Investor Attention and Stock Market Volatility*

Daniel Andrei†   Michael Hasler‡

February 26, 2014

Abstract

We investigate in a theoretical framework the joint role played by investors’ attention to news and learning uncertainty in determining asset prices. The model provides two main predictions. First, stock return variance and risk premia increase with both attention and uncertainty. Second, this increasing relationship is quadratic. We empirically test these two predictions, and we show that the data lend support to the increasing relationship. The evidence for a quadratic relationship is mixed. Overall, our study shows theoretically and empirically that both attention and uncertainty are key determinants of asset prices.

Keywords: Asset pricing, general equilibrium, learning, attention, uncertainty, volatility, risk premium.

*We would like to thank Brian Boyer, Bruce Carlin, Max Croce, Bernard Dumas, David Hirshleifer, Julien Hugonnier, an anonymous referee, and Pietro Veronesi for useful comments. We would also like to acknowledge comments from seminar participants at the Mathematical Finance Days Conference in Montréal, the Princeton-Lausanne Workshop on Quantitative Finance in Lausanne, UCLA Anderson, the 16th SGF Conference in Zurich, the Red Rock Finance Conference 2013, and the Eurofidai Conference in Paris. Financial support from the Swiss Finance Institute, NCCR FINRISK of the Swiss National Science Foundation, UCLA, and the University of Toronto is gratefully acknowledged.

†UCLA, Anderson School of Management, 110 Westwood Plaza, Suite C420, Los Angeles, CA 90095, USA, daniel.andrei@anderson.ucla.edu, www.danielandrei.net

‡University of Toronto, 105 St. George Street, Suite 431, Toronto, ON M5S 2E8, Canada, Michael.Hasler@rotman.utoronto.ca, www.rotman.utoronto.ca/Faculty/Hasler.aspx
1 Introduction

Growing empirical evidence suggests that investor attention fluctuates over time, which impacts asset prices (Da, Engelberg, and Gao, 2011). For example, high levels of attention cause buying pressures and sudden price reactions (Barber and Odean, 2008; Barber, Odean, and Zhu, 2009), whereas low levels generate underreaction to announcements (Dellavigna and Pollet, 2009). In fact, as Huberman and Regev (2001) point out, prices react to new information only when investors pay attention to it.

In this paper, we characterize the mechanism that accounts for these findings and study the relationship between attention, return volatility, and risk premia. We study time-varying attention in a pure exchange economy (Lucas, 1978) populated by a representative investor with recursive preferences (Epstein and Zin, 1989). A single risky asset pays a continuous stream of dividends with an expected growth rate that is unobservable and therefore needs to be estimated. In our model, the investor estimates the expected growth rate by paying fluctuating attention to available information. Fluctuations in attention are governed by changes in the state of the economy.

We show that stock-return volatility and risk premia increase with attention. The intuition for this can be understood as follows. When investors pay little attention to news, information is only gradually incorporated into prices because learning is slow. Therefore, low attention results in low return volatility. In contrast, attentive investors immediately incorporate new information into prices and thus high attention induces high return volatility. In addition, because high attention generates volatile returns, investors require a large risk premium to bear this attention-induced risk. Conversely, low attention generates low return volatility and thus a smaller risk premium.

Through the learning mechanism, fluctuations in attention generate further effects on the

\footnote{Lustig and Verdelhan (2012) provide evidence of time-varying attention. They compute the number of Google Insight searches for the word “recession” and observe huge spikes in the search volume in December 2007 and January 2008. Vlastakis and Markellos (2012), Dimpfl and Jank (2011), and Kita and Wang (2012) find support of counter-cyclical attention to news.}
volatility and risk premium. When attention is low, learning is slow, and thus uncertainty about the future dividend growth rate tends to be high, which implies high levels of return volatility and risk premia. Conversely, when attention is high, learning is fast, and thus uncertainty tends to be low, which generates low levels of volatility and risk premia. Therefore, our model predicts that volatility and risk premia increase not only with attention but also with uncertainty.

Attention and uncertainty, thus, have a joint impact on equilibrium asset prices. Several empirical studies show that uncertainty is indeed a priced risk factor (e.g., Massa and Simonov, 2005; Ozoguz, 2009) and that attention and volatility are strongly related (e.g., Vlastakis and Markellos, 2012; Dimpfl and Jank, 2011; Kita and Wang, 2012). In our own empirical investigation, we validate the above predictions of the model using Google search data to proxy for investors’ attention to news, the dispersion in analyst forecasts to proxy for uncertainty, the GARCH estimation on S&P500 returns to proxy for realized variance, and the fitted values of a predictive regression of future S&P500 returns on current dividend yields to proxy for risk premia. We first calibrate the parameters of our model using the General Method of Moments (Hansen, 1982). Then, regressing the S&P500 variance and risk premium on attention and uncertainty provides positive and significant parameters, which suggests that the data support the increasing relations between attention/uncertainty and variance/risk premia.

Additionally, the model predicts quadratic relationships between attention/uncertainty and variance/risk premia, with positive quadratic coefficients. We further investigate these predictions empirically and find only mixed support for them in the data. Indeed, variance and risk premia increase quadratically with attention only for sufficiently large values of attention, whereas the quadratic coefficients of uncertainty are only weakly significant, though nonetheless positive.

It is natural to expect investors’ attention to react to changes in the state of the economy, as shown in the theoretical work of Detemple and Kihlstrom (1987), Huang and Liu (2007), Kacperczyk, Van Nieuwerburgh, and Veldkamp (2009), and Hasler (2012). The approach that we
adopt in our model is to measure the state of the economy as the weighted average of past dividend
surprises. Since there is no clear evidence as to whether investors’ attention depends on past
 dividend surprises, we propose a variant of our model in which attention depends on past stock
return surprises. We obtain very similar results, aside from a higher unconditional volatility for
the market price of risk, volatility, risk premium, and the risk-free rate.\(^2\)

In addition, we extend our model to account for dividend leverage (see, e.g., Abel, 1999; Bansal, Dittmar, and Lundblad, 2005). This extension preserves the positive relationship between
attention/uncertainty and risk premium/volatility, but it substantially magnifies the average level
of stock-return volatility, price-dividend-ratio volatility, and risk premium; all of these values
become comparable with historical estimates.

Our work contributes to a growing literature shedding light on the importance of attention
allocation in financial markets. The rational inattention model of Sims (2003), which captures the
trade-off between information quality and effort costs, has been particularly influential. Rational
inattention theory can explain a wide array of phenomena, from category learning and return
cou-movement (Peng and Xiong, 2006), to state-dependent attention allocation (Kacperczyk et al.,
2009) and under-diversification in portfolio holdings (Van Nieuwerburgh and Veldkamp, 2009,
2010).

A separate and large body of literature investigates the effect of information quality and learning
on portfolio choice and asset prices (see Pastor and Veronesi (2009) for a recent survey).\(^3\) Our
paper contributes to this literature by providing predictions on the joint impact of time-varying
attention and uncertainty on stock-return volatility and risk premia.

Learning over the course of the business cycle has been extensively analyzed in the literature, and

\(^2\)The high volatility of the market price of risk is puzzling for asset-pricing models. It is rationalized with non-
standard preferences (e.g., Campbell and Cochrane, 1999; Barberis, Huang, and Santos, 2001) or with intermittent
portfolio re-balancing (Chien, Cole, and Lustig, 2012).

\(^3\)Other representative papers are Dothan and Feldman (1986), Detemple (1986), Gennotte (1986), Timmermann
(2003), Dumas, Kurshev, and Uppal (2009), and Ai (2010).
two papers in particular are closely related. The first, Van Nieuwerburgh and Veldkamp (2006), provides a foundation for the observed sharp and sudden downturns, followed by gradual recoveries. Their explanation is based on the fact that agents learn better when productivity is high than when it is low.\footnote{Consistent with their theoretical model, Van Nieuwerburgh and Veldkamp (2006) show that analyst forecast errors on the level of nominal GDP are counter-cyclical. In contrast, we estimate our model on real GDP growth rates and corresponding analyst forecasts and find support for pro-cyclical analyst forecast errors in this case.} The second, Kacperczyk et al. (2009), belongs to the rational inattention literature originated by Sims (2003).\footnote{See, e.g., Peng and Xiong (2006), Mondria (2010), Van Nieuwerburgh and Veldkamp (2010), and Mondria and Quintana-Domeque (2013) for other studies on limited information-processing capacity and their financial consequences.} It builds a static model in which mutual fund managers allocate more attention to aggregate shocks and less attention to idiosyncratic shocks during recessions, which is consistent with our calibration.

Our model features a representative investor with recursive utility who learns about the expected growth rate of consumption. Consequently, our paper is related to the long-run risk literature as a whole (e.g., Bansal and Yaron, 2004; Ai, 2010; Drechsler and Yaron, 2011; Colacito and Croce, 2013; Bansal, Kiku, and Shaliastovich, 2013), but more precisely to Bansal and Shaliastovich (2011), who study an economy with long-run risk in which the expected growth rate of consumption is unobservable. In their setup, the investor optimally chooses to pay a cost and observe the expected growth rate when the volatility of consumption is sufficiently high. This implies jumps in asset prices in bad times, even though the determinants of consumption have smooth dynamics. In contrast, we provide and test predictions linking attention and uncertainty to volatility and risk premia.

The remainder of the paper is organized as follows. Section 2 describes the model and derives the equilibrium, Sections 3 and 4 present the results, Section 5 presents two extensions that help to better fit observed unconditional asset-pricing moments than the regular model, and Section 6 concludes. Derivations and computational details are provided in Appendix A.
2 An Equilibrium Model of Fluctuating Attention

We consider a pure exchange economy (Lucas, 1978) in which the attention paid by a representative investor to news is assumed to be state-dependent. First, we describe the economic setting and the learning problem of the investor. Second, we discuss the dynamics of the investor’s attention to news. Third, we solve for the equilibrium and derive its asset-pricing implications. Fourth, we estimate the parameters of the model by using the Generalized Method of Moments (Hansen, 1982).

2.1 The Economic Setting

The economy is characterized by a single output process (the dividend) having an unobservable expected growth rate (the fundamental). There are two securities, one risky asset in positive supply of one unit and one risk-free asset in zero net supply. The risky asset is defined as the claim to the dividend process $\delta$, with dynamics that are given by

$$\frac{d\delta_t}{\delta_t} = f_t dt + \sigma_\delta dZ^\delta_t.$$

The unobservable fundamental $f$ follows a mean reverting process, as follows:

$$df_t = \lambda (\bar{f} - f_t) dt + \sigma_f Z^f_t.$$

The economy is populated by a representative investor. Given that the fundamental is unobservable, the investor uses the information at hand to estimate it. The investor observes the current dividend $\delta$ and an informative signal $s$ with dynamics

$$ds_t = \Phi_t dZ^\delta_t + \sqrt{1 - \Phi_t^2} dZ^s_t.$$
The vector \((Z^δ, Z^f, Z^s)^\top\) is a 3-dimensional standard Brownian motion under the complete information filtration. The process \(\Phi\) represents the correlation between the signal and the fundamental. Without loss of generality, this correlation is assumed to be positive.

The model belongs to the literature on continuous-time consumption-portfolio decision problems with incomplete information (Detemple, 1986; Gennotte, 1986; Dothan and Feldman, 1986). We adopt, however, a slightly different signal structure. Specifically, in our setup the signal provides information on changes in the fundamental and not on its level.\(^6\) Consequently, even if the signal is perfectly accurate, the investor doesn’t observe the true value of the fundamental. This difference from classic models does not change our results qualitatively.

Let us focus on the correlation \(\Phi\) between the signal and the fundamental. Consistent with Detemple and Kihlstrom (1987), Veldkamp (2006a,b), Huang and Liu (2007), Peng and Xiong (2006), and Hasler (2012), \(\Phi\) can be interpreted as the accuracy of news updates observed by the investor. The case \(\Phi = 0\) is equivalent to no news updates, while \(\Phi = 1\) is equivalent to perfectly accurate news updates. In the references above, the investor can exert control over this accuracy, and thus the parameter \(\Phi\) is called attention to news.\(^7\)

We follow the same interpretation and assume that the investor directly controls this accuracy, although we do not solve for the optimal attention allocation. Instead, we claim that the investor changes her attention whenever she observes changes in the state of the economy (in a way that we will define precisely). Hence, in our model \(\Phi\) is time-varying and is directly determined by the (observable) state of the economy.

\(^6\)To the best of our knowledge, this specification was first adopted by Dumas et al. (2009). The alternative would be to assume that the investor observes a noisy signal of the fundamental, \(ds_t = f_t dt + \frac{1-\Phi_t}{\sqrt{\Phi_t}} dZ^s_t\). In that case, the variance of the noisy signal would be stochastic and would belong to the interval \([0, \infty)\).

\(^7\)Other types of attention have been studied. In the rational inattention literature (Kahneman, 1973; Sims, 2003), Peng and Xiong (2006), Kacperczyk et al. (2009), and Van Nieuwerburgh and Veldkamp (2010) study attention-allocation problems when information-processing capacity is limited. Duffie and Sun (1990), Abel, Eberly, and Panageas (2007), and Rossi (2010) study attention to wealth in the sense that investors optimally choose not to trade in certain periods. Chien et al. (2012), Bacchetta and Wincoop (2010), and Duffie (2010) assume that these periods of inattention are fixed and investigate their impact on asset prices.
Assuming an exogenous process for attention has two advantages. First, it offers the benefit of analytical tractability. Second, by linking information choice (which is often difficult to measure) to the state of the economy (which is observable), we are able to generate testable predictions and bring our model to the data. It is worth mentioning that the main goal in this analysis is not to provide a theoretical foundation of fluctuating attention but to understand how asset prices react when attention fluctuates.\textsuperscript{8} Fluctuating attention is confirmed by empirical research,\textsuperscript{9} and also by our own estimation. As such, our model relies on the premise that investors’ attention is time-varying and thus can be viewed as a simplification of a more complex problem.

2.2 Time-Varying Attention

Our aim is to connect attention movements to variables that represent the state of the economy. In the present model, two observable variables can fulfill this role: the dividend or the stock price. The dividend is exogenous, whereas the stock price is endogenously determined in equilibrium.

Because the dividend is exogenous, assuming that attention depends on it is technically easier than if attention were to depend on the stock price. In the latter case, solving for the equilibrium involves a fixed-point problem. We solve for the equilibrium in both cases and show that results are similar. Therefore, and for ease of exposition, we analyze in this section the case in which movements in dividends drive investors’ attention. We relegate to Section 5.1 a full derivation and discussion of the case in which attention is driven by price movements.

We postulate a measure of the state of the economy that captures two important features. First, it reflects not only the current but also the past performance of dividends. This describes a usual behavior of investors to search for trends in financial data (exponential smoothing for forecasting is a standard practice). Second, broad evidence shows that surprises, rather than news, face increased

\textsuperscript{8}Potential foundations of fluctuating attention to news can be found in Detemple and Kihlstrom (1987), Veldkamp (2006a,b), Kacperczyk et al. (2009), Bansal and Shaliastovich (2011), and Hasler (2012).

scrutiny from investors.\textsuperscript{10} Thus, we rely on surprises in dividend growth rather than just dividend growth. The resulting variable, which we denote the performance index, captures both features. It is defined as follows

$$\phi_t \equiv \int_0^t e^{-\omega(t-u)} \left( \frac{d\delta_u}{\delta_u} - \tilde{f}_u du \right),$$

where $\tilde{f}_u$ represents the investor’s estimate of the fundamental at time $u$. This estimate results from Bayesian learning and is defined in Section 2.3. The parameter $\omega > 0$ represents the weight associated with the present relative to the past. If $\omega$ is large, the past dividend growth influences, to a minimal degree, the performance index, and the latter becomes almost a substitute for the current dividend growth. On the other hand, if $\omega$ is small, the past dividend growth influences the performance index to a greater extent.\textsuperscript{11}

The dynamics of the performance index are derived from the dynamics of the dividend. An application of Itô’s lemma on the performance index yields

$$d\phi_t = -\omega \phi_t dt + \sigma_d W^\delta_t,$$

where $W^\delta$ is the innovation process defined in Section 2.3. This shows that the performance index fluctuates around a long-term mean of zero with a mean-reversion speed $\omega$.

We are now ready to introduce the link between the performance index and investors’ attention. The following definition is the core of our way of modeling time-varying attention.


\textsuperscript{11}Koijen, Rodriguez, and Sbuelz (2009) build a similar performance index in a dynamic asset allocation problem to allow for momentum and mean reversion in stock returns. In our case, this index is built directly from the dividend process in order to capture in a parsimonious way the recent development of the dividend.
Definition 1. Attention $\Phi$ is defined as a function $g$ of the performance index:

$$\Phi_t = g(\phi_t) \equiv \frac{\bar{\Phi}}{\Phi + \left(1 - \bar{\Phi}\right)e^{\Lambda\phi_t}},$$

(1)

where $\Lambda \in \mathbb{R}$ and $0 \leq \bar{\Phi} \leq 1$.

Attention $\Phi$ fluctuates around a long-run mean $\bar{\Phi}$ and lies in the interval $[0, 1]$. According to the sign of the parameter $\Lambda$, attention can either increase ($\Lambda < 0$) or decrease ($\Lambda > 0$) with the performance index $\phi$. The dynamics of investors’ attention are thus explained by three parameters: $\omega$, $\bar{\Phi}$, and $\Lambda$.

2.2.1 Parameters of the Attention Process

The unconditional distribution of the performance index $\phi$ is Gaussian with mean 0 and a variance given by $\sigma^2/2\omega$ (see Appendix A.1 for a proof of this statement). We know from Eq. (1) that, for $\Lambda \neq 0$, $\Phi$ is a strictly monotone function of $\phi$. This monotonicity allows us to compute the density function of attention $\Phi$ by a change-of-variable argument. We provide this density function in Appendix A.1 and proceed here with its discussion.

The parameter $\bar{\Phi}$ dictates the location of the unconditional distribution of attention. Two other parameters govern the shape of this distribution. The first is $\Lambda$, the parameter that dictates the adjustment of attention after changes in the performance index. The second is $\omega$, the parameter that dictates how fast the performance index adjusts after changes in dividends. Figure 1 illustrates the probability density function of attention for different values of these two parameters. The solid black line corresponds to the calibration performed in Section 2.5 on U.S. data, which shows that attention can vary substantially and take extreme values with significant probabilities.

The two additional lines show that a decrease in the parameter $\Lambda$ (dashed blue line) and an increase in the parameter $\omega$ (dotted red line), respectively, have similar effects: both tend to
Figure 1: Probability density function of investors’ attention.

The figure shows the probability density function of $\Phi$ for different values of $\Lambda$ and $\omega$. Other parameters are $\lambda = 0.42$, $\bar{f} = 0.028$, $\Phi = 0.368$, $\sigma_f = 0.029$, and $\sigma_{\delta} = 0.014$. The solid black line illustrates the probability density function for $\Lambda = 286$ and $\omega = 4.74$ (this corresponds to the calibration performed in Section 2.5 on U.S. data). The dashed blue line shows how the distribution changes with a lower $\Lambda$ ($\Lambda = 100$ and $\omega = 4.74$). The dotted red line shows how the distribution changes with a higher $\omega$ ($\Lambda = 286$ and $\omega = 10$).

bring attention closer to its long-run mean. Although these effects are similar, the parameters $\omega$ and $\Lambda$ have different impacts on the process $\Phi$. The parameter $\omega$ dictates the length of the history of dividends taken into account by the investor. If $\omega$ is large, the investor tends to focus more on recent dividend surprises, and attention reverts quickly to its mean. Consequently, the unconditional distribution concentrates more around the long-run mean $\Phi$. On the other hand, the parameter $\Lambda$ governs the amplitude of the attention movements. A parameter $\Lambda$ close to 0 (positive or negative) would keep the attention close to its long-run mean. The distinct roles played by these two parameters help us to calibrate them on U.S. data, which is a task that we undertake in Section 2.5.
2.3 Bayesian Learning

The advantage of linking attention to an observable variable is that attention itself becomes observable. Thus, the setup remains conditionally Gaussian, and the Kalman filter is applicable. The state vector prior to the filtering exercise consists of one unobservable variable (the fundamental $f$) and a vector of two observable variables $\vartheta = (\zeta_s)^\top$, where we define $\zeta \equiv \log \delta$. In other words, the investor observes the dividend and the signal and tries to infer the fundamental. Since the performance index $\phi$ is built entirely from past values of dividends, it does not bring any additional information.

The distribution of $f_t$ conditional on observing the history of the dividend and the signal is Gaussian with mean $\hat{f}_t$ and variance $\gamma_t$ (Liptser and Shiryaev, 2001). Therefore, the investor’s updating rule is defined by

$$
\begin{align*}
    d\zeta_t &= \left(\hat{f}_t - \frac{1}{2} \sigma_\delta^2\right) dt + \begin{pmatrix} \sigma_\delta & 0 \end{pmatrix} dW_t \\
    d\hat{f}_t &= \lambda \left(\bar{f} - \hat{f}_t\right) dt + \begin{pmatrix} \frac{n_t}{\sigma_\delta} & \sigma_f \Phi_t \end{pmatrix} dW_t \\
    d\gamma_t &= \sigma_f^2 - 2\lambda \gamma_t - \left(\sigma_f^2 \Phi_t^2 + \frac{\gamma_t^2}{\sigma_\delta^2}\right)
\end{align*}
$$

where $W \equiv (W^\delta, W^s)^\top$ is a 2-dimensional Brownian motion under the investor’s observation filtration, and $\Phi$ is provided in Definition 1. The estimated fundamental is denoted by $\hat{f}$, and the two Brownian motions governing the system are defined by

$$
\begin{align*}
    dW_t^\delta &= \frac{1}{\sigma_\delta} \left[d\zeta_t - \left(\hat{f}_t - \frac{1}{2} \sigma_\delta^2\right) dt\right] \\
    dW_t^s &= ds_t
\end{align*}
$$

and represent the normalized innovation processes of dividend and signal realizations. The proof of the above statements is provided in Appendix A.2.
Formally, we define uncertainty as the conditional variance of the fundamental given today’s information. In the present setup, this corresponds to \( \gamma \) (usually referred to as the posterior variance, or Bayesian uncertainty). Equation (3) shows the dynamics of \( \gamma \). The first two terms denote the incremental uncertainty induced by the dynamics of the fundamental itself; the smaller the volatility of the fundamental and the larger the mean-reversion speed, the smaller the incremental uncertainty becomes. The term in brackets defines the reduction in uncertainty due to more accurate information; the more attentive the investor and the smaller the volatility of the dividend, the smaller the incremental uncertainty becomes.

Note that, unlike other learning models such as Scheinkman and Xiong (2003) and Dumas et al. (2009), in our framework uncertainty never converges to a steady state, although it is locally deterministic. The reason is that high attention tends to reduce uncertainty, whereas low attention tends to increase uncertainty. A similar channel of fluctuating uncertainty with Gaussian shocks is studied by Xia (2001) in a dynamic portfolio choice problem in which an agent learns about stock return predictability.\(^{12}\)

The dynamics of \( \hat{f} \) in Eq. (2) reveal that two diffusion components drive the overall noise in the fundamental. The first component loads on dividend innovations and the second on news innovations. Since these two types of innovation represent the signals used by the investor to infer the value of the fundamental, the vector \([\gamma_t/\sigma_\delta \sigma_f \Phi_t]\) constitutes the weights assigned by the investor to both signals. When attention changes, these weights move in opposite directions: for high levels of attention, the investor assigns more weight to news shocks, whereas for low levels of attention the investor assigns more weight to dividend shocks. Consequently, the variance of the estimated fundamental, \( \sigma^2(\hat{f}_t) \), is time-varying and comprises two opposing effects that arise from

learning. It follows from Eq. (2) that this variance is written as

\[ \sigma^2(\hat{f}_t) = \frac{\gamma^2}{\sigma^2_{\delta}} + \sigma^2 \Phi^2. \]  

Equation (4) shows how attention drives the volatility of the estimated fundamental. Higher attention incorporates more information into the updating of \( f \) and thus generates a higher instantaneous volatility (a direct impact). But higher attention also reduces uncertainty and thus tends to decrease volatility (an indirect impact).\(^{13}\)

To summarize, we offer a framework to analyze the impact of both attention and uncertainty on asset prices. In particular, our aim is, first, to establish model-implied predictions on the relationships among attention, uncertainty, stock-return volatility, and risk premia, and second, to test these predictions by using real data. We now turn to the computation of the equilibrium.

### 2.4 Equilibrium with Epstein-Zin Preferences

The representative investor’s preferences over the uncertain consumption stream \( \{c_t\} \) are represented by a utility index \( U_t \) that satisfies the following recursive equation:

\[
U_t = \left\{ \left(1 - e^{-\rho dt}\right) c_t^{1-\psi} + e^{-\rho dt} E_t \left[U_{t+dt}^{1-\alpha}\right]^{\frac{1-\psi}{1-\alpha}} \right\}^{\frac{1}{1-\psi}},
\]

where \( \rho \) is the subjective discount factor, \( 1/\psi \) is the elasticity of intertemporal substitution, and \( \alpha \) is the relative risk-aversion coefficient. Replacing \( dt = 1 \) in Eq. (5) gives the discrete time formulation of Kreps and Porteus (1978), Epstein and Zin (1989), and Weil (1989). When the risk-aversion coefficient is equal to the reciprocal of the elasticity of intertemporal substitution,\(^{13}\)

\(^{13}\)Note that the indirect effect is not contemporaneous; since the uncertainty process \( \gamma \) has deterministic dynamics (see Eq. (3) and its interpretation), uncertainty decreases only gradually when attention is high. In other words, because uncertainty is a “\( dt \)” process and attention is driven by a Brownian motion, uncertainty is always “one step behind” (Xia, 2001). This lead-lag relation disconnects uncertainty from short-term movements in attention, and thus sudden spikes in attention can result in states of high attention and high uncertainty.
\( \alpha = \psi \), the recursive utility is reduced to the time-separable power utility with relative risk-aversion \( \alpha \).

Let us define

\[
J_t \equiv \frac{1}{1 - \alpha} U_t^{1 - \alpha} = \mathbb{E}_t \left[ \int_t^{\infty} f(c_s, J_s) \, ds \right],
\]

where

\[
f(c, J) = \frac{\rho \psi^{1 - \psi}}{((1 - \alpha)J)^{1/\nu - 1} - \rho \nu J}
\]

is the normalized aggregator (see Duffie and Epstein, 1992a,b) and \( \nu \equiv \frac{1 - \alpha}{1 - \psi} \). Following Benzoni, Collin-Dufresne, and Goldstein (2011), a state-price density is defined as

\[
\xi_t = e^{\int_0^t f(c_s, J_s) \, ds} f_c(c_t, J_t).
\]

The following proposition, the proof of which is presented in Benzoni et al. (2011), provides the partial differential equation defining the price-dividend ratio.

**Proposition 1.** For \( \psi, \alpha \neq 1 \) we have

\[
J_t = \frac{1}{1 - \alpha} c_t^{1 - \alpha} (\rho I(x_t))^\nu
\]

\[
\xi_t = e^{-\int_0^t (\rho \nu + \frac{1 - \alpha}{\nu} \sigma^2_\delta) \, ds} \rho \nu \, c_t^{-\alpha} I(x_t)^{\nu - 1},
\]

where \( x \equiv \left( \frac{\hat{f}}{\phi} \gamma \right)^\top \) and \( I(x) \) is the price-dividend ratio. The price-dividend ratio \( I(.) \) satisfies the following partial differential equation:

\[
0 = I \left( (1 - \alpha) \left( \frac{\hat{f}}{2 \sigma^{2}_\delta} + (1 - \alpha) \frac{(1 - \nu) \sigma^{2}_\delta}{2} - \rho \nu \right) + \frac{\partial I}{\nabla - 1} + (1 - \alpha) \nu \left( \gamma \frac{\hat{f}}{\phi} + \sigma^{2}_\delta \right) \right) + \nu,
\]
where we define $D h(x) \equiv h'(x) \mu(x) + \frac{1}{2} \text{trace} \left( h''(x) \sigma(x) \sigma(x)^\top \right)$.

The dynamics of the vector of state variables $x$ and price-dividend ratio $I(x)$ are defined by

$$
\begin{align*}
\text{dx}_t & = \mu_x(x_t) dt + \sigma_x(x_t) dW_t \\
\frac{dI(x_t)}{I(x_t)} & = (\ldots) dt + \sigma_I(x_t) dW_t,
\end{align*}
$$

where

$$
\begin{align*}
\mu_x(x_t) & = \left( \lambda(\bar{f} - \tilde{f}_t) - \omega \phi_t, \sigma_f^2(1 - \Phi^2) - 2 \lambda \gamma_t - \frac{\gamma^2}{\sigma^2} \right)^	op \\
\sigma_x(x_t) & = \begin{pmatrix}
\frac{\gamma}{\sigma} & \sigma_f \Phi_t \\
\sigma^2 & 0 \\
0 & 0
\end{pmatrix} \\
\sigma_I(x_t) & \equiv \begin{pmatrix}
\sigma_{1I}(x_t) & \sigma_{2I}(x_t)
\end{pmatrix} = \frac{1}{I(x_t)} I_x(x_t) \sigma_x(x_t).
\end{align*}
$$

The partial differential equation (7) can be rewritten as

$$
0 = I \left( (1 - \alpha) \left( \tilde{f} - \frac{1}{2} \sigma^2 + (1 - \alpha) \frac{\sigma^2}{2} - \rho \nu \right) \\
+ \nu \left( \lambda \left( \tilde{f} - \tilde{f} \right) I_{\tilde{f}} - \omega \phi I_{\phi} + \left( \sigma_f^2(1 - \Phi^2) - 2 \lambda \gamma - \frac{\gamma^2}{\sigma^2} \right) I_{\gamma} \right) \\
+ \frac{1}{2} \nu \left( \left( \frac{\gamma^2}{\sigma^2} + \sigma_f^2 \Phi^2 \right) I_{\tilde{f}} + \sigma^2 I_{\phi} + 2 \gamma I_{\gamma} \right) \right) (8)
$$

As in Benzoni et al. (2011), let us conjecture that the price-dividend ratio $I(x)$ can be approximated
by the following exponential form:

\[ I(x) \approx e^{\beta_0 + \beta_1 x}, \]  

(9)

where \( \beta_0 \) is a scalar and \( \beta_1 = \begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} \end{pmatrix} \). Plugging the exponential form (9) in Eq. (8) and performing a first-order linearization of the PDE around \( x^0 = \begin{pmatrix} \bar{f} & 0 & \gamma_{ss} \end{pmatrix} \) yields a system of the form\(^{14}\)

\[ A + Bx = 0, \]

where the scalar \( A \) and the vector \( B = \begin{pmatrix} B_1 & B_2 & B_3 \end{pmatrix} \) are large expressions. Setting \( A = 0 \) and \( B = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} \) yields a system of four equations with four unknowns \( (\beta_0, \beta_{11}, \beta_{12}, \text{ and } \beta_{13}) \) that can be solved numerically.

As in Benzoni et al. (2011), we assume from now on that the coefficient of risk aversion is \( \alpha = 10 \), the elasticity of intertemporal substitution is \( \frac{1}{\psi} = 2 \), and the subjective discount factor is \( \rho = 0.03 \).

The aforementioned utility parameters, together with the parameters exposed in Table 1, imply the following equilibrium price-dividend ratio:

\[ I(x) \approx e^{\beta_0 + \beta_{11} \bar{f} + \beta_{12} \phi + \beta_{13} \gamma}, \]

where

\[ \beta_0 = 3.6152 \quad \beta_{11} = 1.1207 \quad \beta_{12} = -0.006 \quad \beta_{13} = -11.3741. \]

\(^{14}\)Note that \( \bar{f} \) is the long-term mean of \( \bar{f} \), \( 0 \) is the long-term mean of \( \phi \), and \( \gamma_{ss} = -\lambda \sigma_\delta^2 + \sqrt{\sigma_\delta^2 \left( \lambda^2 \sigma_\delta^2 + \sigma_f^2 (1 - \Phi^2) \right)} \) is the long-term posterior variance under the assumption that \( \phi_t = 0 \).
These parameters show that the level of the price-dividend ratio depends strongly on the estimated fundamental $\hat{f}$, slightly on the posterior variance $\gamma$, and insignificantly on the performance index $\phi$ (or attention $\Phi$). Indeed, since $\hat{f}$, $\gamma$, and $\phi$ are of the order of $10^{-2}$, $10^{-4}$, and $10^{-2}$, respectively, the impacts of these processes on the log price-dividend ratio are of the order of $10^{-2}$, $10^{-3}$, and $10^{-6}$, respectively.

The positive relationship between the estimated fundamental and the price-dividend ratio ($\beta_{11} > 0$) can be understood by analyzing the precautionary savings and the substitution effects. First, an increase in the estimated fundamental implies an increase in current consumption because future consumption is expected to be larger, and investors wish to smooth consumption over time. Hence, the demand for the stock decreases, implying a drop in the price. This precautionary savings effect generates an inverse relationship between prices and expected dividend growth rates. Second, an increase in the estimated fundamental implies an improvement of risky investment opportunities, which pushes investors to increase their demand for the stock. This substitution effect outweighs the precautionary savings effect as long as the elasticity of intertemporal substitution is larger than 1. Therefore, prices increase with the estimated fundamental in our model.

An increase in uncertainty generates a drop in prices ($\beta_{13} < 0$). Intuitively, an increase in uncertainty pushes investors to lower current consumption because expected consumption is more uncertain and investors wish to smooth consumption over time. Hence the demand for the stock rises, increasing its price. But riskier investment opportunities push investors to lower their risky investments. This tends to decrease the stock price. Since the substitution effect dominates the precautionary savings effect as long as the elasticity of intertemporal substitution is larger than 1, uncertainty and prices are inversely related in our model.
2.4.1 Risk Free Rate, Risk Premium, and Volatility

Applying Itô’s lemma to the state-price density $\xi$ provided in Eq. (6) yields the risk free rate $r$ and the vector of market prices of risk $\theta$ defined in Proposition 2 below. Proofs are provided in Benzoni et al. (2011).

**Proposition 2.** The risk-free rate $r$ and market price of risk $\theta$ satisfy

$$r_t = \rho + \psi \hat{f}_t - \frac{1}{2} \alpha (1 + \psi) \sigma^2$$

$$- (1 - \nu) \left( \sigma_{1f}(x_t) \sigma_{\delta} + \frac{1}{2} \sigma_{1f}(x_t)^2 + \frac{1}{2} \sigma_{2f}(x_t)^2 \right)$$

$$\theta_t = \left( \alpha \sigma_{\delta} + (1 - \nu) \sigma_{1f}(x_t) \ (1 - \nu) \sigma_{2f}(x_t) \right)^\top.$$  

(10)

18

The dynamics of the stock price $S = \delta I(x)$ are written as

$$\frac{dS_t}{S_t} = \left( \mu_t - \frac{\delta_t}{S_t} \right) dt + \sigma_t dW_t.$$  

The diffusion vector $\sigma_t$ and the risk premium $\mu_t - r_t$ satisfy

$$\sigma_t = \left( \sigma_{\delta} + \sigma_{1f}(x_t) \quad \sigma_{2f}(x_t) \right)$$

$$\mu_t - r_t = \sigma_t \theta_t = \left( \sigma_{\delta} + \sigma_{1f}(x_t) \quad \sigma_{2f}(x_t) \right) \left( \alpha \sigma_{\delta} + (1 - \nu) \sigma_{1f}(x_t) \ (1 - \nu) \sigma_{2f}(x_t) \right)^\top$$

$$= \alpha \sigma_{\delta}^2 + (1 - \nu + \alpha) \sigma_{\delta} \sigma_{1f}(x_t) + (1 - \nu) \left( \sigma_{1f}(x_t)^2 + \sigma_{2f}(x_t)^2 \right).$$  

Proposition 2 shows that the risk-free rate depends on the estimated fundamental $\hat{f}$ and on the diffusion of the price-dividend ratio as long as $\alpha \neq \psi$. When the risk aversion and the elasticity of intertemporal substitution are larger than one, $1 - \nu$ is positive. Hence, in this case, the second part of Eq. (10) is negative, making the level of the risk-free rate smaller than with the CRRA
utility. In addition, the larger the elasticity of substitution is, the smaller $\psi$ and $1 - \nu$ become, and thus the smaller is the volatility of the risk-free rate.

Equation (11) shows that the risk associated to time-varying fundamentals is priced. As long as $\alpha \neq \psi$, the market price of risk vector $\theta$ consists of two positive terms that depend on the diffusion of the price-dividend ratio. The first term loads on dividend surprises, and the second loads on news surprises. As risk aversion $\alpha$ increases or the elasticity of intertemporal substitution decreases, $1 - \nu$ rises, and hence the prices of risk and risk premia also rise.

2.5 Calibration to the U.S. Economy

Our theoretical model assumes that the investor observes two processes (the dividend stream $\delta$ and the flow of information $s$) and uses them to estimate the evolution of the unobservable variable $f$. This poses a challenge when one is trying to calibrate the model to observed data: in practice, the flow of information is not observable. To manage this difficulty, we follow David (2008) and use 1-quarter-ahead analyst forecasts of real U.S. GDP growth rates as a proxy for the estimated fundamental $\hat{f}$. To be consistent, we use the real U.S. GDP realized growth rate as a proxy for the dividend growth rate. Quarterly data from Q1:1969 to Q4:2012 are obtained from the Federal Reserve Bank of Philadelphia.

An immediate discretization of the stochastic differential equations defining the state variables $\zeta$, $\hat{f}$, $\phi$, and $\gamma$ would provide biased estimators, so we first solve this set of four stochastic differential equations. The solutions are provided in Appendix A.3. We then approximate the integrals pertaining to those solutions using a simple discretization scheme provided in Appendix A.4.

By observing the vectors $\log \frac{\delta_{t+\Delta}}{\delta_t}$ and $\hat{f}_t$ for $\Delta, \ldots, T\Delta$, we can directly infer the value of the Brownian vector $\epsilon^\delta_{t+\Delta} \equiv W^\delta_{t+\Delta} - W^\delta_t$. In addition, because the observed vector $\hat{f}_t$, $\Delta, \ldots, T\Delta$ depends on $\epsilon^\delta_{t+\Delta}$ and $\epsilon^s_{t+\Delta} \equiv W^s_{t+\Delta} - W^s_t$, we obtain a direct characterization of the signal vector $\epsilon^s_{t+\Delta}$ by substitution. This shows that observing $\delta$ and $\hat{f}$ instead of $\delta$ and $s$
provides a well-defined system.

### 2.5.1 Moment Conditions

Our model is calibrated on the two aforementioned time series using the Generalized Method of Moments (Hansen, 1982). The vector of parameters is defined by $\Theta = (\lambda, \tilde{f}, \omega, \Phi, \Lambda, \sigma_f, \sigma_\delta)^\top$. Consequently, we need at least seven moment conditions to infer the vector of parameters $\Theta$. For the sake of brevity, we relegate the moment conditions to Appendix A.4 and proceed here with their interpretation.

The analyst forecasts of the growth rate allow us to build moments that identify $\lambda$ and $\tilde{f}$; the conditional mean of $\hat{f}_{t+\Delta}$ and the unconditional autocovariance of $\hat{f}_t$ help to pin down the long-term mean parameter $\tilde{f}$ and the mean-reversion parameter $\lambda$. The unconditional variance of the observed time series defined by $\log \frac{\delta_{t+\Delta}}{\delta_t} - \hat{f}_t \Delta$ identifies the volatility parameter $\sigma_\delta$. Next, the realized growth rate permits us to construct recursively the performance index $\phi_t$. The unconditional autocovariance and variance of $\phi_t$ as well as its conditional variance are moments that identify the mean-reversion parameter $\omega$. Then, the conditional variance of $\hat{f}_{t+\Delta}$ and the unconditional mean and variance of $\Phi_t$ help us to estimate $\Phi$, $\Lambda$, and $\sigma_f$ (note that here we have to construct recursively $\gamma_t$ and $\Phi_t$). In total, we have 11 moment conditions that help us to estimate 7 parameters.

We match the unconditional variance of the model-implied attention $\Phi_t$ to the unconditional variance of a proxy of investors’ attention. To build this proxy, we follow Da et al. (2011) and use Google search volumes on groups of words with financial or economic content. To avoid any bias, none of the terms considered have positive or negative connotations. We scale this Google attention index between 0 and 1 (as our attention $\Phi$) and compute its unconditional variance, which should match the unconditional variance of the model-implied attention.

---

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Estimate</th>
<th>t-stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persistence growth rate $f$</td>
<td>$\lambda$</td>
<td>0.42</td>
<td>1.3</td>
<td>0.19</td>
</tr>
<tr>
<td>Mean growth rate $f$</td>
<td>$\bar{f}$</td>
<td>0.028***</td>
<td>5.6</td>
<td>0</td>
</tr>
<tr>
<td>Persistence performance index $\phi$</td>
<td>$\omega$</td>
<td>4.74***</td>
<td>188.7</td>
<td>0</td>
</tr>
<tr>
<td>Mean attention $\Phi$</td>
<td>$\bar{\Phi}$</td>
<td>0.368***</td>
<td>4.7</td>
<td>0</td>
</tr>
<tr>
<td>Sensitivity attention to $\phi$</td>
<td>$\Lambda$</td>
<td>286***</td>
<td>9.5</td>
<td>0</td>
</tr>
<tr>
<td>Volatility growth rate $f$</td>
<td>$\sigma_f$</td>
<td>0.029***</td>
<td>5.9</td>
<td>0</td>
</tr>
<tr>
<td>Volatility dividend growth</td>
<td>$\sigma_\delta$</td>
<td>0.014***</td>
<td>10.8</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1: Calibration to the U.S. Economy (GMM Estimation)
Parameter values resulting from a GMM estimation with 11 moment conditions. The Hansen J-test of over-identification cannot be rejected (Pr[Chi-sq.(4) > J] = 0.35). Statistical significance at the 10%, 5%, and 1% levels is labeled with \* / \*\* / \*\*\**, respectively.

2.5.2 Parameter Estimates

The values, t-stats, and p-values of the vector $\Theta$, resulting from the GMM estimation, are provided in Table 1 above. The test of over-identifying restrictions indicates that the model provides a good fit to the GDP realized growth and GDP growth forecast, with a $J$-test $p$-value of 0.35. The mean reversion speed $\lambda$ is the only non-significant parameter, and its estimate is relatively far from what the long-run risk literature assumes. Studies in this literature typically assume that the mean reversion parameter is between 0 and 0.25. In Bansal and Yaron (2004) the AR(1) parameter of the fundamental is worth 0.979 at monthly frequency (this would correspond to $\lambda = -12 \ln(0.979) = 0.25$), whereas Barsky and De Long (1993) assume that the fundamental process is integrated. Although our data set does not confirm the hypothesis of Barsky and De Long (1993) and Bansal and Yaron (2004), only the far future can potentially tell us if this hypothesis is sustainable. Indeed, 43 years of quarterly data are largely insufficient to estimate a parameter implying a half-life of at least 3 years.\(^{16}\)

We obtain a low volatility of real GDP realized growth rate, $\sigma_\delta$ (which is equal to the volatility of consumption in our model), and a low volatility of the fundamental, $\sigma_f$. Both parameters are

\(^{16}\)The half-life is a measure of the speed of mean-reversion, and it is given by $\ln(2)/\lambda$. For the Bansal and Yaron (2004) calibration, the half-life is roughly 33 months.
significant. The volatility of the real GDP realized growth rate is close to 1% and is in line with the estimation of Beeler and Campbell (2012) from postwar data.

We obtain a large and significant value for the parameter $\Lambda$, which suggests that investors’ attention reacts heavily to changes in the performance index. This is coupled with a high parameter $\omega$, which suggests that the performance index changes quickly based on recent information (i.e., investors use mostly data on the last year of dividend growth). According to Beeler and Campbell (2012), there is a significant value for the parameter $\Lambda$, which suggests that investors’ attention reacts heavily to changes in the performance index.

Put differently, investors’ attention is strongly sensitive to recent experience. Coming back to the solid black line in Fig. 1, which shows the probability density function of investors’ attention using the calibration from Table 1, we remark that attention varies substantially, taking values over the entire interval with significant probabilities.

The parameter $\Lambda$ is positive, which indicates that attention is high in bad times ($\phi < 0$) and low in good times ($\phi > 0$). Because the conditional volatility of forecasts is large in bad times, $\gamma$ and/or $\Phi$ need to be high (see Eq. 4). Our GMM estimation tells us that $\Phi$ (the attention) has to be high in bad times in order to satisfy the moment conditions, and this has the following interpretation: when the economy is in an expansionary phase, the output $\delta$ might decrease only with a low probability. Thus, investors do not have the incentive to exert a learning effort. Reciprocally, when the economy enters a recessionary phase, the high probability of a decrease in future consumption grabs investors’ attention and leads them to estimate as accurately as possible the change in the fundamental.

Counter-cyclical attention ($\Lambda > 0$) suggests that we should observe more accurate forecasts in bad times than in good times, but the evidence on this is mixed.

On the one hand, Patton and Timmermann (2008) find that forecasters estimated GDP growth

---

17A value of $\omega = 4.74$ for the exponential moving average means that, at a quarterly frequency, the investor applies a 69% weighting to the most recent output reading. Then the weights decrease as follows: 21%, 6%, 2%, and so on.

18When we split the data in NBER recessions and expansions, the conditional volatilities of forecasts are 1.98% and 1.43%, respectively.
quite well for the recessions in the early 1990s and 2001, but they underestimated the strong GDP
growth in the mid- to late 1990s and consistently overestimated the realized values of GDP growth
after the 2001 recession, which indicates that forecasts are better in bad times. We complement
these findings by using a larger data set (from 1969 to 2012), and we find further evidence
that analyst forecasts are more accurate in recessions than in expansions.\(^{19}\) Next, according to
Da, Gurun, and Warachka (2011), forecast errors are smaller when the past 12-month returns are
negative than when they are positive, suggesting that information gathered by analysts in downturns
is more accurate than information obtained in upturns.\(^{20}\) Finally, Garcia (2013) documents that
investors react strongly to news (good or bad) during recessions, whereas during expansions
investors’ sensitivity to information is much weaker, which is in line with the concept that investors
focus more on information during recessions.

On the other hand, Van Nieuwerburgh and Veldkamp (2006) find contrary evidence, namely
that agents’ forecast precision is lower in bad times than in good times. Additionally,
Karlsson, Loewenstein, and Seppi (2009) show that, when people are emotionally invested in
information, they monitor their portfolios more frequently in rising markets than in falling markets.
In other words, people “put their heads in the sand” when confronted with adverse news.\(^{21}\) There
have also been periods in which the economy did well and attention was high (e.g., the frenzied
media coverage of the stock market in the late 1990s). Ultimately, this mixed evidence suggests
that there is no clear answer as to whether investors really focus more on news in bad times. Our
estimation seems to favor this hypothesis, but further investigation is needed to draw definitive

\(^{19}\) Specifically, we perform Mincer and Zarnowitz (1969) regressions in the two subsamples (recessions and
expansions)—i.e., the realized GDP growth rate is regressed on a constant, and its corresponding analyst forecast
in both subsamples. The null hypothesis is that the intercept is equal to 0 and the slope is equal to 1—i.e., the
forecast error is conditionally unbiased. The test, which is available upon request, shows that the null cannot be
rejected in recession, whereas it is rejected in expansion.

\(^{20}\) This statement holds when investors acquire information continuously.

\(^{21}\) We note that in Karlsson et al. (2009) investors collect information about the value of their portfolios—attention
to wealth—whereas our investors collect information on the fundamental—attention to news. As such, these two
views do not exclude each other but simply suggest that attention to news and attention to wealth might be inversely
related.
conclusions.

Theoretical results seem to suggest that investors optimally allocate more attention to news during bad times. Kacperczyk et al. (2009) show that investors focus on aggregate news in recessions (for market timing) and idiosyncratic news in expansions (for stock picking). In Bansal and Shaliastovich (2011), investors acquire perfect information when the volatility of output growth or the uncertainty is sufficiently large (most likely during recessions). Furthermore, Hasler (2012) uncovers a decreasing relationship between current attention and past returns, providing a foundation for a positive parameter $\Lambda$. Our model simplifies the more complex optimal attention allocation problem by taking the aforementioned theoretical results as primitives.

2.5.3 Model-Implied Attention Dynamics

Our estimation generates a model-implied attention index. We compare this index with the one built from Google search volumes (sampled at quarterly frequency). The two indices are depicted in Fig. 2. The correlation between these two attention indices is worth 0.44 and is significant at the 99% confidence level. Movements in the implied attention seem to be well aligned with movements in the Google attention index. An ordinary least squares regression of the Google attention index on the implied attention index provides a significant slope coefficient of 0.95 and an adjusted R-squared of 0.17 (see Fig. 2). The joint Wald test of zero-intercept and unit-slope cannot be rejected at the 99% confidence level, which provides further support for the belief that we can extract from the data a valid proxy for investors’ attention to news.

---

22 We emphasize that in our estimation procedure we only use Google attention data to fit the unconditional variance of attention. All other moment conditions rely on GDP data.
Figure 2: Google attention index versus implied attention.

The solid black line depicts the attention index implied by our estimation, and the dashed red line depicts the weighted Google search index on financial and economic news (detrended). Indices are divided by their sample average to allow comparison with one another. The data set comprises 36 data points at quarterly frequency from Q1:2004 to Q4:2012. The table on the right presents the results obtained by regressing the Google attention index on the implied attention index. Newey-West standard errors are reported in brackets, and statistical significance at the 10%, 5%, and 1% levels is labeled with */**/*** respectively.

3 Attention, Uncertainty, and Volatility

In this section, we analyze the relationships among attention, uncertainty, and the variance of stock returns. Our model has two main predictions. The first is that the variance of stock returns increases with both attention and uncertainty, and the second is that this increasing relationship is quadratic. We perform an empirical investigation and find that, indeed, stock-return variance increases with attention and uncertainty. The data, however, provide mixed support for the quadratic relationship.

The variance of stock returns follows from Proposition 2:

\[ \|\sigma_t\|^2 = \sigma^2_{2t} + \sigma^2_{1t} = \left(\frac{\hat{I}}{I}\right)^2 \sigma^2_\eta \Phi^2_t + \left[\frac{\hat{I} \gamma_t}{I \sigma_\delta} + \sigma_\delta \left(1 + \frac{I_\phi}{I}\right)\right]^2. \] (12)

Equation (12) shows that stock-return variance is determined by attention, uncertainty, and investors’ price valuations.\(^{23}\) These price valuations are reflected in the price-dividend ratio \(I\) and

its partial derivatives with respect to $\hat{f}$ and $\phi$—i.e., $I_{\hat{f}}/I$ and $I_{\phi}/I$. Given the exponential affine conjecture for the price-dividend ratio provided in Eq. (9), $I_{\hat{f}}/I$ and $I_{\phi}/I$ are constants that are equal to $\beta_{11}$ and $\beta_{12}$, respectively. Furthermore, for a wide range of parameter values, numerical computations show that $I_{\phi}/I$ is rather small and thus deserves no further investigation.

The term $I_{\hat{f}}/I$ mainly depends on the elasticity of intertemporal substitution. If the elasticity of intertemporal substitution is larger than 1, $I_{\hat{f}}/I$ is positive. In this case, the asset price increases with the estimated fundamental. Conversely, if the elasticity of intertemporal substitution is smaller than 1 (e.g., in a CRRA setting with risk aversion larger than 1), $I_{\hat{f}}/I$ is negative. In this case, the asset price decreases with the estimated fundamental. Finally, if the elasticity of intertemporal substitution equals 1, the asset price does not depend on the fundamental. This feature is typically present when investors are log utility maximizers.

Having $I_{\hat{f}}/I$ and $I_{\phi}/I$ constant facilitates our discussion, for the variance of stock returns depends only on attention $\Phi$ and uncertainty $\gamma$. Since the elasticity of intertemporal substitution is larger than 1 in our calibration, stock-return variance increases strictly and quadratically with both attention (through the first term in Eq. 12) and uncertainty (through the second term in Eq. 12). This model-implied relationship is quantified in Fig. 3.

The effect of attention and uncertainty on stock returns can be interpreted in terms of weights assigned to information provided by news and dividend shocks, respectively. Indeed, an increase in attention means that the investor increases the weight assigned to information provided by news. First, this increases stock-return volatility by accelerating the transmission of news into prices. Second, it disconnects prices from dividend shocks and strengthens the connection between prices and news. The latter implication is supported by Garcia (2013), who shows that the predictability of stock returns using news content is concentrated in recessions (according to our calibration,

\[\text{Note that if we had considered an exponential quadratic form for the price-dividend ratio instead of an exponential affine form, then this statement would be altered because } I_{\hat{f}}/I \text{ and } I_{\phi}/I \text{ would become functions of the state variables. Nonetheless, we find in separate calculations that an exponential quadratic form yields exactly the same results.}\]
Figure 3: Attention, uncertainty, and volatility.

This figure depicts the model-implied relationship between attention, uncertainty, and volatility. Uncertainty $\gamma$ is scaled between 0 and 1 for convenience. Parameter values are presented in Table 1.

during times of high attention).

Although our model generates excess volatility—volatility of consumption is close to 1% in our calibration—the model does not match the average level of volatility observed in the data. In Section 3, we further extend the model to account for dividend leverage and obtain an average level of volatility much closer to its empirical counter-part. Furthermore, our model implies counter-cyclical volatility\footnote{See, e.g., Schwert (1989) and Mele (2008) for empirical evidence of counter-cyclical volatility.} because volatility increases with attention and, according to our estimation, attention is counter-cyclical. Several equilibrium models (e.g., Campbell and Cochrane, 1999; Veronesi, 1999; Bansal and Yaron, 2004) are able to explain both the level and the counter-cyclical nature of stock-return volatility. Therefore, we choose to focus on and test the novel prediction of our model, namely the positive and quadratic relationship between attention, uncertainty, and stock-return variance.

We now turn to the empirical evaluation of our main predictions regarding stock-return volatility. First, we test whether the variance of stock return increases with attention and uncertainty. Second, we test whether this increasing relationship, if it exists, is quadratic. With this aim, we use three
time series: (i) the variance of S&P500 returns, obtained through a GARCH(1,1) estimation; (ii) the Google attention index described in Section 2.5; and (iii) a measure of cross-sectional dispersion of 1-quarter-ahead real GDP growth forecasts obtained from the Federal Reserve Bank of Philadelphia. Under reasonable assumptions, the distribution of forecasts matches the distribution of beliefs (see Laster, Bennett, and Geoum, 1999), and thus it is common practice to use the cross-sectional dispersion of analyst forecasts as a proxy for uncertainty. The Google attention index contains data since 2004, while the cross-sectional dispersion of analyst forecasts is available at quarterly frequency. This results in a quarterly data set from Q1:2004 to Q4:2012.

Two concerns have to be addressed when performing this exercise. First, it is possible that investors become more attentive precisely because volatility increases, and not the other way around. To eliminate this concern, we control for lagged variance in our regressions.\textsuperscript{26} Note that adding lagged variance to the regressions also controls for the persistence in variance. Second, uncertainty depends heavily on attention—when investors pay more attention, they observe a more informative signal, which reduces the amount of uncertainty about fundamentals. In the context of Eq. (12), $\gamma_t$ depends on recent values of $\Phi$.\textsuperscript{27} To control for this effect, we include lagged attention and lagged uncertainty in the regressions. It is worth mentioning that adding these explanatory variables also controls for the persistence in attention and uncertainty.

The results are presented in Table 2. We perform four ordinary least squares regressions. We start by regressing stock return variance on attention and uncertainty (“linear” regressions—columns 1 and 2), then we add quadratic terms to analyze the non linear relationship (“quadratic” regressions—columns 3 and 4). Both the linear and quadratic regressions are conducted in two specifications. Columns 1 and 3 do not control for the relationship between attention and

\textsuperscript{26}As an additional test, we regress attention on lagged variance. The slope coefficient obtained is not significant, and the adjusted R-squared is 0.003. This provides evidence that lagged volatility does not seem to drive attention (at least at a quarterly frequency).

\textsuperscript{27}We thank an anonymous referee for drawing our attention to this. A regression of current uncertainty on lagged attention (up to three lags) and lagged uncertainty (one lag) provides significant coefficients, thus suggesting that attention is indeed a strong driver of future uncertainty.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.003</td>
<td>0</td>
<td>0.01**</td>
<td>0.011***</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.002)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>( \text{Att}_t )</td>
<td>0.026</td>
<td>0.059***</td>
<td>-0.044***</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.004)</td>
<td>(0.011)</td>
<td>(0.034)</td>
</tr>
<tr>
<td>( \text{Unc}_t )</td>
<td>0.021***</td>
<td>0.017**</td>
<td>0.012</td>
<td>-0.019</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.009)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>( \text{Att}_t^2 )</td>
<td></td>
<td></td>
<td></td>
<td>0.086***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.031)</td>
</tr>
<tr>
<td>( \text{Unc}_t^2 )</td>
<td>0.008</td>
<td>0.048*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.025)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{Att}_{t-1} )</td>
<td>-0.029***</td>
<td>-0.022**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.011)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{Att}_{t-2} )</td>
<td>-0.02***</td>
<td>-0.031**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.012)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{Att}_{t-3} )</td>
<td>-0.003</td>
<td>0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{Unc}_{t-1} )</td>
<td>0</td>
<td>0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{Var}_{t-1} )</td>
<td>0.594***</td>
<td>0.795***</td>
<td>0.548***</td>
<td>0.617***</td>
</tr>
<tr>
<td></td>
<td>(0.091)</td>
<td>(0.081)</td>
<td>(0.066)</td>
<td>(0.083)</td>
</tr>
</tbody>
</table>

| Adj. \( R^2 \) | 0.712 | 0.898 | 0.851 | 0.905 |
| N             | 35    | 33    | 35    | 33    |

Table 2: OLS Regressions of Stock-Return Variance on Attention and Uncertainty

The table provides the outputs of four regressions: (1) variance on attention and uncertainty, controlling for lagged variance; (2) variance on attention and uncertainty, controlling for lagged variance, lagged attention, and lagged uncertainty; (3) same as (1) but with quadratic terms added; and (4) same as (2) but with quadratic terms added. The data set comprises 36 data points at quarterly frequency from Q1:2004 to Q4:2012. The Newey-West standard errors are reported in brackets and statistical significance at the 10%, 5%, and 1% levels is labeled with */**/***,

respectively.

uncertainty, whereas columns 2 and 4 do. We control for the persistence in variance in each regression.

The loadings on attention and uncertainty are positive and significant in the full linear regression (column 2), and this lends support to our prediction that stock return variance increases with both attention and uncertainty. However, the non-linear relationship receives somewhat mixed support in the data. The loading on squared attention is positive and significant in specification (3) only, and the loading on squared uncertainty is positive and significant in specification (4) only.
In addition, the loading on attention itself is negative and significant in specification (3), which suggests that variance increases quadratically with attention only for sufficiently large values of attention.\footnote{Attention needs to be larger than 0.26 for the increasing quadratic relationship to apply. When attention is lower than 0.26, variance decreases quadratically with attention.}

Overall, the data confirm that the variance of stock returns increases with attention and uncertainty. The non-linear relationship presented in Eq. (12), however, finds only partial support since none of the regressions confirm that variance increases strictly and quadratically with both attention and uncertainty.

4 Attention, Uncertainty, and Risk Premium

Expectations matter for the risk premium for two reasons. First, they represent investors’ beliefs about the speed of economic growth. This is reflected in the estimated fundamental, $\hat{f}_t$, which enters price valuations. Second, if expectations are volatile, then investors require a higher risk premium. This is reflected in the volatility of the filtered fundamental, $\sigma(\hat{f}_t)$, which, in turn, is a mixture of attention and uncertainty (see Eq. 4).

It follows from Proposition 2 that the risk premium is written as

$$\mu_t - r_t = \alpha \sigma_\delta \left[ \frac{I_\gamma_t}{I} \sigma_\delta + \sigma_\delta \left( 1 + \frac{I_\phi}{I} \right) \right] + \left( 1 - \nu \right) \left[ \sigma_\delta \left( \frac{I_\gamma_t}{I} \sigma_\delta + \sigma_\delta \right) + \left( \frac{I_\gamma_t}{I} \sigma_\delta + \sigma_\delta \right)^2 + \left( \frac{I_\gamma_t}{I} \right)^2 \sigma_\delta^2 \Phi_t \right]. \tag{13}$$

In the CRRA case, the risk premium is summarized by the first term in Eq. (13) because $\nu = 1$. Furthermore, if risk aversion is larger than 1, then prices decrease with the filtered fundamental, and thus $I_\gamma/I < 0$. This implies that an increase in uncertainty decreases the risk premium (Veronesi, 2000). It is also worth noting that attention does not directly affect the risk premium in this case.
Figure 4: Attention, uncertainty, and risk premium.
This graph depicts the model-implied relationships among attention, uncertainty, and risk premium. Uncertainty $\gamma$ is scaled between 0 and 1 for convenience. Parameter values are presented in Table 1.

When the elasticity of intertemporal substitution and the coefficient of relative risk aversion are treated separately ($\nu \neq 1$), the second term in Eq. (13) becomes relevant and complements the uncertainty channel analyzed by Veronesi (2000). If both the coefficient of risk aversion and the elasticity of intertemporal substitution are larger than 1, then $I_d/I > 0$ and $(1 - \nu) > 0$. This in turn implies that both the first and the second terms in Eq. (13) contribute positively to the risk premium. More precisely, the model predicts, first, that the risk premium increases with both attention and uncertainty, and second, that this relationship is quadratic. The model-implied relationships among attention, uncertainty, and risk premium are quantified in Fig. 4.

The intuition as to why the risk premium depends positively on attention and uncertainty is the following: high uncertainty means that the investor has an inaccurate forecast of the future dividend growth rate, whereas high attention means fast incorporation of news into prices and consequently volatile returns. In order to bear these two risks, investors require a large risk premium.

Consistent with empirical findings in Campbell and Shiller (1988) and Fama and French (1989), our model of fluctuating attention generates counter-cyclical risk premia. The reason is that the risk premium increases with attention, and attention is counter-cyclical. Several models are able
to explain the counter-cyclical nature of risk premia (see, e.g., Campbell and Cochrane, 1999; Bansal and Yaron, 2004). We therefore choose to focus on and test the novel prediction of our model, namely the increasing and quadratic relationship between attention, uncertainty, and risk premia.

We test the two main predictions of the model regarding risk premia. The first and most fundamental prediction suggests that the risk premium increases with attention and uncertainty. The second and more precise prediction says that the risk premium increases quadratically with attention and uncertainty. Our empirical analysis considers two different measures of equity risk premia: (i) the adjusted fitted values obtained by performing a predictive regression of future returns on current dividend yields,29 and (ii) the quarterly survey responses reported in Graham and Harvey (2013). This survey was conducted over more than ten years, from June 2000 to December 2012. The survey data consist in a total of 17,500 survey responses, with an average of 352 individual responses each quarter. The empirical measures of attention and uncertainty are the same as those discussed in the previous section, resulting in a quarterly data set from Q1:2004 to Q4:2012.

As in Section 3, we perform a regression analysis and control for the persistence in the risk premium, for the relationship between attention and uncertainty, and for the persistence in attention and uncertainty. The results are shown in Table 3, which is divided in two parts. The left part shows the regression coefficients when the equity risk premium results from the predictive regression, while the right part uses Graham and Harvey’s survey responses.

Turning to the “linear” regressions (columns 1 – 2 and 5 – 6), the loading on attention is positive and significant in the full specifications (2) and (6), in line with the prediction of the model. The loading on uncertainty, however, is positive and significant in specification (2) only.

29 Regressing 1-quarter-ahead S&P500 returns on current dividend yields between Q2:1950 and Q4:2012 provides a positive and significant slope coefficient at the 95% confidence level. Our measure of risk premia is then obtained by subtracting the 3-month risk-free rates from the fitted values.
<table>
<thead>
<tr>
<th></th>
<th>RP from Predictive Regressions</th>
<th></th>
<th>RP from Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.009 (0.006)</td>
<td>0.001 (0.003)</td>
<td>0.003 (0.007)</td>
</tr>
<tr>
<td>Att&lt;sub&gt;t&lt;/sub&gt;</td>
<td>0.026 (0.019)</td>
<td>0.06*** (0.004)</td>
<td>-0.06*** (0.025)</td>
</tr>
<tr>
<td>Unc&lt;sub&gt;t&lt;/sub&gt;</td>
<td>0.014 (0.016)</td>
<td>0.027** (0.012)</td>
<td>0.028 (0.026)</td>
</tr>
<tr>
<td>Att&lt;sup&gt;2&lt;/sup&gt;&lt;sub&gt;t&lt;/sub&gt;</td>
<td>0.105*** (0.027)</td>
<td>-0.006 (0.027)</td>
<td></td>
</tr>
<tr>
<td>Unc&lt;sup&gt;2&lt;/sup&gt;&lt;sub&gt;t&lt;/sub&gt;</td>
<td>-0.008 (0.027)</td>
<td>0.09** (0.043)</td>
<td></td>
</tr>
<tr>
<td>Att&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>-0.005 (0.007)</td>
<td>-0.011 (0.009)</td>
<td></td>
</tr>
<tr>
<td>Att&lt;sub&gt;t-2&lt;/sub&gt;</td>
<td>-0.046*** (0.006)</td>
<td>-0.072*** (0.012)</td>
<td></td>
</tr>
<tr>
<td>Att&lt;sub&gt;t-3&lt;/sub&gt;</td>
<td>-0.016*** (0.005)</td>
<td>-0.002 (0.007)</td>
<td></td>
</tr>
<tr>
<td>Unc&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>-0.01 (0.009)</td>
<td>-0.011 (0.01)</td>
<td></td>
</tr>
<tr>
<td>RP&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>0.924*** (0.085)</td>
<td>0.925*** (0.055)</td>
<td>0.809*** (0.089)</td>
</tr>
<tr>
<td>Adj. R&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.874</td>
<td>0.945</td>
<td>0.905</td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>33</td>
<td>35</td>
</tr>
</tbody>
</table>

Table 3: OLS Regressions of Risk Premium on Attention and Uncertainty

The table is divided into two parts that differ in the measure of the risk premium (predictive regression or survey). Each part shows the results from four regressions: (1) risk premium on attention and uncertainty, controlling for lagged risk premium; (2) risk premium on attention and uncertainty, controlling for lagged risk premium, lagged attention, and lagged uncertainty; (3) same as (1) but with quadratic terms added; and (4) same as (2) but with quadratic terms added. The data set comprises 36 data points at quarterly frequency from Q1:2004 to Q4:2012. The Newey-West standard errors are reported in brackets, and statistical significance at the 10%, 5%, and 1% levels is labeled with */**/*** respectively.

These results show that the first prediction is better supported by the data when risk premia are measured via the predictive regression. The quadratic specifications show that the data are only weakly supportive of the quadratic relationships among attention, uncertainty, and the risk premium. Indeed, the loading on squared attention is positive and significant in specifications (3) and (7) only, and the loading on squared uncertainty is positive and significant in specification (4) only. Furthermore, the loading on attention itself is negative and significant in specifications (3)
and (7), implying an increasing quadratic relationship between attention and risk premia only for attention values larger than 0.32 and 0.39, respectively. Once all controls are present, attention loses all of its quadratic explanatory power, whereas risk premia increase strictly and quadratically with uncertainty in specification (4) only.

Overall, the data offer support for positive relationships among attention, uncertainty, and the risk premium. These results are in line with Massa and Simonov (2005) and Ozoguz (2009), who find that investors require a risk premium in order to be compensated for high uncertainty. We complement these studies by showing that, besides uncertainty, investors’ attention is also a priced risk factor.

5 Extensions

This section considers two extensions of our model. The first part describes a model in which attention depends on stock-return surprises. The second part describes a model that accounts for dividend leverage (Abel, 1999). Overall, these extensions help our model to better fit observed asset pricing moments while maintaining the positive relationship between attention/uncertainty and risk premium/volatility.

5.1 Attention Driven by Return Surprises

Assume that the performance index depends on past return surprises, as follows:

$$\phi_t = \int_0^t \phi_{\epsilon(t-u)} \left( \frac{dS_u + \delta_u du}{S_u} - \mu_u du \right).$$

(14)
Applying Itô’s lemma to Eq. (14) yields the following dynamics:

\[ d\phi_t = -\omega \phi dt + \sigma_t dW_t, \]

where \( \sigma \equiv \begin{pmatrix} \sigma_{1t} \\ \sigma_{2t} \end{pmatrix} \) is the stock-return diffusion vector.

The one-to-one mapping between attention and the performance index is provided in Definition 1 above. As before, we approximate the price-dividend ratio by the exponential form presented in Eq. (9). Applying Itô’s lemma to that equation yields the following functional form for the stock-return diffusion:

\[ \sigma_t \equiv \begin{pmatrix} \sigma_{1t} \\ \sigma_{2t} \end{pmatrix} = \begin{pmatrix} \sigma_\delta + \beta_{11} \gamma_t + \beta_{12} \sigma_{1t} \\ \beta_{11} \sigma_f \Phi_t + \beta_{12} \sigma_{2t} \end{pmatrix}. \]

Solving for \( \sigma_1 \) and \( \sigma_2 \) yields

\[ \sigma_t = \begin{pmatrix} \beta_{11} \gamma_t + \sigma_\delta^2 \\ \beta_{11} \sigma_f \Phi_t \end{pmatrix}, \]

\[ \begin{pmatrix} (1-\beta_{12}) \sigma_\delta \\ 1-\beta_{12} \end{pmatrix}. \]

Note that now the performance index features stochastic volatility, as returns do. Substituting the above expression in the dynamics of the performance index \( \phi \) and proceeding exactly as in Section 2.4 yields the price-dividend ratio. When attention depends on return surprises, the price-dividend ratio satisfies

\[ I(x) \approx e^{\beta_0 + \beta_{11} \tilde{f} + \beta_{12} \phi + \beta_{13} \gamma}, \]

where \( x \) is the vector of state variables and

\[
\beta_0 = 3.6156 \quad \beta_{11} = 1.1207 \quad \beta_{12} = -0.0006 \quad \beta_{13} = -11.3618. \quad (15)
\]
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>(i)</th>
<th>(ii)</th>
<th>(iii)</th>
<th>(iv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean price-dividend</td>
<td>E(I)</td>
<td>38.24</td>
<td>38.26</td>
<td>28.23</td>
<td>28.29</td>
</tr>
<tr>
<td>Volatility price-dividend</td>
<td>Vol(dI/I)</td>
<td>2.73%</td>
<td>2.74%</td>
<td>8.02%</td>
<td>8.06%</td>
</tr>
<tr>
<td>Mean risk premium</td>
<td>E(µ−r)</td>
<td>2.56%</td>
<td>2.46%</td>
<td>6.76%</td>
<td>6.72%</td>
</tr>
<tr>
<td>Volatility risk premium</td>
<td>Vol(d(µ−r))</td>
<td>1.27%</td>
<td>2.52%</td>
<td>3.72%</td>
<td>5.58%</td>
</tr>
<tr>
<td>Mean risk-free rate</td>
<td>E(r)</td>
<td>2.93%</td>
<td>2.96%</td>
<td>2.93%</td>
<td>2.96%</td>
</tr>
<tr>
<td>Volatility risk-free rate</td>
<td>Vol(d(µ−r))</td>
<td>1.77%</td>
<td>2.24%</td>
<td>1.77%</td>
<td>2.04%</td>
</tr>
<tr>
<td>Mean return volatility</td>
<td>E(</td>
<td></td>
<td>σ</td>
<td></td>
<td>)</td>
</tr>
<tr>
<td>Volatility of return volatility</td>
<td>Vol(d</td>
<td></td>
<td>σ</td>
<td></td>
<td>)</td>
</tr>
</tbody>
</table>

Table 4: Unconditional Moments from Simulations

Columns (i) and (ii) show the unconditional asset-pricing moments when attention depends on dividend surprises and return surprises, respectively. Dividend is equal to consumption in these two cases. In columns (iii) and (iv), attention depends on consumption surprises and dividend surprises, respectively. The dividend satisfies $D \equiv \delta \eta$, where $\eta = 2$ is the dividend leverage parameter. 20,000 simulations of daily data are performed over a 100-year period. First, averages are computed over each time series to provide 20,000 estimates. Then, the median of the 20,000 estimates is reported above. Percentages are annualized.

The loadings in the regular model are $\beta_0 = 3.6152$, $\beta_{11} = 1.1207$, $\beta_{12} = -0.0006$, and $\beta_{13} = -11.3741$. Comparing them with the loadings presented in (15) shows that the price-dividend ratio does not depend on whether attention is driven by dividend or return surprises. In addition, because $\beta_{12}$ is small, the diffusion of the price-dividend ratio is similar in both cases. Consequently, the levels of the price-dividend ratio, risk-free rate, stock-return volatility, market prices of risks, and risk premia are very close to one another in both cases.

Table 4 presents the unconditional asset pricing moments resulting from 20,000 simulations of daily data considered over a period of 100 years. Columns (i) and (ii) confirm that the average levels of the price-dividend, risk-free rate, stock-return volatility, and risk premia are basically the same for both attention specifications.

Since returns are more volatile than dividend growth rates, the return-driven performance index is more volatile than the dividend-driven performance index. This implies that the return-driven attention is more volatile than the dividend-driven attention. Consequently, the variables that
depend significantly on attention inherit a larger unconditional volatility in the return-driven attention case than in the dividend-driven attention case. As shown in Proposition 2 and Eqs. (12) and (13), these variables are the market prices of risk, the risk-free rate, the stock-return volatility, and the risk premium. Columns (i) and (ii) of Table 4 show that, indeed, the volatility of risk premia doubles when we move from the dividend-driven attention case to the return-driven attention case.

A leading explanation of the high volatility of risk premia is offered by Chien et al. (2012), who show that the presence of traders who re-balance their portfolios infrequently amplifies the effects of aggregate shocks on the time variation in risk premia by a factor of 3. Our potential explanation resides in the fact that investors become more attentive to news when return surprises (as opposed to dividend surprises) decline, and vice versa. Therefore, the volatility of risk premia increases by a factor of 2, the volatility of volatility increases by a factor of 1.65, but the volatility of the risk-free rate is not excessively amplified (it increases by a factor of only 1.27).

5.2 Dividend Leverage

Another extension of our model consists of pricing a dividend stream that accounts for dividend leverage. Abel (1999) introduces dividend leverage (levered equity) in an equilibrium model and obtains a low variability of the riskless rate along with a large equity premium. In the context of our model, we introduce levered equity by defining the dividend paid by the stock to be $D \equiv \delta^\eta$. The parameter $\eta$ thus provides a convenient way to model dividend leverage. Consistent with the estimations performed in Bansal et al. (2005), we set the dividend leverage parameter to $\eta = 2$.

We solve for the price-dividend ratio in two different cases. In the first case, attention depends on consumption surprises. In the second case, attention depends on dividend surprises. Note that the only difference between these two cases resides in the fact that the volatility of dividend surprises is equal to $\eta$ times the volatility of consumption surprises.
The price-dividend ratio satisfies

\[ I^i(x) \approx e^{\beta_0^i + \beta_{11}^i \hat{r} + \beta_{12}^i \phi + \beta_{13}^i \gamma}, \]  

(16)

where \( x \) is the vector of state variables and \( i \in \{\delta, D\} \) stands for the case in which attention depends on surprises in state \( i \). The loadings appearing in (16) are obtained by linearizing and solving the following equation:

\[
 r^i dt = \mathbb{E}^Q_t \left( \frac{dS^i + D dt}{S^i} \right) 
= \mathbb{E}^Q_t \left( \frac{dI^i}{I^i} + \frac{dD}{D} + \frac{dI^i}{I^i} \frac{dD}{D} \right) + \frac{1}{I^i} dt,
\]

where \( Q \) is the risk-neutral measure, and \( S^i \) and \( r^i \) are the equilibrium stock price and risk-free rate obtained when attention is driven by surprises in state \( i \in \{\delta, D\} \), respectively.

The values of the loadings are as follows:

\[
\beta_0^\delta = 3.2595 \quad \beta_0^D = 3.2602 \\
\beta_{11}^\delta = 3.2924 \quad \beta_{11}^D = 3.2926 \\
\beta_{12}^\delta = -0.0030 \quad \beta_{12}^D = -0.0028 \\
\beta_{13}^\delta = -53.0946 \quad \beta_{13}^D = -53.0804. \tag{17}
\]

The loadings in the regular model are \( \beta_0 = 3.6152 \), \( \beta_{11} = 1.1207 \), \( \beta_{12} = -0.0006 \), and \( \beta_{13} = -11.3741 \). Comparing them with the loading presented in (17) shows that leverage significantly increases the sensitivity of the price-dividend ratio to changes in the state variables. Indeed, leverage increases the sensitivity to the estimated fundamental by a factor 3, and the sensitivity to attention and uncertainty by a factor 5. In addition, the loadings provided in (17) show that
the level of the price-dividend ratio in both the consumption-driven and dividend-driven attention cases is sensibly the same.

Columns (iii) and (iv) of Table 4 provide the asset-pricing moments obtained in the consumption-driven and dividend-driven attention cases with leverage, respectively. Since leverage significantly increases the sensitivity of the price-dividend ratio to state variables, the average level of stock-return volatility, price-dividend-ratio volatility, and risk premium are significantly magnified. Leverage increases these quantities by a factor of roughly 3, which helps our model to better fit observed levels of volatilities and risk premia. Note also that risk-premium volatility, risk-free rate volatility, and volatility of return volatility are larger in the dividend-driven attention case than in the consumption-driven attention case. The reason is that the volatility of dividend surprises is twice as large as the volatility of consumption surprises.

6 Concluding Remarks

We have developed an asset-pricing model in which investors’ attention and learning uncertainty affect simultaneously the dynamics of asset returns. The model predicts that volatility and the risk premium increase with attention and uncertainty. These predictions are supported by our empirical analysis. Also, the theoretical relationships among attention, uncertainty, and the volatility/risk premium are quadratic. This prediction finds mixed support in the data.

We hope that this paper represents a useful step forward in the important task of understanding the effect of learning on asset prices. Several extensions of the present model deserve to be addressed in future research. We summarize here a few of them.

The model can be adapted to include the dispersion of beliefs. Massa and Simonov (2005) show that the dispersion of beliefs is a priced factor, whereas Carlin, Longstaff, and Matoba (2014) find that disagreement is time-varying and is associated with higher expected returns, higher return
volatility, higher uncertainty, and larger trading volume. It is therefore of interest to study the impact of fluctuating attention on disagreement and ultimately on asset prices. Our conjecture is that if investors learn from different sources of information, or if they have different priors, spikes in investors’ attention might contribute to polarization of beliefs, which could further amplify the volatility and the risk premium in the market.

Another possible extension is concerned with overconfidence: higher attention to news might exacerbate the subjective confidence in investors’ judgments, further increasing the volatility of asset prices, as in Dumas et al. (2009). It would also be worthwhile to disentangle attention to news and attention to wealth: evidence (Karlsson et al., 2009) suggests that retail investors temporarily ignore their portfolios in bad times, and thus liquidity dries up when the market suffers downturns—examples are the Asian crisis of 1997, the Russian debt default in 1998, and the “credit crunch” of 2008. This effect can be reinforced by fluctuations in investors’ attention to news. Finally, our model assumes that a single factor (the performance index) drives investors’ attention. Other exogenous events unrelated to the economy can temporarily grab investors’ attention, generating further fluctuations in volatility and trading volume: see, for example, the anecdote about Tiger Woods and the New York Stock Exchange volume, in the Presidential Address of Duffie (2010). Trading on all U.S. stock exchanges declined during Wood’s live televised speech and shot up once the speech ended, suggesting that investors do not pay constant attention to financial news.
References


[11, 47]

[1, 5]


A Appendix

A.1 Unconditional moments of the performance index $\phi$ and probability density function of the attention $\Phi$

Consider

$$Y_t = \begin{bmatrix} f_t \\ \phi_t \end{bmatrix}, \quad dY_t = (A - BY_t) dt + C \begin{bmatrix} dZ^f_t \\ dZ^\delta_t \end{bmatrix}$$

with

$$B = \begin{bmatrix} \lambda & 0 \\ 0 & \omega \end{bmatrix}$$

and

$$C = \begin{bmatrix} \sigma_f & 0 \\ 0 & \sigma_\delta \end{bmatrix}.$$ 

The solution is found by applying Itô’s lemma to

$$F_t = e^{Bt} Y_t = \begin{bmatrix} e^{\lambda t} f_t \\ e^{\omega t} \phi_t \end{bmatrix}.$$ 

After integrating from 0 to $t$ we obtain

$$F_t - F_0 = \left[ \int_0^t \lambda \bar{f} e^{\lambda u} du + \int_0^t \sigma_f e^{\lambda u} dZ^f_u \right].$$

Thus, the first moments of $f$ and $\phi$ solve the following system of equations:

$$\begin{cases} 
  e^{\lambda t} \mathbb{E}[f_t] - f_0 = \bar{f} (e^{\lambda t} - 1) \\
  e^{\omega t} \mathbb{E}[\phi_t] - \phi_0 = 0.
\end{cases}$$

It follows that the long-term mean of $f$ is $\bar{f}$ and the long-term mean of $\phi$ is 0. The variance of $f$ is found with the standard formula

$$\text{Var}[f_t] = \mathbb{E}[(f_t - \mathbb{E}[f_t]) (f_t - \mathbb{E}[f_t])]
= \mathbb{E} \left[ (\int_0^t \sigma_f e^{\lambda u} dZ^f_u)^2 \right]
= \frac{\sigma_f^2 (1 - e^{-2\lambda t})}{2\lambda}.$$ 

The long-term variance of $f$ is then $\frac{\sigma_f^2}{2\lambda}$. Similarly, the long-term variance of $\phi$ is $\frac{\sigma_\delta^2}{2\omega}$. The density function of attention $\Phi$ is written as

$$f_{\Phi}(\Phi_t) = \frac{1}{g'(g^{-1}(\Phi_t))} f_{\phi} \left( g^{-1}(\Phi_t) \right) = \frac{\exp \left( -\frac{\omega \log^2 \left( \frac{\Phi_t - 1}{\phi^{-1}(\Phi_t)} \right)}{(\Lambda^2 \sigma_\delta^2)} \right)}{\sqrt{\pi} (\Lambda \Phi_t - \Lambda \Phi_t^2)^{\frac{\sigma_\delta^2}{2\omega}}}.$$
A.2 Details on $\zeta$, $\hat{f}$, $\phi$, and $\gamma$

We have

$$df_t = \left(\lambda \hat{f} + (-\lambda) f_t\right) dt + \sigma f_t dZ_t^f + \left[ \begin{array}{c} 0 \\ 0 \end{array} \right] \left[ \begin{array}{c} dZ_t^\delta \\ dZ_t^s \end{array} \right]$$

or (as in Liptser and Shiryaev (2001))

$$df_t = [a_0(t,\vartheta) + a_1(t,\vartheta) f_t] dt + b_1(t,\vartheta) dZ_t^f + b_2(t,\vartheta) \left[ \begin{array}{c} dZ_t^\delta \\ dZ_t^s \end{array} \right].$$

In addition, the observable process is given by

$$d\vartheta_t = \left(\frac{1}{2} \sigma^2 \vartheta\right) dt + \left[ \begin{array}{c} 0 \\ \Phi \end{array} \right] dZ_t^f + \left[ \begin{array}{c} \sigma^\delta \\ 0 \sqrt{1 - \Phi^2} \end{array} \right] \left[ \begin{array}{c} dZ_t^\delta \\ dZ_t^s \end{array} \right],$$

or

$$d\vartheta_t = [A_0(t,\vartheta) + A_1(t,\vartheta) f_t] dt + B_1(t,\vartheta) dZ_t^f + B_2(t,\vartheta) \left[ \begin{array}{c} dZ_t^\delta \\ dZ_t^s \end{array} \right].$$

Using Liptser and Shiryaev’s notations, we get

$$b \circ b = b_1 b'_1 + b_2 b'_2 = \sigma_f^2$$

$$B \circ B = B_1 B'_1 + B_2 B'_2 = \left[ \begin{array}{cc} \sigma_\delta^2 & 0 \\ 0 & 1 \end{array} \right]$$

$$b \circ B = \left[ \begin{array}{c} 0 \\ \sigma_f \Phi \end{array} \right].$$

Then, Theorem 12.7 (Liptser and Shiryaev, 2001) shows that the filter evolves according to

$$d\hat{f}_t = [a_0 + a_1 \hat{f}_t] dt + [(b \circ B) + \gamma_t A'_1] (B \circ B)^{-1} \left[ \begin{array}{c} d\vartheta_t - \left( A_0 + A_1 \hat{f}_t \right) dt \end{array} \right]$$

$$\dot{\gamma}_t = a_1 \gamma_t + \gamma a_1' + (b \circ b) + [(b \circ B) + \gamma_t A'_1] (B \circ B)^{-1} [(b \circ B) + \gamma_t A'_1],$$

where $\gamma$ represents the posterior variance. Notice that the dynamics of $\gamma$ depend on $\phi$ through the term $b \circ B$. Consequently, we cannot follow Scheinkman and Xiong (2003) and solve for the steady state. We have no other choice than to include the posterior variance $\gamma$ in the state space.

A.3 Solutions for $\zeta$, $\hat{f}$, $\phi$, and $\gamma$

Since the dividend process $\delta$ is a geometric Brownian motion, its solution is immediately given by

$$\delta_t = \delta_v e^{\int_v^t f_u du - \frac{1}{2} \sigma^2 (t-v) + \sigma_\delta (W_t^\delta - W_v^\delta)}, \quad t \geq v.$$

In order to solve for $\hat{f}$ and $\phi$, we have to notice that the vector defined by

$$Y_t = \left[ \begin{array}{c} \hat{f}_t \\ \phi_t \end{array} \right], \quad dY_t = (A - BY_t) dt + C \left[ \begin{array}{c} dW_t^\delta \\ dW_t^s \end{array} \right]$$

with

$$B = \left[ \begin{array}{cc} \lambda & 0 \\ 0 & \omega \end{array} \right]$$

47
is a bivariate Ornstein-Uhlenbeck process. The solution is found by applying Itô’s lemma to

\[ F_t = e^{Bt} = \begin{bmatrix} e^{\lambda t} \hat{f}_t \\ e^{\omega t} \hat{\phi}_t \end{bmatrix}. \]

The dynamics of \( F \) obey

\[
dF_t = \begin{bmatrix} e^{t\lambda} (\sigma_f \sigma_\delta dW_t^\delta + \gamma_t dW_t^\delta + dt \lambda \bar{\sigma}_f) \\ e^{t\omega} \bar{\sigma}_\delta dW_t^\delta \end{bmatrix}.
\]

After integrating from \( v \) to \( t \) and rearranging we obtain

\[
\hat{f}_t = e^{-\lambda (t-v)} \hat{f}_v + \hat{f} \left( 1 - e^{-\lambda (t-v)} \right) + \frac{1}{\sigma_\delta} \int_v^t e^{-\lambda (t-u)} \gamma_u dW_u^\delta + \sigma_f \int_v^t e^{-\lambda (t-u)} \Phi_u dW_u^s
\]

\[
\hat{\phi}_t = e^{-\omega (t-v)} \hat{\phi}_v + \int_v^t \sigma_\delta e^{-\omega (t-u)} dW_u^\delta.
\]

The dynamics of the posterior variance \( \gamma \) can be rewritten as

\[
\frac{\partial}{\partial t} \begin{bmatrix} G_t & F_t \end{bmatrix} = \begin{bmatrix} G_t & F_t \end{bmatrix} \begin{bmatrix} -2\lambda \\ \sigma_f^2 (1 - \Phi_t^2) \end{bmatrix},
\]

where \( \gamma_t = \frac{\gamma}{\bar{\lambda}} \). The solution is obtained through exponentiation and is given by

\[
\gamma_t = \frac{\sigma_\delta}{\sigma_\delta (i_{v,t} - \Delta \gamma_v)} \sinh \left( \sqrt{\Delta} \sqrt{i_{v,t} + \Delta \lambda^2} \sigma_\delta^2 \right) + \sqrt{\Delta} \gamma_v \sqrt{i_{v,t} + \Delta \lambda^2} \sigma_\delta^2 \cosh \left( \sqrt{\Delta} \sqrt{i_{v,t} + \Delta \lambda^2} \sigma_\delta^2 \right),
\]

where

\[
\Delta = t - v
\]

\[
i_{v,t} = \sigma_f^2 \int_v^t (1 - \Phi_u^2) du.
\]

### A.4 Moment Conditions

Note that \( \gamma \) depends on the attention \( \Phi \), \( \Phi \) depends on the dividend performance \( \phi \), and \( \phi \) is driven by surprises in the dividend growth. Hence, the posterior variance \( \gamma_t \), the attention \( \Phi_t \), and the dividend performance index \( \phi_t \), for \( t = 0, \Delta, \ldots, T\Delta \), have to be constructed recursively. These (implicit) time series are used for some of the moment conditions that follow.

We approximate the continuous-time processes by using the following simple discretization scheme

\[
\int_{t_1}^{t_2} \kappa_{1,u} du \approx \kappa_{1,t_1} \Delta
\]

\[
\int_{t_1}^{t_2} \kappa_{2,u} dW_u \approx \kappa_{2,t_1} \epsilon_{t_1 + \Delta},
\]

where \( \kappa_1 \) and \( \kappa_2 \) are some arbitrary processes, \( \Delta = t_2 - t_1 = \frac{1}{4} \), and \( \epsilon_{t_1 + \Delta} \sim N(0, \Delta) \).
A.4.1 Conditional mean of $\hat{f}_{t+\Delta}$

We have

$$
\hat{f}_{t+\Delta} = e^{-\lambda \Delta} \hat{f}_t + \bar{f} \left(1 - e^{-\lambda \Delta}\right) + \frac{1}{\sigma_{\delta}} \int_t^{t+\Delta} e^{-\lambda(t+\Delta-u)} \gamma_u dW_u^\delta + \sigma_f \int_t^{t+\Delta} e^{-\lambda(t+\Delta-u)} \Phi_u dW_u^s.
$$

(18)

The following moment condition must hold

$$
\mathbb{E}_t \left[ \hat{f}_{t+\Delta} \right] = e^{-\lambda \Delta} \hat{f}_t + \bar{f} \left(1 - e^{-\lambda \Delta}\right).
$$

The empirical counterpart is

$$
0 = \frac{1}{T} \sum_{i=1}^T \left[ \hat{f}_{i,\Delta} - e^{-\lambda \Delta} \hat{f}_{(i-1)\Delta} - \bar{f} \left(1 - e^{-\lambda \Delta}\right) \right].
$$

The Federal Reserve Bank of Philadelphia provides not only the 1-quarter-ahead forecast (which in our case is denoted by $\hat{f}_t$), but also 2-quarter-ahead and 3-quarter ahead forecasts—i.e., $g_{1,t} \equiv \mathbb{E}_t[\hat{f}_{t+\Delta}]$ and $g_{2,t} \equiv \mathbb{E}_t[\hat{f}_{t+2\Delta}]$. This establishes two additional moment conditions, as follows

$$
g_{1,t} = e^{-\lambda \Delta} \hat{f}_t + \bar{f} \left(1 - e^{-\lambda \Delta}\right)
$$

and

$$
g_{2,t} = e^{-\lambda \Delta} g_{1,t} + \bar{f} \left(1 - e^{-\lambda \Delta}\right).
$$

Their empirical counterparts are

$$
0 = \frac{1}{T} \sum_{i=1}^T \left[ g_{1,i,\Delta} - e^{-\lambda \Delta} \hat{f}_{i,\Delta} - \bar{f} \left(1 - e^{-\lambda \Delta}\right) \right]
$$

and

$$
0 = \frac{1}{T} \sum_{i=1}^T \left[ g_{2,i,\Delta} - e^{-\lambda \Delta} g_{1,i,\Delta} - \bar{f} \left(1 - e^{-\lambda \Delta}\right) \right].
$$

A.4.2 Unconditional autocovariance of $\hat{f}_t$

We have

$$
\text{Cov} \left( \hat{f}_{t+\Delta}, \hat{f}_t \right) = e^{-\lambda \Delta} \text{Var} \left( \hat{f}_t \right).
$$

The empirical counterpart is

$$
0 = \frac{1}{T} \sum_{i=1}^T \left[ \left( \hat{f}_{i,\Delta} - \mu_{\hat{f}_{1:T}} \right) \left( \hat{f}_{(i-1)\Delta} - \mu_{\hat{f}_{0:T}} \right) - e^{-\lambda \Delta} \left( \hat{f}_{(i-1)\Delta} - \mu_{\hat{f}_{0:T-1}} \right)^2 \right],
$$

where $\mu(\cdot)$ represents the sample average.
A.4.3 Unconditional autocovariance of $\phi_t$

We have

$$\text{Cov}(\phi_{t+\Delta}, \phi_t) = e^{-\omega \Delta} \Var(\phi_t).$$

The empirical counterpart is

$$0 = \frac{1}{T} \sum_{i=1}^{T} \left[ (\phi_{i\Delta} - \mu_{\phi,1:T}) (\phi_{(i-1)\Delta} - \mu_{\phi,0:T-1}) - e^{-\omega \Delta} (\phi_{(i-1)\Delta} - \mu_{\phi,0:T-1})^2 \right].$$

A.4.4 Unconditional variance of $\phi_t$

We have

$$\Var(\phi_t) = \frac{\sigma_\delta^2}{2\omega}.$$

The empirical counterpart is

$$0 = \frac{1}{T} \sum_{i=1}^{T} \left[ (\phi_{i\Delta} - \mu_{\phi,1:T})^2 \right] - \frac{\sigma_\delta^2}{2\omega}.$$

A.4.5 Conditional variance of $\hat{f}_{t+\Delta}$

Take the first diffusion term in Eq. (18):

$$\Var_t \left( \frac{1}{\sigma_\delta} \int_t^{t+\Delta} e^{-\lambda(t+\Delta-u)} \gamma_u dW_u^\delta \right) = \frac{1}{\sigma_\delta^2} e^{-2\lambda(t+\Delta)} \mathbb{E}_t \left[ \left( \int_t^{t+\Delta} e^{\lambda u} \gamma_u dW_u^\delta \right)^2 \right]$$

$$= \frac{1}{\sigma_\delta^2} e^{-2\lambda(t+\Delta)} \mathbb{E}_t \left[ \int_t^{t+\Delta} e^{2\lambda u} \gamma_u^2 du \right]$$

$$\approx \frac{1}{2\lambda \sigma_\delta^2} \gamma_t^2 \left( 1 - e^{-2\lambda \Delta} \right).$$

The second equality in Eq. (19) results from Itô isometry, whereas the third equality comes from the approximation $\gamma_u \approx \gamma_t$. The variance of the second diffusion term in Eq. (18) is obtained similarly. The conditional variance of $\hat{f}_{t+\Delta}$ is then

$$\Var_t (\hat{f}_{t+\Delta}) = \left( \frac{\gamma_t^2}{\sigma_\delta^2} + \sigma_f^2 \Phi_t^2 \right) \frac{1 - e^{-2\lambda \Delta}}{2\lambda},$$

which represents a moment condition. Its empirical counterpart is

$$0 = \frac{1}{T} \sum_{i=1}^{T} \left[ (\hat{f}_{i\Delta} - e^{-\lambda \Delta} \hat{f}_{(i-1)\Delta} - \bar{f} (1 - e^{-\lambda \Delta}))^2 - \left( \frac{\gamma_{(i-1)\Delta}}{\sigma_\delta^2} + \sigma_f^2 \Phi_{(i-1)\Delta} \right) \frac{1 - e^{-2\lambda \Delta}}{2\lambda} \right].$$

A.4.6 Conditional variance of $\phi_{t+\Delta}$

We know that

$$\phi_{t+\Delta} = \phi_t e^{-\omega \Delta} + \sigma_\delta \int_t^{t+\Delta} e^{-\omega(t+\Delta-u)} dW_u^\delta.$$
Thus,

\[
\Var_t (\phi_{t+\Delta}) = \sigma^2 \delta e^{-2\omega(t+\Delta)} \mathbb{E}_t \left[ \left( \int_t^{t+\Delta} e^{\omega u} dW_u \right)^2 \right] \\
= \sigma^2 \delta e^{-2\omega(t+\Delta)} \mathbb{E}_t \left[ \int_t^{t+\Delta} e^{2\omega u} du \right] \\
= \frac{\sigma^2}{2\omega} \left( 1 - e^{-2\omega\Delta} \right). 
\]

(20)

The empirical counterpart is

\[
0 = \frac{1}{T} \sum_{i=1}^{T} \left( \phi_{i\Delta} - e^{-\omega\Delta} \phi_{(i-1)\Delta} \right)^2 - \frac{\sigma^2}{2\omega} \left( 1 - e^{-2\omega\Delta} \right). 
\]

The second equality in Eq. (20) results from Ito isometry.

A.4.7 Unconditional mean of \( \Phi_t \)

The process \( \Phi \) has a long-term mean that can be approximated by \( \bar{\Phi} \). Therefore, one can write

\[
\mathbb{E} [\Phi_t] \approx \bar{\Phi}. 
\]

The empirical counterpart is

\[
0 = \frac{1}{T} \sum_{i=1}^{T} \left( \phi_{i\Delta} - \bar{\Phi} \right)^2. 
\]

A.4.8 Unconditional variance of \( \Phi_t \)

The unconditional variance of attention implied by our model should match the unconditional variance of the Google attention index (which has been adjusted to take values between 0 and 1). That is,

\[
\Var [\Phi_t] = 0.054. 
\]

The empirical counterpart is

\[
0 = \frac{1}{T} \sum_{i=1}^{T} \left( (\phi_{i\Delta} - \mu_{\Phi,1:T})^2 - 0.054 \right). 
\]

A.4.9 Unconditional variance of \( \frac{d\delta_t}{\delta_t} - \hat{f}_t dt \)

We define the observable process \( d \) as follows

\[
d_{t+\Delta} \equiv \log \frac{\delta_{t+\Delta}}{\delta_t} - \hat{f}_t \Delta = -\frac{1}{2} \sigma^2 \delta \Delta + \sigma \epsilon^\delta_{t+\Delta}. 
\]

We can thus write

\[
\Var (d_{t+\Delta}) = \sigma^2 \delta \Delta. 
\]
The empirical counterpart is

\[
0 = \frac{1}{T} \sum_{i=1}^{T} \left[ (d_i \Delta - \mu_{d,1:T})^2 - \sigma_{\delta}^2 \Delta \right].
\]

To summarize, there are 11 boxed equations defining a system of 11 moment conditions used to estimate the 7-dimensional vector of parameters \( \Theta = (\lambda, \bar{f}, \omega, \bar{\Phi}, \Lambda, \sigma_f, \sigma_{\delta})^\top \).