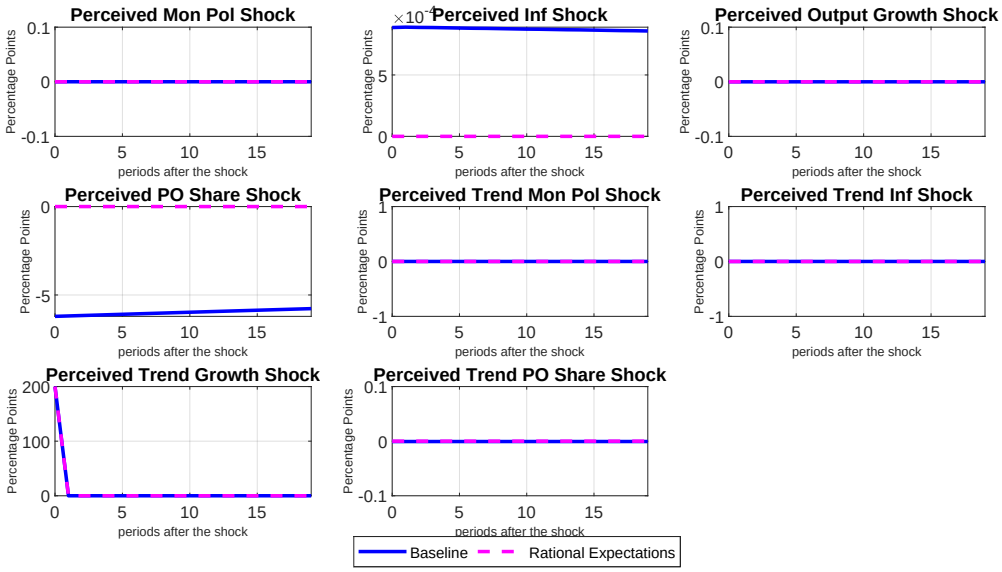


Notes: See Figure 2. The figure reports shock decompositions for the four biggest corporate earnings news events based on absolute jumps in the S&P 500 in the high-frequency event window.

Figure 4 shows shock decompositions for the stock market's reaction to news about

corporate earnings, measured from the market close the day before to the market open the day after big corporate earnings news days. Consider January 20, 2009, a news day in the wake of the financial crisis when the market declined 5.2% amid extensive reports about unrealized losses in the portfolio of asset manager State Street and in the portfolios of large banks–panel (b). The baseline model shows that the market declined largely because investors again overreacted to an upwardly revised perceived payout share shock, which was only partially offset by a downward revision in the subjective equity premium. Under RE, the market would have declined just 1%, largely because there is no overreaction in this case to the payout share shock. However, there is an additional effect that also presents itself in panel (c) of Figure 2: under RE, an upward revision in the perceived shock to trend growth $\overline{\Delta y}_t$ (yellow bar) makes a large negative contribution to the market response that is, although present, very small and thus barely visible in the plot for the baseline model. To understand this result, recall that the estimates imply $0 > \widetilde{\beta}_{k,\Delta y} > \beta_{k,\Delta y}$, which means that investors underestimate the negative impact of higher $\overline{\Delta y}_t$ on k_t and, as a consequence, perceived declines in k_t originating from increases in $\varepsilon_{\overline{\Delta y},t}$ will be partly misattributed to another shock.

Figure 5: Responses of Perceived Shocks to Actual Trend Growth Shock



Notes: This figure plots estimated impulse responses at the posterior model parameter values, in deviations from steady-state, of perceived shocks to an actual (positive) trend growth shock.

To see which shocks this misattribution maps into, we report in Figure 5 estimated impulse responses of all perceived shocks in $\widetilde{\varepsilon}_t^M$ to a 2 standard deviation increase in the

actual trend growth shock $\varepsilon_{\overline{\Delta y},t}$. Under RE, only the perceived trend growth shock responds to an actual trend growth shock, as all perceptions are accurate. By contrast, in the baseline model, an increase in $\varepsilon_{\overline{\Delta y},t}$ not only causes $\widetilde{\varepsilon}_{\overline{\Delta y},t}$ to increase, it also causes $\widetilde{\varepsilon}_{k,t}$ to decrease strongly and persistently, and causes $\widetilde{\varepsilon}_{\pi,t}$ to increase by a smaller absolute magnitude. The confounding negative effect on $\widetilde{\varepsilon}_{k,t}$ of a positive impulse to $\widetilde{\varepsilon}_{\overline{\Delta y},t}$ creates the false expectation of catch-up growth in payout next period, with pricing effects that are amplified by DE (top panel, Figure 3). The over-optimism about catch-up growth almost entirely counteracts the negative effects on perceived future payouts originating from $\varepsilon_{\overline{\Delta y},t}$. This explains why a shock to trend growth (yellow bar) has a very small affect on stock prices in the baseline model in Figure 4. Note that the overall effect of this misattribution is necessarily amplified by DE, since diagnostic expectations operates on the shocks investors perceive rather than those that actually occurred.

Table 4: Average Jump Differentials

All Events	Biggest Jumps	Smallest Jumps
Macro News		
−12.1%	24.5%	−26.7%
CorporateEarnings News		
14.2%	42.8%	5.3%
FOMC News		
−13.2%	12.9%	−18.1%

This table reports $\left(|J^{Base}| - |J^{RE}|\right)/|J^{Market}|$ the average difference between the pre-/post- news event jump (in absolute value) for the baseline model $|J^{Base}|$ and that for the counterfactual RE case $|J^{RE}|$ divided by the absolute market jump $|J^{Market}|$. For macro and Fed news, "Biggest" ("Smallest") refers to the top (bottom) 10% of all events based on the absolute change in the S&P 500 over the news window. For earnings news "Biggest" ("Smallest") refers to the top (bottom) 3 events.

Table 4 summarizes the magnitude of over- or underreaction across all news event of a given type in our sample. We compute the difference between the absolute value of the pre-/post- news-event jump in the S&P 500 implied by the baseline model and the RE counterfactual, then average these differences across all events in a given category and express it as a fraction of the absolute jump in the market. Positive values for this difference indicate overreaction on average, while negative values indicate underreaction.[27] We repeat the computation for news that generated the "Biggest" and "Smallest" absolute jumps in the S&P 500 during the news window. For Fed and macro news (where we have hundreds of events)

[27] We average across all events in which the baseline and non-distorted jumps are in the same direction. Jumps in the opposite direction happen infrequently in our sample, but also can't be readily categorized as either over- or underreaction, as opposed to simple "wrong" reaction.

"Biggest" ("Smallest") refers to the top (bottom) 10% of all events based on the absolute change in the S&P 500. For the corporate earnings news events, where we have only 16 event-days, we define "Biggest" ("Smallest) as the top (bottom) 3 events according to the same criteria.

Table 4 shows that, averaged across all events, we find negative differentials in the categories of Fed and macro news, i.e., underreaction, a result driven by the smallest market events. The biggest market events are characterized by overreaction in all news categories. The largest of these is for earnings news, where the market overreacted by an average of 43% of the total market change. To points are worth noting. First, the corporate earnings news events are only large events, as BBDS focus on days with large stock market movements. Second, many of the macroeconomic data and FOMC press releases convey little if any information that was not already anticipated. Naturally, these events reside in our "Smallest" events category because they generate little to no reaction in the market and thus little to no over- or underreaction in absolute terms, even though the latter can still be large as a percentage of a tiny market change.

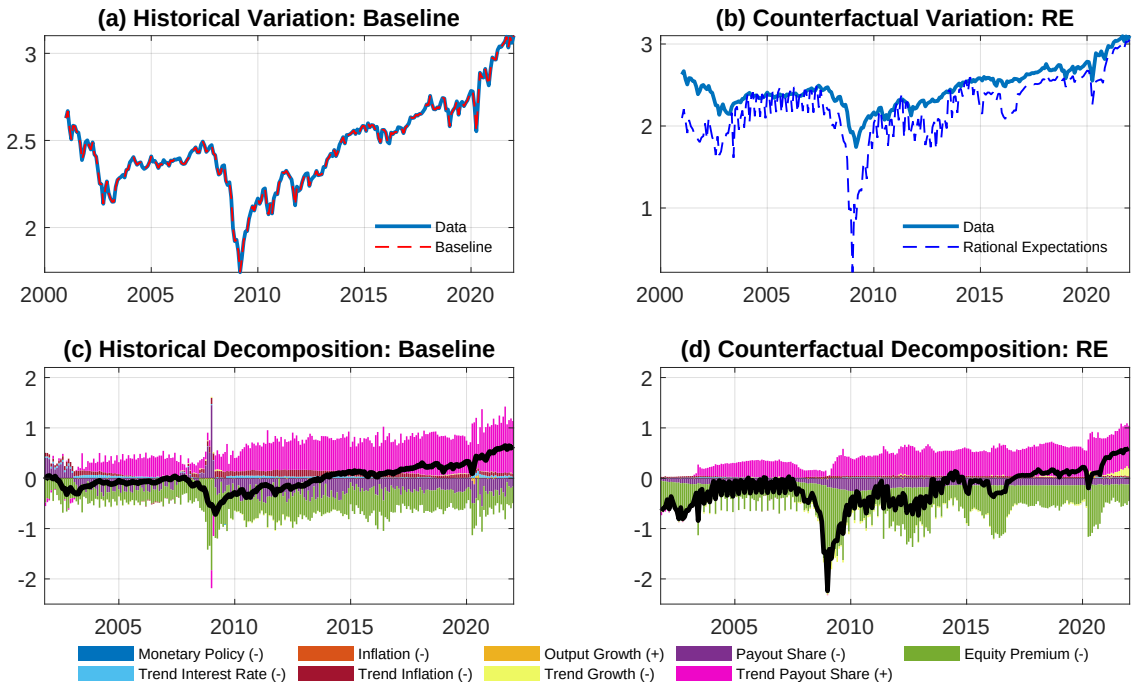
We can also use the structural model to categorize news events by whether the market exhibited over- or underreaction. Table A.5 in Appendix D shows that market jumps in tight windows around news events in a given month to which the stock market is categorized as having overreacted are followed by lower subsequent returns, while those categorized as underreactions are followed by higher future returns. This additional validation is encouraging, given that we measure only a handful of events in any given month, whereas the market as a whole is buffeted by many types of other unmeasured news that would create noise.

Market Valuation: Historical Analysis We now turn to a longer term analysis to study the model's predictions outside of tight windows around news events. Panel (a) of Figure 6 reports the log ratio of market equity to last month's output, $p_t^D - y_{t-1}$, for both the data[28] and the baseline model, where the latter is computed at the modal values of all parameters and latent states. (Because the model fits the observed series effectively without error, two lines lie on top of each other.) Panel (b) reports the data once more, along with our estimate of the market evolution under a counterfactual simulation in which parameters consistent with RE prevailed.[29] The plots span 2001:01-2021:12. The bottom row of Figure 6 reports

[28] We use the interpolation method of (Stock and Watson (2010)) to obtain a monthly GDP series for estimation.

[29] A counterfactual simulation feeds in the shocks implied by the baseline model estimates, and only changes the parameter values, a procedure that isolates the strength of the mechanism in the baseline model compared to some counterfactual mechanism. This differs from the implications of a counterfactual *model*, in which the shocks would be re-estimated under an alternative set of parameters not chosen by the baseline

Figure 6: Counterfactual Simulations of S&P 500-GDP Ratio



Notes: Panel (a) plots the log S&P500-to-lagged GDP ratio along with the the model-implied series based on the modal parameter and latent state values. Panel (b) plots the data and the counterfactual RE series in dashed blue. Panel (c) plots a historical decomposition of log S&P 500-lagged GDP ratio, normalized to have a mean of zero over the sample period. Panel (d) plots the counterfactual historical decomposition under RE of the same series. The sample spans 2001:M1 - 2021:M12

historical decompositions of the variation in $p_t^D - y_{t-1}$, for the baseline model in panel (c) and for the RE counterfactual in panel (d). These report cumulative month-to-month changes in $p_t^D - y_{t-1}$ decomposed into cumulative fundamental shocks and premia.

While panel (a) says that the baseline model perfectly explains the market's fluctuations, panel (b) tells us that the fit of the counterfactual RE case is far worse. Indeed RE implies a counterfactually volatile stock market, demonstrating the extent to which distorted beliefs with behavioral overreaction were a strong stabilizing force over the post-millennial period, substantially cushioning declines during the global financial crisis, among other episodes.

The historical decompositions in the bottom row help to explain why. First consider the sharp decline in the stock market in the financial crisis (around 2008 in the plot). This episode is characterized by a sharp decline in the payout share shock $\varepsilon_{k,t}$, to which in the baseline model the the investor strongly overreacts, leading to excessive optimism about

estimation. The latter may be of interest in some contexts, but it cannot isolate the strength of a baseline model mechanism, since both the mechanism and the shocks change.

catch-up growth in payouts. That over-optimism makes a large positive contribution to the market (purple bar), partially offsetting the predominating negative contributions due to other shocks that were still overreacted to but to a lesser degree. The market declines more under RE in this episode because there is no overreaction to the decline in $\varepsilon_{k,t}$ and thus no excessive optimism about catch-up growth.

Following the outbreak of the Covid-19 pandemic, panel (c) of Figure 6 shows that the market declines mainly because the predominating negative forces for the market coming from a lower perceived output growth and upward revisions in the subjective return premia more than offset overreaction to better-than-expected trend payout news and a lower trend interest rate, slightly dampening the market's decline relative to RE in panel (d). A subsequent rebound occurs in both cases, but for different reasons. Panel (d) shows that lower values for the trend growth shock $\varepsilon_{\overline{\Delta y},t}$ (yellow bar) make a series of strong positive contributions to expected future payouts and thus the market under RE that are missing in panel (c) for the baseline model. This happens because, as discussed above, the investor erroneously attributes part of the movement in $\varepsilon_{\overline{\Delta y},t}$ to an impulse in $\widetilde{\varepsilon}_{k,t}$, which counteracts any influence on perceived future payouts originating from the objective impulse in $\varepsilon_{\overline{\Delta y},t}$.

7 Conclusion

We measure the nature and severity of a variety of belief distortions in market reactions to hundreds of economic news events. To do so, we use a new methodology developed here that synthesizes estimation of a structural asset pricing model with algorithmic machine learning to quantify bias. The structural model allows for beliefs that could either over- or underreact to multiple perceived fundamental shocks (depending on parameter values), and for the perceived law of motion driving macroeconomic dynamics to differ from the actual law of motion.

Our estimates imply that investors overreact to perceptions about multiple fundamental shocks, as well as to numerous real-world news events. Yet we show that behavioral overreaction can sometimes be a force for market stability in a multivariate setting. This happens because the sum total of different individual overreactions to multiple simultaneous shocks can dampen rather than amplify their combined market impact relative to a fully rational counterfactual. We find this phenomenon playing out over the course of our sample around many news events and during several major episodes of post-millennial stock market history, substantially cushioning market declines during the global financial crisis. In other events, our estimates imply that investors underreacted to specific shocks because they systemat-

ically misperceived the true origin of the impulse, a phenomenon necessarily amplified by diagnostic expectations.

A transformative idea of 20th century economic thought is that financial markets are "excessively volatile" vis-a-vis predictions of canonical theory in which stock prices are the rational expectation of future cash flow fundamentals, discounted at a constant rate (Shiller 1981, 2000). We showed here that an arguably more plausible rational expectations benchmark that allows discount rates to vary implies highly volatile markets–more so than in the data–creating a puzzle of excess stability rather than excess volatility. Yet we find that a macro-dynamic model with belief overreaction in the spirit of diagnostic expectations can perfectly explain the data, not because it creates more volatility but because it creates less.

References

Adam, K., and S. Nagel (2023): "Expectations data in asset pricing," in *Handbook of Economic Expectations*: Elsevier, 477–506.

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

Ang, A., and M. Piazzesi (2003): "A No-Arbitrage Vector Autoregression of Term Structure Dynamics With Macroeconomic and Latent Variables," *Journal of Monetary Economics*, 50, 745–787.

Baker, S., N. Bloom, S. J. Davis, and M. Sammon (2019): "What triggers stock market jumps?", Unpublilshed manuscript, Stanford University.

Bansal, R., and H. Zhou (2002): "Term structure of interest rates with regime shifts," *The Journal of Finance*, 57, 1997–2043.

Barberis, N., A. Shleifer, and R. W. Vishny (1998): "A Model of Investor Sentiment," *Journal of Financial Economics*, 49, 307–43.

Bastianello, F., and P. Fontanier (2022): "Expectations and Learning from Prices."

Bauer, M. D., and E. T. Swanson (2023): "An Alternative Explanation for the "Fed Information Effect"," *American Economic Review*, 113, 664–700.

Bianchi, F., T. Kind, and H. Kung (2019): "Threats to central bank independence: High-frequency identification with Twitter,"Technical report, National Bureau of Economic Research.

Bianchi, F., H. Kung, and M. Tirskikh (2018): "The origins and effects of macroeconomic uncertainty,"Technical report, National Bureau of Economic Research.

Bianchi, F., S. C. Ludvigson, and S. Ma (2022a): "Belief Distortions and macroeconomic fluctuations," *American Economic Review*, 112, 2269–2315.

——— (2022b): "Monetary-Based Asset Pricing: A Mixed-Frequency Structural Approach,"Technical report, National Bureau of Economic Research, No. w30072.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003): "Latent dirichlet allocation," *Journal of machine Learning research*, 3, 993–1022.

Bordalo, P., N. Gennaioli, R. LaPorta, and A. Shleifer (2019): "Diagnostic expectations and stock returns," *The Journal of Finance*, 74, 2839–2874.

——— (2022): "Belief Overreaction and Stock Market Puzzles," *Journal of Polticial Economy,* forthcoming.

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

Bordalo, P., N. Gennaioli, and A. Shleifer (2018): "Diagnostic expectations and credit cycles," *The Journal of Finance*, 73, 199–227.

Bybee, L., B. T. Kelly, A. Manela, and D. Xiu (2021): "Business news and business cycles,"Technical report, National Bureau of Economic Research.

Campbell, J. Y., and J. H. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.

Campbell, J. Y., A. W. Lo, and C. MacKinlay (1997): *The Econometrics of Financial Markets*, Princeton, NJ: Princeton University Press.

Campbell, J. Y., and R. J. Shiller (1989): "The dividend-price ratio and expectations of future dividends and discount factors," *Review of Financial Studies*, 1, 195–228.

Chaudhry, A. (2023): *The Impact of Prices on Analyst Cash Flow Expectations*, Ph.D. dissertation, The Ohio State University.

Chen, L., Z. Da, and X. Zhao (2013): "What drives stock price movements?" *The Review of Financial Studies*, 26, 841–876.

Cochrane, J. H. (2005): *Asset Pricing:* Revised Edition, Princeton, NJ: Princeton University Press.

Coibion, O., and Y. Gorodnichenko (2015): "Information rigidity and the expectations formation process: A simple framework and new facts," *American Economic Review*, 105, 2644–78.

Cong, L. W., K. Tang, J. Wang, and Y. Zhang (2021): "Deep sequence modeling: Development and applications in asset pricing," *The Journal of Financial Data Science*, 3, 28–42.

De La O, R., and S. Myers (2021): "Subjective cash flow and discount rate expectations," *The Journal of Finance*, 76, 1339–1387.

——— (2023): "Which Subjective Expectations Explain Asset Prices?", Manuscript, Wharton.

Farmer, R. E., D. F. Waggoner, and T. Zha (2011): "Minimal state variable solutions to Markov-switching rational expectations models," *Journal of Economic Dynamics and Control*, 35, 2150–2166.

Gabaix, X. (2019): "Behavioral Inattention," in *Handbook of Behavioral Economics, Vol. 2* ed. by Bernheim, D., DellaVigna, S., and Laibson, D. New York, NY: Elsivier, 261–343.

Gabaix, X., and R. S. Koijen (2021): "In search of the origins of financial fluctuations: The inelastic markets hypothesis,"Technical report, National Bureau of Economic Research.

Gormsen, N. J., and R. S. Koijen (2020): "Coronavirus: Impact on stock prices and growth expectations," *The Review of Asset Pricing Studies*, 10, 574–597.

Goyal, A., and I. Welch (2008): "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, 21, 1455–1508.

Greenwald, D., M. Lettau, and S. C. Ludvigson (2019): "How the wealth was won: factor shares as market fundamentals," NBER Working paper No. 25769.

Gu, S., B. Kelly, and D. Xiu (2020): "Empirical asset pricing via machine learning," *The Review of Financial Studies*, 33, 2223–2273.

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

Hartzmark, S. M., and D. H. Solomon (2022): "Predictable price pressure,"Technical report, National Bureau of Economic Research.

Hecht-Nielsen, R. (1987): "Kolmogorovś mapping neural network existence theorem," in *Proceedings of the international conference on Neural Networks* Volume 3, 11–14, IEEE press New York, NY, USA.

Hillenbrand, S., and O. McCarthy (2021): "Heterogeneous investors and stock market fluctuations," *Available at SSRN 3944887*.

Ismailov, V. E. (2023): "A three layer neural network can represent any multivariate function," *Journal of Mathematical Analysis and Applications*, 523, 127096.

Jin, L. J., and J. Li (2023): "Overreaction of Long-Term Cash Flow Expectations to Past Price Movements," Working paper, Cornell University.

Kim, C.-J. (1994): "Dynamic Linear Models with Markov-Switching," *Journal of Econometrics*, 60, 1–22.

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

Krishnamurthy, A., and A. Vissing-Jorgensen (2012): "The aggregate demand for Treasury debt," *Journal of Political Economy*, 120, 233–267.

Lettau, M., and J. Wachter (2007): "The Term Structures of Equity and Interest Rates," Unpublished paper.

Ludvigson, S. C., and S. Ng (2007): "The Empirical Risk-Return Relation: A Factor Analysis Approach," *Journal of Financial Economics*, 83, 171–222.

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

Nagel, S., and Z. XU (2021): "Dynamics of Subjective Risk Premia,"Technical report, Working Paper, University of Chicago.

Nagel, S., and Z. Xu (2022): "Asset pricing with fading memory," *The Review of Financial Studies*, 35, 2190–2245.

Piazzesi, M., and E. Swanson (2008): "Futures Prices as Risk-Adjusted Forecasts of Monetary Policy," *Journal of Monetary Economics*, 55, 677–691.

Shiller, R. J. (1981): "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review*, 71, 421–436.

——— (2000): *Irrational Exurberance*, Princeton, NJ: Princeton University Press.

Sims, C. A. (2003): "Implications of rational inattention," *Journal of monetary Economics*, 50, 665–690.

Stock, J. H., and M. W. Watson (2010): "Research Memorandum," `https://www.princeton.edu/~mwatson/mgdp_gdi/Monthly_GDP_GDI_Sept20.pdf`.

Woodford, M. (2002): "Imperfect Common Knowledge and the Effects of Monetary Policy," in *In Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps* ed. by Aghion, P., Frydman, R., Stiglitz, J., and Woodford, M. Cambridge MA: Princeton University Press, 25–58.

Online Appendix

A Data Used in Structural Estimation and Filtering Around News Events

S&P Market Cap, Indexes, S&P Futures, S&P Dividends, Stock Market Returns, S&P Earnings, and Treasury Bill Data

For the structural estimation we use data on both stock market returns (price growth plus a dividend yield) and on stock market price growth. Monthly data on stock returns are obtained from the Center for Research in Security Prices (CRSP) downloaded from WRDS `https://wrds.wharton.upenn.edu/wrdsauth/members.cgi`. We use the CRSP value-weighted monthly return series $VWRETD$ (includes dividends) and compute annualized log returns $\ln CRSPD = 12\ln(1 + VWRETD)$. For machine forecasts of returns or price growth we take the difference between the price growth measure or return, e.g., $\ln(VWRETD)$, and the lagged log of the 3-month T-bill rate ($3MTB$). Since the $3MTB$ is reported at an annual rate in percent, we compute the annualized (raw unit) log of future returns less the current short rate, i.e., $\ln(VWREX_{t+12}) \equiv 12\ln(1 + VWRETD_{t+12}) - \ln(1 + 3MTB_t/100)$. The structural estimation uses monthly data (or higher frequency), so we map the annualized monthly stock return onto the one-year return in the model. Both series were downloaded from WRDS on February 12, 2023.

When evaluating the MSEs ratios of the machine relative to that of a benchmark survey, we compute machine forecasts for either the annual CRSP return, or S&P 500 price growth depending on which value most closely aligns with the concept that survey respondents are asked to predict. See below. To measure one-year stock market price growth we obtain a monthly series on the S&P 500 market capitalization, obtained as the end-of-month series from Ycharts.com available at `https://ycharts.com/indicators/sp\_500\_market\_cap`. This series span the periods 1959:01 to 2021:12 and were downloaded on March 13, 2022. This series is used to measure the monthly stock price to output ratio for the structural estimation. Below we also use the one-year log cumulative growth rate of the S&P 500 index, $\log\left(\frac{P^{S\&P}_{t+12}}{P^{S\&P}_{t}}\right)$. The monthly S&P index series spans the period 1957:03 to 2022:12 and was downloaded from WRDS on January 24, 2024 from the Annual Update data of the Index File on the S&P 500. To measure the one-year log CRSP return, we compute $\sum_{j=1}^{12} \ln(1 + VWRETD(t+j)) - ln(1 + 3MTB(t)/100$.

We use S&P 500 earnings divided by GDP as a noisy signal on the payout share K_t in

the structural estimation. To map into a monthly estimation, we ideally would use monthly earnings data. Instead, we have quarterly S&P 500 earnings per share (EPS) data that starts in 1988:Q2 from S&P Global `https://www.spglobal.com/spdji/en/documents/additional-material/sp-500-eps-est.xlsx`. To extend our sample backward, we use monthly EPS data on the S&P 500 from Robert Shiller's data depository at URL: `http://www.econ.yale.edu/~shiller/data/ie_data.xls`. These are monthly EPS data equal to the sum over the trailing 12 months, computed from the S&P four-quarter trailing totals. We use data for this series from 1959:01 to 1988:03. (There are no quarterly EPS observations recorded publicly prior to 1988:03.) To obtain a single quarterly earnings series extending backward to 1959:Q1, we employ a recursive process that combines these two series. Specifically, let time t be measured in months. Shiller's series provides a monthly series of earnings over the past 12 months:

$$em_t = e_t + e_{t-1} + ... + e_{t-11}$$

Starting from 1988:Q2, we also have a quarterly series of earnings over the quarter:

$$eq_t = e_t + e_{t-1} + e_{t-2}$$

if $t = \{3, 6, 9, 12\}$, $eq_t = 0$ otherwise. Suppose that we are interested in earnings for 1988:Q1. We have:

$$\begin{aligned} em_{1988:M12} &= \underbrace{e_{1988:M12} + e_{1988:M11} + e_{1988:M10}}_{eq_{1988:M12}} + ... + \underbrace{e_{1988:M3} + e_{1988:M2} + e_{1988:M1}}_{eq_{1988:M3}} \\ &= eq_{1988:M12} + eq_{1988:M9} + eq_{1986:M6} + eq_{1988:M3} \end{aligned}$$

We can then compute implied earnings for 1988:Q1 as

$$eq_{1988:M3} = em_{1988:M12} - \left[eq_{1988:M12} + eq_{1988:M9} + eq_{1988:M6}\right].$$

We can then use the same formula recursively to obtain earnings before 1988, i.e., with

$$em_{1988:M9} = \underbrace{e_{1988:M9} + e_{1988:M8} + e_{1988:M7}}_{eq_{1988:M9}} + ... + \underbrace{e_{1987:M12} + e_{1987:M11} + e_{1987:M10}}_{eq_{1987:M12}},$$

which gives the implied earnings for 1987:Q4. We continue recursively, working backward to the beginning of our sample in 1959:Q1. In the S&P Global dataset there is one observation in 2008:Q4 with a negative EPS. Since we need to ultimately compute earnings growth rates, we remove this single negative observation by replacing the 2008:Q4 EPS observation with the Shiller 12-month trailing total EPS observation for 2008:Q4. To convert EPS to total earnings, we next multiply the resulting quarterly EPS series by the quarterly S&P 500 divisor available at URL: `https://ycharts.com/indicators/sp\_500\_divisor`. Finally, to obtain a monthly S&P 500 earnings series, we linearly interpolate the resulting quarterly total earnings series. The final monthly total earnings series spans the period 1959:03 to 2021:12. The EPS data from S&P Global, Shiller, and the divisor data were downloaded on March 13, 2022.

We obtain S&P 500 Dividend from Robert Shiller's online data depository at URL: `http://www.econ.yale.edu/~shiller/data.htm`. The series is computed from the S&P four-quarter trailing totals and linearly interpolated to monthly data sampling intervals for the period 1959:01 to 2021:12 and was downloaded on February 15, 2022.

For the high-frequency filtering, we use tick-by-tick data on S&P 500 index from tickdata.com. The series was purchased and downloaded on July 2, 2022 from `https://www.tickdata.com/`. We create the minutely data using the close price within each minute. Within trading hours, we construct S&P 500 market capitalization by multiplying the minutely S&P 500 index value by last month's S&P 500 Divisor. The S&P 500 Divisor is available at the URL: `https://ycharts.com/indicators/sp\_500\_divisor`. We supplement S&P 500 index using S&P 500 E-mini futures for events that occur in off-market hours. We use the current-quarter contract futures. We purchased the S&P 500 E-mini futures from CME group at URL: urlhttps://datamine.cmegroup.com/. Our sample spans January 2nd 1986 to June 30th, 2022. The S&P 500 futures data were downloaded on July 2, 2022.

Net Dividends Plus Net Repurchases (Equity Payout)

We use an eight quarter moving average of equity payout divided by GDP as a noisy signal on K_t. Equity payout for the corporate sector is quarterly and measured as net dividends minus net equity issuance is computed using flow of funds data. Net dividends ("netdiv") is the series named for corporate business; net dividends paid (FA096121073.Q). Net repurchases are repurchases net of share issuance, so net repurchases is the negative of net equity issuance. Net equity issuance ("netequi") is the sum of Equity Issuance for Non-financial corporate business; corporate equities; liability (Table F.103, series FA103164103) and Eq-

uity Issuance for domestic financial sectors; corporate equities; liability (Table F.108, series FA793164105). Since netdiv and netequi are annualized, the quarterly payout is computed as payout=(netdiv-netequi)/4. The units are in millions of dollars. Source: Federal Reserve Board. We map the quarterly observation into the model implications for the share k_t in the last month of each quarter. The quarterly data span the period 1951:Q4 to 2022:Q3.

Survey Data on Stock Market Expectations

We use the surveys listed below in our structural estimation. Following Nagel and XU (2021), we use the mean values of the Gallup/UBS, CFO survey, and Livingston forecasts. For the University of Michigan Survey of Consumers (SOC), which are qualitative up/down forecasts, the structural estimation maps this onto model-implied investor expectations of one-year-ahead stock returns using the method described in Section 17 below. For comparison purposes, we compute machine forecasts for either the annual CRSP return, or S&P 500 price growth depending on which variable most closely aligns with the concept that survey respondents are asked to predict.

UBS/Gallup Survey Stock Return Forecasts The UBS/Gallup is a monthly survey of one-year-ahead stock market return expectations, obtained from `https://www.ubs.com/global/en.html`. We use the mean point forecast in our estimation and compare these to machine forecasts of the annual CRSP return. Gallup conducted 1,000 interviews of investors during the first two weeks of every month and results were reported on the last Monday of the month. The UBS/Gallup survey respondents report only the return they expect on their own portfolio. Following Nagel and Xu (2022) we assume this return includes dividends and impute market return expectations by regressing expected market returns on own portfolio expectations using the part of the sample where both are available and using the fitted value. We take a stand on the information set of respondents when each forecast was made, and we assume that respondents could have used all data released before they completed the survey. Since interviews are in the first two weeks of a month (e.g., February), we conservatively set the response deadline for the machine forecast to be the first day of the survey month (e.g., February 1), implying that we allow the machine to use information only up through the end of the previous month (e.g., through January 31st). This ensures that the machine only sees information that would have been available to all UBS/Gallup respondents for that survey month (February). This approach is conservative in the sense that it handicaps the machine, since all survey respondents who are being interviewed during the next month would have access to more timely information than the machine. Since the survey asks about the "one-year-ahead" we interpret the question to be asking about the forecast period spanning from

the current survey month to the same month one year ahead. The data spans the periods 1998:01 to 2007:12. The data were downloaded on August 8th, 2022.

Michigan Survey of Consumers (SOC) The SOC contains approximately 50 core questions, and a minimum of 500 interviews are conducted by telephone over the course of the entire month, each month. Table 20 of the Michigan Survey of Consumers (Soc) reports the probability of an increase in stock market in next year. The survey question was "*The next question is about investing in the stock market. Please think about the type of mutual fund known as a diversified stock fund. This type of mutual fund holds stock in many different companies engaged in a wide variety of business activities. Suppose that tomorrow someone were to invest one thousand dollars in such a mutual fund. Please think about how much money this investment would be worth one year from now. What do you think the percent chance that this one thousand dollar investment will increase in value in the year ahead, so that it is worth more than one thousand dollars one year from now?*"

When we use this survey forecast to compare to machine forecasts, we impute a point forecast for stock market returns using the method described in Section A below. We compare the imputed point forecast to machine forecasts of CRSP returns. When we use this survey in the structural estimation, we map the survey answer on probability onto model-implied investor expectations of one-year-ahead stock returns using the method described in Section 17 below.

For the SOC, interviews are conducted monthly typically over the course of an entire month. (In rare cases, interviews may commence at the end of the previous month, as in February 2018 when interviews began on January 31st 2018.) We take a stand on the information set of respondents when each forecast was made, and we assume that respondents could have used all data released before they completed the survey. Since interviews are almost always conducted over the course of an entire month (e.g., February), we conservatively set the response deadline for the machine forecast to be the first day of the survey month (e.g., February 1), implying that we allow the machine to use information only up through the end of the previous month (e.g., through January 31st). This ensures that the machine only sees information that would have been available to all respondents for that survey month (February). This approach is conservative in the sense that it handicaps the machine, since all survey respondents who are being interviewed during the next month would have access to more timely information than the machine. Since the survey asks about the "year ahead" we interpret the question to be asking about the forecast period spanning the period running from the current survey month to the same month one year ahead. The data spans 2002:06 to 2021:12. The SOC responses were obtained from https://data.sca.isr.umich.edu/data-

archive/mine.php and downloaded on August 13th, 2022.

The CFO Survey Stock Return Forecasts The CFO survey is a quarterly survey that asks respondents about their expectations for the S&P 500 return over the next 12 months, obtained from `https://www.richmondfed.org/research/national_economy/cfo_survey`. We use the mean point forecast in our estimation and compare these to machine forecasts of the annual CRSP return. The survey asks the respondent "*over the next 12 months, I expect the average annual S&P 500 return will be:* ___". Following Nagel and XU (2021), we assume the forecasted S&P 500 return includes dividends. The CFO survey panel includes firms that range from small operations to Fortune 500 companies across all major industries. Respondents include chief financial officers, owner-operators, vice presidents and directors of finance, and others with financial decision-making roles. The CFO panel has 1,600 members as of December 2022. As for the SOC, we take a stand on the information set of respondents when each forecast was made, and we assume that respondents could have used all data released before they completed the survey. Because the CFO survey releases quarterly forecasts at the end of each quarter, we conservatively set the response deadline for the machine forecast to be the first day of the last month of each quarter (e.g., March 1). The data spans the periods 2001Q4 to 2021Q1. The data were downloaded on August 8th, 2022.

Livingston Survey Stock Index Forecast We obtained the Livingston Survey S&P500 index forecast (SPIF) from the Federal Reserve Bank of Philadelphia, URL: `https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/livingston-historical-data`, and use the mean values in our structural and forecasting models. We compare the one-year growth in these forecasts to machine forecasts of S&P 500 price growth. Our sample spans 1947:06 to 2021:06. The forecast series were downloaded on September 20, 2021.

The survey provides semi-annual forecasts on the level of the S&P 500 index. Participants are asked to provide forecasts for the level of the S&P 500 index for the end of the current survey month, 6 months ahead, and 12 months ahead. We use the mean of the respondents' forecasts each period, where the sample is based on about 50 observations. Most of the survey participants are professional forecasters with "formal and advanced training in economic theory and forecasting and use econometric models to generate their forecasts." Participants receive questionnaires for the survey in May and November, after the Consumer Price Index (CPI) data release for the previous month. All forecasts are typically submitted by the end of the respective month of May and November. The results of the survey are released near the end of the following month, on June and December of each calendar year. The

exact release dates are available on the Philadelphia Fed website, at the header of each news release. We take a stand on the information set of the respondents when each forecast was made by assuming that respondents could have used all data released before they completed the survey. Since all forecasts are typically submitted by the end of May and November of each calendar year, we set the response deadline for the machine forecast to be the first day of the last month of June and December, implying that we allow the machine to use information only up through the end of the May and November.

We follow Nagel and Xu (2022) in constructing one-year stock price growth expectations from the level forecasts. Starting from June 1992, we use the ratio between the 12-month level forecast ($\texttt{SPIF_12M}_t$) and 0-month level nowcasts ($\texttt{SPIF_ZM}_t$) of the S&P 500 index. Before June 1992, the 0-month nowcast is not available. Therefore we use the annualized ratio between the 12-month ($\texttt{spi12}_t$) and 6-month ($\texttt{spi6}_t$) level forecast of the S&P 500 index

$$\mathbb{F}_t^{(Liv)}\left[\frac{P_{t+12}}{P_t}\right] \approx \begin{cases} \frac{\mathbb{F}_t^{(Liv)}[P_{t+12}]}{\mathbb{F}_t^{(Liv)}[P_t]} = \frac{\texttt{SPIF_12M}_t}{\texttt{SPIF_ZM}_t} & \text{if } t \geq 1992M6 \\ \left(\frac{\mathbb{F}_t^{(Liv)}[P_{t+12}]}{\mathbb{F}_t^{(Liv)}[P_{t+6}]}\right)^2 = \left(\frac{\texttt{spi12}_t}{\texttt{spi6}_t}\right)^2 & \text{if } t < 1992M6 \end{cases} \tag{A.1}$$

where P_t is the S&P 500 index and t indexes the survey's response deadline.

Bloomberg Consensus Survey Stock Index Forecasts As an additional signal of investor expectations, we use the Bloomberg (BBG) Consensus Forecast of the stock market. Survey respondents are asked to forecast the "end-of-year" closing value of the S&P 500 index on the last trading day of the calendar year (id: SPXSFRCS). The forecast horizon therefore changes depending on when in the year panelists are answering the survey. Surveys conducted between January through November forecast the index for the end-of-current-year (EOCY). Surveys held in December forecast the index for the end-of-next-year (EONY). For example: On January 2021, the survey forecasts the S&P 500 index 11 months ahead for the end of 2021. On November 2021, the survey forecasts the S&P 500 index 1 month ahead for the end of 2021. On December 2021, the survey forecasts the S&P 500 index 12 months ahead for the end of 2022. The data span the period from 16-Apr-1999 to 15-Jun-2022 and were downloaded from the Bloomberg terminal on July 8, 2022. The survey has been conducted irregularly over time. It was conducted roughly once a week from 1999 to 2014, roughly two to three times per month from 2014 to 2016, and once each month since 2017. We construct a monthly dataset of these survey observations by taking the last observation for a month as our monthly observation for the years 1999 to 2014.

We use these data to augment the estimation as an additional signal on stock market return expectations. This requires a mapping into the monthly subjective return expectation

counterpart from the model. Our procedure is to treat survey forecasts for month $M = 1, 2, 3, ..., 12$ as a signal on the 12-M month underlying expectations process for returns. Thus, for surveys conducted in January of a given year, we take the BBG forecast in January of the EOCY S&P 500 index value and divide it by the observed S&P 500 index value on December 31 of the immediately previous year. This observation is mapped into the model implications for 11-month-ahead subjective return expectations of investors. For surveys conducted in February of a given year, we take the BBG forecast of the EOCY S&P index value and divide it by the observed S&P value for January 31 of the current year and map that into the model implications for 10 month ahead subjective return expectations. We follow this procedure for all surveys conducted between January through November of each year. For surveys conducted in December, the BBG forecasts are for the end of the next year. Thus, for surveys conducted in December of a given year, we take the BBG forecast for the EONY S&P 500 index value and divide it by the observed S&P 500 index value on November 31 of the current year. This is mapped into the model implications for subjectively expected 12 month ahead returns. In all cases if the observation needed for the S&P 500 index value used in the divisor fell on a day in which the market was closed, we instead use the value for the index from the last trading day prior to this date.

Finally, we convert the end-of-year S&P 500 return forecasts to annualized units. For example, for all forecasts conducted in May, we raise our gross return forecasts to the power 12/7; for all forecasts conducted in June, we raise gross return forecasts by 12/6, and so on. For mapping to log returns, we instead multiply by 12/7, 12/6, and so on.

Converting Qualitative Forecasts to Point Forecasts (SOC) We use the SOC probability to impute a quantitative point forecast of stock returns using a linear regression of CFO point forecasts for returns onto the SOC probablity of a price increase. The SOC asks respondents about the percent chance that an investment will "increase in value in the year ahead." We interpret this as asking about the ex dividend value, i.e., about price price growth. The CFO survey is conducted quarterly, where the survey quarters span 2001:Q4 to 2021:Q1. The SOC survey is conducted monthly, where survey months span 2002:06 to 2021:12. Since the CFO is a quarterly survey, the regression is estimated in real-time over a quarterly overlapping sample. Since the CFO survey is conducted during the last month of the quarter while the SOC is conducted monthly, we align the survey months between CFO and SOC by regressing the quarterly CFO survey point forecast with the qualitative SOC survey response during the last month of the quarter. Since the SOC survey question is interpreted as asking about S&P 500 price growth while the CFO survey question asks about stock returns including dividends, we follow Nagel and Xu (2022) in subtracting the current

dividend yield of the CRSP value weighted index from the CFO variable before running the regression. After estimating the regression, we then add back the dividend yield to the fitted value to obtain an imputed SOC point forecast of stock returns including dividends.

Specifically, at time t, we assume that the CFO forecast of stock returns, $\mathbb{F}_t^{\text{CFO}}[r_{t,t+4}]$, minus the current dividend yield, D_t/P_t, is related to the contemporaneous SOC probability of an increase in the stock market next year, $P_{t,t+4}^{\text{SOC}}$, by:

$$\mathbb{F}_t^{\text{CFO}}[r_{t,t+4}] - D_t/P_t = \beta_0 + \beta_1 P_{t,t+4}^{\text{SOC}} + \epsilon_t.$$

The final imputed SOC point forecast is constructed as $\mathbb{F}_t^{\text{SOC}}[r_{t,t+4}] = \hat{\beta}_0 + \hat{\beta}_1 P_{t,t+4}^{\text{SOC}} + D_t/P_t$. We first estimate the coefficients of the above regression over an initial overlapping sample of 2002:Q2 to 2004:Q4, where the quarterly observations from the CFO survey is regressed on the SOC survey responses from the last month of each calendar quarter. Using the estimated coefficients and the SOC probability from 2005:03 gives us the point forecast of the one-year stock return from 2005:Q1 to 2006:Q1. We then re-estimate this equation, recursively, adding one quarterly observation to the end of the sample at a time, and storing the fitted values. This results in a time series of SOC point forecasts $\mathbb{F}_t^{\text{SOC}}[r_{t,t+4}]$ spanning 2005:Q1 to 2021:Q1.

Earnings Expectations

We obtained the monthly survey data for the median analyst earnings per share forecast and actual earnings per share from the Institutional Brokers Estimate System (IBES) via Wharton Research Data Services (WRDS). The data spans the period 1976:01 to 2021:12. All data were downloaded in October 2022.

We build measures of aggregate S&P 500 earnings expectations growth using the constituents of the S&P 500 at each point in time following De La O and Myers (2021). We first construct expected earnings expectations for aggregate earnings h-months-ahead as

$$\mathbb{F}_t[E_{t+h}] = \Omega_t \left[\sum_{i \in x_{t+h}} \mathbb{F}_t \left[EPS_{i,t+h} \right] S_{i,t} \right] / Divisor_t,$$

where $\mathbb{F}$ is the median analyst survey forecast, E is aggregate S&P 500 earnings, EPS_i is earning per share of firm i among all S&P 500 firms x_{t+h} for which we have forecasts in IBES for $t+h$, S_i is shares outstanding of firm i, and $Divisor_t$ is calculated as the S&P 500 market capitalization divided by the S&P 500 index. We obtain the number of outstanding shares for all companies in the S&P500 from Compustat. (All data from Compustat were

downloaded on November 17th, 2022.) IBES estimates are available for most but not all S&P 500 companies. Following De La O and Myers (2021), we multiply this aggregate by Ω_{t+h}, a ratio of total S&P 500 market value to the market value of the forecasted companies at $t+h$ to account for the fact that IBES does not provide earnings forecasts for all firms in the S&P 500 in every period.

IBES database contains earning forecasts up to five annual fiscal periods (FY1 to FY5) and as a result, we interpolate across the different horizons to obtain the expectation over the next 12 months. This procedure has been used in the literature, including De La O and Myers (2021). Specifically, if the fiscal year of firm XYZ ends nine months after the survey date, we have a 9-month earning forecast $\mathbb{F}_t[E_{t+9}]$ from FY1 and a 21-month forecast $\mathbb{F}_t[E_{t+21}]$ from FY2. We then obtain the 12-month ahead forecast by interpolating these two forecasts as follows,

$$\mathbb{F}_t[E_{t+12}] = \frac{9}{12}\mathbb{F}_t[E_{t+9}] + \frac{3}{12}\mathbb{F}_t[E_{t+21}].$$

For the forecasting performance estimates, we use quarterly data. To convert the monthly forecast to quarterly frequency, we use the forecast made in the middle month of each quarter, and construct one-year earnings expectations from 1976Q1 to 2021Q4 and the earning expectation growth is calculated as an approximation following following De La O and Myers (2021):

$$\mathbb{F}_t\left(\Delta e_{t+12}\right) \approx \ln\left(\mathbb{F}_t[E_{t+12}]\right) - e_t$$

where e_t is log earnings for S&P 500 at time t calculated as $e_t = \ln\left(EPS_t \cdot Divisor_t\right)$, where EPS_t is the earnings per share for the S&P 500 obtained from Shiller's data depository and S&P Global, as described above.

We constructed long term expected earnings growth (LTG) for the S&P 500 following Bordalo et al. (2019). Specifically, we obtained the median firm-level LTG forecast from IBES, and aggregate the value-weighted firm-level forecasts,

$$LTG_t = \sum_{i=1}^{S} LTG_{i,t} \frac{P_{i,t}Q_{i,t}}{\sum_{i=1}^{S} P_{i,t}Q_{i,t}}$$

where S is the number of firms in the S&P 500 index, and where $P_{i,t}$ and $Q_{i,t}$ are the stock price and the number of shares outsanding of firm i at time t, respectively. $LTG_{i,t}$ is the median forecast of firm i's long term expected earnings growth. The data spans the periods from 1981:12 to 2021:12. All data were downloaded in February 2023.

To estimate any biases in IBES analyst forecasts, our dynamic machine algorithm takes as an input a likely date corresponding to information analysts could have known at the

time of their forecast. IBES does not provide an explicit deadline for their forecasts to be returned. Therefore we instead use the "statistical period" day (the day when the set of summary statistics was calculated) as a proxy for the deadline. We set the machine deadline to be the day before this date. The statistical period date is typically between day 14 and day 20 of a given month, implying that the machine deadline varies from month to month. As the machine learning algorithm uses mixed-frequency techniques adapted to quarterly sampling intervals, while the IBES forecasts are monthly, we compare machine and IBES analyst forecasts as of the middle month of each quarter, considering 12-month ahead forecasts from that month.

Daily Earnings Nowcasts We obtain daily nowcasts of S&P 500 earnings per share (EPS) from BBG. The survey respondents are equity strategists that are asked to provide nowcasts of earnings per share (EPS) for the constituents of the S&P 500. For each S&P 500 constituent, BBG provides the mean nowcast across survey respondents as well as a bottom-up aggregate nowcast of EPS for the S&P 500 by aggregating the EPS nowcasts across the S&P 500 constituents. We construct a mean respondent nowcast for the level of S&P 500 earnings by multiplying this aggregate with the S&P 500 index divisor. (The index is the market capitalization of the 500 companies covered by the index divided by the S&P 500 divisor, roughly the number of shares outsanding across all companies.) The S&P 500 divisor is available at the URL: `https://ycharts.com/indicators/sp_500_divisor`. These nowcasts are available daily for the current standard financial quarter (Jan-Mar, Apr-June, Jul-Sep, Oct-Dec). For example, the observation for July 10, 2024 (which falls under the standard financial quarter of 2024:Q3) would contain nowcasts for 2024:Q3 EPS for the S&P 500. Bloomberg does not require respondents to submit their nowcasts on a specific timeline or frequency. Instead, respondents voluntarily decide how often to update their nowcasts. To ensure that consensus nowcasts are not heavily influenced by outdated information, Bloomberg excludes stale nowcasts submitted before the most recent earnings announcement date. The data was downloaded from the Bloomberg terminal on July 24, 2024, using the Earnings & Estimates (EE) function on the S&P 500 index (SPX Index). The aggregated consensus nowcasts are available daily, except weekends and holidays, spanning the period from Jan 2, 2006 to July 24, 2024. The divisor series series span the period 1959:03 to 2021:12 and was downloaded on March 13, 2022.

Dividend Growth Expectations

We obtained the S&P dividend futures from Bloomberg terminal and obtained data on S&P dividends from S&P (via Bloomberg terminal). The data spans the periods from 2015:01 to

2021:12 and are expressed in annual units. The series were downloaded on April 18th, 2023. We constructed estimates of S&P 500 dividend growth expectation following the procedure of Gormsen and Koijen (2020) by first calculating the equity yields as

$$e_t^{(n)} = \frac{1}{n}\ln\left(\frac{D_t}{F_t^{(n)}}\right)$$

where D_t is the S&P dividend, $F_t^{(n)}$ is the dividend futures with contact length of n years, where t is measured in quarters. We then run a regression of realized dividend growth rates on the S&P500 onto the 2-year equity yield

$$\Delta D_{t,t+8} = \beta_0^D + \beta_1^D e_t^{(2)} + \varepsilon_t.$$

We use the parameter estimates from this quarterly regression to estimate expected two-year-ahead dividend growth at daily frequency based on the fitted values and daily observations on $e_t^{(2)}$. To do this, since we have quarterly observations on D_t, we use the 2019 year-end value of dividends D_t for all days in 2020 as the numerator value for $e_t^{(2)}$. For the denominator, since the futures contracts always mature in December, to have a 2-year price in, for example, May 1 of 2020, we interpolate the futures price of the December contract of that year and the following year as $F_{2020,May01}^{INT,2} \equiv \frac{19}{24}F_{2020,May01}^{(19)} + \frac{5}{24}F_{2020,May01}^{(31)}$. Thus the daily observation for the yield on May 1, 2020 is the 2019 year-end value for D_t divided by $F_{2020,May01}^{INT,2}$.

Fed Funds Futures and Eurodollar Data

We use tick-by-tick data on Fed funds futures (FFF) and Eurodollar futures obtained from the CME Group. Our sample spans January 3, 1995 to June 30, 2022. FFF contracts settle based on the average federal funds rate that prevails over a given calendar month. Fed funds futures are priced at $100 - f_t^{(n)}$, where $f_t^{(n)}$ is the time-t contracted federal funds futures market rate that investors lock in. Contracts are monthly and expire at month-end, with maturities ranging up to 60 months. For the buyer of the futures contract, the amount of $\left(f_t^{(n)} - r_{t+n}\right) \times$ \$D, where r_{t+n} is the ex post realized value of the federal funds rate for month $t+n$ calculated as the average of the daily Fed funds rates in month $t+n$, and \$D is a dollar "deposit", represents the payoff of a zero-cost portfolio.

Eurodollar futures contracts are quarterly, expiring two business days before the third Wednesday in the last month of the quarter. Eurodollar futures are similarly quoted, where $f_t^{(q)}$ is the average 3-month LIBOR in quarter q of contract expiry. Maturities range up to 40

quarters. For both types of contracts, the implied contract rate is recovered by subtracting 100 from the price and multiplying by -1.

Both types of contracts are cleaned following the same procedure following communication with the CME Group. First, trades with zero volume, which indicate a canceled order, are excluded. Floor trades, which do not require a volume on record, are included. Next, trades with a recorded expiry (in YYMM format) of 9900 indicate bad data and are excluded (Only 1390 trades, or less than 0.01% of the raw Fed funds data, have contract delivery dates of 9900). For trades time stamped to the same second, we follow Bianchi, Kind and Kung (2019) and keep the trade with the lowest sequence number, corresponding to the first trade that second.

Fed funds futures data require additional cleaning. Trade prices were quoted in different units prior to August 2008. To standardize units across our sample, we start by noting that Fed funds futures are priced to the average effective Fed funds rate realized in the contract month. And in our sample, we expect a reasonable effective Fed funds rate to correspond to prices in the 90 to 100 range. As such, we rescale prices to be less than 100 in the pre-August 2008 subsample.[1] After rescaling, a small number of trades still appear to have prices that are far away from the effective Fed funds rates at both trade day and contract expiry, along with trades in the immediate transactions. The CME Group could not explain this data issue, so following Bianchi et al. (2019) and others in the high frequency equity literature, we apply an additional filter to exclude trades with such non-sensible prices. Specifically, for each maturity contract, we only keep trades where

$$|p_t - \overline{p}_t(k, \delta)| < 3\sigma_t(k, \delta) + \gamma,$$

where p_t denotes the trade price (where t corresponds to a second), and $\overline{p}_t(k, \delta)$ and $\sigma_t(k, \delta)$ denote the average price and standard deviation, respectively, centered with $k/2$ observations on each side of t excluding $\delta k/2$ trades with highest price and excluding $\delta k/2$ trades with lowest price. Finally, γ is a positive constant to account for the cases where prices are constant within the window. Our main specification uses $k = 30$, $\delta = 0.05$ and $\gamma = 0.4$, and alternative parameters produce similar results.

[1] For trades with prices significantly greater than 100, we repeatedly divide by 10 until prices are in the range of 90 to 100. We exclude all trades otherwise.

Historical Macro data (GDP and inflation)

Real Gross Domestic Product is obtained from the US Bureau of Economic Analysis. It is in billions of chained 2012 dollars, quarterly frequency, seasonally adjusted, and at annual rate. The source is from Bureau of Economic Analysis (BEA code: A191RX). The sample spans 1959:Q1 to 2021:Q4. The quarterly series was interpolated to monthly frequency using the method in Stock and Watson (2010).The quarterly series was downloaded on June 15th, 2022. Monthly inflation is measured as the log difference in the Consumer Price Index for all urban consumers, all items, seasonally adjusted, 1982=100, from FRED (CPIAUCSL). The sample spans 1959:01 to 2022:06. The monthly series was downloaded on August 17, 2022.

Real Time Macro Data (GDP and inflation)

At each forecast date in the sample, we construct a dataset of macro variables that could have been observed on or before the day of the survey deadline. We use the Philadelphia Fed's Real-Time Data Set to obtain vintages of macro variables.[2] These vintages capture changes to historical data due to periodic revisions made by government statistical agencies. We use the real time vintages of the same variables for GDP and inflation used for the historical data stipulated above. For real time GDP data we linearly interpolate the quarterly series to monthly values. For a complete list of the the details on variables used in real time, see the subsection below "Data Inputs for Machine Learning Algorithm."

Baa Spread, 20-yr T-bond, Long-term US government securities

We obtained daily Moody's Baa Corporate Bond Yield from FRED (series ID: DBAA) at URL: `https://fred.stlouisfed.org/series/BAA`, US Treasury securities at 20-year constant maturity from FRED (series ID: DGS20) at URL: `https://fred.stlouisfed.org/series/DGS20`, and long-term US government securities from FRED (series ID: LTGOVTBD) at URL: `https://fred.stlouisfed.org/series/LTGOVTBD`. The sample for Baa spans the periods 1986:01 to 2021:06. To construct the long term bond yields, we use LTGOVTBD before 2000 (1959:01 to 1999:12) and use DGS20 after 2000 (2000:01 to 2021:06). The Baa spread is the difference between the Moody's Corporate bond yield and the 20-year US government yield. The excess bond premium is obtained at URL: `https://www.federalreserve.gov/econres/notes/feds-notes/ebp_csv.csv`. All series

[2]The real-time data sets are available at https://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/data-files.

were downloaded on Feb 21, 2022.

Bloomberg Consensus Inflation and GDP forecasts

We obtain the Bloomberg (BBG) US GDP (id: ECGDUS) and inflation (id: ECPIUS) consensus mean forecast from the Bloomberg Terminal available on a daily basis up to a few days before the release of GDP and inflation data. The Bloomberg (BBG) US consensus forecasts are updated daily (except for weekends and holidays) and reports daily quarter-over-quarter real GDP growth and CPI forecasts from 2003:Q1 to 2021Q2. These forecasts provide more high-frequency information on the professional outlook for economic indicators. Both forecast series were downloaded on October 21, 2021.

Livingston Survey Inflation Forecast

We obtained the Livingston Survey mean 1-year and 10-year CPI inflation forecast from the Federal Reserve Bank of Philadelphia, URL: `https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/livingston-historical-data` and use the median values in our structural and forecasting models. Our sample spans 1947:06 to 2021:06. The forecast series were downloaded on September 20, 2021.

Bluechip Inflation and GDP Forecasts

We obtain Blue Chip expectation data from Blue Chip Financial Forecasts from Wolters Kluwer. The surveys are conducted each month by sending out surveys to forecasters in around 50 financial firms such as Bank of America, Goldman Sachs & Co., Swiss Re, Loomis, Sayles & Company, and J.P. Morgan Chase. The participants are surveyed around the 25th of each month and the results published a few days later on the 1st of the following month. The forecasters are asked to forecast the average of the level of U.S. interest rates over a particular calendar quarter, e.g. the federal funds rate and the set of H.15 Constant Maturity Treasuries (CMT) of the following maturities: 3-month, 6-month, 1-year, 2-year, 5-year and 10-year, and the quarter over quarter percentage changes in Real GDP, the GDP Price Index and the Consumer Price Index, beginning with the current quarter and extending 4 to 5 quarters into the future.

In this study, we look at a subset of the forecasted variables. Specifically, we use the Blue Chip micro data on individual forecasts of the quarter-over-quarter (Q/Q) percentage change in the Real GDP, the GDP Price Index and the CPI, and convert to quarterly observations

as explained below. In our estimation we use the median survey forecast from the micro data.

1. CPI inflation: We use quarter-over-quarter percentage change in the consumer price index, which is defined as

   "Forecasts for the quarter-over-quarter percentage change in the CPI (consumer prices for all urban consumers). Seasonally adjusted, annual rate."

   Quarterly and annual CPI inflation are constructed the same way as for PGDP inflation, except CPI replaces PGDP.

2. For real GDP growth, We use quarter-over-quarter percentage change in the Real GDP, which is defined as

   "Forecasts for the quarter-over-quarter percentage change in the level of chain-weighted real GDP. Seasonally adjusted, annual rate. Prior to 1992, Q/Q % change (SAAR) in real GNP."

The surveys are conducted right before the publication of the newsletter. Each issue is always dated the 1st of the month and the actual survey conducted over a two-day period almost always between 24th and 28th of the month. The major exception is the January issue when the survey is conducted a few days earlier to avoid conflict with the Christmas holiday. Therefore, we assume that the end of the last month (equivalently beginning of current month) is when the forecast is made. For example, for the report in 2008 Feb, we assume that the forecast is made on Feb 1, 2008. We obtained Blue Chip Financial Forecasts from Wolters Kluver in several stages starting in 2017 and with the last update purchased in June of 2022 and received on June 22, 2022. URL:https://law-store.wolterskluwer.com/s/product/blue-chip-financial-forecast-print/ 01tG000000LuDUCIA3.

Survey of Professional Forecasters (SPF)

The SPF is conducted each quarter by sending out surveys to professional forecasters, defined as forecasters. The number of surveys sent varies over time, but recent waves sent around 50 surveys each quarter according to officials at the Federal Reserve Bank of Philadelphia. Only forecasters with sufficient academic training and experience as macroeconomic forecasters are eligible to participate. Over the course of our sample, the number of respondents ranges from a minimum of 9, to a maximum of 83, and the mean number of respondents is 37. The surveys are sent out at the end of the first month of each quarter, and they are collected in the second or third week of the middle month of each quarter. Each survey asks respondents to provide

nowcasts and quarterly forecasts from one to four quarters ahead for a variety of variables. Specifically, we use the SPF micro data on individual forecasts of the price level, long-run inflation, and real GDP.[3] Below we provide the exact definitions of these variables as well as our method for constructing nowcasts and forecasts of quarterly and annual inflation for each respondent.[4]

We use the median values of the following variables in our structural estimation and forecasting models:

1. Quarterly and annual inflation (1968:Q4 - present): We use survey responses for the level of the GDP price index (PGDP), defined as

   "Forecasts for the quarterly and annual level of the chain-weighted GDP price index. Seasonally adjusted, index, base year varies. 1992-1995, GDP implicit deflator. Prior to 1992, GNP implicit deflator. Annual forecasts are for the annual average of the quarterly levels."

   Since advance BEA estimates of these variables for the current quarter are unavailable at the time SPF respondents turn in their forecasts, four quarter-ahead inflation and GDP growth forecasts are constructed by dividing the forecasted level by the survey respondent-type's nowcast. Let $\mathbb{F}_t^{(i)}[P_{t+h}]$ be forecaster i's prediction of PGDP h quarters ahead and $\mathbb{N}_t^{(i)}[P_t]$ be forecaster i's nowcast of PGDP for the current quarter. Annualized inflation forecasts for forecaster i are

   $$\mathbb{F}_t^{(i)}[\pi_{t+h,t}] = (400/h) \times \ln\left(\frac{\mathbb{F}_t^{(i)}[P_{t+h}]}{\mathbb{N}_t^{(i)}[P_t]}\right),$$

   where $h = 1$ for quarterly inflation and $h = 4$ for annual inflation. Similarly, we construct quarterly and annual nowcasts of inflation as

   $$\mathbb{N}_t^{(i)}[\pi_{t,t-h}] = (400/h) \times \ln\left(\frac{\mathbb{N}_t^{(i)}[P_t]}{P_{t-h}}\right),$$

   where $h = 1$ for quarterly inflation and $h = 4$ for annual inflation, and where P_{t-1} is the BEA's advance estimate of PGDP in the previous quarter observed by the respondent in time t, and P_{t-4} is the BEA's most accurate estimate of PGDP four quarters back.

[3] Individual forecasts for all variables can be downloaded at https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/historical-data/individual-forecasts.

[4] The SPF documentation file can be found at https://www.philadelphiafed.org/-/media/research-and-data/real-time-center/survey-of-professional-forecasters/spf-documentation.pdf?la=en.

After computing inflation for each survey respondent, we calculate the 5th through the 95th percentiles as well as the average, variance, and skewness of inflation forecasts across respondents.

2. Long-run inflation (1991:Q4 - present): We use survey responses for 10-year-ahead CPI inflation (CPI10), which is defined as

   "Forecasts for the annual average rate of headline CPI inflation over the next 10 years. Seasonally adjusted, annualized percentage points. The "next 10 years" includes the year in which we conducted the survey and the following nine years. Conceptually, the calculation of inflation is one that runs from the fourth quarter of the year before the survey to the fourth quarter of the year that is ten years beyond the survey year, representing a total of 40 quarters or 10 years. The fourth-quarter level is the quarterly average of the underlying monthly levels."

   Only the median response is provided for CPI10, and it is already reported as an inflation rate, so we do not make any adjustments and cannot compute other moments or percentiles.

3. Real GDP growth (1968:Q4 - present): We use the level of real GDP (RGDP), which is defined as

   *"Forecasts for the quarterly and annual level of chain-weighted real GDP. Seasonally adjusted, annual rate, base year varies. 1992-1995, fixed-weighted real GDP. Prior to 1992, fixed-weighted real GNP. Annual forecasts are for the annual average of the quarterly levels. Prior to 1981:Q3, RGDP is computed by using the formula NGDP / PGDP * 100."*

Source: Federal Reserve Bank of Philadelphia.All series were downloaded on September 17th, 2021.

Data used for News Events

Federal Reserve News Events

Federal Reserve news events are taken from Federal Open Market Committee news releases. We compile dates and times of FOMC meetings from 1994 to 2004 from Gürkaynak, Sack and Swanson (2005). The dates of the remaining FOMC meetings are collected from the Federal Reserve Board website. The times of statement releases were coalesced in the following priority: the Federal Reserve Board calendar, the Federal Reserve Board minutes,

Bloomberg's FOMC page, and the first news article to appear on Bloomberg. We only include scheduled meetings and unscheduled meetings where a statement was released. Our final database covers the period 1994:02 - 2021:12 and consists of 219 Fed news events.

Macro News Events

Macroeconomic data releases are news events cover news about GDP, CPI, employment data, and payroll data. To pin down the timing of when the macro news is released, we rely on published tables of releases from the Bureau of Labor Statistics (BLS), obtained from https://www.bls.gov/bls/archived_sched.htm. The published tables of releases for GDP are from the Bureau of Economic Analysis (BEA), obtained from https://www.bea.gov/news/archive. (A complete list of the release dates is available from the authors of each news release or through the Money Market Service Survey.) For GDP, the advance releases typically occur at 8:30AM EST on the last Thursday of the first month in the quarter following the quarter to which the data pertain. The 2nd and 3rd releases typically occur at 8:30am EST on the last Thursday of the second and third month in the quarter following the quarter to which the data pertain, respectively. For example, the advance release of real GDP for 2021:Q2 occurred on Thursday July 29, 2021. The advance release for 2021:Q2 was later revised in the second and third releases on Thursday August 26, 2021 and Thursday September 30, 2021, respectively. For core CPI, the releases occur monthly at 8:30AM EST around the 15th of each month following the month to which the data pertain. The releases typically occur during the second week of the month, either on a Tuesday, Wednesday, or Thursday. For example, the release of the core CPI of June 2021 occurred on Tuesday July 13th, 2021. For employment data (including the unemployment rate and nonfarm payroll), the releases typically occur at 8:30AM EST on the first Friday of the month following the month to which the data pertain. For example, the release of the unemployment rate for June 2021 occurred on Friday July 2nd, 2021. Our final database covers the period 1980:01- 2021:12 and consists of a total of 1482 macro news events.

Corporate Earnings News Events

We obtained days of big stock return jumps primarily attributable to corporate earnings news from Baker et al. (2019) (BBDS). In assigning days to categories of proximate causes for jumps, BBDS focus on articles from the Wall Street Journal (WSJ). To isolate events with stock market jumps that were attributable to corporate earnings news with high confidence, we choose events from BBDS that have (i) journalist confidence at or above a confidence score of 2.5 and (ii) weights on corporate topic of at least 0.75. A confidence score of

2.5 is about halfway between the median and 75th percentile of confidence scores given to category classifications over the full sample of WSJ articles studied by BBDS. The data were provided by the authors on March 12, 2023. Table A.1 shows the dates and daily change in the S&P 500 stock market index for our database, which covers the period 1985:09-2020:09 and consists of a total of 16 corporate earnings news events.

Table A.1: List of Corporate News Events

Date	Daily Δ
1999/03/23	-2.65%
2000/03/07	-2.67%
2000/10/19	3.49%
2001/04/05	4.39%
2002/01/29	-2.86%
2008/07/16	2.51%
2008/09/09	-3.36%
2008/09/15	-4.64%
2008/10/21	-3.06%
2008/10/22	-5.91%
2009/01/07	-3.00%
2009/01/20	-5.22%
2009/03/12	4.07%
2009/04/09	3.71%
2009/07/15	2.95%
2020/05/01	-2.80%

Data Inputs for Machine Learning Algorithm

Macro Data Surprises

These data are used as inputs into the machine learning forecasts. We obtain median forecasts for GDP growth (Q/Q percentage change), core CPI (Month/Month change), unemployment rate (percentage point), and nonfarm payroll (month/month change) from the Money Market Service Survey. The median market survey forecasts are compiled and published by the Money Market Services (MMS) the Friday before each release. We apply the approach used in Bauer and Swanson (2023) and define macroeconomic data surprise as the actual value of the data release minus the median expectation from MMS on the Friday immediately prior to that data release. The GDP growth forecasts are available quarterly

from 1990Q1 to 2022Q1. The core CPI forecast is available monthly from July 1989 to April 2022. The median forecasts for the unemployment rate and nonfarm payrolls are available monthly from Jan 1980 to May 2022, and Jan. 1985 to May 2022, respectively. All survey forecasts were downloaded from Haver Analytics on December 17, 2022. To pin down the timing of when the news was actually released we follow the published tables of releases from the Bureau of Labor Statistics (BLS), discussed below.

FOMC Surprises

FOMC surprises are defined as the changes in the current-month, 1, 2, 6, 12, and 24 month-ahead federal funds futures (FFF) contract rate and changes in the 1, 2, 4, and 8 quarter-ahead Eurodollar (ED) futures contract rate, from 10 minutes before to 20 minutes after each U.S. Federal Reserve Federal Open Market Committee (FOMC) announcement. The data on FFF and ED were downloaded on June 3rd 2022. When benchmarking against a survey, we use the last FOMC meeting before the survey deadline to compute surprises. For surveys that do not have a clear deadline, we compute surprises using from the last FOMC in the first month of the quarter. When benchmarking against moving average, we use the last FOMC meeting before the end of the first month in each quarter to compute surprises.

Real-Time Macro Data

This section gives details on the real time macro data inputs used in the machine learning forecasts. A subset of these series are used in the structural estimation. At each forecast date in the sample, we construct a dataset of macro variables that could have been observed on or before the day of the survey deadline. We use the Philadelphia Fed's Real-Time Data Set to obtain vintages of macro variables.[5] These vintages capture changes to historical data due to periodic revisions made by government statistical agencies. The vintages for a particular series can be available at the monthly and/or quarterly frequencies, and the series have monthly and/or quarterly observations. In cases where a variable has both frequencies available for its vintages and/or its observations, we choose one format of the variable. For instance, nominal personal consumption expenditures on goods is quarterly data with both monthly and quarterly vintages available; in this case, we use the version with monthly vintages.

Table A.2 gives the complete list of real-time macro variables. Included in the table is the first available vintages for each variable that has multiple vintages. We do not include the last

[5]The real-time data sets are available at https://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/data-files.

vintage because most variables have vintages through the present.[6] Table A.2 also lists the transformation applied to each variable to make them stationary before generating factors. Let X_{it} denote variable i at time t after the transformation, and let X_{it}^A be the untransformed series. Let $\Delta = (1 - L)$ with $LX_{it} = X_{it-1}$. There are seven possible transformations with the following codes:

1 Code lv: $X_{it} = X_{it}^A$
2 Code Δlv: $X_{it} = X_{it}^A - X_{it-1}^A$
3 Code $\Delta^2 lv$: $X_{it} = \Delta^2 X_{it}^A$
4 Code ln: $X_{it} = \ln(X_{it}^A)$
5 Code Δln: $X_{it} = \ln(X_{it}^A) - \ln(X_{it-1}^A)$
6 Code $\Delta^2 ln$: $X_{it} = \Delta^2 \ln(X_{it}^A)$
7 Code $\Delta lv/lv$: $X_{it} = (X_{it}^A - X_{it-1}^A)/X_{it-1}^A$

Table A.2: List of Macro Dataset Variables

No.	Short Name	Source	Tran	Description	First Vintage
			Group 1: Output and Income		
1	IPMMVMD	Philly Fed	Δln	Ind. production index - Manufacturing	1962:M11
2	IPTMVMD	Philly Fed	Δln	Ind. production index - Total	1962:M11
3	CUMMVMD	Philly Fed	lv	Capacity utilization - Manufacturing	1979:M8
4	CUTMVMD	Philly Fed	lv	Capacity utilization - Total	1983:M7
5	NCPROFATMVQD	Philly Fed	Δln	Nom. corp. profits after tax without IVA/CCAdj	1965:Q4
6	NCPROFATWMVQD	Philly Fed	Δln	Nom. corp. profits after tax with IVA/CCAdj	1981:Q1
7	OPHMVQD	Philly Fed	Δln	Output per hour - Business sector	1998:Q4
8	NDPIQVQD	Philly Fed	Δln	Nom. disposable personal income	1965:Q4
9	NOUTPUTQVQD	Philly Fed	Δln	Nom. GNP/GDP	1965:Q4
10	NPIQVQD	Philly Fed	Δln	Nom. personal income	1965:Q4
11	NPSAVQVQD	Philly Fed	Δlv	Nom. personal saving	1965:Q4
12	OLIQVQD	Philly Fed	Δln	Other labor income	1965:Q4

[6]For variables BASEBASAQVMD, NBRBASAQVMD, NBRECBASAQVMD, and TRBASAQVMD, the last available vintage is 2013:Q2.

Table A.2 (Cont'd)

No.	Short Name	Source	Tran	Description	First Vintage
13	PINTIQVQD	Philly Fed	Δln	Personal interest income	1965:Q4
14	PINTPAIDQVQD	Philly Fed	Δln	Interest paid by consumers	1965:Q4
15	PROPIQVQD	Philly Fed	Δln	Proprietors' income	1965:Q4
16	PTAXQVQD	Philly Fed	Δln	Personal tax and nontax payments	1965:Q4
17	RATESAVQVQD	Philly Fed	Δlv	Personal saving rate	1965:Q4
18	RENTIQVQD	Philly Fed	Δlv	Rental income of persons	1965:Q4
19	ROUTPUTQVQD	Philly Fed	Δln	Real GNP/GDP	1965:Q4
20	SSCONTRIBQVQD	Philly Fed	Δln	Personal contributions for social insurance	1965:Q4
21	TRANPFQVQD	Philly Fed	Δln	Personal transfer payments to foreigners	1965:Q4
22	TRANRQVQD	Philly Fed	Δln	Transfer payments	1965:Q4
23	CUUR0000SA0E	BLS	$\Delta^2 ln$	Energy in U.S. city avg., all urban consumers, not seasonally adj	
				Group 2: Employment	
24	EMPLOYMVMD	Philly Fed	Δln	Nonfarm payroll	1946:M12
25	HMVMD	Philly Fed	lv	Aggregate weekly hours - Total	1971:M9
26	HGMVMD	Philly Fed	lv	Agg. weekly hours - Goods-producing	1971:M9
27	HSMVMD	Philly Fed	lv	Agg. weekly hours - Service-producing	1971:M9
28	LFCMVMD	Philly Fed	Δln	Civilian labor force	1998:M11
29	LFPARTMVMD	Philly Fed	lv	Civilian participation rate	1998:M11
30	POPMVMD	Philly Fed	Δln	Civilian noninstitutional population	1998:M11
31	ULCMVQD	Philly Fed	Δln	Unit labor costs - Business sector	1998:Q4
32	RUCQVMD	Philly Fed	Δlv	Unemployment rate	1965:Q4
33	WSDQVQD	Philly Fed	Δln	Wage and salary disbursements	1965:Q4
				Group 3: Orders, Investment, Housing	
34	HSTARTSMVMD	Philly Fed	Δln	Housing starts	1968:M2
35	RINVBFMVQD	Philly Fed	Δln	Real gross private domestic inv. - Nonresidential	1965:Q4
36	RINVCHIMVQD	Philly Fed	Δlv	Real gross private domestic inv. - Change in private inventories	1965:Q4

Table A.2 (Cont'd)

No.	Short Name	Source	Tran	Description	First Vintage
37	RINVRESIDMVQD	Philly Fed	Δln	Real gross private domestic inv. - Residential	1965:Q4
38	CASESHILLER	S&P	Δln	Case-Shiller US National Home Price index/CPI	1987:M1
			Group 4: Consumption		
39	NCONGMMVMD	Philly Fed	Δln	Nom. personal cons. exp. - Goods	2009:M8
40	NCONHHMMVMD	Philly Fed	Δln	Nom. hh. cons. exp.	2009:M8
41	NCONSHHMMVMD	Philly Fed	Δln	Nom. hh. cons. exp. - Services	2009:M8
42	NCONSNPMMVMD	Philly Fed	Δln	Nom. final cons. exp. of NPISH	2009:M8
43	RCONDMMVMD	Philly Fed	Δln	Real personal cons. exp. - Durables	1998:M11
44	RCONGMMVMD	Philly Fed	Δln	Real personal cons. exp. - Goods	2009:M8
45	RCONHHMMVMD	Philly Fed	Δln	Real hh. cons. exp.	2009:M8
46	RCONMMVMD	Philly Fed	Δln	Real personal cons. exp. - Total	1998:M11
47	RCONNDMVMD	Philly Fed	Δln	Real personal cons. exp. - Nondurables	1998:M11
48	RCONSHHMMVMD	Philly Fed	Δln	Real hh. cons. exp. - Services	2009:M8
49	RCONSMMVMD	Philly Fed	Δln	Real personal cons. exp. - Services	1998:M11
50	RCONSNPMMVMD	Philly Fed	Δln	Real final cons. exp. of NPISH	2009:M8
51	NCONGMVQD	Philly Fed	Δln	Nom. personal cons. exp. - Goods	2009:Q3
52	NCONHHMVQD	Philly Fed	Δln	Nom. hh. cons. exp.	0209:Q3
53	NCONSHHMVQD	Philly Fed	Δln	Nom. hh. cons. exp. - Services	2009:Q3
54	NCONSNPMVQD	Philly Fed	Δln	Nom. final cons. exp. of NPISH	2009:Q3
55	RCONDMVQD	Philly Fed	Δln	Real personal cons. exp. - Durable goods	1965:Q4
56	RCONGMVQD	Philly Fed	Δln	Real personal cons. exp. - Goods	2009:Q3
57	RCONHHMVQD	Philly Fed	Δln	Real hh. cons. exp.	2009:Q3
58	RCONMVQD	Philly Fed	Δln	Real personal cons. exp. - Total	1965:Q4
59	RCONNDMVQD	Philly Fed	Δln	Real pesonal cons. exp. - Nondurable goods	1965:Q4
60	RCONSHHMVQD	Philly Fed	Δln	Real hh. cons. exp. - Services	2009:Q3
61	RCONSMVQD	Philly Fed	Δln	Real personal cons. exp. - Services	1965:Q4
62	RCONSNPMVQD	Philly Fed	Δln	Real final cons. exp. of NPISH	2009:Q3
63	NCONQVQD	Philly Fed	Δln	Nom. personal cons. exp.	1965:Q4

Table A.2 (Cont'd)

No.	Short Name	Source	Tran	Description	First Vintage
			Group 5: Prices		
64	PCONGMMVMD	Philly Fed	$\Delta^2 ln$	Price index for personal cons. exp. - Goods	2009:M8
65	PCONHHMMVMD	Philly Fed	$\Delta^2 ln$	Price index for hh. cons. exp.	2009:M8
66	PCONSHHMMVMD	Philly Fed	$\Delta^2 ln$	Price index for hh. cons. exp. - Services	2009:M8
67	PCONSNPMMVMD	Philly Fed	$\Delta^2 ln$	Price index for final cons. exp. of NPISH	2009:M8
68	PCPIMVMD	Philly Fed	$\Delta^2 ln$	Consumer price index	1998:M11
69	PCPIXMVMD	Philly Fed	$\Delta^2 ln$	Core consumer price index	1998:M11
70	PPPIMVMD	Philly Fed	$\Delta^2 ln$	Producer price index	1998:M11
71	PPPIXMVMD	Philly Fed	$\Delta^2 ln$	Core producer price index	1998:M11
72	PCONGMVQD	Philly Fed	$\Delta^2 ln$	Price index for personal. cons. exp. - Goods	2009:Q3
73	PCONHHMVQD	Philly Fed	$\Delta^2 ln$	Price index for hh. cons. exp.	2009:Q3
74	PCONSHHMVQD	Philly Fed	$\Delta^2 ln$	Price index for hh. cons. exp. - Services	2009:Q3
75	PCONSNPMVQD	Philly Fed	$\Delta^2 ln$	Price index for final cons. exp. of NPISH	2009:Q3
76	PCONXMVQD	Philly Fed	$\Delta^2 ln$	Core price index for personal cons. exp.	1996:Q1
77	CPIQVMD	Philly Fed	$\Delta^2 ln$	Consumer price index	1994:Q3
78	PQVQD	Philly Fed	$\Delta^2 ln$	Price index for GNP/GDP	1965:Q4
79	PCONQVQD	Philly Fed	$\Delta^2 ln$	Price index for personal cons. exp.	1965:Q4
80	PIMPQVQD	Philly Fed	$\Delta^2 ln$	Price index for imports of goods and services	1965:Q4
			Group 6: Trade and Government		
81	REXMVQD	Philly Fed	Δln	Real exports of goods and services	1965:Q4
82	RGMVQD	Philly Fed	Δln	Real government cons. and gross inv. - Total	1965:Q4
83	RGFMVQD	Philly Fed	Δln	Real government cons. and gross inv. - Federal	1965:Q4
84	RGSLMVQD	Philly Fed	Δln	Real government cons. and gross. inv. - State and local	1965:Q4
85	RIMPMVQD	Philly Fed	Δln	Real imports of goods and services	1965:Q4
86	RNXMVQD	Philly Fed	Δlv	Real net exports of goods and services	1965:Q4
			Group 7: Money and Credit		
87	BASEBASAQVMD	Philly Fed	$\Delta^2 ln$	Monetary base	1980:Q2

Table A.2 (Cont'd)

No.	Short Name	Source	Tran	Description	First Vintage
88	M1QVMD	Philly Fed	$\Delta^2 ln$	M1 money stock	1965:Q4
89	M2QVMD	Philly Fed	$\Delta^2 ln$	M2 money stock	1971:Q2
90	NBRBASAQVMD	Philly Fed	$\Delta lv/lv$	Nonborrowed reserves	1967:Q3
91	NBRECBASAQVMD	Philly Fed	$\Delta lv/lv$	Nonborrowed reserves plus extended credit	1984:Q2
92	TRBASAQVMD	Philly Fed	$\Delta^2 ln$	Total reserves	1967:Q3
93	DIVQVQD	Philly Fed	Δln	Dividends	1965:Q4

Daily Financial Data

Daily Data and construction of daily factors These data are used in the machine learning forecasts. The daily financial series in this data set are from the daily financial dataset used in Andreou, Ghysels and Kourtellos (2013). We create a smaller daily database which is a subset of the large cross-section of 991 daily series in their dataset. Our dataset covers five classes of financial assets: (i) the Commodities class; (ii) the Corporate Risk category; (iii) the Equities class; (iv) the Foreign Exchange Rates class and (v) the Government Securities.

The dataset includes up to 87 daily predictors in a daily frequency from 23-Oct-1959 to 24-Oct-2021 (14852 trading days) from the above five categories of financial assets. We remove series with fewer than ten years of data and time periods with no variables observed, which occurs for some series in the early part of the sample. For those years, we have less than 87 series. There are 39 commodity variables which include commodity indices, prices and futures, 16 corporate risk series, 9 equity series which include major US stock market indices and the 500 Implied Volatility, 16 government securities which include the federal funds rate, government treasury bills of securities from three months to ten years, and 7 foreign exchange variables which include the individual foreign exchange rates of major five US trading partners and two effective exchange rate. We choose these daily predictors because they are proposed in the literature as good predictors of economic growth.

We construct daily financial factors in a quarterly frequency in two steps. First, we use these daily financial time series to form factors at a daily frequency. The raw data used to form factors are always transformed to achieve stationarity and standardized before performing factor estimation (see generic description below). We re-estimate factors at each date in the sample recursively over time using the entire history of data available in real time

prior to each out-of-sample forecast.

In the second step, we convert these daily financial indicators to quarterly weighted variables to form quarterly factors by selecting an optimal weighting scheme according to the method described below (see the weighting scheme section).

The data series used in this dataset are listed below in Table A.3 by data source. The tables also list the transformation applied to each variable to make them stationary before generating factors. The transformations used to stationarize a time series are the same as those explained in the section "Monthly financial factor data".

Table A.3: List of Daily Financial Dataset Variables

No.	Short Name	Source	Tran	Description
			Group 1: Commodities	
1	GSIZSPT	Data Stream	Δln	S&P GSCI Zinc Spot - PRICE INDEX
2	GSSBSPT	Data Stream	Δln	S&P GSCI Sugar Spot - PRICE INDEX
3	GSSOSPT	Data Stream	Δln	S&P GSCI Soybeans Spot - PRICE INDEX
4	GSSISPT	Data Stream	Δln	S&P GSCI Silver Spot - PRICE INDEX
5	GSIKSPT	Data Stream	Δln	S&P GSCI Nickel Spot - PRICE INDEX
6	GSLCSPT	Data Stream	Δln	S&P GSCI Live Cattle Spot - PRICE INDEX
7	GSLHSPT	Data Stream	Δln	S&P GSCI Lean Hogs Index Spot - PRICE INDEX
8	GSILSPT	Data Stream	Δln	S&P GSCI Lead Spot - PRICE INDEX
9	GSGCSPT	Data Stream	Δln	S&P GSCI Gold Spot - PRICE INDEX
10	GSCTSPT	Data Stream	Δln	S&P GSCI Cotton Spot - PRICE INDEX
11	GSKCSPT	Data Stream	Δln	S&P GSCI Coffee Spot - PRICE INDEX
12	GSCCSPT	Data Stream	Δln	S&P GSCI Cocoa Index Spot - PRICE INDEX
13	GSIASPT	Data Stream	Δln	S&P GSCI Aluminum Spot - PRICE INDEX
14	SGWTSPT	Data Stream	Δln	S&P GSCI All Wheat Spot - PRICE INDEX
15	EIAEBRT	Data Stream	Δln	Europe Brent Spot FOB U$/BBL Daily
16	CRUDOIL	Data Stream	Δln	Crude Oil-WTI Spot Cushing U$/BBL - MID PRICE
17	LTICASH	Data Stream	Δln	LME-Tin 99.85% Cash U$/MT
18	CWFCS00	Data Stream	Δln	CBT-WHEAT COMPOSITE FUTURES CONT. - SETT. PRICE

Table A.3 (Cont'd)

No.	Short Name	Source	Tran	Description
19	CCFCS00	Data Stream	Δln	CBT-CORN COMP. CONTINUOUS - SETT. PRICE
20	CSYCS00	Data Stream	Δln	CBT-SOYBEANS COMP. CONT. - SETT. PRICE
21	NCTCS20	Data Stream	Δln	CSCE-COTTON #2 CONT.2ND FUT - SETT. PRICE
22	NSBCS00	Data Stream	Δln	CSCE-SUGAR #11 CONTINUOUS - SETT. PRICE
23	NKCCS00	Data Stream	Δln	CSCE-COFFEE C CONTINUOUS - SETT. PRICE
24	NCCCS00	Data Stream	Δln	CSCE-COCOA CONTINUOUS - SETT. PRICE
25	CZLCS00	Data Stream	Δln	ECBOT-SOYBEAN OIL CONTINUOUS - SETT. PRICE
26	COFC01	Data Stream	Δln	CBT-OATS COMP. TRc1 - SETT. PRICE
27	CLDCS00	Data Stream	Δln	CME-LIVE CATTLE COMP. CONTINUOUS - SETT. PRICE
28	CLGC01	Data Stream	Δln	CME-LEAN HOGS COMP. TRc1 - SETT. PRICE
29	NGCCS00	Data Stream	Δln	CMX-GOLD 100 OZ CONTINUOUS - SETT. PRICE
30	LAH3MTH	Data Stream	Δln	LME-Aluminium 99.7% 3 Months U$/MT
31	LED3MTH	Data Stream	Δln	LME-Lead 3 Months U$/MT
32	LNI3MTH	Data Stream	Δln	LME-Nickel 3 Months U$/MT
33	LTI3MTH	Data Stream	Δln	LME-Tin 99.85% 3 Months U$/MT
34	PLNYD	www.macrotrends.net	Δln	Platinum Cash Price (U$ per troy ounce)
35	XPDD	www.macrotrends.net	Δln	Palladium (U$ per troy ounce)
36	CUS2D	www.macrotrends.net	Δln	Corn Spot Price (U$/Bushel)
37	SoybOil	www.macrotrends.net	Δln	Soybean Oil Price (U$/Pound)
38	OATSD	www.macrotrends.net	Δln	Oat Spot Price (US$/Bushel)
39	WTIOilFut	US EIA	Δln	Light Sweet Crude Oil Futures Price: 1St Expiring Contract Settlement ($/Bbl)
			Group 2: Equities	
40	S&PCOMP	Data Stream	Δln	S&P 500 COMPOSITE - PRICE INDEX
41	ISPCS00	Data Stream	Δln	CME-S&P 500 INDEX CONTINUOUS - SETT. PRICE
42	SP5EIND	Data Stream	Δln	S&P500 ES INDUSTRIALS - PRICE INDEX
43	DJINDUS	Data Stream	Δln	DOW JONES INDUSTRIALS - PRICE INDEX

Table A.3 (Cont'd)

No.	Short Name	Source	Tran	Description
44	CYMCS00	Data Stream	Δln	CBT-MINI DOW JONES CONTINUOUS - SETT. PRICE
45	NASCOMP	Data Stream	Δln	NASDAQ COMPOSITE - PRICE INDEX
46	NASA100	Data Stream	Δln	NASDAQ 100 - PRICE INDEX
47	CBOEVIX	Data Stream	lv	CBOE SPX VOLATILITY VIX (NEW) - PRICE INDEX
48	S&P500toVIX	Data Stream	Δln	S&P500/VIX
			Group 3: Corporate Risk	
49	LIBOR	FRED	Δlv	Overnight London Interbank Offered Rate (%)
50	1MLIBOR	FRED	Δlv	1-Month London Interbank Offered Rate (%)
51	3MLIBOR	FRED	Δlv	3-Month London Interbank Offered Rate (%)
52	6MLIBOR	FRED	Δlv	6-Month London Interbank Offered Rate (%)
53	1YLIBOR	FRED	Δlv	One-Year London Interbank Offered Rate (%)
54	1MEuro-FF	FRED	lv	1-Month Eurodollar Deposits (London Bid) (% P.A.) minus Fed Funds
55	3MEuro-FF	FRED	lv	3-Month Eurodollar Deposits (London Bid) (% P.A.) minus Fed Funds
56	6MEuro-FF	FRED	lv	6-Month Eurodollar Deposits (London Bid) (% P.A.) minus Fed Funds
57	APFNF-AANF	Data Stream	lv	1-Month A2/P2/F2 Nonfinancial Commercial Paper (NCP) (% P. A.) minus 1-Month Aa NCP (% P.A.)
58	APFNF-AAF	Data Stream	lv	1-Month A2/P2/F2 NCP (% P.A.) minus 1-Month Aa Financial Commercial Paper (% P.A.)
59	TED	Data Stream, FRED	lv	3Month Tbill minus 3-Month London Interbank Offered Rate (%)
60	MAaa-10YTB	Data Stream	lv	Moody Seasoned Aaa Corporate Bond Yield (% P.A.) minus Y10-Tbond
61	MBaa-10YTB	Data Stream	lv	Moody Seasoned Baa Corporate Bond Yield (% P.A.) minus Y10-Tbond

Table A.3 (Cont'd)

No.	Short Name	Source	Tran	Description
62	MLA-10YTB	Data Stream, FRED	lv	Merrill Lynch Corporate Bonds: A Rated: Effective Yield (%) minus Y10-Tbond
63	MLAA-10YTB	Data Stream, FRED	lv	Merrill Lynch Corporate Bonds: Aa Rated: Effective Yield (%) minus Y10-Tbond
64	MLAAA-10YTB	Data Stream, FRED	lv	Merrill Lynch Corporate Bonds: Aaa Rated: Effective Yield (%) minus Y10-Tbond
			Group 4: Treasuries	
65	FRFEDFD	Data Stream	Δlv	US FED FUNDS EFF RATE (D) - MIDDLE RATE
66	FRTBS3M	Data Stream	Δlv	US T-BILL SEC MARKET 3 MONTH (D) - MIDDLE RATE
67	FRTBS6M	Data Stream	Δlv	US T-BILL SEC MARKET 6 MONTH (D) - MIDDLE RATE
68	FRTCM1Y	Data Stream	Δlv	US TREASURY CONST MAT 1 YEAR (D) - MIDDLE RATE
69	FRTCM10	Data Stream	Δlv	US TREASURY CONST MAT 10 YEAR (D) - MIDDLE RATE
70	6MTB-FF	Data Stream	lv	6-month treasury bill market bid yield at constant maturity (%) minus Fed Funds
71	1YTB-FF	Data Stream	lv	1-year treasury bill yield at constant maturity (% P.A.) minus Fed Funds
72	10YTB-FF	Data Stream	lv	10-year treasury bond yield at constant maturity (% P.A.) minus Fed Funds
73	6MTB-3MTB	Data Stream	lv	6-month treasury bill yield at constant maturity (% P.A.) minus 3M-Tbills
74	1YTB-3MTB	Data Stream	lv	1-year treasury bill yield at constant maturity (% P.A.) minus 3M-Tbills
75	10YTB-3MTB	Data Stream	lv	10-year treasury bond yield at constant maturity (% P.A.) minus 3M-Tbills
76	BKEVEN05	FRB	lv	US Inflation compensation: continuously compounded zero-coupon yield: 5-year (%)

Table A.3 (Cont'd)

No.	Short Name	Source	Tran	Description
77	BKEVEN10	FRB	lv	US Inflation compensation: continuously compounded zero-coupon yield: 10-year (%)
78	BKEVEN1F4	FRB	lv	BKEVEN1F4
79	BKEVEN1F9	FRB	lv	BKEVEN1F9
80	BKEVEN5F5	FRB	lv	US Inflation compensation: coupon equivalent forward rate: 5-10 years (%)
		Group 5: Foreign Exchange (FX)		
81	US_CWBN	Data Stream	Δln	US NOMINAL DOLLAR BROAD INDEX - EXCHANGE INDEX
82	US_CWMN	Data Stream	Δln	US NOMINAL DOLLAR MAJOR CURR INDEX - EXCHANGE INDEX
83	US_CSFR2	Data Stream	Δln	CANADIAN $ TO US $ NOON NY - EXCHANGE RATE
84	EU_USFR2	Data Stream	Δln	EURO TO US$ NOON NY - EXCHANGE RATE
85	US_YFR2	Data Stream	Δln	JAPANESE YEN TO US $ NOON NY - EXCHANGE RATE
86	US_SFFR2	Data Stream	Δln	SWISS FRANC TO US $ NOON NY - EXCHANGE RATE
87	US_UKFR2	Data Stream	Δln	UK POUND TO US $ NOON NY - EXCHANGE RATE

LDA Data

These data are used as inputs into the machine learning forecasts. The database for our Latent Dirichlet Allocation (LDA) analysis contains around one million articles published in *Wall Street Journal* between January 1984 to June 2022. The current vintage of the results reported here is based a randomly selected sub-sample of 200,000 articles over the same period, one-fifth size of the entire database. The sample selection procedures follows Bybee et al. (2021). First, we remove all articles prior to January 1984 and after June 2022 and exclude articles published in weekends. Second, we exclude articles with subject tags associated with obviously non-economic content such as sports. Third, we exclude articles with the certain headline patterns, such as those associated with data tables or those corresponding to regular sports, leisure, or books columns. We filter the articles using the same list of exclusions provided by Bybee et al. (2021). Last, we exclude articles with

less than 100 words.

Processing of texts The processing of the texts can be summarized in the following five steps.

1. Tokenization: parse each article's text into a white-space-separated word list retaining the article's word ordering.

2. We drop all non-alphabetical characters and set the remaining characters to lower-case, remove words with less than 3 letters, and remove common stop words and URL-based terms. We use a standard list of stop words from the Python library *gensim.parsing.preprocessing*.

3. Lemmatization and Stemming: lemmatization returns the original form of a word using external dictionary *Textblob.Word* in Python and based on the context of the word. For instance, as a verb, "went" is converted to"go". Stemming usually refers to a heuristic process that remove the trailing letters at the end of the words, such as from "assesses" to "assess', and "really" to "real". We use the Python library *Textblob.Word* to implement the lemmatization and *SnowballStemmer* for the stemming. The results are not very sensitive to the particular Python packages being used.

4. From the first three steps, we obtain a list of uni-grams which are a list of singular words. For example, "united" and "states" are uni-grams from "united states". From the list of uni-grams, we generate a set of bi-grams as all pairs of (ordered) adjacent uni-grams. For example, "united states" together is one bi-gram. We then exclude uni-grams and bi-grams appearing in less than 0.1% of articles.

5. Last, we convert an article's word list into a vector of counts for each uni-gram and bi-gram. For example, the vector of counts $[5, 7, 2]$ corresponds to the number of times the words $[\text{"}federal\text{"}, \text{"}reserve\text{"}, \text{"}bank\text{"}]$ appear in the article.

The LDA Model The LDA model Blei, Ng and Jordan (2003) essentially achieves substantial dimension reduction of the word distribution of each article using the following assumptions. We assume a factor structure on the vectors of word counts. Each factor is a topic and each article is a parametric distribution of topics, specified as follows,

$$\underbrace{\overbrace{w_i}^{V\times 1}}_{\text{word dist of article } i} \sim \text{Mult}\left(\underbrace{\overbrace{\Phi'}^{V\times K}}_{\text{topic-word dist.}}, \underbrace{\overbrace{\theta_i}^{K\times 1}}_{\text{topic dist.}}, \underbrace{N_i}_{\text{\# of words}}\right) \tag{A.2}$$

where Mult is the multinomial distribution. In the above equation, w_i is a vector of word counts of each unique term (uni-gram or bi-gram) in article i, whose size is equal to the number of unique terms V. K is the number of factors in article i. In the estimation, we assume $K = 180$ following Bybee et al. (2021). Φ is a matrix sized $K \times V$, whose kth row and vth column is equal to the probability of the unique term v showing up in topic k. θ_i stores the weights of all k topics contained in article i, which sum up to one. Dimension reduction is achieved as long as $K << V$ (the number of topics are significantly smaller than the number of unique terms). More specifically, it reduces the dimension from $T \times V$ to $T \times K$ (the size of θ) + $K \times V$ (the size of Φ).

Real-time news factors. We also generate real-time news factors for each month t starting from January 1991. In theory, we could train the LDA model using each real-time monthly vintage but it is computationally challenging. Instead, we simplify the procedure by training the LDA model using quarterly vintages t, t+3, t+6, etc, and use the LDA model parameters estimated at t to filter news paper articles within the quarter and generate news factors for those months. More specifically, given every article's word distribution $w_{i,t+s}$,for $s = 0, 1, 2$, and the estimated real-time topic-word distribution parameters $\hat{\Phi}_t$ using articles till date t, one can obtain the filtered topic distribution of each article $\hat{\theta}_{i,t+s}$, as follows,

$$\underbrace{\overbrace{w_{i,t+s}}^{V\times 1}}_{\text{word dist of article } i \text{ at time } t+s} \sim \text{Mult}\left(\underbrace{\overbrace{\hat{\Phi}'}^{V\times K}}_{\text{topic-word dist.}}, \underbrace{\overbrace{\hat{\theta}_{i,t+s}}^{K\times 1}}_{\text{topic dist.}}, \underbrace{N_{i,t+s}}_{\#\text{ of words}}\right). \tag{A.3}$$

LDA Estimation We use the built-in LDA model estimation toolbox in the Python library `https://pypi.org/project/gensim/`Gensim to implement the model estimation. The model requires following initial inputs and parameters and it is estimated using Bayesian methods.[7]

1. We create a document-term matrix $\mathbf{W}$ as a collection of w_i for all articles i in the sample. The number of rows in $\mathbf{W}$ is equal to the number of articles in our sample and the number of columns in $\mathbf{W}$ is equal to the number of unique uni-gram and bi-grams (after being filtered) across all articles. The matrix $\mathbf{W}$ is used as an input for the LDA model estimation. We then follow Bybee et al. (2021) and set the number of topics K to be 180.[8]

2. In the Python library Gensim, the key parameters of the LDA estim are α and β.With

[7] In theory, maximum-likelihood estimation is possible but it is computationally challenging.
[8] The authors used Bayesian criteria to find 180 to be an optimal number of topics.

a higher value of α, the documents are composed of more topics. With a higher values of β, each topic contains more terms (uni- or bi-grams). In the implementations, we do not impose any explicit restrictions on initial values of those parameters and set them to be "auto". These two parameters, alongside Φ' and $\{\theta_i\}_i$, are estimated by the toolbox from Python library `https://pypi.org/project/gensim/`Gensim.

Real-time LDA Factors With the estimated topic weights $\theta_{i,t}$ of each article i from the LDA model, we fruther construct time series of the overall news attention to each topic, or a news factor. The value of the topic k at time t is the average weights of topic k of all articles published at t, specified as follows,

$$F_{k,t} = \frac{\sum_i \hat{\theta}_{i,k,t}}{\#\text{ of articles at } t} \tag{A.4}$$

for all topics k.

B Machine Learning

Machine Algorithm Details

The basic dynamic algorithm follows the six step approach of Bianchi et al. (2022a) of 1. Sample partitioning, 2. In-sample estimation, 3. Training and cross-validation, 4. Grid reoptimization, 5. Out-of-sample prediction, and 6. Roll forward and repeat. We refer the interested reader to that paper for details and discuss details of the implementation here only insofar as they differ.

At time t, a prior training sample of size $\dot{T}$ is partitioned into two subsample windows: an "estimation" subsample consisting of the first T_E observations, and a hold-out "validation" sample of T_V subsequent observations so that $\dot{T} = T_E + T_V$. The model to be estimated over the estimation subsample is

$$y_{j,t+h} = G^e\left(\mathcal{X}_t, \boldsymbol{\beta}_{j,h,t}\right) + \epsilon_{jt+h}.$$

where $y_{j,t+h}$ is a time series indexed by j whose value in period $h \geq 1$ the machine is asked to predict at time t, $\mathcal{X}_t$ is a large input dataset of right-hand-side variables including the intercept, and $G^e(\cdot)$ is a machine learning estimator that can be represented by a (potentially) high dimensional set of finite-valued parameters $\boldsymbol{\beta}^e_{j,h,t}$. We consider two estimators for $G^e(\cdot)$: Elastic Net $G^{\text{EN}}(\mathcal{X}_t, \boldsymbol{\beta}^{\text{EN}}_{j,h})$, and Long Short-Term Memory (LSTM) network

$G^{\text{LSTM}}(\mathcal{X}_t, \boldsymbol{\beta}^{\text{LSTM}}_{j,h})$. The $e \in \{\text{EN}, \text{LSTM}\}$ superscripts on $\boldsymbol{\beta}$ indicate that the parameters depend on the estimator being used (See the next section for a description of EN and LSTM). $\mathcal{X}_t$ always denotes the most recent data that would have been in real time prior to the date on which the forecast was submitted. To ensure that the effect of each variable in the input vector is regularized fairly during the estimation, we standardize the elements of $\mathcal{X}_t$ such that sample means are zero and sample standard deviations are unity.It should be noted that the most recent observation on the left-hand-side is generally available in real time only with a one-period lag, thus the forecasting estimations can only be run with data over a sample that stops one period later than today in real time.

The parameters $\boldsymbol{\beta}^e_{j,h,t}$ are estimated by minimizing the mean-square loss function with L_1 and L_2 penalties

$$L(\boldsymbol{\beta}^e_{j,h,t}, \mathbf{X}_{T_E}, \boldsymbol{\lambda}^e_t) \equiv \underbrace{\frac{1}{T_E}\sum_{\tau=1}^{T_E}\left(y_{j,\tau+h} - G^e\left(\mathcal{X}_t, \boldsymbol{\beta}^e_{j,h,t}\right)\right)^2}_{\text{Mean Square Error}} + \underbrace{\lambda^e_{1,t}\sum_{k=1}^{K}\left|\boldsymbol{\beta}^e_{j,h,t,k}\right|}_{L_1 \text{ Penalty}} + \underbrace{\lambda^e_{2,t}\sum_{k=1}^{K}(\boldsymbol{\beta}^e_{j,h,t,k})^2}_{L_2 \text{ Penalty}}$$

where $\mathbf{X}_{T_E} = (y_{j,t-T_E}, \ldots, y_{j,t}, \mathcal{X}'_{t-T_E}, \ldots, \mathcal{X}'_t)'$ is the vector containing all observations in an estimation subsample of size T_E. The estimated $\boldsymbol{\beta}^e_{j,h,t}$ is a function of the data $\boldsymbol{X}_{T_E}$ and a non-negative regularization parameter vector $\boldsymbol{\lambda}^e_t = \left(\lambda^e_{1,t}, \lambda^e_{2,t}, \boldsymbol{\lambda}^{LSTM}_t\right)'$ where $\boldsymbol{\lambda}^{LSTM}_t$ is a set of hyperparameters only relevant when using the LSTM estimator for $G^e(\cdot)$ (see below). For the EN case there are only two hyperparameters, which determine the optimal shrinkage and sparsity of the time t machine specification. The regularization parameters $\boldsymbol{\lambda}^e_t$ are estimated by minimizing the mean-square loss over pseudo-out-of-sample forecast errors generated from

rolling regressions through the validation sample:

$$\widehat{\boldsymbol{\lambda}}_t^e, \widehat{T}_E, \widehat{T}_V = \underset{\boldsymbol{\lambda}_t^e, T_E, T_V}{\operatorname{argmin}} \left\{ \frac{1}{T_V - h} \sum_{\tau=T_E}^{T_E+T_V-h} \left(y_{j,\tau+h} - G^{EN}(\mathcal{X}_\tau, \widehat{\boldsymbol{\beta}}_{j,h,\tau}^{EN}(\mathbf{X}_{T_E})) \right)^2 + \underbrace{\lambda_{1,t}^{EN} \sum_{k=1}^{K} \boldsymbol{\beta}_{j,h,t,k}^{EN}}_{L_1 \text{ Penalty}} + \underbrace{\lambda_{2,t}^{EN} \sum_{k=1}^{K} (\boldsymbol{\beta}_{j,h,t,k}^{EN})^2}_{L_2 \text{ Penalty}} \right\}$$

$$\widehat{\boldsymbol{\lambda}}_t^e, \widehat{T}_E, \widehat{T}_V = \underset{\boldsymbol{\lambda}_t^e, T_E, T_V}{\operatorname{argmin}} \left\{ \frac{1}{T_V - h} \sum_{\tau=T_E}^{T_E+T_V-h} \left(y_{j,\tau+h} - G^{LSTM}(\mathcal{X}_\tau, \widehat{\boldsymbol{\beta}}_{j,h,\tau}^{LSTM}(\mathbf{X}_{T_E}, \boldsymbol{\lambda}_t^{LSTM})) \right)^2 + \underbrace{\lambda_{1,t}^{LSTM} \sum_{k=1}^{K} \left| \boldsymbol{\beta}_{j,h,t,k}^{LSTM} \right|}_{L_1 \text{ Penalty}} + \underbrace{\lambda_{2,t}^{LSTM} \sum_{k=1}^{K} (\boldsymbol{\beta}_{j,h,t,k}^{LSTM})^2}_{L_2 \text{ Penalty}} \right\}$$

where $\widehat{\boldsymbol{\beta}}_{j,h,\tau}^e(\cdot)$, $e \in \{\text{EN}, \text{LSTM}\}$, is the time τ estimate of $\boldsymbol{\beta}_{j,h}^e$ given $\boldsymbol{\lambda}_t^e$ and data through time τ in a sample of size T_E. Denote the combined final estimator $\widehat{\boldsymbol{\beta}}_{j,h,t}^e(\boldsymbol{X}_{\widehat{T}_E}, \widehat{\boldsymbol{\lambda}}_t^e)$, where the regularization parameter $\widehat{\boldsymbol{\lambda}}_t^e$ is estimated using cross-validation dynamically over time. Note that the algorithm also asks the machine to dynamically choose both the optimal estimation window $\widehat{T}_E$ and the optimal validation window $\widehat{T}_V$ by minimizing the pseudo-out-of-sample MSE.

The estimation of $\widehat{\boldsymbol{\beta}}_{j,h,t}^e(\boldsymbol{X}_{\widehat{T}_E}, \widehat{\boldsymbol{\lambda}}_t^e)$ is repeated sequentially in rolling subsamples, with parameters estimated from information known at time t. Note that the time t subscripts of $\widehat{\boldsymbol{\beta}}_{j,h,t}^e$ and $\widehat{\boldsymbol{\lambda}}_t^e$ denote one in a sequence of time-invariant parameter estimates obtained from rolling subsamples, rather than estimates that vary over time within a sample. Likewise, we denote the time t machine belief about $y_{j,t+h}$ as $\mathbb{E}_t^e[y_{j,t+h}]$, defined by

$$\mathbb{E}_t^e[y_{j,t+h}] \equiv G^e \left(\mathcal{X}_t, \widehat{\boldsymbol{\beta}}_{j,h,t}^e(\boldsymbol{X}_{\widehat{T}_E}, \widehat{\boldsymbol{\lambda}}_t^e) \right)$$

Finally, the machine MSE is computed by averaging across the sequence of squared forecast errors in the true out-of-sample forecasts for periods $t = (\dot{T} + h), \ldots, T$ where T is the last period of our sample. We refer to this subperiod as the external forecast *evaluation sample*.

On rare occasions, one or more of the explanatory variables used in the machine forecast specification assumes a value that is order of magnitudes different from its historical value. This is usually indicative of a measurement problem in the raw data. We therefore program the machine to detect in real-time whether its forecast is an extreme outlier, and in that case to discard the forecast replacing it with the historical mean. Specifically, at each t, the

machine forecast $\mathbb{E}_t^e[y_{j,t+h}]$ is set to be the historical mean calculated up to time t whenever the former is five or more standard deviations above its own rolling mean over the most recent 20 quarters.

We include the contemporaneous survey forecasts $\mathbb{F}_t[y_{j,t+h}]$ for the median respondent only for inflation and GDP forecasts, following BLM1. This procedure allows the machine to capture intangible information due to judgement or private signals. Specifically, for these forecasts of inflation and GDP growth, we consider the following machine learning empirical specification for forecasting $y_{j,t+h}$ given information at time t, to be benchmarked against the time t survey forecast of respondent-type X, where this type is the median here:

$$y_{j,t+h} = G_{jh}^e(\mathbf{Z}_t) + \gamma_{jh\mathbb{M}}\mathbb{F}_t[y_{j,t+h}] + \epsilon_{jt+h}, \qquad h \geq 1 \tag{A.5}$$

where $\gamma_{jh\mathbb{M}}$ is a parameter to be estimated, and where $G_{jh\mathbb{M}}(\mathbf{Z}_t)$ represents a ML estimator as function of big data. Note that the intercept α_{jh} from BLM gets absorbed into the $G_{jh}^e(\mathbf{Z}_t)$ in LSTM via the outermost bias term. 2.

Elastic Net Estimator

We use the Elastic Net (EN) estimator, which combines Least Absolute Shrinkage and Selection Operator (LASSO) and ridge type penalties. The model can be written as:

$$y_{j,t+h} = \mathcal{X}_{tj}'\boldsymbol{\beta}_{j,h}^{\text{EN}} + \epsilon_{j,t+h}$$

where $\mathcal{X}_t = (1, \mathcal{X}_{1t,...,}\mathcal{X}_{Kt})'$ include the independent variable observations $(\mathbb{F}_t[y_{j,t+h}], \mathcal{Z}_{jt})$ into a vector with "1" and $\boldsymbol{\beta}_{j,h}^{\text{EN}} = \left(\alpha_{j,h}, \beta_{j,h\mathbb{F}}, vec(\mathbf{B}_{j,h\mathcal{Z}})\right)' \equiv (\beta_0, \beta_1, ...\beta_K)'$ collects all the coefficients.

It is customary to standardize the elements of $\mathcal{X}_t$ such that sample means are zero and sample standard deviations are unity. The coefficient estimates are then put back in their original scale by multiplying the slope coefficients by their respective standard deviations, and adding back the mean (scaled by slope coefficient over standard deviation.) The EN estimator incorporates both an L_1 and L_2 penalty:

$$\widehat{\boldsymbol{\beta}}_{j,h}^{\text{EN}} = \underset{\beta_0,\beta_1,...,\beta_K}{argmin} \left\{ \frac{1}{T_E}\sum_{\tau=1}^{T_E}\left(y_{j,\tau+h} - \mathcal{X}_\tau'\boldsymbol{\beta}_{j,h}\right)^2 + \underbrace{\lambda_1\sum_{k=1}^{K}\left|\boldsymbol{\beta}_{j,h,k}\right|}_{\text{LASSO}} + \underbrace{\lambda_2\sum_{k=1}^{K}(\boldsymbol{\beta}_{j,h,k})^2}_{\text{ridge}} \right\}$$

By minimizing the MSE over the training samples, we choose the optimal λ_1 and λ_2 values simultaneously.

In the implementation, the EN estimator is sometimes used as an imput into the algorithm using the LSTM estimator. Specifically, we ensure that the machine forecast can only differ from the relevant benchmark if it demonstrably improves the pseudo out-of-sample prediction in the training samples *prior* to making a true out-of-sample forecast. Otherwise, the machine is replaced by the benchmark calculated up to time t. In some cases the benchmark is a survey forecast, in others it could be a historical mean value for the variable. However, for the implementation using LSTM, we also use the EN forecast as a benchmark.

Long Short-Term Memory (LSTM) Network

An LSTM network is a type of Recurrent Neural Network (RNN), which are neural networks used to learn about sequential data such as time series or natural language. In particular, LSTM networks can learn long-term dependencies between across time periods by introducing hidden layers and memory cells to control the flow of information over longer time periods. The general case of the LSTM network with up to N hidden layers is defined as

$$\underbrace{G^{\text{LSTM}}(\mathcal{X}_t, \boldsymbol{\beta}^{\text{LSTM}}_{j,h})}_{1\times 1} = \underbrace{W^{(yh^N)}}_{1\times D_{h^N}} \underbrace{h_t^N}_{D_{h^N}\times 1} + \underbrace{b_y}_{1\times 1} \qquad \text{(Output layer)}$$

$$\underbrace{h_t^n}_{D_{h^n}\times 1} = \underbrace{o_t^n}_{D_{h^n}\times 1} \odot \tanh(\underbrace{c_t^n}_{D_{h^n}\times 1}) \qquad \text{(Hidden layer)}$$

$$\underbrace{c_t^n}_{D_{h^n}\times 1} = \underbrace{f_t^n}_{D_{h^n}\times 1} \odot \underbrace{c_{t-1}^n}_{D_{h^n}\times 1} + \underbrace{i_t^n}_{D_{h^n}\times 1} \odot \underbrace{\widetilde{c}_t^n}_{D_{h^n}\times 1} \qquad \text{(Final memory)}$$

$$\underbrace{\widetilde{c}_t^n}_{D_{h^n}\times 1} = \tanh(\underbrace{W^{(c^n h^{n-1})}}_{D_{h^n}\times D_{h^{n-1}}} \underbrace{h_t^{n-1}}_{D_{h^{n-1}}\times 1} + \underbrace{W^{(c^n h^n)}}_{D_{h^n}\times D_{h^n}} \underbrace{h_{t-1}^n}_{D_{h^n}\times 1} + \underbrace{b_{c^n}}_{D_{h^n}\times 1}) \qquad \text{(New memory)}$$

$$\underbrace{f_t^n}_{D_{h^n}\times 1} = \sigma(\underbrace{W^{(f^n h^{n-1})}}_{D_{h^n}\times D_{h^{n-1}}} \underbrace{h_t^{n-1}}_{D_{h^{n-1}}\times 1} + \underbrace{W^{(f^n h^n)}}_{D_{h^n}\times D_{h^n}} \underbrace{h_{t-1}^n}_{D_{h^n}\times 1} + \underbrace{b_{f^n}}_{D_{h^n}\times 1}) \qquad \text{(Forget gate)}$$

$$\underbrace{i_t^n}_{D_{h^n}\times 1} = \sigma(\underbrace{W^{(i^n h^{n-1})}}_{D_{h^n}\times D_{h^{n-1}}} \underbrace{h_t^{n-1}}_{D_{h^{n-1}}\times 1} + \underbrace{W^{(i^n h^n)}}_{D_{h^n}\times D_{h^n}} \underbrace{h_{t-1}^n}_{D_{h^n}\times 1} + \underbrace{b_{i^n}}_{D_{h^n}\times 1}) \qquad \text{(Input gate)}$$

$$\underbrace{o_t^n}_{D_{h^n}\times 1} = \sigma(\underbrace{W^{(o^n h^{n-1})}}_{D_{h^n}\times D_{h^{n-1}}} \underbrace{h_t^{n-1}}_{D_{h^{n-1}}\times 1} + \underbrace{W^{(o^n h^n)}}_{D_{h^n}\times D_{h^n}} \underbrace{h_{t-1}^n}_{D_{h^n}\times 1} + \underbrace{b_{o^n}}_{D_{h^n}\times 1}) \qquad \text{(Output gate)}$$

where $n = 1, \ldots, N$ indexes each hidden layer. $h_t^n \in \mathbb{R}^{D_{h^n}}$ is the n-th *hidden layer*, where D_{h^n} is the number of *neurons* or *nodes* in the hidden layer. The 0-th layer is defined as the input data: $h_t^0 \equiv \mathcal{X}_t$. The memory cell c_t^n allows the LSTM network to retain information over longer time periods. The output gate o_t^n controls the extent to which the memory cell c_t^n maps to the hidden layer h_t^n. The forget gate f_t^n controls the flow of information carried over from the final memory in the previous timestep c_{t-1}^n. The input gate i_t^n controls the flow of information from the new memory cell $\widetilde{c}_t^n$. The initial states for the hidden layers $(h_0^n)_{n=1}^N$ and memory cells $(c_0^n)_{n=1}^N$ are set to zeros.

$\sigma(\cdot)$ and $\tanh(\cdot)$ are *activation functions* that introduce non-linearities in the LSTM network, applied elementwise. $\sigma : \mathbb{R} \to \mathbb{R}$ is the sigmoid function: $\sigma(x) = (1 + e^{-x})^{-1}$. $\tanh : \mathbb{R} \to \mathbb{R}$ is the hyperbolic tangent function: $\tanh(x) = \frac{e^{2x}-1}{e^{2x}+1}$. The $\odot$ operator refers to elementwise multiplication.

$\boldsymbol{\beta}_{j,h}^{\text{LSTM}} \equiv (((vec(W^{(g^n h^{n-1})})', vec(W^{(g^n h^n)})', b'_{g^n})_{g \in \{c,f,i,o\}}, h_0^{n\prime}, c_0^{n\prime})_{n=1}^N, vec(W^{(y h^N)})', b_y)'$ are parameters to be estimated. We will refer to parameters indexed with W as *weights*; parameters indexed with b are *biases*. We estimate the parameters $\boldsymbol{\beta}_{j,h}^{\text{LSTM}}$ for the LSTM network using Stochastic Gradient Decent (SGD), which is an iterative algorithm for minimizing the loss function and proceeds as follows:

1. *Initialization.* Fix a random seed R and draw a starting value of the parameters $\boldsymbol{\beta}_{j,h}^{(0)}$ randomly, where the superscript (0) in parentheses indexes the iteration for an estimate of $\boldsymbol{\beta}_{j,h}^{\text{LSTM}}$.

   (a) Initialize the input weights $W^{(g^n h^{n-1})} \in \mathbb{R}^{D_{h^n} \times D_{h^{n-1}}}$ for $g \in \{c, f, i, o\}$ using the *Glorot* initializer by drawing randomly from a uniform distribution with zero mean and a variance that depends on the dimensions of the matrix:

   $$W_{ij}^{(g^n h^{n-1})} \overset{iid}{\sim} U\left[-\sqrt{\frac{6}{D_{h^n} + D_{h^{n-1}}}}, \sqrt{\frac{6}{D_{h^n} + D_{h^{n-1}}}}\right]$$

   for each $i = 1, \ldots, D_{h^n}$ and $j = 1, \ldots, D_{h^{n-1}}$.

   (b) Initialize the recurrent weights $W^{(g^n h^n)} \in \mathbb{R}^{D_{h^n} \times D_{h^n}}$ for $g \in \{c, f, i, o\}$ using the *Orthogonal* initializer by using the orthogonal matrix obtained from the QR decomposition of a $D_{h^n} \times D_{h^n}$ matrix of random numbers drawn from a standard normal distribution.

   (c) Biases $(b_{g^n})_{g \in \{c,f,i,o\}}$, hidden layers h_0^n, and memory cells c_0^n are initialized with zeros.

2. Repeat until the stopping condition is satisfied ($k = 1, 2, 3, \ldots$):

(a) *Mini-Batch.* Randomly draw a subset $I^{(k)}$ of size M from the estimation sample:

$$I^{(k)} = [I_1^{(k)}, \ldots, I_M^{(k)}]' \text{ where } I_j^{(k)} \overset{iid}{\sim} Uniform(\{1, \ldots, T_E\})$$

The subset $I^{(k)}$ is called a "mini-batch" and contains the time indices of the randonly sampled observations. The size of the mini-batch M is a hyper-parameter. Smaller batch sizes M can have a regularizing effect by adding noise to the learning process.

(b) *Dropout.* Apply dropout to the mini-batch by multiplying (elementwise) each layer h_t^n with a binary mask r_t^n such that

$$\widehat{h}_t^n = r_t^n \odot h_t^n \text{ where } r_{i,t}^n \overset{iid}{\sim} Bernoulli(1-p)$$

for each $t \in I^{(k)}$ and $i = 1, \ldots, D_{h^n}$. The probability of dropout p is a hyper-parameter to be cross-validated. Dropout adds noise to the learning process by randomly selecting which weight to update during each iteration. Intuitively, dropout is similar to averaging the predictions of many smaller sub-networks.

(c) *Stochastic Gradient.* Average the gradient over observations in the mini-batch

$$\nabla L(\boldsymbol{\beta}_{j,h}^{(k-1)}, \mathbf{X}_{I^{(k)}}, \boldsymbol{\lambda}^{\text{LSTM}}) = \frac{1}{M} \sum_{t \in I^{(k)}} \nabla L(\boldsymbol{\beta}_{j,h}^{(k-1)}, \mathbf{X}_t, \boldsymbol{\lambda}^{\text{LSTM}})$$

where $\nabla L(\boldsymbol{\beta}_{j,h}^{(k-1)}, \mathbf{X}_t, \boldsymbol{\lambda}^{\text{LSTM}})$ is the gradient of the loss function with respect to the parameters $\boldsymbol{\beta}_{j,h}^{(k-1)}$, evaluated at the time t observation $\mathbf{X}_t = (y_{j,t+h}, \widehat{\mathcal{X}}_t')'$ after applying dropout.

(d) *Learning rate shrinkage.* Update the parameters to $\boldsymbol{\beta}_{j,h}^{(k)}$ using the Adaptive Moment Estimation (Adam) algorithm. The method uses the first and second moments of the gradients to shrink the overall learning rate to zero as the gradient approaches zero.

$$\boldsymbol{\beta}_{j,h}^{(k)} = \boldsymbol{\beta}_{j,h}^{(k-1)} - \gamma \frac{m^{(k)}}{\sqrt{v^{(k)}} + \varepsilon}$$

where $m^{(k)}$ and $v^{(k)}$ are weighted averages of first two moments of past gradients:

$$m^{(k)} = \frac{1}{1 - \pi_1^k}(\pi_1 m^{(k-1)} + (1 - \pi_1) \nabla L(\boldsymbol{\beta}_{j,h}^{(k-1)}, \mathbf{X}_{I^{(k)}}, \boldsymbol{\lambda}^{\text{LSTM}}))$$

$$v^{(k)} = \frac{1}{1-\pi_2^k}(\pi_2 v^{(k-1)} + (1-\pi_2)\nabla L(\boldsymbol{\beta}_{j,h}^{(k-1)}, \mathbf{X}_{I^{(k)}}, \boldsymbol{\lambda}^{\text{LSTM}})^2)$$

π^k denotes the k-the power of $\pi \in (0,1)$, and $/$, $\sqrt{\cdot}$, and $(\cdot)^2$ are applied elementwise. The default values of the hyper-parameters are $m^{(0)} = v^{(0)} = 0$ (initial moment vectors), $\gamma = 0.001$ (initial learning rate), $(\pi_1, \pi_2) = (0.9, 0.999)$ (decay rates), and $\varepsilon = 10^{-8}$ (prevent zero denominators).

(e) *Stopping Critera.* Stop iterating and return $\boldsymbol{\beta}_{j,h}^{(k)}$ if one of the following holds:

- *Early stopping.* At each iteration, use the updated $\boldsymbol{\beta}_{j,h}^{(k)}$ to calculate the loss from the validation sample. Stop when the validation loss has not improved for S steps, where S is a "patience" hyper-parameter. By updating the parameters for fewer iterations, early stopping shrinks the final parameters $\boldsymbol{\beta}_{j,h}$ towards the initial guess $\boldsymbol{\beta}_{j,h}^{(0)}$, and at a lower computational cost than ℓ_2 regularization.
- *Maximum number of epochs.* Stop if the number of iterations reaches the maximum number of epochs E. An epoch happens when the full set of the estimation sample has been used to update the parameters. If the estimation sample has T_E observations and each mini-batch has M observations, then each epoch would contain $\lceil T_E/M \rceil$ iterations (after rounding up as needed). So the maximum number of iterations is bounded by $E \times \lceil T_E/M \rceil$.

3. *Ensemble forecasts.* Repeat steps 1. and 2. over different random seeds R and save each of the estimated parameters $\widehat{\boldsymbol{\beta}}_{j,h,T_E}^{LSTM}(\boldsymbol{X}_{T_E}, \boldsymbol{\lambda}^{\text{LSTM}}, R)$. Then construct out-of-sample forecasts using the top 10 out of 20 starting values with the best performance in the validation sample. Ensemble can be considered as a regularization method because it aims to guard against overfitting by shrinking the forecasts toward the average across different random seeds. The random seed affects the random draws of the parameter's initial starting value $\boldsymbol{\beta}_{j,h}^{(0)}$, the mini-batches $I^{(k)}$, and the dropout mask $r_t^{(k)}$.

Hyperparameters Let $\boldsymbol{\lambda}^{\text{LSTM}} \equiv [\lambda_1, \lambda_2, \gamma, \pi_1, \pi_2, p, N, (D_{h^n})_{n=1}^N, M, E, S]'$ collect all the hyper-parameters that control the LSTM network's complexity and prevent the model from overfitting the data. The number of hidden layers N and the number of neurons $D_{h^1}, \ldots, D_{h^N}$ in each hidden layer are hyper-parameters that characterize the network's *architecture*. To choose the number of neurons in each layer, we apply a geometric pyramid rule where the dimension of each additional hidden layer is half that of the previous hidden layer.

Table A.4 reports the hyper-parameters for the LSTM network and its estimation. Hyper-parameters reported as a range or a set of values are cross-validated. The hyper-parameters

Table A.4: Candidate hyper-parameters for the LSTM network

Hyper-parameter	Description
[0.5ex] $\lambda_1 \in [10^{-6}, 10^{-2}]$	L_1 penalty
$\lambda_2 \in [10^{-6}, 10^{-2}]$	L_2 penalty
$\gamma \in \{0.01, 0.001\}$	Initial learning rate
$(\pi_1, \pi_2) = (0.9, 0.999)$	Gradient decay rates
$p = 0.2$	Probability of dropout
$N = 1$	Number of hidden layers
$(D_{h^n})_{n=1}^N = (4, 8, 32)$	Number of neurons in each hidden layer
$M = 4$	Mini-batch size
$E = 10,000$	Maximum number of epochs
$S = 20$	Patience for early stopping
$R \in \{1, 2, 3, \ldots, 20\}$	Random seeds
$T_E \in \{6, 9, 12, 15\}$	Window length of estimation sample
$T_V \in \{3, 5, 7\}$	Window length of validation sample

are estimated by minimizing the mean-square loss over pseudo out-of-sample forecast errors generated from rolling regressions through the validation sample. The pseudo out-of-sample forecasts are ensemble averages implied by parameters based on different random seeds R.

Machine Variables to Be forecast

Returns and price growth When evaluating the MSE ratio of the machine relative to that of a benchmark survey, we use the machine forecast for the return or price growth measure that most closely corresponds to the concept that survey respondents are asked to predict:

1. CFO survey asks respondents about their expectations for the S&P 500 return over the next 12 months. Following Nagel and Xu (2022), we interpret the survey to be asking about $r^d_{t,t+12}$, the one-year CRSP value-weighted return (including dividends) from the current survey month to the same month one year ahead.

2. Gallup/UBS survey respondents report the return (including dividends) they expect on their own portfolio one year ahead. We interpret the survey to be asking about $r^d_{t,t+12}$, the one-year CRSP value-weighted return(including dividends) from the current survey month to the same month one year ahead.

3. Livingston survey respondents provide 12-month ahead forecasts of the S&P 500 index. We convert the level forecast to price growth forecast by taking the log difference

between the 12-month ahead level forecast and the nowcast of the S&P 500 index for the current survey month. Therefore, we interpret the survey to be asking about the one-year price growth in the S&P 500 index.

4. Bloomberg Consensus Forecasts asks survey respondents about the end-of-year closing value of the S&P 500 index. We interpret the survey to be asking about the h-month price growth in the S&P 500 index. The horizon of the forecast changes depending on when in the year the panelists are answering the survey.

5. Michigan Survey of Consumers (SOC) asks respondents about their perceived probability that an investment in a diversified stock fund would increase in value in the year ahead. We interpret the question to be asking about the one-year price growth in the S&P 500 index.

6. Conference Board (CB) survey asks respondents about their categorical belief on whether they expect stock prices to increase, decrease, or stay the same over the next year. We interpret the question to be asking about the one-year price growth in the S&P 500 index.

Earnings growth For earning growth forecasts, we use a quarterly S&P 500 total earnings series that combines data from S&P Global, Shiller, and the S&P 500 divisor, as described above. Since the machine learning algorithm has been adapted to a quarterly forecasting frequency, we use the quarterly series before the monthly interpolation. The quarterly series spans the period 1959:Q1 to 2021:Q4.

Inflation We construct forecasts of annual inflation defined as

$$\pi_{t+4,t} = \ln\left(\frac{PGDP_{t+4}}{PGDP_t}\right)$$

where $PGDP_t$ is the quarterly level of the chain-weighted GDP price index. Following Coibion and Gorodnichenko (2015), we use the vintage of inflation data that is available four quarters after the period being forecast.

GDP growth We construct forecasts of annual real GDP growth defined as

$$y_{t+4,t} = \ln\left(\frac{RGDP_{t+4}}{RGDP_t}\right)$$

where $RGDP_t$ is the quarterly level of chain-weighted real GDP. Following Coibion and Gorodnichenko (2015), we use the vintage of inflation data that is available four quarters after the period being forecast.

Machine Input Data: Predictor Variables

The vector $\mathbf{Z}_{jt} \equiv \left(y_{j,t}, \hat{\mathbf{G}}'_t, \mathbf{W}'_{jt}\right)'$ is an $r = 1 + r_G + r_W$ vector which collects the data at time t with $\mathcal{Z}_{jt} \equiv \left(y_{j,t}, ..., y_{j,t-p_y}, \hat{\mathbf{G}}'_t, ..., \hat{\mathbf{G}}'_{t-p_G}, \mathbf{W}'_{jt}, ..., \mathbf{W}'_{jt-p_W}\right)'$ a vector of contemporaneous and lagged values of $\mathbf{Z}_{jt}$, where p_y, p_G, p_W denote the total number of lags of $y_{j,t}$, $\hat{\mathbf{G}}'_t$, $\mathbf{W}'_{jt}$, respectively. The predictors below are listed as elements of $y_{j,t}, \hat{\mathbf{G}}'_{jt}$, or $\mathbf{W}'_{jt}$ for variables.

Stock return and price growth predictor variables For y_j equal to CRSP value-weighted returns or S&P 500 price index, the forecasting model considers the following variables.

In $\mathbf{W}'_{jt}$:

1. $\mathbf{G}_{M,t-k}$, for $k = 0, 1$ are factors formed from a real-time macro dataset $\mathcal{D}^M$ with 92 real-time macro series; includes both monthly and quarterly series, with monthly series converted to quarterly according to the method described in the data appendix.

2. $\mathbf{G}_{F,t-k}$, for $k = 0, 1$ are factors formed from a financial data set $\mathcal{D}^F$ with 147 monthly financial series.

3. $\mathbf{G}^Q_{D,t-k}$, for $k = 0$ are quarterly factors formed from a daily financial dataset $\mathcal{D}^D$ of 87 daily financial indicators. The raw daily series are first converted to daily factors $\mathbf{G}_{D,t}(\mathbf{w})$ and the daily factors are aggregated up to quarterly observations $\mathbf{G}^Q_{D,t}(\mathbf{w})$ using a weighted average of daily factors, with the weights $\mathbf{w}$ dependent on two free parameters that are chosen to minimize the sum of squared residuals in a regression of $y_{j,t+h}$ on $\mathbf{G}_{D,t}(\mathbf{w})$.

4. *LDA topics* $F_{k,t-j}$, for topic $k = 1, 2, ...50$ and $j = 0, 1$. The value of the topic k at time t is the average weights of topic k of all articles published at t.

5. *Macro data surprises* from the money market survey. The macro news include, GDP growth (Q/Q percentage change), core CPI (Month/Month change), unemployment rate (percentage point), and nonfarm payroll (month/month change). We include first release, second release, and final release for GDP growth. This constitutes six macro data surprises per quarter.

6. *FOMC surprises* are defined as the changes in the current-month, 1, 2, 6, 12, and 24 month-ahead federal funds futures (FFF) contract rate and the changes in the 1, 2, 4, and 8 quarter-ahead Eurodollar (ED) futures contracts, from 10 minutes before to 20 minutes after each FOMC announcement. When benchmarking against a survey, we use the last FOMC meeting before the survey deadline to compute surprises. For surveys that do not have a clear deadline, we compute surprises using from the last FOMC in the first month of the quarter. When benchmarking against moving average, we use the last FOMC meeting before the end of the first month in each quarter to compute surprises. This leaves 10 FOMC surprise variables per quarter.

7. *Stock market jumps* are accumulated 30-minute window negative and positive jumps in the S&P 500 around news events over the previous quarter.

8. $\bar{\mu}_{t-k}$ for $k = 0, 1, 2$ is the historical mean of returns calculated up to time t. The initial period is 1959Q1.

9. *Long-term growth of earnings*: 5-year growth of the SP500 earnings per share.

10. *Short rates.* When forecasting returns or price growth, the machine controls for the current nominal short rate, $\ln(1 + 3MTB_t/100)$, imposing a unit coefficient. This is equivalent to forecasting the future return minus the current short rate.

The 92 macro series in $\mathcal{D}^M$ are selected to represent broad categories of macroeconomic time series. The majority of these are real activity measures: real output and income, employment and hours, consumer spending, housing starts, orders and unfilled orders, compensation and labor costs, and capacity utilization measures. The dataset also includes commodity and price indexes and a handful of bond and stock market indexes, and foreign exchange measures. The financial dataset $\mathcal{D}^f$ is an updated monthly version of the of 147 variables comprised solely of financial market time series used in Ludvigson and Ng (2007). These data include valuation ratios such as the dividend-price ratio and earnings-price ratio, growth rates of aggregate dividends and prices, default and term spreads, yields on corporate bonds of different ratings grades, yields on Treasuries and yield spreads, and a broad cross-section of industry, size, book-market, and momentum portfolio equity returns.[9] The 87 daily financial indicators in $\mathcal{D}^D$ include daily time series on commodities spot prices and futures prices, aggregate stock market indexes, volatility indexes, credit spreads and yield spreads, and exchange rates.

[9]A detailed description of the series is given in the Data Appendix of the online supplementary file at www.sydneyludvigson.com/s/ucc_data_appendix.pdf

Earning growth predictor variables and specifications For earning growth forecasts, we first detrend the (log) earnings level in real time by, starting with an initial sample, recursively running the following regression at each point in time t

$$\log\left(earnings_t\right) = \alpha_t + \beta_t t + y_t$$

For y_t equal to the detrended (log) earning level, we construct a forecasted value for y_t, denoted $\hat{y}_{t|t-h}$, based on information known up to time t using the following variables:

1. $\mathbf{G}_{M,t-k}$, for $k = 0, 1$ are factors formed from a real-time macro dataset $\mathcal{D}^M$ with 92 real-time macro series; includes both monthly and quarterly series, with monthly series converted to quarterly according to the method described in the data appendix.

2. $\mathbf{G}_{F,t-k}$, for $k = 0, 1$ are factors formed from a financial data set $\mathcal{D}^F$ with 147 monthly financial series.

3. $\mathbf{G}^Q_{D,t-k}$, for $k = 0$ are quarterly factors formed from a daily financial dataset $\mathcal{D}^D$ of 87 daily financial indicators. The raw daily series are first converted to daily factors $\mathbf{G}_{D,t}(\mathbf{w})$ and the daily factors are aggregated up to quarterly observations $\mathbf{G}^Q_{D,t}(\mathbf{w})$ using a weighted average of daily factors, with the weights $\mathbf{w}$ dependent on two free parameters that are chosen to minimize the sum of squared residuals in a regression of $y_{j,t}$ on $\mathbf{G}_{D,t}(\mathbf{w})$.

4. *LDA factors* $F_{k,t-j}$, for topic $k = 1, 2, ...50$ and $j = 0, 1$. The value of the topic k at time t is the average weights of topic k of all articles published at t.

5. *Macro data surprises* from the money market survey. The macro news include, GDP growth (Q/Q percentage change), core CPI (Month/Month change), unemployment rate (percentage point), and nonfarm payroll (month/month change). We include first release, second release, and final release for GDP growth. This constitutes six macro data surprises per quarter.

6. *FOMC surprises* are defined as the changes in the current-month, 1, 2, 6, 12, and 24 month-ahead federal funds futures (FFF) contract rate and the changes in the 1, 2, 4, and 8 quarter-ahead Eurodollar (ED) futures contracts, from 10 minutes before to 20 minutes after each FOMC announcement. When benchmarking against a survey, we use the last FOMC meeting before the survey deadline to compute surprises. For surveys that do not have a clear deadline, we compute surprises using from the last FOMC in the first month of the quarter. When benchmarking against moving average,

we use the last FOMC meeting before the end of the first month in each quarter to compute surprises. This leaves 10 FOMC surprise variables per quarter.

7. *Stock market jumps* are accumulated 30-minute window negative and positive jumps in the S&P 500 around news events over the previous quarter.

After we obtain the machine forecast for the detrended level of earnings, y, we obtain the h-horizon machine earnings growth forecast (from $t-h$ to t denoted $\mathbb{E}_{t-h}\left[\Delta \log\left(earnings_t^M\right)\right]$) by constructing

$$\mathbb{E}_{t-h}\left[\Delta \log\left(earnings_t^M\right)\right] \equiv \hat{\alpha}_{t-h} + \hat{\beta}_{t-h} t + \hat{y}^M_{t|t-h} - \log\left(earnings_{t-h}\right)$$

where $\log\left(earnings_{t-h}\right)$ is the realized log earning level at time $t-h$, and $\hat{y}^M_{t|t-h}$ is the machine forecast of the detrended log earnings based on information up to time $t-h$. To use this approach to forecast the 20-quarter ahead annual forward earnings i.e., (from $t-4$ to t on basis of information at $t-20$), we would construct

$$\mathbb{E}_{t-20}\left[\log\left(earnings_t^M\right)\right] = \hat{\alpha}_{t-20} + \hat{\beta}_{t-20} t + \hat{y}^M_{t|t-20}.$$

To construct 20-quarter ahead annual earnings growth forecast we compute

$$\mathbb{E}_{t-20}\left[\log\left(earnings_{t-4}^M\right)\right] = \hat{\alpha}_{t-20} + \hat{\beta}_{t-20}(t-4) + \hat{y}^M_{t-4|t-20}$$

to get the machine forecast of 20-quarter forward annual earnings log growth as

$$\mathbb{E}_{t-20}\left[\log\left(earnings_t^M\right) - \log\left(earnings_{t-4}^M\right)\right] = \hat{\beta}_{t-20} 4 + \hat{y}^M_{t|t-20} - \hat{y}^M_{t-4|t-20}.$$

An alternative is to use the machine inputs to directly forecast 20-quarter forward annual earnings log growth $\mathbb{E}_{t-20}\left[\log\left(earnings_t^M\right) - \log\left(earnings_{t-4}^M\right)\right]$.

Inflation predictor variables For y_j equal to inflation, the forecasting model considers the following variables.

In $\mathbf{W}'_{jt}$:

1. $\mathbb{F}^{(i)}_{jt-k}[y_{jt+h-k}]$, lagged values of the ith type's forecast, where $k = 1, 2$
2. $\mathbb{F}^{(s\neq i)}_{jt-1}[y_{jt+h-1}]$, lagged values of other type's forecasts, $s \neq i$

3. $var_N\left(\mathbb{F}_{t-1}^{(\cdot)}[y_{jt+h-1}]\right)$, where $var_N(\cdot)$ denotes the cross-sectional variance of lagged survey forecasts

4. $skew_N\left(\mathbb{F}_{t-1}^{(\cdot)}[y_{jt+h-1}]\right)$, where $skew_N(\cdot)$ denotes the cross-sectional skewness of lagged survey forecasts

5. Trend inflation measured as $\overline{\pi}_{t-1} = \begin{cases} \rho\overline{\pi}_{t-2} + (1-\rho)\pi_{t-1}, \rho = 0.95 & \text{if } t < \text{1991:Q4} \\ \text{CPI10}_{t-1} & \text{if } t \geq \text{1991:Q4} \end{cases}$, where CPI10 is the median SPF forecast of annualized average inflation over the current and next nine years. Trend inflation is intended to capture long-run trends. When long-run forecasts of inflation are not available, as is the case pre-1991:Q4, we us a moving average of past inflation.

6. $\dot{GDP}_{t-1}$ = detrended gross domestic product, defined as the residual from a regression of GDP_{t-1} on a constant and the four most recent values of GDP as of date $t-8$. See Hamilton (2018).

7. $\dot{EMP}_{t-1}$ = detrended employment, defined as the residual from a regression of EMP_{t-1} on a constant and the four most recent values of EMP as of date $t-8$. See Hamilton (2018).

8. $\mathbb{N}_t^{(i)}[\pi_{t,t-h}]$ = Nowcast as of time t of the ith percentile of inflation over the period $t-h$ to t.

Lags of the dependent variable:

1. $y_{t-1,t-h-1}$ one quarter lagged inflation.

The factors in $\hat{\mathbf{G}}'_{jt}$ include factors formed from three large datasets separately:

1. $\mathbf{G}_{M,t-k}$, for $k = 0, 1$ are factors formed from a real-time macro dataset $\mathcal{D}^M$ with 92 real-time macro series; includes both monthly and quarterly series, with monthly series converted to quarterly according to the method described in the data appendix.

2. $\mathbf{G}_{F,t-k}$, for $k = 0, 1$ are factors formed from a financial data set $\mathcal{D}^F$ with 147 monthly financial series.

3. $\mathbf{G}_{D,t-k}^Q$, for $k = 0$ are quarterly factors formed from a daily financial dataset $\mathcal{D}^D$ of 87 daily financial indicators. The raw daily series are first converted to daily factors $\mathbf{G}_{D,t}(\mathbf{w})$ and the daily factors are aggregated up to quarterly observations $\mathbf{G}_{D,t}^Q(\mathbf{w})$ using a weighted average of daily factors, with the weights $\mathbf{w}$ dependent on two free

parameters that are chosen to minimize the sum of squared residuals in a regression of $y_{j,t+h}$ on $\mathbf{G}_{D,t}(\mathbf{w})$.

The 92 macro series in $\mathcal{D}^M$ are selected to represent broad categories of macroeconomic time series. The majority of these are real activity measures: real output and income, employment and hours, consumer spending, housing starts, orders and unfilled orders, compensation and labor costs, and capacity utilization measures. The dataset also includes commodity and price indexes and a handful of bond and stock market indexes, and foreign exchange measures. The financial dataset $\mathcal{D}^f$ is an updated monthly version of the of 147 variables comprised solely of financial market time series used in Ludvigson and Ng (2007). These data include valuation ratios such as the dividend-price ratio and earnings-price ratio, growth rates of aggregate dividends and prices, default and term spreads, yields on corporate bonds of different ratings grades, yields on Treasuries and yield spreads, and a broad cross-section of industry, size, book-market, and momentum portfolio equity returns.[10] The 87 daily financial indicators in $\mathcal{D}^D$ include daily time series on commodities spot prices and futures prices, aggregate stock market indexes, volatility indexes, credit spreads and yield spreads, and exchange rates.

GDP growth predictor variables For y_j equal to GDP growth, the forecasting model considers the following variables.

In $\mathbf{W}'_{jt}$:

1. $\mathbb{F}^{(i)}_{jt-k}[y_{jt+h-k}]$, lagged values of the ith type's forecast, where $k = 1, 2$
2. $\mathbb{F}^{(s\neq i)}_{jt-1}[y_{jt+h-1}]$, lagged values of other type's forecasts, $s \neq i$
3. $var_N\left(\mathbb{F}^{(\cdot)}_{t-1}[y_{jt+h-1}]\right)$, where $var_N(\cdot)$ denotes the cross-sectional variance of lagged survey forecasts
4. $skew_N\left(\mathbb{F}^{(\cdot)}_{t-1}[y_{jt+h-1}]\right)$, where $skew_N(\cdot)$ denotes the cross-sectional skewness of lagged survey forecasts
5. $\mathbb{N}^{(i)}_t[\pi_{t,t-h}]$ = Nowcast as of time t of the ith percentile of inflation over the period $t-h$ to t.

Lags of the dependent variable:

1. $y_{t-1,t-h-1}$ one quarter lagged annual GDP growth.

[10] A detailed description of the series is given in the Data Appendix of the online supplementary file at www.sydneyludvigson.com/s/ucc_data_appendix.pdf

The factors in $\hat{\mathbf{G}}'_{jt}$ include factors formed from three large datasets separately:

1. $\mathbf{G}_{M,t-k}$, for $k = 0, 1$ are factors formed from a real-time macro dataset $\mathcal{D}^M$ with 92 real-time macro series; includes both monthly and quarterly series, with monthly series converted to quarterly according to the method described in the data appendix.

2. $\mathbf{G}_{F,t-k}$, for $k = 0, 1$ are factors formed from a financial data set $\mathcal{D}^F$ with 147 monthly financial series.

3. $\mathbf{G}^Q_{D,t-k}$, for $k = 0$ are quarterly factors formed from a daily financial dataset $\mathcal{D}^D$ of 87 daily financial indicators. The raw daily series are first converted to daily factors $\mathbf{G}_{D,t}(\mathbf{w})$ and the daily factors are aggregated up to quarterly observations $\mathbf{G}^Q_{D,t}(\mathbf{w})$ using a weighted average of daily factors, with the weights $\mathbf{w}$ dependent on two free parameters that are chosen to minimize the sum of squared residuals in a regression of $y_{j,t+h}$ on $\mathbf{G}_{D,t}(\mathbf{w})$.

The 92 macro series in $\mathcal{D}^M$ are selected to represent broad categories of macroeconomic time series. The majority of these are real activity measures: real output and income, employment and hours, consumer spending, housing starts, orders and unfilled orders, compensation and labor costs, and capacity utilization measures. The dataset also includes commodity and price indexes and a handful of bond and stock market indexes, and foreign exchange measures. The financial dataset $\mathcal{D}^f$ is an updated monthly version of the of 147 variables comprised solely of financial market time series used in Ludvigson and Ng (2007). These data include valuation ratios such as the dividend-price ratio and earnings-price ratio, growth rates of aggregate dividends and prices, default and term spreads, yields on corporate bonds of different ratings grades, yields on Treasuries and yield spreads, and a broad cross-section of industry, size, book-market, and momentum portfolio equity returns.[11] The 87 daily financial indicators in $\mathcal{D}^D$ include daily time series on commodities spot prices and futures prices, aggregate stock market indexes, volatility indexes, credit spreads and yield spreads, and exchange rates. Once converted into factors the total number of series used as inputs into the machine learning specifications is given below.

C Model Solution

We use the algorithm of Farmer, Waggoner and Zha (2011) to solve the system of structural model equations that must hold in equilibrium, where agents form expectations taking into

[11] A detailed description of the series is given in the Data Appendix of the online supplementary file at www.sydneyludvigson.com/s/ucc_data_appendix.pdf

Table A.5: Number of RHS Variables

	Stock Return	Earnings	GDP	Inflation
Macro Factors	10 (0-1 lag)	10 (0-1 lag)	10 (0-1 lag)	10 (0-1 lag)
Financial Factors	10 (0-1 lag)	10 (0-1 lag)	10 (0-1 lag)	10 (0-1 lag)
Daily Factors	10 (0-1 lag)	10 (0-1 lag)	10 (0-1 lag)	10 (0-1 lag)
LDA Factors	50	50	0	0
FOMC Surprises	10	10	0	0
Macro Data Surprises	6 (0-1 lag)	6 (0-1 lag)	0	0
Other predictors	0	3	10	11
Total	132	135	60	61

This table shows the number of predictors using for each forecast

account the probability of regime change ξ_t in the future. This solution is obtained in three steps.

1. Solve for the true law of motion of S_t^M in (16) such that (12)-(15) are satisfied and for the perceived law of motion of S_t^{M*} in (17) such that perceived versions of (12)-(15) are satisfied.

2. Solve for the law of motion for $S_t^A \equiv \left[m_t, pd_t, lp_t, \widetilde{\mathbb{E}}_t\left(m_{t+1}\right), \widetilde{\mathbb{E}}_t\left(pd_{t+1}\right)\right]$ such that (22)-(??) are satisfied. The resulting solution takes the form:

$$S_t^A = \widetilde{C}_{A,\xi_t} + \widetilde{T}_{A,M} S_{t-1}^{M*} + \widetilde{T}_{A,A} S_{t-1}^A + \widetilde{R}_{A,\eta}\eta_t + \widetilde{R}_{A,M}\widetilde{Q}_{M,\xi_t}\widetilde{\varepsilon}_t^M + \widetilde{R}_{A,A}\sigma_{lp,\xi_t}\varepsilon_{lp,t}, \quad \text{(A.6)}$$

   where $\widetilde{C}_{A,\xi_t}$,$\widetilde{T}_{A,M}$, etc., are matrices involving the perceived parameters $\widetilde{\theta}^M$ from (17). Since (22)-(??) involve conditional subjective second moment terms $\widetilde{\mathbb{V}}_t$ and $\widetilde{\mathbb{COV}}_t$ that are affected by ξ_t, we follow Bansal and Zhou (2002), Bianchi, Kung and Tirskikh (2018), and BLM2 in using a "Risk Adjustment with Lognormal Approximation," to preserve log-normality of the entire system. This implies that $\widetilde{C}_{A,\xi_t}$depends on ξ_t.

3. Let $S_t \equiv \left[S_t^M, S_t^{M*}, S_t^A, \widetilde{\varepsilon}_t^M, \eta_t\right]'$ and $\varepsilon_t^M = \left[\varepsilon_{\Delta y,t}, \varepsilon_{i,t}, \varepsilon_{\pi,t}, \varepsilon_{k,t}, \varepsilon_{\overline{\Delta y},t}, \varepsilon_{\overline{i},t}, \varepsilon_{\overline{\pi},t}, \varepsilon_{\overline{k},t}\right]'$. The third and final step is to combine the equations from steps 1 and 2 into a single system representing the complete structural model:

$$S_t = \bar{C}\left(\theta_{\xi_t}, \widetilde{\theta}_{\xi_t}\right) + \bar{T}\left(\theta_{\xi_t}, \widetilde{\theta}_{\xi_t}\right) S_{t-1} + \bar{R}\left(\theta_{\xi_t}, \widetilde{\theta}_{\xi_t}\right) Q_{\xi_t}\varepsilon_t, \quad \text{(A.7)}$$

   where $\bar{C}(\cdot)$, $\bar{T}(\cdot)$, $\bar{R}(\cdot)$ are matrices of primitive parameters involving elements of θ_{ξ_t}

and $\widetilde{\theta}_{\xi_t}$, some of which vary with the Markov-switching variable ξ_t, and $Q_{\xi_t}(\cdot)$ is a matrix of shock volatilities that vary stochastically with ξ_t. The structural shocks of the full model are contained in $\varepsilon_t = \left(\varepsilon_t^M, \varepsilon_{lp,t}, \varepsilon_{v,t}\right)'$, which stacks the primitive macro shocks ε_t^M, the liquidity premium shock $\varepsilon_{lp,t}$ (a feature of preferences), and the vintage errors $\varepsilon_{v,t}$. Neither $\widetilde{\varepsilon}_t^M$ or η_t appear separately in ε_t because $\widetilde{\varepsilon}_t^M = \left(\widetilde{R}^M\widetilde{Q}^M\right)^{-1}\left(S_t^{M*} - \widetilde{C}^M - \widetilde{T}^M S_{t-1}^{M*}\right)$ is entirely pinned down S_t^{M*} (and thus by ε_t^M and $\varepsilon_{v,t}$), while η_t has an innovation that is proportional to $\widetilde{\varepsilon}_t^M$.

Observation Equation

The mapping from the variables of the model to the observables in the data can be written using matrix algebra to obtain the observation equation $X_t = D_{\xi_t,t} + Z_{\xi_t,t}S_t' + U_t v_t$, where $S_t \equiv \left[S_t^M, S_t^{M*}, S_t^A, \eta_t, \widetilde{\varepsilon}_t^M\right]'$, and where

$$
\begin{aligned}
S_t^A &\equiv \left[m_t, pd_t, lp_t, \widetilde{\mathbb{E}}_t\left(m_{t+1}\right), \widetilde{\mathbb{E}}_t\left(pd_{t+1}\right)\right] \\
S_t^M &\equiv \left[\Delta y_t, \overline{\Delta y}_t, \Delta d_t \pi_t, \overline{\pi}_t, i_t, \bar{i}_t, k_t, \overline{k}_t\right]' \\
S_t^{M*} &\equiv \left[\Delta y_t^*, \overline{\Delta y}_t^*, \Delta d_t^*, \pi_t^*, \overline{\pi}_t^*, i_t^*, \bar{i}_t^*, k_t^*, \overline{k}_t^*\right]'.
\end{aligned}
$$

Annualizing the monthly growth rates to get annualized GDP growth we have $\Delta\ln\left(GDP_t\right) \equiv 12\Delta\ln\left(Y_t\right) = 12\Delta y_t$. For quarterly GDP growth we interpolate to monthly frequency. For our other quarterly variables we drop these from the observation vector in the months for which they aren't available. Machine forecasts and investor forecasts load on different subvectors of S_t. Let the subjector relevant for the machine forecasts be denoted $S_t^{MF} \equiv \left[S_t^{M*}, S_t^A, \eta_t, \widetilde{\varepsilon}_t^M\right]'$ and the subvector relevant for the investor forecasts be $S_t^I = \left[S_t^{M*}, S_t^A, \eta_t\right]'$. Let matrices with a subscript, e.g., Z_x, denote the subvector of Z that when multiplied by the appropriate subvector of S_t and added to $D_x + U_x v_{x,t}$ picks out the appropriate theoretical concept to map into empirical observations on an element x_t of X_t. For time t expected values in the model, we construct formulas for computing e.g., the expected value of a variable x_t over the next h periods under the assumption that $\xi_t = j$, such that $\widetilde{\mathbb{E}}_t\left(x_{t,t+h}\right) = D^I_{\xi_t,x_{t,t+h}} + Z^I_{\xi_t,x_{t,t+h}}S_t^I$, where $Z^I_{\xi_t,x_{t,t+h}}$are row vectors that load the subvector S_t^I, and $D^I_{\xi_t,x_{t,t+h}}$is a comformable intercept that applies to investor forecasts. These formulas are mapped into survey forecasts and machine forecasts for variables h periods ahead, respectively. We also construct formulas for computing the expected value of a variable x_t *in* h periods under the assumption that $\xi_t = j$, denoted by mapping vectors taking the form $Z^I_{\xi_t,x_{t+h}}$. Analogous mappings for the

machine expectation are denoted with "ML" superscripts, i.e.., $Z^{MF}_{\xi_t,}$, and load on S^{MF}_t. These loadings differ because investor forecasts use their perceived law of motion for the macro block, while machine forecasts use the true law of motion and in addition take into account the AR(1) evolution of η_t that varies with perceived news $\widetilde{\varepsilon}^M_t$.
The observation equation when all variables in X_t are available takes the form:

$$
\begin{bmatrix}
\Delta \ln\left(GDP_t\right) \\
Inflation_t \\
\Delta \ln\left(GDP_t^*\right) \\
Inflation_t^* \\
FFR_t \\
f_t^{(0)} \\
\mathbb{F}_{t,h}^{(s)}\left(Inflation\right) \\
\mathbb{E}_{t,h}\left(Inflation\right) \\
\mathbb{F}_{t,h}^{(s)}\left(\Delta GDP\right) \\
\mathbb{E}_{t,h}\left(\Delta GDP\right) \\
\mathbb{F}_{t,h}^{(BC)}\left(FFR\right) \\
f_t^{(n)} \\
ED_t^{(n)} \\
Baa_t \\
pgdp_t \\
EGDP_{t,t-1} \\
BBG(EGDP_{t,t-1}) \\
POGDP_{t,t-1} \\
DGDP_{t,t-1} \\
r_t^D - i_{t-1} \\
\mathbb{F}_{t,h}^{(s)}(\Delta \ln P_t^D - i_{t-1}) \\
\\
\mathbb{E}_t(\Delta \ln P_t^D - i_{t-1}) \\
\\
\mathbb{F}_{t,h}^{(s)}(r_t^D - i_{t-1}) \\
\\
\mathbb{E}_t(r_t^D - i_{t-1}) \\
\\
F_t^{(n)}(\Delta d) \\
\mathbb{F}_{t,h}^{(IBES)}(\Delta e_t) \\
\mathbb{F}_t(LTG) \\
\\
\mathbb{E}_{t,h}(\Delta e) \\
\mathbb{E}_t(LTG) \\
\\
\end{bmatrix}
=
\begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
D^I_{\xi_t,\pi_{t,t+h}} \\
D^{ML}_{\xi_t,\pi_{t,t+h}} \\
D^I_{\xi_t,y_{t,t+h}} \\
D^{ML}_{\xi_t,y_{t,t+h}} \\
D^I_{\xi_t,i_{t,t+h}} \\
D^I_{\xi_t,i_{t+n}} \\
D^I_{\xi_t,i_{t+n}} \\
C_{Baa} \\
k \\
K \\
K \\
K \\
K \\
\kappa_{pd,0} \\
\alpha_{\widetilde{\mathbb{E}}} \\
\\
\alpha_{\mathbb{E}} \\
\\
\kappa_{pd,0} \\
\\
\kappa_{pd,0} \\
\\
(1-\widetilde{\rho})\mu \\
(1-\widetilde{\rho})\mu \\
(1-\widetilde{\rho})\mu \\
\\
(1-\rho)\mu \\
(1-\rho)\mu \\
\\
\end{bmatrix}
+
\begin{bmatrix}
12\Delta y_t \\
12\pi_t \\
12\Delta y_t^* \\
12\pi_t^* \\
12i_t \\
12i_t \\
Z^I_{\xi_t,\pi_{t,t+h}} S_t^I \\
Z^{ML}_{\xi_t,\pi_{t,t+h}} S_t^{MF} \\
Z^I_{\xi_t,\Delta y_{t,t+h}} S_t^I \\
Z^{ML}_{\xi_t,\Delta y_{t,t+h}} S_t^{MF} \\
Z^I_{\xi_t,i_{t,t+h}} S_t^I \\
Z^I_{\xi_t,i_{t+n}} S_t \\
Z^I_{\xi_t,i_{t+n}}, S_t \\
B\widehat{lp}_t \\
(k_t - k) + pd_t + \Delta y_t \\
K(k_t - k + \Delta y_t) \\
K(k_t - k + \Delta y_t) \\
K(k_t - k + \Delta y_t) \\
K(k_t - k + \Delta y_t) \\
\beta Z_{pd} S_t - Z_{pd} S_{t-1} + Z_k\left(S_t - S_{t-1}\right) + Z_{\Delta y} S_t - Z_i S_{t-1} + Z_\pi S_t \\
Z^I_{\xi_t,pd_{t+h}} S_t^I - Z^I_{\xi_t,pd_t} S_t^I + Z^I_{\xi_t,k_{t,t+h}} S_t^I \\
+ Z^I_{\xi_t,\Delta y_{t,t+h}} S_t^I - Z^I_{\xi_t,i_{t,t+h}} S_t^I + Z^I_{\xi_t,\pi_{t,t+h}} S_t^I \\
Z^{ML}_{\xi_t,pd_{t+h}} S_t^{ML} - Z^{ML}_{\xi_t,pd_t} S_t^{ML} + Z^{ML}_{\xi_t,k_{t,t+h}} S_t^{ML} \\
+ Z^{ML}_{\xi_t,\Delta y_{t,t+h}} S_t^{ML} - Z^{ML}_{\xi_t,i_{t,t+h}} S_t^{ML} + Z^{ML}_{\xi_t,\pi_{t,t+h}} S_t^{ML} \\
\beta Z^I_{\xi_t,pd_{t+h}} S_t^I - Z^I_{\xi_t,pd_t} S_t^I + Z^I_{\xi_t,k_{t,t+h}} S_t^I \\
+ Z^I_{\xi_t,\Delta y_{t,t+h}} S_t^I - Z^I_{\xi_t,i_{t,t+h}} S_t^I + Z^I_{\xi_t,\pi_{t,t+h}} S_t^I \\
\beta Z^{ML}_{\xi_t,pd_{t+h}} S_t^{ML} - Z^{ML}_{\xi_t,pd_t} S_t^{ML} + Z^{ML}_{\xi_t,k_{t,t+h}} S_t^{ML} \\
+ Z^{ML}_{\xi_t,\Delta y_{t,t+h}} S_t^{ML} - Z^{ML}_{\xi_t,i_{t,t+h}} S_t^{ML} + Z^{ML}_{\xi_t,\pi_{t,t+h}} S_t^{ML} \\
\widetilde{\rho}\left(Z^I_{\xi_t,\Delta d_{t,t+n}} + \zeta Z^I_{\xi_t,\eta_t}\right) S_t^I + Z^I_{\xi_t,\pi_{t,t+n}} S_t^I \\
\widetilde{\rho}\left(Z^I_{\xi_t,\Delta d_{t,t+h}} + \zeta Z^I_{\xi_t,\eta_t}\right) S_t^I + Z^I_{\xi_t,\pi_{t,t+h}} S_t^I \\
\widetilde{\rho}\left(Z^I_{\xi_t,\Delta d_{t,t+20}} + \zeta Z^I_{\xi_t,\eta_t}\right) S_t^I + Z^I_{\xi_t,\pi_{t,t+20}} S_t^I \\
-\widetilde{\rho}\left(Z^I_{\xi_t,\Delta d_{t,t+16}} + \zeta Z^I_{\xi_t,\eta_t}\right) S_t^I + Z^I_{\xi_t,\pi_{t,t+16}} S_t^I \\
\rho Z^{ML}_{\xi_t,\Delta d,t,t+h} S_t^{ML} + Z^{ML}_{\xi_t,\pi_{t,t+h}} S_t^{ML} \\
\rho Z^{ML}_{\xi_t,\Delta d,t,t+20} S_t^{ML} + Z^{ML}_{\xi_t,\pi_{t,t+20}} S_t^{ML} \\
-\rho Z^{ML}_{\xi_t,\Delta d,t,t+16} S_t^{ML} + Z^{ML}_{\xi_t,\pi_{t,t+16}} S_t^{ML} \\
\end{bmatrix}
+ U_t v_t
$$

The term GDP_t refers to real gross domestic product, with GDP_t^* the real-time version available at time t. The term $Inflation_t$ in the above stands for 12-monoth ahead CPI inflation with $Inflation_t^*$ the real-time version available at time t. $f_t^{(0)}$ refers to the FFF

contract rate that expires in the current month. FFR is the annualized nominal federal funds rate. $\mathbb{F}_{t,h}^{(s)}$ refers to h-period ahead survey forecast at time t for survey s. For inflation and real GDP growth, surveys s include one-year ahead forecasts from Blue Chip (BC, 12 months ahead), Livingston (LIV, 2 biannual periods), Bloomberg (BBG, 12-months ahead), and Survey of Professional Forecasters (SPF, 4 quarters ahead). For inflation, we also include 10-year ahead forecast from LIV. $\mathbb{F}_{t,h}^{(BC)}(FFR)$ refers to h-period ahead Blue Chip forecast for Fed Funds Rate, with $h = 12$ months. $\mathbb{E}_{t,h}(x)$ refers to h-period ahead machine forecasts of variable x at time t. $f_t^{(n)}$ refers to the time-t contracted federal funds futures market rate, expiring in n months. Here we use $n = \{0, 6, 10, 20, 35\}$, where 0 refers the contract that expires in the current month. $ED_t^{(n)}$ refers to the time-t contracted Eurodollar rate, expiring in n quarters. Here we use $n = \{1, 2, 4, 8\}$. Baa_t is the Baa spread described above, where C_{Baa} and B are scalar parameters. To allow for the fact that the true liquidity premium is only a function of Baa_t, we add a constant C_{Baa} to our model-implied lp_t and scale it by the parameter B to be estimated. The variable $pgdp_{t,t-1}$ is the log of the SP500 capitalization-to-lagged nominal GDP (NGDP) ratio, i.e., $\ln(P_t/NGDP_{t-1})$; $EGDP_{t,t-1}$ is the level of the S&P 500 earnings-to-lagged NGDP ratio (nominal earnings divided by lagged nominal GDP); $POGDP_{t,t-1}$ is the eight quarter moving average of U.S. corporate sector nominal payout relative to lagged NGDP; $DGDP_{t,t-1}$ is the monthly S&P 500 nominal dividend-to- lagged NGDP ratio. These variables are mapped into the model implications for K_t, with $EGDP_{t,t-1} \approx K + K(k_t - k + \Delta y_t)$ and likewise for $POGDP_{t,t-1}, DGDP_{t,t-1}$, where K is the steady state level of $K_t = \exp(k_t)$. To obtain high-frequency information on $EGDP_{t,t-1}$, we use the BBG earnings nowcasts divided by one-month lagged real-time GDP. For announcements that occur during the time frame for which these nowcasts are available (Jan 2, 2006 to July 24, 2024), we use the pre- and post- announcement BBG earnings nowcast-to-lagged GDP ratio. For announcements that pre-date this time frame, we use the end-of-month earnings-lagged GDP ratio allowing for observation error as a noisy proxy for these nowcasts. The variable $r_t^D - i_{t-1}$ is the CRSP-VW excess stock return including dividend distributions. $\mathbb{F}_{t,h}^{(s)}(\Delta \ln P_t^D - i_{t-1})$ refers to survey forecasts of S&P 500 price index growth in excess of the lagged short rate, which corresponds to the LIV and BBG surveys point survey forecasts of the index. $\mathbb{F}_{t,h}^{(s)}(r_t^D - i_{t-1})$ refers to point survey forecasts of returns, which corresponds Gallup/UBS and CFO. The LIV, Gallup/UBS, and CFO surveys are mapped onto annual price growth or return expectations, as appropriate, in the model. The BBG survey is mapped into multi-month returns, depending on the month of the year (see data description above). For SOC, which is not a point forecast but instead a subjective probability of an increase in stock market in next year, we map it onto the investor expectation of one-year ahead returns, allowing for a freely estimated slope

$C_{SOC}$and intercept to account for the change in units to indicate that this is a measure that moves with point forecasts while not being identical to them. $\mathbb{E}_{t,h}\left(\Delta \ln P_t^D - i_{t-1}\right)$ refers to machine forecasts of price growth in excess of the lagged short rate. $\mathbb{E}_{t,h}\left(r_t^D - i_{t-1}\right)$ refers to machine forecasts of returns in excess of the lagged short rate. $F_t^{(n)}(\Delta d)$ refers to the expectations of future dividends constructed from dividend futures markets for $n = 8$ quarters ahead. $\mathbb{F}_{t,h}^{(IBES)}(\Delta e_t)$ refers to $IBES$ analysts forecasts of earnings for $h = 4$ quarters ahead. $\mathbb{E}_{t,h}(\Delta e)$ is the h-quarter ahead machine forecast for earnings growth from IBES with $h = 4$, a noisy signal on rational expectations of Δd_{t+h}.$\mathbb{F}_t(LTG)$ refers to the LTG forecasts. For the mapping to the structrual model, we treat LTG as measuring annual five-year forward growth expectations, i.e., annual earnings growth from four to five years ahead. Machine forecasts for the five-year forward earning growth are denoted $\mathbb{E}_t(LTG)$.

Two points about the mapping bear noting. First, the observation equation often uses multiple measures of observables on a single variable, e.g., investor expectations of inflation 12 months ahead are measured by four different surveys (BC, SPF, LIV, and BBG). Likewise, dividend futures and survey expectations $F_t^{(n)}(\Delta d)$, $\mathbb{F}_{t,h}^{(IBES)}(\Delta e_t)$, and $\mathbb{F}_t(LTG)$ are all taken as noisy signals on the underlying investor expectations process for Δd. In the filtering algorithm above, these provide four noisy signals on the same latent variable. Second, a number of different surveys are used to gauge expectations for multiple variables. These surveys have different deadlines for respondents to turn in their forecasts. Whether monthly or quarterly, the different surveys conduct interviews or have response deadlines that happen somewhere during the course of a specific month. We therefore conservatively set the "response deadline" for the machine forecast to be the first day of every month, implying that we allow the machine to use information only up through the end of the previous month (e.g., through January 31st for an interview or response deadline in February). This ensures that the machine only sees information that would have been available to survey respondents in the months for which that survey is conducted. This approach is conservative in the sense that it handicaps the machine, since all survey respondents who are being interviewed during the next month would have access to more timely information than the machine.

D Additional Figures and Tables

Table A.5 reports the results of regressing excess returns on the S&P 500 on past news in monthly data. Past news is measured as the high-frequency jump in the stock market due to a news event. We first sort all news events by whether the market over- or underreacted based on the structural model estimates. We then sum all the high-frequency jumps in the market around news events in month t in a given reaction category. This aggregated

$Jumps_t$ variable is our measure of past news. We regress future excess returns on $Jumps_t$. We find that news events characterized by overreaction predicts lower future excess returns, while those characterized by underreaction predict higher future excess returns. The results for events where the news was bad, as indicated by a downward jump in the market, are marginally more significant than those where the news was good and are reported separately.

Table A.5: Predicting Returns Using Reactions to News

$rx_{t+h} = \alpha + \beta_J Jumps_t, + \beta_r rx_t + \varepsilon_{t+h}$				
	$h = 12$	$h = 24$	$h = 36$	$h = 60$
Panel (a): **Overreaction**				
All overreaction events				
β_J	$\mathbf{-0.129}^*$	$\mathbf{-0.175}^*$	$\mathbf{-0.214}^*$	$\mathbf{-0.229}^*$
(t-stats)	(-1.75)	(-1.81)	(-1.92)	(-1.77)
Bad market news				
β_J	$\mathbf{-0.228}^*$	$\mathbf{-0.237}^{**}$	$\mathbf{-0.254}^{**}$	$\mathbf{-0.231}^*$
(t-stats)	(-1.84)	(-2.08)	(-2.02)	(-1.93)
Panel (b): **Underreaction**				
All underreaction events				
β_J	0.105	$\mathbf{0.193}^*$	$\mathbf{0.199}^*$	$\mathbf{0.230}^*$
(t-stats)	(1.21)	(1.71)	(1.89)	(1.72)
Bad market news				
β_J	$\mathbf{0.109}^*$	$\mathbf{0.215}^*$	$\mathbf{0.191}^{**}$	$\mathbf{0.159}^{**}$
(t-stats)	(1.79)	(1.94)	(2.04)	(1.98)

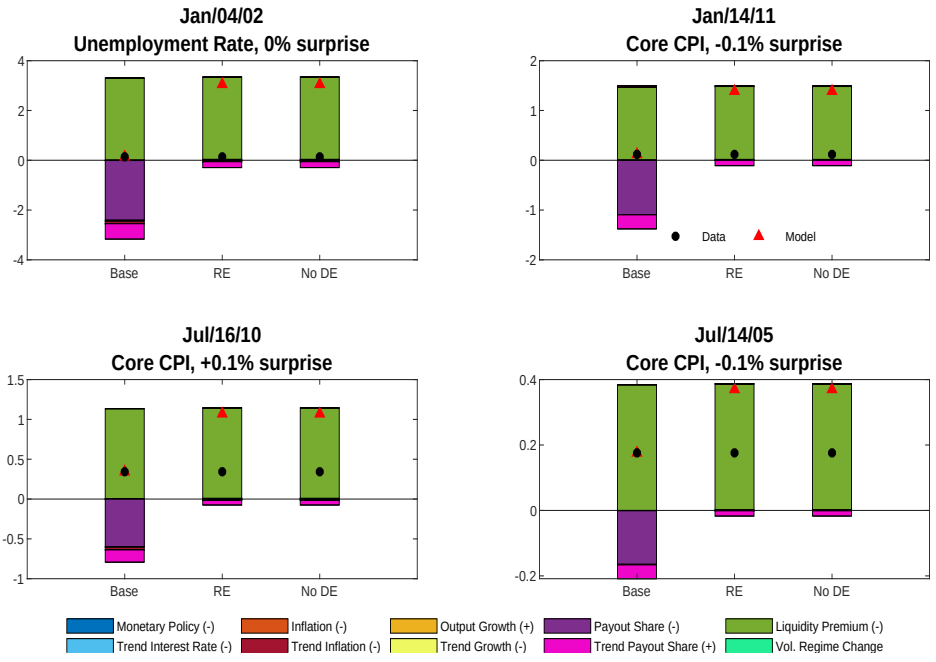
Notes: This table reports results of monthly regressions of the h-month ahead log S\&P 500 stock market return (measured as the log difference in the S\&P 500 market cap) in excess of the 1-month Treasury bill rate ("rx_{t+h}") on the sum of high-frequency changes in the S\&P 500 around all news events in month t in a specific reaction category ("$Jumps_t$"). To obtain a reaction category, we first sort all news events by whether the market over- or underreacted based on the structural model estimates. We then sum the high-frequency jumps in the S\&P 500 around all news events in that reaction category for month t to obtain $Jumps_t$. The results for the subset of events in which $Jumps_t < 0$ are reported under the panel labeled "Bad market news". Newey-west t-statistics are reported in brackets. Bolded numbers indicate significance at 10%$ level. * = sig 10%, ** = sig 5%. The sample spans 1986:M2 - 2021:M12.

Table A.6: Parameter Estimates

	Regime 1			Regime 2	
	Actual	Perceived		Actual	Perceived
σ_i	0.0015	0.0015	σ_i	0.0033	0.0033
σ_π	0.0019	0.0019	σ_π	0.0036	0.0036
$\sigma_{\Delta y}$	0.0054	0.0054	$\sigma_{\Delta y}$	0.0243	0.0244
σ_k	0.1446	0.1448	σ_k	0.3888	0.3891
σ_{lp}	0.0221	–	σ_{lp}	0.0302	–
$\sigma_{\bar{i}}$	0.0047	0.0047	$\sigma_{\bar{i}}$	0.0193	0.0193
$\sigma_{\Delta\bar{y}}$	0.0083	0.0083	$\sigma_{\Delta\bar{y}}$	0.0498	0.0499
$\sigma_{\bar{\pi}}$	0.0364	0.0365	$\sigma_{\bar{\pi}}$	0.0724	0.0725
$\sigma_{\bar{k}}$	0.0430	0.0431	$\sigma_{\bar{k}}$	0.0981	0.0982

Notes: Posterior mode values of the parameters. The estimation sample spans 1961:M1-2021:M12.

Figure A.1: Largest Underreaction Events (%)



Notes: The figure reports shock decompositions of pre-/post- FOMC announcement changes in S&P 500 attributable to revisions in the perceived macro shocks and the subjective equity premium (the combined effect of shocks to lp_t and stochastic volatility). The specific FOMC events reported on are those for which the absolute difference between the market's jump under the RE counterfactual and the baseline model as a fraction of the market jump is largest. The modifiers (+) or (-) refer to the sign of the baseline response to a positive increment in the fundamental shock labeled in the legend. The sample is 2001:M1-2021:M12.