

# Investor Attention and Stock Market Volatility\*

Daniel Andrei<sup>†</sup>

Michael Hasler<sup>‡</sup>

June 3, 2013

## Abstract

We offer a framework to simultaneously study the role played by investors' attention to news and learning uncertainty in the determination of asset prices. We show that asset return volatility and risk premia increase quadratically with both attention and uncertainty. Our empirical investigation lends support to these theoretical predictions. Moreover, learning yields a lead-lag relation between attention and uncertainty; this relation is found to enable “panic states,” featuring spikes in volatilities and risk premia. During these panic states asset prices are very sensitive to news, consistent with recent empirical findings.

**Keywords:** Asset Pricing, General Equilibrium, Learning, Attention, Uncertainty, Volatility, and Risk Premia

---

\*We would like to thank Bernard Dumas, Julien Hugonnier, an anonymous referee, and the editor 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, and UCLA Anderson. Financial support from the Swiss Finance Institute and from NCCR FINRISK of the Swiss National Science Foundation is gratefully acknowledged.

<sup>†</sup>UCLA, Anderson School of Management, 110 Westwood Plaza, Suite C4.20, Los Angeles, CA 90095, USA, [daniel.andrei@anderson.ucla.edu](mailto:daniel.andrei@anderson.ucla.edu), [www.danielandrei.net](http://www.danielandrei.net)

<sup>‡</sup>SFI, EPFL, Extranef 129, Quartier UNIL-Dorigny, CH-1015, Lausanne, Switzerland, [michael.hasler@epfl.ch](mailto:michael.hasler@epfl.ch), [people.epfl.ch/michael.hasler](http://people.epfl.ch/michael.hasler)

# 1 Introduction

Investors' processing of information, or learning, is a broad generic term that is difficult to define with specificity. It comprises several dimensions forming a coherent whole, with each dimension raising separate important questions. How much *attention* do investors pay to information? How much *learning uncertainty* is generated through investors' processing of information? Does the process of learning yield *dispersion of beliefs* across investors? Are there any *systematic biases* in investors' cognitive processing of information?

Whether together or separate, these questions have received great attention in the literature. In this paper we focus on two of them, namely, how much attention is allocated to learning, and how much uncertainty results from learning. Our motivation to work on these two factors, i.e., attention and uncertainty, stems from two reasons. First, at an empirical level, there is agreement that attention and uncertainty impact asset returns. Convincing empirical work is conducted on two separate battlefronts. On one side, investor attention, measured from Google search volumes, is strongly time-varying and higher in periods of high volatility (Da, Engelberg, and Gao, 2011; Vlastakis and Markellos, 2012).<sup>1</sup> On the other side, uncertainty is a priced risk factor and is associated to high volatility (Massa and Simonov 2005, Ozoguz 2009).<sup>2</sup> Yet, there is neither empirical nor theoretical work that assesses the relative importance of these two factors for risk premium and stock market volatility. Here we tackle this task.

The second reason we study attention and uncertainty simultaneously is because, in fact, they are closely related. If investors pay a lot of attention to learn about the fundamental structure of the economy, then uncertainty is expected to decrease. In the opposite case uncertainty is expected to increase. The overall consequence of attention and uncertainty is obscured by this intuitive synchronicity. Most theoretical models lack a clear-cut separation between attention and uncertainty. Here we deal with this important theoretical challenge.

We build the simplest setting in which the effects of attention and uncertainty emerge. We show precisely how attention and uncertainty drive risk premia and volatilities and we confirm empirically our findings. Although our model is greatly simplified (single and rational agent, Bayesian learning, normally distributed variables), it has a rich set of testable implications. This simplicity also creates many possibilities for extensions to richer models.

The following example, adapted from a standard setup in the literature of learning in finance, conveys our intuition. An investor tries to learn the fundamental structure of

---

<sup>1</sup>Other papers using Google search volumes on companies names, tickers, or other economic terms to gauge investors' attention are Dimpfl and Jank (2011), and Kita and Wang (2012).

<sup>2</sup>For a comprehensive review of the literature of learning and uncertainty in financial markets see Pastor and Veronesi (2009).

the economy; precisely, the investor observes the history of total output generated by this economy and tries to guess its time-varying growth rate. Various news—*signals*—from TV, newspapers, Internet, or other sources provide incremental information about this growth rate.<sup>3</sup> If the investor is attentive to news, then volatility of the estimated growth rate is high (because news observed are informative), but uncertainty is low (because the investor learns well). Conversely, if the investor pays no attention to news and learns by observing the output only, the volatility of the estimated growth rate is relatively low (because no news are observed), but uncertainty tends to be high (because the investor learns badly). Put differently, two antagonistic effects arise endogenously from learning matter for expectations. The first comes from attention—or how *sure* the investor is about the likely growth rate. The second comes from uncertainty—or how *unsure* the investor is about the likely growth rate.

Our paper embeds this simple but powerful insight within an elementary pure exchange economy and studies its implications for fluctuations in stock market volatility and risk premium. We build a dynamic general equilibrium model in which an investor collects information—with *fluctuating attention*—on the unobservable state of the economy. The key consequence of our model structure is a dynamic interaction between attention and uncertainty, which naturally captures the above insight.

We uncover a subtle relationship between attention, uncertainty, and volatility. Our main prediction is that the variance of stock returns increases quadratically with attention and uncertainty. We perform an empirical investigation which lends support to this prediction. Furthermore, we run a horse race between attention and uncertainty. We find that attention is a more powerful driver of volatility, especially when controlling for lagged volatility.

Then, we show that risk premia are also determined by both attention and uncertainty. Disentangling the effects of these two drivers, we highlight a quadratic relationship between risk premia and attention as well as between risk premia and uncertainty. Our empirical analysis shows that our theoretical predictions are sustained by the data.

A fortuitous byproduct of our analysis arises from the transparent relation between attention and uncertainty and adds a novel result to the literature. Intuitively, one minute of extreme attention does not necessarily push uncertainty to zero and one minute of complete inattention does not necessarily push it to the highest level. That is, as attention moves, uncertainty adjusts gradually. This generates a lead-lag relation, a decoupling between attention and uncertainty. As a result, the economy can enter at times in states of high attention and high uncertainty, or “panic states,” characterized by spikes in volatilities and risk premia. Once the economy shifts to a panic state, the asset price becomes

---

<sup>3</sup>Notable works in the literature of learning in financial markets adopting this setup are Detemple (1986), Timmermann (1993), David (1997), Veronesi (1999, 2000), Brennan and Xia (2001), Scheinkman and Xiong (2003) and Dumas, Kurshev, and Uppal (2009) among others.

much more sensitive to news. This result connects our model—at a qualitative level—with events that we witnessed during the turmoil in the Fall of 2008.

The relation and differences between our work and the rest of the literature will be highlighted throughout much of this paper, but it would be helpful at this stage to offer a taste of some of them. Our paper is related to the broad literature of learning and uncertainty in finance (Detemple 1986, Genotte 1986, Dothan and Feldman 1986, David 1997, Veronesi 1999, 2000, Brennan and Xia 2001). However, our analysis differs from this literature in three significant ways. First, our model clearly separates the effects of attention and uncertainty on stock market volatility and risk premium, whereas the existing literature focuses mainly on uncertainty. Attention is an important ingredient, because uncertainty inherently depends on how much attention investors pay to information. Second, while most of the literature usually performs comparative statics to demonstrate the effect of learning (see, e.g., Veronesi 2000, or Brennan and Xia 2001), our analysis is fully dynamic. This extra degree of freedom allows us to spot states of high volatility and high risk premium overlooked by static comparisons. We believe that this prediction has not been formulated before. Still, this prediction seems consistent with what we observe during financial panics. Third, in order to obtain stochastic uncertainty, the previous literature had to resort to non-Gaussian distributions (Detemple 1991, David 1997, and Veronesi 1999).<sup>4</sup> By assuming fluctuating attention, we can still use Gaussian-type diffusion processes and benefit from their mathematical simplicity.

We adopt an infinite horizon economy with a representative agent with Kreps-Porteus preferences, as in the long-run risk literature (Bansal and Yaron, 2004). This literature introduces exogenous time-varying risk in the fundamentals, whereas here the risk is endogenous and comes from learning. Moreover, our calibration is different from the long-run risk calibration, in that expected consumption growth is less persistent. We also do not consider dividend and consumption separately. Adding this feature would improve the fit to the level of observed asset pricing moments, while keeping our main results unchanged. Since our aim is to study the dynamic relationship between attention, uncertainty, volatility, and risk premium and not to match asset pricing moments, we let consumption be equal to dividend in equilibrium.

Because our model features learning, it is closely related to the study of Ai (2010), who introduces learning in a long-run risk model with linear production technology. This latter assumption is crucial in the sense that the direction in which information quality affects the equity premium depends only on the risk aversion parameter. In our pure exchange economy, the direction and the magnitude of the effect of learning depend crucially on both parameters.

---

<sup>4</sup>If investors have Gaussian priors and all variables are normally distributed, the conditional variance of investors' estimates—the *learning uncertainty*—follows a deterministic path and quickly converges to its steady-state value.

## 2 An Equilibrium Model of Fluctuating Attention

The novelty of our approach is to incorporate state-dependent attention in a pure exchange economy *à la* Lucas (1978). This is only a minimal extension, but the set of implications is rich. First, we characterize the output process and the learning problem of the representative agent. Next, we define fluctuating attention and explain the dynamic properties of attention and uncertainty. Before solving for the equilibrium, it is useful to estimate the parameters associated with the learning system that we bring forth. We do this by using the Generalized Method of Moments (Hansen, 1982) and find that the learning problem with fluctuating attention is a reasonable description of reality. Then, we solve for the equilibrium and we move on to break down the predictions of the theory in Sections 3, 4, and 5. Finally, we show that these predictions are consistent with actual observations.

Our argument relies on the premise that investors' attention is time-varying, a feature confirmed by empirical research (Da et al. 2011, Vlastakis and Markellos 2012, Dimpfl and Jank 2011, Kita and Wang 2012, Chauvet, Gabriel, and Lutz 2012, Schmidt 2013) and ultimately confirmed by our own estimation. Although we connect attention movements either to surprises in output readings or even in price returns, all our results go through no matter how attention moves.

### 2.1 The Economic Setting

The economy is characterized by a single output process (*the dividend*) having an unobservable 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  $\delta$ , whose dynamics are given by

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

The unobservable fundamental  $f$  follows a mean reverting process

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

Our 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

$$ds_t = \Phi_t dZ_t^f + \sqrt{1 - \Phi_t^2} dZ_t^s$$

The vector  $(Z^\delta, Z^f, Z^s)^\top$  is a 3-dimensional standard Brownian motion under the complete information filtration. The three Brownians are uncorrelated. 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 signals provide information on changes in the fundamental and not on its level. Consequently, even if the information is infinitely precise, the investor could never learn the true value of the fundamental. Although this difference with well-known models does not change qualitatively our results, we consider this setup because, as explained precisely in Section 2.3, it more cleanly disentangles the effect of attention on the one hand and uncertainty on the other hand.<sup>5</sup>

Let us focus on the correlation  $\Phi$  between the signal and the fundamental. In the spirit of Detemple and Kihlstrom (1987), Veldkamp (2006a,b), Huang and Liu (2007), and Hasler (2012),  $\Phi$  can be interpreted as the accuracy of news updates observed by the investor. In the references above, the investor can exert control on this accuracy. If  $\Phi = 0$ , it is equivalent to no news updates, whereas if  $\Phi = 1$  it is equivalent to perfectly accurate news updates. Since the investor exerts control on the parameter  $\Phi$ , this parameter is called *attention to news*<sup>6</sup>.

We follow the same interpretation and assume in our model that the investor directly controls this accuracy. We, however, do not assume that she does so optimally. Instead, we claim that she changes her attention whenever she observes changes in the general state of the economy, or the current economic conditions (a term that will be defined in the next section). Hence, in our model  $\Phi$  is time-varying and is determined by current economic conditions. By adopting such a reduced form approach we are able to build a full-fledged general equilibrium model. In the next section we provide a complete characterization of the dynamics of attention  $\Phi$  and we justify our assumption.

It is worth mentioning that our focus is the impact of fluctuating attention to news on asset prices and not its foundation. In our model, fluctuating attention is considered because it is sustained by empirical observations (e.g., Da et al., 2011; Vlastakis and Markellos, 2012). Potential foundations of fluctuating attention to news can be found in Detemple and Kihlstrom (1987), Veldkamp (2006a,b), Bansal and Shaliastovich (2011), and Hasler (2012).

---

<sup>5</sup>To the best of our knowledge, this specification has first been adopted by Scheinkman and Xiong (2003), Dumas et al. (2009), and Xiong and Yan (2010). We could also assume that the investor observes a noisy signal of the fundamental  $f_t dt$ . In that case, the variance of the noisy signal would be stochastic and would belong to the interval  $[0, \infty)$ .

<sup>6</sup>Note that other types of attention have also been studied. Peng and Xiong (2006), Kacperczyk, Van Nieuwerburgh, and Veldkamp (2009), and Van Nieuwerburgh and Veldkamp (2010) study attention allocation problems when information-processing capacity is limited. Abel, Eberly, and Panageas (2007) solve for optimal no-trade periods under the assumption that observing the value of the portfolio is costly. These no-trade periods are called inattention periods, as investors do not observe their wealth in these time intervals.

## 2.2 Time-Varying Attention

It is reasonable to assume that investors' attention  $\Phi$  is time-varying. But how do we define current economic conditions and how to make attention depend on them? In the present model there are two observable variables that could fill this role: the dividend and the stock price. The dividend is exogenous, whereas the stock price is endogenously determined in equilibrium.

Since the dividend is exogenous, assuming that attention depends on it is technically easier than if attention depended on the stock price. In the latter case, solving for the equilibrium comprises a challenge. Because prices are endogenously determined and depend on attention, when prices drive attention we are facing a fixed-point problem. Despite this difficulty, we are able to solve this particular model. Most important, we obtain very similar results in both cases. Given this, and for ease of exposition, we analyze in this section the first case and assume that movements in dividend drive investors' attention. Some important differences with the other case (which perhaps could be considered more intuitive) are explained along the way. We relegate to Section 6.1 a full derivation and discussion of the case when attention is driven by price movements.

We postulate a measure of economic conditions 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, it is well known that surprises rather than news themselves face increased scrutiny from investors. Thus, we rely on surprises in dividend growth rather than just dividend growth. The resulting variable, which we denote *performance index*, captures both features. It is defined as follows

$$\phi_t \equiv \int_0^t e^{-\omega(t-u)} \left( \frac{d\delta_u}{\delta_u} - \hat{f}_u du \right), \quad (1)$$

where  $\hat{f}_u$  represents investors' estimate of the fundamental at time  $u$ . The stochastic process of this estimate results from Bayesian learning and will be defined in Section 2.3. The parameter  $\omega > 0$  represents the weight associated to the present relative to the past. If  $\omega$  is large, the past dividend growth influences to a minimal degree the performance index—the latter becomes a substitute of the current dividend growth. On the other hand, if  $\omega$  is small, the past dividend growth influences to a greater extent the performance index. By assuming that  $\omega \in [0, \infty)$  we let the investor decide how much of the history of past dividends to consider.<sup>7</sup>

A similar performance index could be built from current and past surprises in stock

---

<sup>7</sup>Koijen, Rodriguez, and Sbuelz (2009) build a similar performance index in a partial equilibrium setting to allow for momentum and mean reversion in stock returns. In our case this index is built directly from the dividend process, to capture in a parsimonious way the recent development of the dividend.



returns. In that case, we would have

$$\phi_t \equiv \int_0^t e^{-\omega(t-u)} \left( \frac{dS_u + \delta_u du}{S_u} - \mu_u du \right), \quad (2)$$

where  $\mu$  is the expected return on the stock (to be determined in equilibrium) and  $S$  is the stock price. This would correspond to a situation where investors' attention reacts to current and past surprises in stock returns. As previously stated, we relegate this case to Section 6.1 and focus the rest of this section to the case where investors' attention depends on current and past surprises in dividends, as defined in Equation (1).

The dynamics of the performance index can be derived from the dynamics of the dividend. An application of Itô's lemma on the performance index yields

$$d\phi_t = -\omega\phi_t dt + \sigma_\delta dZ_t^\delta,$$

which shows that the performance index fluctuates around a long-term mean of zero with a mean-reversion speed  $\omega$ .

Having built a measure of current economic conditions, we are now ready to introduce the link between this measure and investors' attention. The following definition is the core of our way to model time-varying attention.

**Definition 1.** *Attention  $\Phi$  is defined as a function  $g$  of the performance index:*

$$\Phi_t = g(\phi_t) \equiv \frac{\bar{\Phi}}{\bar{\Phi} + (1 - \bar{\Phi}) e^{\Lambda\phi_t}}, \quad (3)$$

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 3 parameters:  $\omega$ ,  $\bar{\Phi}$ , and  $\Lambda$ .

We stress that the only assumptions made so far are that investors' attention is time varying and depends on the performance index. This performance index can be built either from current and past surprises in dividend growth rates or from current and past surprises in stock returns. How much weight is given to the past is decided by the parameter  $\omega$ . Furthermore, the attention can either increase, decrease, or remain constant, according to the parameter  $\Lambda$ . Finally, the long-run mean of the attention,  $\bar{\Phi}$ , can take any value between 0 and 1. Therefore, our specification is quite comprehensive and encompasses a wide range of plausible cases.

While the sign of the parameter  $\Lambda$  crucially dictates the negative or positive dependence of attention on past dividend growth, our main results obtain regardless which



sign is used. The single key ingredient for our results is that  $\Lambda \neq 0$ , i.e., that attention fluctuates. Fluctuating attention has been studied theoretically by Duffie and Sun (1990). They propose a model featuring slowness of individual portfolio adjustments, where the investor sets an optimal “time-out” during which she focuses on other activities.<sup>8</sup> Chien, Cole, and Lustig (2012), Bacchetta and Wincoop (2010), and Duffie (2010) assume that the periods of inattention are exogenously fixed and show that this feature helps understand the volatility of the market price of risk, the forward discount puzzle, and stock price over-reaction and reversal. These studies focus on investors’ attention to wealth, whereas we focus on investors’ attention to financial news.

### 2.2.1 The Role of the 3 Parameters in the Dynamics of the Attention Process

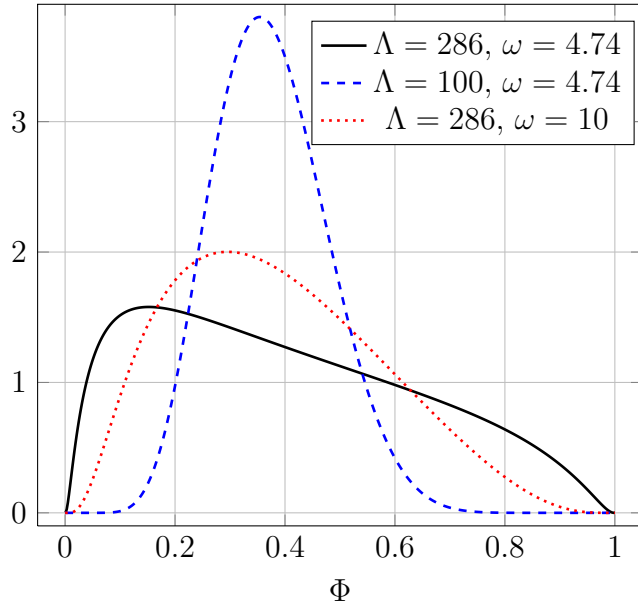
The unconditional distribution of the performance index  $\phi$  is Gaussian with mean 0 and variance given by  $\frac{\sigma_\delta^2}{2\omega}$  (see Appendix A.1 for a proof of this statement). We know from Equation (3) that, for  $\Lambda \neq 0$ ,  $\Phi$  is a strictly monotone function of  $\phi$ . This monotonicity allows 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.

While the parameter  $\bar{\Phi}$  dictates the location of the unconditional distribution of attention, two other important parameters govern the shape of this distribution. The first is  $\Lambda$ , the parameter which dictates the adjustment of attention after changes in the performance index. The second is  $\omega$ , the parameter which 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 black solid line corresponds to the calibration performed in Section 2.4 on US data. It shows that attention can vary substantially, as it can take very large and very low values with significant probabilities. The two additional lines show that a decrease in the parameter  $\Lambda$  (dashed blue line) and respectively an increase in the parameter  $\omega$  (dotted red line) have similar effects: both tend to 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  $\bar{\Phi}$ . 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 will help us to calibrate them on US data, task that we undertake in Section 2.4.

---

<sup>8</sup>Other studies focusing on optimal “time-outs” are Abel et al. (2007) and Rossi (2010).



**Figure 1: Probability density function of investors' attention**

Probability density function of  $\Phi$  for different values of  $\Lambda$  and  $\omega$ . Other parameters are  $\lambda = 0.42$ ,  $\bar{f} = 0.028$ ,  $\bar{\Phi} = 0.368$ ,  $\sigma_f = 0.029$ ,  $\sigma_\delta = 0.014$ . The black solid line illustrates the probability density function for  $\Lambda = 286$  and  $\omega = 4.74$  (this corresponds to the calibration performed in Section 2.4 on US data). The blue dashed line shows how the distribution changes with a lower  $\Lambda$  ( $\Lambda = 100$  and  $\omega = 4.74$ ). The red dotted line shows how the distribution changes with a higher  $\omega$  ( $\Lambda = 286$  and  $\omega = 10$ ).

### 2.3 Bayesian Learning

Since attention  $\Phi$  is observable by construction, our setup remains conditionally Gaussian and the Kalman filter is applicable for the purpose of learning. The state vector prior to the filtering exercise consists in 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 the past values of dividends, it does not bring any additional information.

Because the conditional correlation between the signal and the fundamental—attention  $\Phi$ —is time-varying and is a function of the performance index, the assessed (filtered) fundamental takes a non-standard form. The major change is that the conditional variance of the unobserved fundamental given today's information (simply referred to as the posterior variance, or *Bayesian uncertainty*) is time-varying. Intuitively, when attention is high uncertainty is low, whereas the opposite occurs when attention is low. Following this reasoning, the vector of filtered state variables includes two additional terms: the performance index, which dictates the level of attention, and the uncertainty that we

denote by  $\gamma$ . Hence, the dynamics of the observed state variables become

$$\begin{aligned}
d\zeta_t &= \left( \hat{f}_t - \frac{1}{2}\sigma_\delta^2 \right) dt + \begin{pmatrix} \sigma_\delta & 0 \end{pmatrix} dW_t \\
d\hat{f}_t &= \lambda (\bar{f} - \hat{f}_t) dt + \begin{pmatrix} \frac{\gamma_t}{\sigma_\delta} & \sigma_f \Phi_t \end{pmatrix} dW_t \\
d\phi_t &= -\omega \phi_t dt + \begin{pmatrix} \sigma_\delta & 0 \end{pmatrix} dW_t \\
d\gamma_t &= \left( \sigma_f^2 (1 - \Phi_t^2) - 2\lambda \gamma_t - \frac{\gamma_t^2}{\sigma_\delta^2} \right) dt,
\end{aligned} \tag{4}$$

where  $W \equiv (W^\delta, W^s)^\top$  is a 2-dimensional Brownian motion under the investor's observation filtration, and  $\Phi$  is given by the functional form (3). The assessed fundamental is denoted by  $\hat{f}$ . The two Brownian motions governing this system are defined by

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

and represent the normalized innovation processes of dividend and signal realizations. The proof of the above statements is provided in Appendix A.2.

A notable difference arises between our model and other models of learning with similar structures (e.g., Scheinkman and Xiong 2003, Dumas et al. 2009). In the latter models it is usually assumed that uncertainty converged to its steady-state value. The deterministic nature of the uncertainty process obtained in the latter references makes this assumption plausible, as  $\gamma$  converges quickly to its steady-state. In our case, although the process of the posterior variance remains locally deterministic, we cannot assume a constant uncertainty, as it depends on attention, which itself is time-varying, as shown in system (4). Thus, uncertainty must be included in the state space.

Exactly this feature turns out to be the novelty of our contribution to the literature, for two reasons. First, in most of the existing models of learning, such as Scheinkman and Xiong (2003) and Dumas et al. (2009), uncertainty is constant. Only a few learning models feature endogenous fluctuating uncertainty. Typically, fluctuating uncertainty shows up when distributions are non-Gaussian or when fundamentals follow regime-switching processes. Second, our model is clearly distinct from the long run risk models (Bansal and Yaron, 2004) in which fundamentals are observable and feature stochastic volatility. Specifically, in our model the dynamics of  $\hat{f}$  in system (4) reveal that two diffusion components drive the overall noise in the fundamental, and these two components are inversely interconnected.

The first component loads on dividend innovations and the second on news innovations. As these two innovations represent the signals used by the investor to infer the fundamental, the vector  $\begin{pmatrix} \frac{\gamma_t}{\sigma_\delta} & \sigma_f \Phi_t \end{pmatrix}$  constitutes the weights assigned by the agent to

both signals. As attention changes, these weights move in opposite direction: higher attention pushes the investor to give more weight to news shocks, whereas lower attention pushes the investor to give more weight to dividend shocks. Consequently, the variance of the filtered fundamental, denoted henceforth by  $\sigma^2(\hat{f}_t)$ , is time-varying and comprises two antagonistic effects which arise endogenously from learning. While our aim is to understand how these antagonistic attention and uncertainty effects drive asset prices, long run risk models ignore these two effects and focus on the persistence of the volatility of the fundamental. This in turn clearly makes our model different and complementary to long run risk studies.

The variance of  $\hat{f}$  satisfies

$$\sigma^2(\hat{f}_t) = \frac{\gamma_t^2}{\sigma_\delta^2} + \sigma_f^2 \Phi_t^2. \quad (5)$$

Equation (5) reveals a clean separation of attention and uncertainty. This results from our signal specification, i.e., because signals provide information on changes in the fundamental and not on its level. This clear separation provides therefore a justification of our choice of signal structure.<sup>9</sup> Furthermore, Equation (5) tells us that observing a high volatility of the forecasted growth rate might result from two different sources. It might reflect how *unsure* investors are about the likely path of future dividend growth—the first term in Equation (5). But it might also reflect how *sure* investors are about the likely path of future dividend growth—the second term in Equation (5). This is a key consequence of our modeling strategy. In other words, there are two forces driving the variance of the filtered fundamental. Fluctuating attention increases the variance through better learning (a direct impact). Better learning, in turn, decreases uncertainty, thus dampening the initial effect (an indirect impact).

This brings us to another key consequence of our endogenous structure of information flow. The deterministic dynamics of the uncertainty process  $\gamma$  outlined in the last equation of the system (4) shows that there is no instantaneous correlation between attention and uncertainty. Indeed, there is no Brownian motion in the dynamics of  $\gamma$ . Instead, uncertainty decreases (increases) gradually when attention is high (low). But the lack of instantaneous correlation gives birth to a lead-lag relation between attention and uncertainty. Thus, a decoupling arises between attention and uncertainty. A spike in attention can result in a state of high attention and high uncertainty—a panic state. Yet this state will not last long; high attention will inevitably reduce uncertainty. This lead-lag relation has important consequences for the dynamics of asset returns, as we will describe in Section 5.

---

<sup>9</sup>If signals provided information on the level of the fundamental ( $ds_t = f_t dt + \frac{1}{\Phi_t} dZ_t^s$ ,  $\Phi \in [0, \infty)$ ), Equation (5) would have a more ambiguous form. The second term would contain the product of  $\gamma$  and  $\Phi$ , which would not change our results but would make the analysis less clear.

To summarize, we offer a framework to study the simultaneous dynamics of attention and uncertainty and their impact on asset prices. This analysis could obviously not be carried out in a long run risk model because the antagonistic effects of attention and uncertainty would be bypassed.<sup>10</sup> Finally, if attention is constant, then uncertainty is constant, and thus neither of the two would affect the dynamics of asset returns.

## 2.4 Calibration to the U.S. Economy

Does our theoretical structure reasonably describe investors' learning behavior? We address this question by calibrating the model to the U.S. Economy. The reader can skip most of this section without compromising understanding of subsequent discussions. Yet, the purpose of this section is to convince him that the economic setting that we propose does a good job at explaining the evolution of investors' beliefs. In addition, performing the calibration before solving for the equilibrium gives more confidence to the predictions of our model.

The investor is able to observe two processes: the dividend stream  $\delta$  and a flow of information  $s$ . Hence, the investor uses  $\delta$  and  $s$  to estimate the evolution of the non-observable variable  $f$ . Since we don't know which variable corresponds to the signal  $s$ , calibrating our model to observed data is challenging—although our theoretical model assumes that the flow of information  $s$  is observable, it is hard to find a proxy for this variable in practice. To manage this problem, we follow David (2008) and use the analyst 1-quarter ahead forecasts on real US GDP growth rate as a proxy for the filtered fundamental  $\hat{f}$ . To be consistent, we use the real US GDP realized growth rate as a proxy for the output growth rate. Quarterly data from Q1:1969 to Q4:2012 are obtained from the Federal Reserve Bank of Philadelphia.

Since we work with quarterly data, an immediate discretization of the stochastic differential equations exposed in the system (4) would provide biased estimators. Hence, we first solve this set of four stochastic differential equations. The solutions are provided in Appendix A.3. We then approximate the 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  $t = 0, \Delta, \dots, T\Delta$ , we can directly infer the value of the Brownian vector  $\epsilon_{t+\Delta}^\delta \equiv W_{t+\Delta}^\delta - W_t^\delta$ . Moreover, because the observed vector  $\hat{f}_t$ ,  $t = 0, \Delta, \dots, T\Delta$  depends on  $\epsilon_{t+\Delta}^\delta$  and  $\epsilon_{t+\Delta}^s \equiv W_{t+\Delta}^s - W_t^s$ , we obtain a direct characterization of the signal vector  $\epsilon_{t+\Delta}^s$  by substitution. This shows that observing  $\delta$  and  $\hat{f}$ , instead of  $\delta$  and  $s$ , also provides a well defined system.

---

<sup>10</sup>It is straightforward to extend the model and assume additional exogenous fluctuations in uncertainty. We do not follow this route, since our aim is to analyse the endogenous effect of investors' attention on uncertainty.

### 2.4.1 Moment Conditions

Our model is calibrated on the 2 time-series discussed previously using the Generalized Method of Moments (Hansen, 1982). The vector of parameters is defined by  $\Theta = (\lambda, \bar{f}, \omega, \bar{\Phi}, \Lambda, \sigma_f, \sigma_\delta)^\top$ . Consequently, we need at least 7 moment conditions to infer the vector of parameters  $\Theta$ . For the sake of brevity, the moment conditions are exposed in Appendix A.4. We proceed here briefly with their interpretation.

The analyst forecasts of the growth rate allow us to build moments that identify  $\lambda$  and  $\bar{f}$ . Indeed, 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  $\bar{f}$  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 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 estimate  $\bar{\Phi}$ ,  $\Lambda$ , and  $\sigma_f$  (note that here we have to construct recursively  $\gamma_t$  and  $\Phi_t$ ). Per total, we have eleven moment conditions helping us to estimate seven parameters.

It is worth emphasizing that we match the unconditional variance of our implied attention  $\Phi_t$  to the unconditional variance of some proxy of investors' attention. In order to do that, we follow Da et al. (2011) and build an empirical measure of attention. We use Google search volumes on groups of words with financial or economic content. To avoid any bias, none of the terms used have positive or negative connotations.<sup>11</sup> We adjust this Google attention index to be between 0 and 1 (as our attention  $\Phi$ ) and compute its unconditional variance. The unconditional variance of our model implied attention index should match the empirical unconditional variance. Given that Google search volume data are only available since 2004, we do not use them for any other moment conditions.

### 2.4.2 Parameter Estimates

The values, t-stats, and p-values of the vector  $\Theta$ , resulting from the GMM estimation, are provided in Table 1. The test of over-identifying restrictions indicates that the model provides a good fit to the GDP realized growth and GDP growth forecast, with the  $J$ -test  $p$ -value being 0.35. The sole insignificant parameter is the mean reversion speed parameter  $\lambda$ . We want to point out that our estimated value of  $\lambda$  is relatively far from what the long run risk literature assumes. Studies dealing with long run risk typically assume that the mean reversion parameter is between 0 and 0.25. In Bansal and Yaron (2004) the AR(1)

---

<sup>11</sup>More precisely, the Google attention index is built based on the following combination of words: "financial news," "economic news," "Wall Street Journal," "Financial Times," "CNN Money," "Bloomberg News," "S&P500," "us economy," "stock prices," "stock market," "NYSE," "NASDAQ," "DAX," and "FTSE." Using other similar words in several combinations provides very similar empirical measures of attention.

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

**Table 1: Calibration to the U.S. economy (GMM estimation)**

Parameter values resulted from a GMM estimation with 11 moment conditions. The parameters identify the information setup described in Equations (1), (3), and (4). The Hansen J-test of overidentification cannot be rejected ( $\Pr[\text{Chi-sq.}(4) > J] = 0.35$ ). Statistical significance at 10%, 5%, and 1% is labeled respectively with \*/\*\*/\*\*\*.

parameter of the fundamental is worth 0.979 at monthly frequency. This parameter would correspond to  $\lambda = -12 \ln(0.979) = 0.25$ . Barsky and De Long (1993) go even further by assuming that the fundamental is an integrated process. Although our dataset 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.<sup>12</sup>

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 significant. The volatility of real GDP realized growth rate is close to 1% and in line with the estimation of Beeler and Campbell (2012) from postwar data. Our estimate is much smaller than 2.4% which would result from a simulation of the Bansal and Yaron (2004) model (see discussion in Beeler and Campbell, 2012, where the authors call this mismatch a “serious difficulty” of the long-run risk model).

Given this, our model deviates from the long-run risk model with respect to both  $\lambda$  and  $\sigma_\delta$ . Nonetheless, we are still able to obtain considerable risk premium and excess volatility, as we will show in Sections 3 and 4. This added with the facts that in our model dividends equal consumption and the fundamental is unobservable, clearly distinguishes our setup from the long-run risk literature.

We obtain a large positive and significant value for the parameter  $\Lambda$ , suggesting 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 the last year of dividend growth data).<sup>13</sup>

<sup>12</sup>The half-life is a measure of the speed of mean-reversion. It is given by  $\ln(2)/\lambda$ . For the Bansal and Yaron (2004) calibration, the half-life is roughly 33 months.

<sup>13</sup>A value of  $\omega = 4.74$  means that, at 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.



Put differently, investors' attention is strongly sensitive to recent experience. Coming back to the black solid line in Figure 1, we remark that attention varies substantially, taking values in the entire interval with significant probabilities.

In addition, a positive  $\Lambda$  means that attention is high in bad aggregate economic states ( $\phi < 0$ ) and low in good aggregate economic states ( $\phi > 0$ ). This can be interpreted as follows. 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. On the other hand, when the economy enters a recessionary phase, the high probability of a decrease in future consumption grabs investors' attention. This leads investors to estimate as accurately as possible the change in the fundamental.

Although we emphasize that our results go through no matter the sign of  $\Lambda$ , a discussion of this sign is of interest here. If attention is higher in bad times, i.e.,  $\Lambda > 0$ , we should observe better learning and thus better forecasts exactly in those times. This observation is consistent with empirical findings by Patton and Timmermann (2008). Using consensus forecasts of US GDP growth over 1991-2004, Patton and Timmermann (2008) find that forecasters estimated GDP growth quite well for the recessions in early 1990s and 2001. They, however, underestimated the strong GDP growth in the mid to late 1990s and consistently overestimated the realized values of GDP growth after the 2001 recession. We add to this evidence by using our larger sample of data (from 1969 to 2012) in Section 6.2. In our separate calculations, we also observe better forecasts in bad times, further strengthening support for a positive  $\Lambda$ .

Furthermore, this implication is in line with two additional pieces of empirical evidence. First, Da, Gurun, and Warachka (2011) show that analyst forecast errors are smaller when past 12 months returns are negative than when they are positive, suggesting that information gathered by analysts in downturns is more accurate than in expansionary phases.<sup>14</sup> Second, Garcia (2013) documents that investors react strongly to good and bad news during recessions, whereas during expansions investors' sensitivity to information is much weaker.

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" given adverse prior news. Karlsson et al. (2009) find support for this effect by examining account monitoring behavior (number of logins) of Scandinavian and American investors. This, a priori, seems at odds with our findings, but we note that in Karlsson et al. (2009) investors collect information *about the value of their portfolios*, whereas our investors collect information *about the fundamental structure of the economy*. We can imagine a situation where in bad times, overwhelmed by adverse information about the economy, investors

---

<sup>14</sup>This claim holds under the assumption of continuous information, as in our case. Under discrete information, the reverse assertion is verified.

are scared to check their portfolios. However, in rising markets, investors find it sufficient to check their portfolios regularly and not listen to information about the economy. Note also that [Karlsson et al. \(2009\)](#) measure number of logins minus number of trades, that is, the number of times people look at their portfolios *without* trading—it is quite obvious that this number is large in good times, since people do not have reasons to trade in those times. Also, a rational investor has low risky positions in bad times, hence one less reason to check his wealth. As such, these two views are not excluding each other. They simply suggest that attention to news and attention to wealth are inversely related.

A natural question arises, whether theoretical models of endogenous attention confirm the finding of higher attention to news in recessionary phases than in expansionary phases. In a portfolio choice problem of optimal attention allocation, [Hasler \(2012\)](#) finds that forecast accuracy is decreasing with past returns, which provides foundation for a positive parameter  $\Lambda$ . This result obtains because the investor optimally acquires more information when it is expected to be more valuable, i.e., the utility gain from the extra information is higher in bad states. Our model, thus, takes as a primitive [Hasler \(2012\)](#)'s result.<sup>15</sup>

### 2.4.3 Time Series Dynamics of Implied Attention

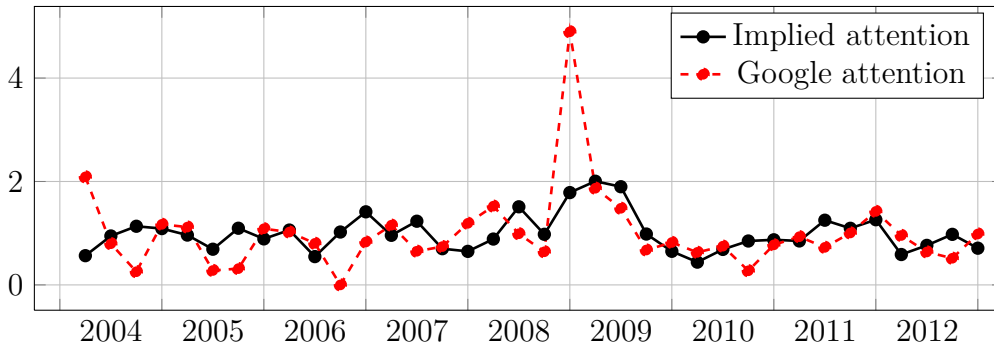
How does our implied measure of attention correlate with other attention proxies? To answer this question, we compare the attention implied by our estimation with the index of attention built from Google search volume (sampled at quarterly frequency). The two indices are depicted in [Figure 2](#). The coefficient of correlation between them is 0.44. Movements in our implied attention seems to be well aligned with movements in the same direction of the Google attention index. An ordinary least squares regression results in a significant coefficient of 0.95 and an R-squared of 0.19. This provides support that our information setup captures well the learning behavior of investors. We emphasize that, in our estimation procedure, we only used Google attention index data to fit the unconditional variance of attention. All the other moment conditions relied on GDP data. We are thus confident that our learning structure is a reasonable description of reality.

The model could be extended on several dimensions in order to match the data even better. For example, factors other than the performance index might affect investors' attention.<sup>16</sup> This can be done by inserting an extra exogenous noise in the dynamics of

---

<sup>15</sup>Other optimal attention allocation problems are mentioned in what follows. [Detemple and Kihlstrom \(1987\)](#) study an optimal information acquisition problem in an economy à la [Cox, Ingersoll, and Ross \(1985\)](#). The solution procedure is discussed but no explicit solution is provided. [Peng and Xiong \(2006\)](#) show that limited information-processing capacity and overconfidence leads to category-learning and excess correlation. In [Veldkamp \(2006b\)](#) costly information yields excess co-movement among asset prices. [Kacperczyk et al. \(2009\)](#) show that investors optimally concentrate on macro news in bad times and idiosyncratic news in good times. This generates market-timing strategies in recessions and stock picking strategies in expansions. In [Bansal and Shaliastovich \(2011\)](#) investors choose to acquire perfect information when the volatility of output or the uncertainty is sufficiently large, implying jumps in asset prices.

<sup>16</sup>See, for example, the anecdote about Tiger Woods and the New York Stock Exchange volume, in the



**Figure 2: Implied attention versus Google attention index**

The black solid line depicts the attention index implied by our estimation. The red dashed line depicts the weighted search index on financial and economic news from 2004 to 2012, at quarterly frequency. Indices are divided by their sample average to allow comparison with each other.

attention. Or, consider the case when attention depends on past stock returns rather than past dividend growth rates, as in Equation (2). In this case the attention index will inherit the time varying volatility of stock returns. Indeed, we observe larger fluctuations in investors' attention when stock returns are more volatile.

## 2.5 Equilibrium with Epstein-Zin Preferences

Our model features a representative agent with Epstein-Zin preferences. The advantage over CRRA is that asset prices increase with the drift of consumption, while they decrease with the volatility of consumption.

### 2.5.1 Optimization Problem

The representative agent's preferences over the uncertain consumption stream  $\{c_t\}$  are represented by a utility index  $U_t$  that satisfies the following recursive equation

$$U_t = \left\{ \left(1 - e^{-\rho dt}\right) c_t^{1-\psi} + e^{-\rho dt} \mathbb{E}_t \left[ U_{t+dt}^{1-\alpha} \right]^{\frac{1-\psi}{1-\alpha}} \right\}^{\frac{1}{1-\psi}}, \quad (6)$$

where  $\rho$  is the subjective discount factor,  $1/\psi$  is the intertemporal elasticity of substitution and  $\alpha$  is the local risk aversion coefficient. Replacing  $dt = 1$  in Equation (6) 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 intertemporal elasticity of substitution,  $\alpha = \psi$ , the recursive utility reduces to the standard time-separable power utility with relative risk aversion  $\alpha$  and intertemporal elasticity of substitution  $1/\alpha$ .

---

Presidential Address of [Duffie \(2010\)](#).

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  $f(c, J)$  is the normalized aggregator (see Duffie and Epstein, 1992a,b). Following Benzoni, Collin-Dufresne, and Goldstein (2011), a state-price density is defined as

$$\xi_t = e^{\int_0^t f_J(c_s, J_s) ds} f_c(c_t, J_t).$$

The following proposition, whose proof is provided in Benzoni et al. (2011), provides the partial differential equation for the price-dividend ratio.

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

$$\begin{aligned} J_t &= \frac{1}{1-\alpha} c_t^{1-\alpha} (\rho I(x_t))^\nu \\ \xi_t &= e^{-\int_0^t (\rho\nu + \frac{1-\nu}{I(x_s)}) ds} c_t^{-\alpha} I(x_t)^{\nu-1}, \end{aligned} \quad (7)$$

where  $\nu \equiv \frac{1-\alpha}{1-\psi}$ ,  $x \equiv (\hat{f} \ \phi \ \gamma)^\top$ , and  $I(x)$  is the price-dividend ratio. The price-dividend ratio  $I(\cdot)$  satisfies the following partial differential equation

$$0 = I \left( (1-\alpha) \left( \hat{f} - \frac{1}{2} \sigma_\delta^2 \right) + (1-\alpha)^2 \frac{\sigma_\delta^2}{2} - \rho\nu \right) + \frac{\mathcal{D}I^\nu}{I^{\nu-1}} + (1-\alpha)\nu (\gamma I_{\hat{f}} + \sigma_\delta^2 I_\phi) + \nu, \quad (8)$$

where we define  $\mathcal{D}h(x) \equiv h_x(x)\mu_x(x) + \frac{1}{2} \text{trace} \left( h_{xx} \sigma_x(x) \sigma_x(x)^\top \right)$ .

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

$$\begin{aligned} dx_t &= \mu_x(x_t) dt + \sigma_x(x_t) dW_t \\ \frac{dI(x_t)}{I(x_t)} &= (\dots) dt + \sigma_I(x_t) dW_t, \end{aligned}$$

where

$$\begin{aligned} \mu_x(x_t) &= \left( \lambda(\bar{f} - \hat{f}_t) \quad -\omega\phi_t \quad \sigma_f^2(1 - \Phi_t^2) - 2\lambda\gamma_t - \frac{\gamma_t^2}{\sigma_\delta^2} \right)^\top \\ \sigma_x(x_t) &= \begin{pmatrix} \frac{\gamma_t}{\sigma_\delta} & \sigma_f \Phi_t \\ \sigma_\delta & 0 \\ 0 & 0 \end{pmatrix} \\ \sigma_I(x_t) &\equiv \left( \sigma_{1I}(x_t) \quad \sigma_{2I}(x_t) \right) = \frac{1}{I(x_t)} I_x(x_t)^\top \sigma_x(x_t). \end{aligned}$$

The partial differential equation (8) can be rewritten as

$$\begin{aligned}
0 = & I \left( (1 - \alpha) \left( \hat{f} - \frac{1}{2} \sigma_{\delta}^2 \right) + (1 - \alpha)^2 \frac{\sigma_{\delta}^2}{2} - \rho \nu \right) \\
& + \nu \left( \lambda (\bar{f} - \hat{f}) I_{\hat{f}} - \omega \phi I_{\phi} + \left( \sigma_f^2 (1 - \Phi^2) - 2\lambda\gamma - \frac{\gamma^2}{\sigma_{\delta}^2} \right) I_{\gamma} \right) \\
& + \frac{1}{2} \nu \left( \left( \frac{\gamma^2}{\sigma_{\delta}^2} + \sigma_f^2 \Phi^2 \right) I_{\hat{f}\hat{f}} + \sigma_{\delta}^2 I_{\phi\phi} + 2\gamma I_{\hat{f}\phi} \right) \\
& + \frac{1}{2I} \nu (\nu - 1) \left( \left( \frac{\gamma^2}{\sigma_{\delta}^2} + \sigma_f^2 \Phi^2 \right) I_{\hat{f}}^2 + 2\gamma I_{\hat{f}} I_{\phi} + \sigma_{\delta}^2 I_{\phi}^2 \right) \\
& + (1 - \alpha) \nu (\gamma I_{\hat{f}} + \sigma_{\delta}^2 I_{\phi}) + \nu.
\end{aligned} \tag{9}$$

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

$$I(x) \approx e^{\beta_0 + \beta_1 x}, \tag{10}$$

where  $\beta_0$  is a scalar and  $\beta_1 = (\beta_{11} \ \beta_{12} \ \beta_{13})$ . Plugging the exponential form (10) in Equation (9) and performing a first order linearization of the PDE around  $x^0 = (\bar{f} \ 0 \ \gamma_{ss})$ <sup>17</sup> yields a system of the form

$$A + Bx = 0,$$

where the scalar  $A$  and the vector  $B = (B_1 \ B_2 \ B_3)$  are large expressions that can be provided upon request. Setting  $A = 0$  and  $B = (0 \ 0 \ 0)$  yields a system of four equations with four unknowns ( $\beta_0, \beta_{11}, \beta_{12}$ , 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

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

where

$$\begin{aligned}
\beta_0 &= 3.6152 & \beta_{11} &= 1.1207 \\
\beta_{12} &= -0.0006 & \beta_{13} &= -11.3741.
\end{aligned} \tag{11}$$

---

<sup>17</sup>Note that  $\bar{f}$  is the long term mean of  $\hat{f}$ , 0 is the long term mean of  $\phi$ , and  $\gamma_{ss} = -\lambda\sigma_{\delta}^2 + \sqrt{\sigma_{\delta}^2 (\lambda^2\sigma_{\delta}^2 + \sigma_f^2(1 - \Phi^2))}$  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  $10^{-2}$ , the impact of these processes on the log price-dividend ratio are of the order of  $10^{-2}$ ,  $10^{-3}$ , and  $10^{-6}$ , respectively.

Since  $\beta_{11}$  is positive, there is a positive relationship between the estimated fundamental and the price-dividend ratio. The reason is as follows. Let us consider an increase in the estimated fundamental. First, this 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 future dividend growth rates. Second, an increase in the estimated fundamental implies an improvement of risky investment opportunities, pushing investors to demand more of the stock. This substitution effect outweighs the precautionary savings effect as long as the elasticity of intertemporal substitution is larger than 1. Thus, prices are positively related to estimated fundamentals in that case.

Because  $\beta_{13}$  is negative, an increase in uncertainty generates a drop in prices. Intuitively, an increase in uncertainty pushes investors to lower current consumption because expected consumption is more uncertain and investors again want to smooth consumption over time. Hence the demand for the stock rises, increasing its price. Also, risky investment opportunities become more uncertain and consequently push investors to lower their risky investments. This tends to push the price of the stock down. Again, since the substitution effect dominates the precautionary savings effect when the elasticity of intertemporal substitution is larger than 1, uncertainty and prices are inversely related.

### 2.5.2 Risk Free Rate, Risk Premium, and Volatility

Applying Itô's lemma to the state-price density  $\xi$  provided in Equation (7) yields the risk free rate  $r$  and the vector of market prices of risk  $\theta$  defined in Proposition 2, whose proof is provided in Benzoni et al. (2011).

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

$$r_t = \rho + \psi \hat{f}_t - \frac{1}{2} \alpha (1 + \psi) \sigma_\delta^2 - (1 - \nu) \left( \sigma_{1I}(x_t) \sigma_\delta + \frac{1}{2} \sigma_{1I}(x_t)^2 + \frac{1}{2} \sigma_{2I}(x_t)^2 \right) \quad (12)$$

$$\theta_t = \left( \alpha \sigma_\delta + (1 - \nu) \sigma_{1I}(x_t) \quad (1 - \nu) \sigma_{2I}(x_t) \right)^\top. \quad (13)$$

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

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

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

$$\begin{aligned} \sigma_t &= \left( \sigma_\delta + \sigma_{1I}(x_t) \quad \sigma_{2I}(x_t) \right) \\ \mu_t - r_t &= \sigma_t \theta_t = \left( \sigma_\delta + \sigma_{1I}(x_t) \quad \sigma_{2I}(x_t) \right) \left( \alpha \sigma_\delta + (1 - \nu) \sigma_{1I}(x_t) \quad (1 - \nu) \sigma_{2I}(x_t) \right)^\top \\ &= \alpha \sigma_\delta^2 + (1 - \nu + \alpha) \sigma_\delta \sigma_{1I}(x_t) + (1 - \nu) \left( \sigma_{1I}(x_t)^2 + \sigma_{2I}(x_t)^2 \right). \end{aligned}$$

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 negative. Hence, in this case, the second part of Equation (12) is negative, making the level of the risk free rate smaller than with CRRA utility. Moreover, the larger the elasticity of substitution is, the smaller  $\psi$  and  $1 - \nu$  become, and thus the smaller the volatility of the risk free rate is.

Equation (13) shows that the risk associated to time-varying fundamentals is priced. Indeed, as long as  $\alpha \neq \psi$  the market price of risk vector  $\theta$  consists in two positive terms that depend on the diffusion of the price-dividend ratio. The first term loads on dividend surprises, whereas the second loads on news surprises. As risk aversion  $\alpha$  increases or the elasticity of intertemporal substitution  $\frac{1}{\psi}$  decreases,  $1 - \nu$  rises and hence prices of risk too. To summarize, an increase in the elasticity of intertemporal substitution lowers prices of risk as well as the level and the volatility of the risk free rate. In contrast, an increase in risk aversion implies an increase in prices of risk, a drop in the level of the risk free rate, and an increase in its volatility.

Detailed discussions on stock return volatility and equity risk premia are exposed in Sections 3 and 4, respectively. In Section 6.1 we discuss the average asset pricing moments obtained when attention depends on dividend surprises on the one hand and on return surprises on the other hand.

### 3 Attention, Uncertainty, and Volatility

In the theoretical literature, spikes in volatility have often been related to spikes in uncertainty (Veronesi, 1999; Timmermann, 1993, 2001; Bloom, 2009). In this Section we illuminate a second powerful driver of volatility, namely, investor attention. We uncover a subtle relationship between attention, uncertainty, and volatility. Our main predictions are that the variance of stock returns increases in both attention and uncertainty and de-



pendes quadratically on them. We perform an empirical investigation which lends support to these predictions. Finally, we show that, after controlling for lagged volatility, uncertainty alone cannot explain fluctuations in volatility. Yet, investor attention remains a powerful driver of volatility in that case.

The variance of stock returns follows from Proposition 2:

$$\|\sigma_t\|^2 = \sigma_{2t}^2 + \sigma_{1t}^2 = \left(\frac{I_{\hat{f}}}{I}\right)^2 \sigma_f^2 \Phi_t^2 + \left[\frac{I_{\hat{f}}}{I} \frac{\gamma_t}{\sigma_\delta} + \sigma_\delta \left(1 + \frac{I_\phi}{I}\right)\right]^2. \quad (14)$$

Stock return variance depends on a complex interaction between attention, uncertainty, and investors' price valuations. 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 Equation (10),  $I_{\hat{f}}/I$  and  $I_\phi/I$  are constants—they 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 its case does not deserve further investigation. A discussion is necessary, however, for the term  $I_{\hat{f}}/I$ .

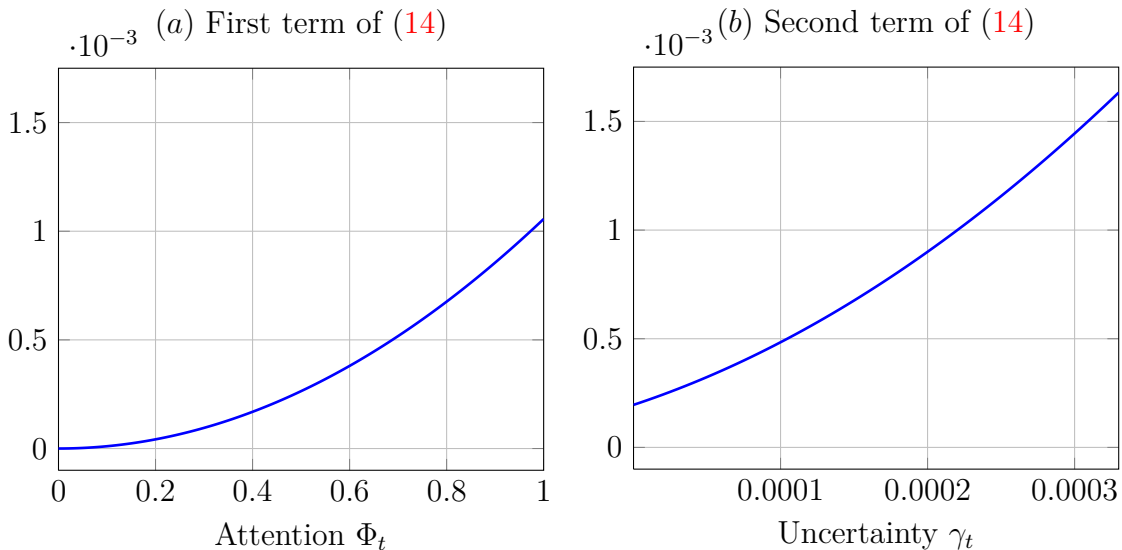
This term mainly depends on the intertemporal elasticity of substitution. If the intertemporal elasticity of substitution is higher than 1, i.e., the investor has preference for early resolution of uncertainty,  $I_{\hat{f}}/I$  is positive. In this case, the asset price increases with an increase in the fundamental. This is the case that we study here, as the price reacts in a plausible fashion to changes in the fundamental. If the intertemporal elasticity of substitution is lower than 1, which would be obtained in a CRRA setting with risk aversion higher than 1,  $I_{\hat{f}}/I$  is negative. In this case, the asset price decreases with an increase in the fundamental. Finally, if the intertemporal elasticity of substitution equals 1, then the asset price does not depend on the fundamental. A similar result obtains with log utility.

The fact that  $I_{\hat{f}}/I$  and  $I_\phi/I$  are constants greatly facilitates our discussion—for the variance of stock returns depends only on attention  $\Phi$  and uncertainty  $\gamma$ .<sup>18</sup> More important, Equation (14) clearly separates the effect of attention (through the first term) and uncertainty (through the second term) on the variance. Finally, there is an obvious quadratic relationship between attention, uncertainty, and stock return variance. Figure 3 depicts the two terms of the variance of stock returns, using parameter values obtained in Section 2.4. As expected, the variance increases in both attention and uncertainty.

As with the variance of the estimated fundamental, the effect of attention on the stock return volatility can also be interpreted in terms of weights. First, as attention increases,

---

<sup>18</sup>Note that if we had considered an exponential quadratic form for the price-dividend ratio instead of an exponential affine form, this statement would be altered as  $I_{\hat{f}}/I$  and  $I_\phi/I$  would become functions of the state variables. Nonetheless, we find in separate calculations that an exponential quadratic form yields exactly the same results.

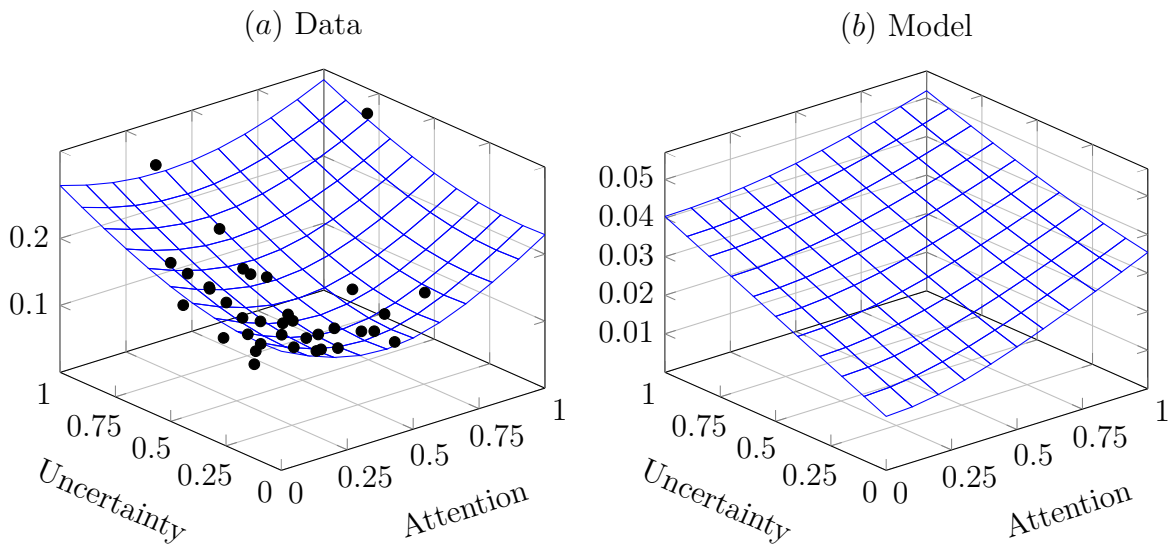


**Figure 3: Decomposition of stock return variance**

Panel (a) depicts the first term of Equation (14), i.e., the quadratic effect of attention  $\Phi$  on the variance of stock returns. Panel (b) depicts the second term of Equation (14), i.e., the quadratic effect of uncertainty  $\gamma$  on the variance of stock returns. The parameter values used for these plots are presented in Table 1.

the investor assigns a higher weight to news, hence the stock return volatility increases by accelerating revelation of news into prices. Second, as attention increases, the investor assigns a lower weight to the dividend, thus decreasing the stock return volatility by incorporating less of the dividend shock into prices. Hence, periods of relatively high attention have the tendency to disconnect the price from dividend shocks and relate it strongly to news. This implication is supported by the recent empirical work of Garcia (2013), who shows that the predictability of stock returns using news' content is concentrated during recessions (i.e., during times of high attention). More precisely, Garcia (2013) finds that one standard deviation shock to news during recessions predicts a change in the conditional average return on the Dow Jones Industrial Average of twelve basis points over one day.

Do we actually observe a quadratic contemporaneous relationship between stock return variance, attention, and uncertainty? We perform an empirical evaluation of this prediction. To 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 2.4, and (iii) a measure of cross-sectional dispersion for quarterly forecasts for real GDP, obtained from the Federal Reserve Bank of Philadelphia. Under reasonable assumptions, the distribution of forecasts will match the distribution of beliefs (see Laster, Bennett, and Geoum, 1999), hence it is common practice to use the cross-sectional dispersion of analyst forecasts as proxy for uncertainty. The Google attention index is available only since 2004, while the cross-sectional dispersion of analyst forecasts is available only at quarterly frequency. This results in a quarterly data set from Q1:2004 to



**Figure 4: Attention, uncertainty, and volatility**

Panel (a) plots data points and quadratic fit for attention, uncertainty, and volatility, resulting from a quarterly data set from Q1:2004 to Q4:2012. Uncertainty (proxied by the cross-sectional dispersion of analyst estimates of real GDP growth) is scaled between 0 and 1. Panel (b) plots the relationship between attention, uncertainty, and volatility resulted from our theoretical model. Uncertainty  $\gamma$  is scaled between 0 and 1, to allow comparison of the two plots. The parameter values for the plot in Panel (b) are presented in Table 1.

Q4:2012.

Panel (a) of Figure 4 plots the data points and a fitted quadratic regression of the variance of stock returns on attention and uncertainty. We represent data in volatility terms for ease of interpretation. Additionally, we scale the values of uncertainty between 0 and 1, to facilitate comparison with the relationship resulting from our theoretical model, which is depicted in Panel (b). Our model does a good job in describing the quadratic relation between attention, uncertainty, and volatility. While our aim is to qualitatively explain this relation, it is worth noting that we do not match the level of the volatility observed in the data. Still, our model generates substantial excess volatility, almost four times higher than the volatility of consumption, which is close to 1% in our calibration.<sup>19</sup>

The coefficients of the quadratic regression between the empirical counterparts of attention, uncertainty, and return variance are presented in Table 2. Specifically, we perform four ordinary least square regressions. We start by regressing stock return variance on attention and uncertainty alone (columns 1 and 2), then we carry out the full regression in

<sup>19</sup>A few extensions of the model could help to match the level of the volatility. For example, we could assume different dividend and consumption dynamics, as in the long-run risk literature, with dividends having a volatility several orders of magnitude higher than the consumption volatility (an assumption backed by the data: the post-war dividend growth volatility has averaged 5% per year). Alternatively, we can also assume that the volatility of consumption growth is stochastic and has a persistent component, as in the long-run risk literature, a desirable property according to Beeler and Campbell (2012). Inevitably, though, these extensions would unnecessarily increase the complexity of our model and the number of parameters, without altering our qualitative predictions.

	(1)	(2)	(3)	(4)
Nb. Obs	36	36	36	35
Intercept	0.036*** (8.13)	0.022*** (4.98)	0.032*** (5.71)	0.008* (1.89)
Attention <sub>t</sub>	-0.074** (-2.64)		-0.073*** (-4.35)	-0.044*** (-3.75)
Attention <sub>t</sub> <sup>2</sup>	0.110*** (3.59)		0.094*** (4.82)	0.086*** (7.28)
Uncertainty <sub>t</sub>		-0.011 (-0.41)	-0.011 (-0.47)	0.014 (0.88)
Uncertainty <sub>t</sub> <sup>2</sup>		0.062** (2.37)	0.057** (2.54)	0.001 (0.04)
Variance <sub>t-1</sub>				0.633*** (10.55)
Adj. R-Squared	0.238	0.442	0.598	0.839

**Table 2: OLS regressions of variance on attention and uncertainty**

There are 4 regressions: (1) variance on attention, (2) variance on uncertainty, (3) variance on attention and uncertainty, and (4) variance on attention, uncertainty, and lagged variance. The dataset comprises 36 datapoints at quarterly frequency from Q1:2004 to Q4:2012. The Newey-West t-statistics are reported in brackets and statistical significance at 10%, 5%, and 1% is labeled respectively with \*/\*\*/\*\*\*.

column 3. Focusing on this full specification, the quadratic coefficients for both attention and uncertainty are significant. The adjusted R-squared, which summarizes the fit while taking into account the number of variables in the model, is close to 0.6. Indeed, attention and uncertainty seem to explain much of the variation of the variance of stock returns.

One might argue that investors become more attentive to the market exactly because volatility increased, and not the other way around. If this is true, then controlling for the lagged variance would eliminate the relationship between attention and variance. We do not find support for this argument. When controlling for the lagged variance (column 4 of Table 2), the coefficients associated to attention are minimally altered and, most important, remain strongly significant. It is worth mentioning, however, that uncertainty loses its explanatory power. This shows that uncertainty alone does not seem to have supplementary explanatory power with respect to lagged variance, whereas attention does.<sup>20</sup>

When regressing return variance on attention alone (column 1 of Table 2), the linear coefficient is negative and significant. This might seem puzzling, but the reason is simple. The lead-lag relationship between attention and uncertainty suggests that attention has two opposite effects on the return variance. First, the variance increases quadratically with attention by speeding up revelation of news into prices. Second, higher attention

<sup>20</sup>As an additional test, we also regressed attention on lagged variance alone. The slope coefficient obtained is not significant and the adjusted R-squared is 0.003. Although our aim is to explain the *contemporaneous* relationship between attention and volatility, we believe we have sufficient evidence to conclude that lagged volatility is not driving attention.

means better learning (lower uncertainty), which tends to decrease the variance. Hence the negative sign for the linear coefficient. We conclude then that the signs of the coefficients in the first regression have a meaningful economic interpretation.

Our decomposition of the variance from Equation (14) bears similarities with the studies of Brennan and Xia (2001) and Veronesi (2000). Brennan and Xia (2001) show in an equilibrium model with incomplete information that learning increases market volatility beyond what one obtains in an economy with complete information. We share the same result; in our case, complete information yields zero uncertainty which, in turn, decreases volatility. One can see this in Panel (b) of Figure 4: volatility for high attention and low uncertainty (complete information) is below the volatility for low attention and high uncertainty (incomplete information). But while in Brennan and Xia (2001) the variance of stock returns is constant—because attention and uncertainty are themselves constant—in our case it fluctuates with movements in attention and uncertainty. And we actually obtain higher volatilities when attention can fluctuate; see, for instance, the volatility implied by our model in states of high attention and high uncertainty. Thus, complementary to Brennan and Xia (2001), our focus is on how much attention is devoted to learning and how this impacts the dynamics of the volatility of stock returns.

Veronesi (2000) builds an equilibrium model with learning in which the unobservable fundamental is driven by a continuous-time Markov chain. Investors' degree of uncertainty is reflected in the term  $V_\theta$ , which is equivalent to  $I_{\hat{f}}\gamma/I$  in our case. Similar to our study, Veronesi (2000) obtains a quadratic relationship between stock return variance and uncertainty, and he goes on to derive—albeit only through comparative statics—asset pricing implications of information quality. While Veronesi (2000) keeps the *quality* of information constant, our main focus is exactly on the dynamics of this variable (investor attention in our case). Not only we show how it affects uncertainty, but, more important, how both attention and uncertainty explain variations in stock market volatility.

David (1997) and Veronesi (1999) develop dynamic rational expectation models in which uncertainty is fluctuating because the unobserved fundamental is assumed to shift between high-growth and low-growth states. This discreteness of states implies a stochastic variance for the estimated growth rate. We also obtain a stochastic variance for the estimated growth rate. Our contribution with respect to David (1997) and Veronesi (1999) is to decompose this variance in two components. The first component is driven by attention and higher attention implies faster learning. The second component is driven by uncertainty, which itself depends on the learning speed. How much of this stochastic variance is driven by attention and how much by uncertainty depends endogenously on our specific intertemporal behavior of investors' beliefs.

To conclude, our study is—to the best of our knowledge—the first to show how attention and uncertainty simultaneously drive stock market volatility. Our theoretical model predicts a clear quadratic relationship between attention, uncertainty, and stock return

variance, and the data lends support to this prediction. Furthermore, we show that uncertainty alone cannot explain fluctuations in volatility when controlling for lagged volatility, yet attention remains a powerful driver .

## 4 Attention, Uncertainty, and Risk Premium

Expectations matter for the risk premium for two reasons. First, they show how fast investors believe the economy will grow. This is reflected in how the assessed fundamental,  $\hat{f}_t$ , enters in investors' price valuations. Second, expectations can be volatile—if investors can't see a clear road ahead, they are most likely going to require a higher risk premium. This is reflected in the variance of the assessed fundamental,  $\sigma^2(\hat{f}_t)$ , which, in turn, is a mixture of attention and uncertainty. Our model offers the advantage of conveniently separating all these effects.

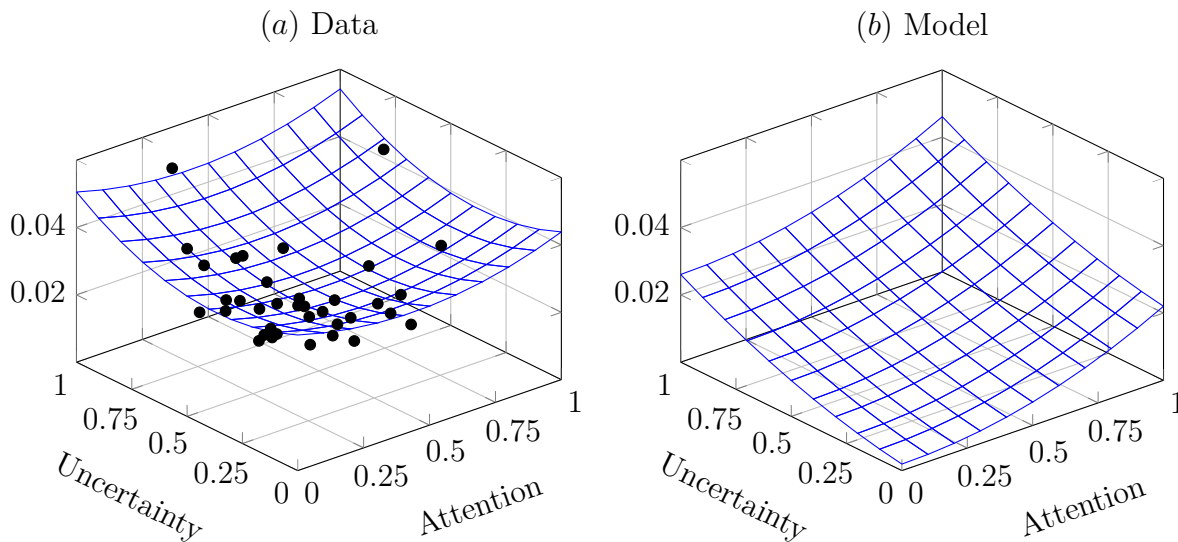
The risk premium follows from Proposition 2

$$\begin{aligned} \mu_t - r_t = & \alpha \sigma_\delta \left[ \frac{I_{\hat{f}} \gamma_t}{I \sigma_\delta} + \sigma_\delta \left( 1 + \frac{I_\phi}{I} \right) \right] \\ & + (1 - \nu) \left[ \sigma_\delta \left( \frac{I_{\hat{f}} \gamma_t}{I \sigma_\delta} + \sigma_\delta \right) + \left( \frac{I_{\hat{f}} \gamma_t}{I \sigma_\delta} + \sigma_\delta \right)^2 + \left( \frac{I_{\hat{f}}}{I} \right)^2 \sigma_f^2 \Phi_t^2 \right]. \end{aligned} \quad (15)$$

Equation (15) has two terms. Only the first term matters in the CRRA case, i.e., when  $\nu = 1$ . Still in the CRRA case, if the risk aversion coefficient is higher than 1, stock prices fall with expected dividend growth, and thus  $I_{\hat{f}}/I < 0$ . Hence, as uncertainty gets higher, the first term of (15) gets lower, and thus the risk premium gets lower. Moreover, a higher risk aversion amplifies this effect. Veronesi (2000) obtains a similar result: for a coefficient of risk aversion higher than 1, lower uncertainty increases the risk premium, whereas the opposite holds for a coefficient of risk aversion lower than 1.

A different story emerges when the intertemporal elasticity of substitution and the coefficient of relative risk aversion are treated separately. First, as soon as  $\nu \neq 1$ , the second term in Equation (15) becomes relevant. Second, if both the coefficient of risk aversion and the intertemporal elasticity of substitution are higher than 1, then  $I_{\hat{f}}/I > 0$  and  $(1 - \nu) > 0$ . This in turn implies that the risk premium increases with uncertainty, and both the first and the second term in Equation (15) contribute positively to the risk premium. In other words, the risk premium increases with both attention and uncertainty. The relationship between risk premium, attention, and uncertainty is, once again, quadratic.

We perform an empirical evaluation of these predictions. The analysis is similar with the one from Section 3, with the main difference that we replace the volatility with risk premium. For equity risk premium values, we rely on quarterly surveys reported



**Figure 5: Attention, uncertainty, and risk premium**

Panel (a) plots data points and quadratic fit for attention, uncertainty, and risk premium, resulting from a quarterly data set from Q1:2004 to Q4:2012. Uncertainty (proxied by the cross-sectional dispersion of analyst estimates of real GDP growth) is scaled between 0 and 1. Panel (b) plots the relationship between attention, uncertainty, and risk premium resulted from our theoretical model. Uncertainty  $\gamma$  is scaled between 0 and 1, to allow comparison of the two plots. The parameter values for the plot in Panel (b) are presented in Table 1.

by [Graham and Harvey \(2013\)](#). These 17,500 survey responses over more than 10 years (averaging 352 individual responses of U.S. CFOs each quarter) offer a clear measure of expectations. And this dataset perfectly fits our needs: the surveys were conducted from June 2000 to December 2012. For the attention and uncertainty we use the same proxies as in previous section, resulting in a quarterly data set from Q1:2004 to Q4:2012.

Panel (a) of Figure 5 plots the data points and a fitted quadratic regression of risk premium on attention and uncertainty. As in the previous section, we scale the values of uncertainty between 0 and 1, to facilitate comparison with the relationship resulting from our theoretical model, which is depicted in Panel (b). There is indeed a quadratic relation between attention, uncertainty, and risk premia, and our model is able to capture it.

Two things are worth mentioning about the level of risk premia resulting from our model. First, we are able to match the overall level of risk premia collected by [Graham and Harvey \(2013\)](#). Second, both the case of perfect learning, i.e., point  $\{\Phi = 1, \gamma = 0\}$  on the plot, and the case of no learning, i.e., point  $\{\Phi = 0, \gamma = 1\}$  on the plot, generate lower risk premia than a wide range of cases characterized by relatively high attention and high uncertainty. Not only this feature will prove to be crucial for our analysis of the dynamic properties of risk premia and volatility, in Section 5, but also it clearly distinguishes our paper from the rest of the theoretical literature. Specifically, while other papers usually perform comparative statics of cases of learning vs perfect information (see, for example, [Brennan and Xia, 2001](#), [Veronesi, 2000](#), or [Ai, 2010](#)), we



	(1)	(2)	(3)
Nb. Obs	36	36	36
Intercept	0.036*** (13.19)	0.037*** (11.48)	0.040*** (9.91)
Attention <sub>t</sub>	-0.022 (-1.56)		-0.023* (-1.87)
Attention <sub>t</sub> <sup>2</sup>	0.030** (2.21)		0.027** (2.15)
Uncertainty <sub>t</sub>		-0.028** (-2.08)	-0.029** (-2.25)
Uncertainty <sub>t</sub> <sup>2</sup>		0.039*** (3.29)	0.039*** (3.27)
Adj. R-Squared	0.086	0.210	0.278

**Table 3: OLS regressions of risk premia on attention and uncertainty**

There are 3 regressions: (1) risk premia on attention, (2) risk premia on uncertainty, and (3) risk premia on attention and uncertainty. The dataset comprises 36 datapoints at quarterly frequency from Q1:2004 to Q4:2012. The Newey-West t-statistics are reported in brackets and statistical significance at 10%, 5%, and 1% is labeled respectively with \*/\*\*/\*\*\*.

are able to show that learning with fluctuating attention can generate risk premia and volatilities beyond levels obtained with constant attention.

Table 3 confirms that the relationship between risk premia, attention, and uncertainty is indeed quadratic. Precisely, column 3 shows that all coefficients are significant and the adjusted R-squared is close to 0.3. Uncertainty seems to be a stronger driver of risk premium, yet attention adds explanatory power. The coefficients of  $\Phi$  from columns 1 and 3 (i.e., the linear coefficients of attention) reveal an interesting result. The first is not significant while the second is only marginally significant at 10%. Close inspection of Equation (15) suggests why this is the case. Attention enters in the definition of risk premium only with a quadratic term. The practical implication of this is that one should expect to obtain a not significant coefficient, which turns out to be the case.

Building on David (1997) and Veronesi (1999), Ozoguz (2009) tests whether investor require a risk premium to be compensated for high uncertainty and finds support for this prediction, but she does not find a significant relationship between uncertainty and market volatility. Massa and Simonov (2005) confirm that uncertainty is priced. Our paper differs from Ozoguz (2009) and Massa and Simonov (2005) in its emphasis, as our main focus is the fluctuating attention and its effects on risk premium and volatility. We find that investor attention is the main driver of volatility, which could explain the weak relationship between uncertainty and volatility documented by Ozoguz (2009). Concerning the risk premium, we find that uncertainty indeed is an important factor, but we also find that attention adds explanatory power. Most important, the relationship between attention,

uncertainty, and risk premium is not linear but quadratic, and we find strong support for this in the data.

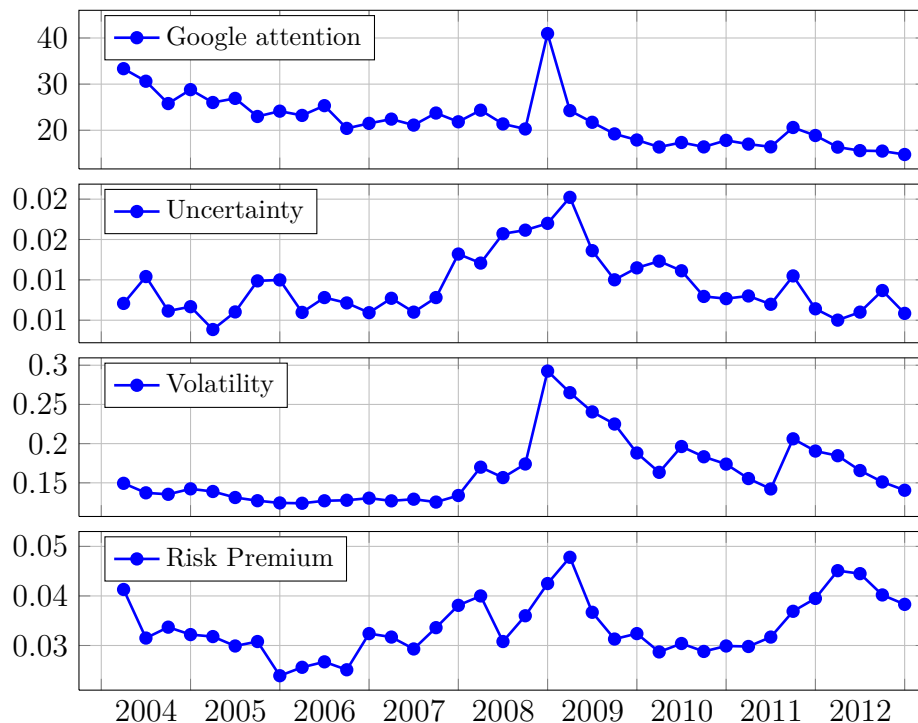
Finally, we would like to emphasize that our paper investigates only the “induced uncertainty” generated by learning. Real life is probably messier than that; there is uncertainty about government policy (Pastor and Veronesi, 2012), uncertainty about monetary policy, uncertainty about regulatory policy, uncertainty about foreign policy, uncertainty about U.S. fiscal policy, uncertainty about the national debt, and the list goes on. Plausible extensions of our model could incorporate exogenous movements in uncertainty, exactly as the long-run risk model has introduced exogenous time-varying risk in the fundamentals (Bansal and Yaron, 2004; Bloom, 2009), while maintaining fluctuating attention as a critical component. Although these extensions are expected to provide additional insights, we leave them for future work and keep the focus of our paper on the clean effect of learning uncertainty on asset prices.

## 5 Accumulation of Uncertainty and Panic States

Our model generates strong effects on the dynamics of volatilities and risk premia, especially during times of crisis. These effects are the result of the joint dynamics of investor attention and uncertainty, which give rise to the lead-lag relation described in Section 2.3. We show that the economy can go at times in “panic states,” or states of high attention and high uncertainty. These states are characterized by spikes in volatility and risk premium, as we witnessed during the turmoil in the Fall of 2008.

To begin with, consider a relatively long period of low investor attention. During this period, investors are less concerned about the economy and therefore pay less attention to news. This, in turn, propagates an accumulation of uncertainty, through the lead-lag relation between  $\Phi$  and  $\gamma$ : the less attentive investors are, the more uncertain the economy becomes. It suffices a few notable events grabbing investor attention to bring fragility into focus. The economy switches suddenly to a panic state, i.e., a state of high attention and high uncertainty. Consequently, volatility and risk premium spike. Since now investors pay attention to news, uncertainty necessarily goes down through the lead-lag relation between  $\Phi$  and  $\gamma$ . This gradual reduction of uncertainty lowers the volatility and risk premium. Ultimately, as the crisis unwinds and recovery takes hold, investors become less concerned about the economy, and we are back to square one.

We illustrate this intuition in Figure 6, where we plot four time series: (i) investor attention, measured by the Google search volume index, (ii) uncertainty, measured by the dispersion of analyst forecasts of real GDP growth, (iii) stock market volatility, obtained through a GARCH(1,1) estimation from the S&P500 returns, and (iv) risk premium obtained from quarterly surveys reported by Graham and Harvey (2013). These quarterly time series are aligned at same dates in order to facilitate interpretation.



**Figure 6: Accumulation of uncertainty and panic states**

The figure plots 4 time series: (i) investor attention, measured by the Google search volume index, (ii) uncertainty, measured by the dispersion of analyst forecasts of real GDP growth, (iii) stock market volatility, obtained through a GARCH(1,1) estimation from the S&P500 returns, and (iv) risk premium obtained from quarterly surveys reported by Graham and Harvey (2013). Series are at quarterly frequency and aligned at same dates.

Uncertainty and attention were both low before 2007, consistent with what economists called the “great moderation.” In 2007, however, uncertainty rises gradually, while attention remains low. During most of 2008, attention remains low but uncertainty keeps rising, bringing up both volatilities and risk premia. In the Fall of 2008, a spike in investor attention brings volatilities and risk premia to all time highs. Then, attentive investors gradually solve the uncertainty, which brings back down volatilities and risk premia. New spikes, although of smaller magnitude, occur in 2011 and 2012 in connection with the European debt crisis.

An additional point to make here is that volatility and risk premium depend differently on attention and uncertainty. For example, in Section 3 we show that attention is the main driver of volatility, whereas in Section 4 we show that uncertainty is the main driver of risk premium. This should explain why we observe high risk premia and low volatility at the end of 2007 and the end of 2012: both times are characterized by high uncertainty and low attention, leading to an apparent disconnect between volatility and risk premium.

Although our model is far less complex than actual financial markets, at a qualitative level it does connect to events observed in recent years. After all, our model is only a pure exchange economy with a representative agent who learns with a variable attention—an

accurate description of financial markets would probably be too much to ask for such a standard model. What we want to emphasize is that investor attention is indeed a critical component and our model shows exactly why this is the case. A reasonable extension, which probably would make the model even more realistic, is to assume that attention does not only depend on economic conditions, but also on important events, such as the Lehman Brothers bankruptcy. One way or another, the message remains the same: spikes in investor attention in highly uncertain times might be the trigger of financial panics.

Finally, assuming a brief disconnect between investor attention and economic conditions could also help explain the events that occurred in the mid- to late 1990s. While the stock market was rising, considerable attention was paid to the internet sector, fuelling the technology boom. But this period of “irrational exuberance,” characterized by highly volatile returns, was quite uncertain given the large unpredictability of new technologies (Pastor and Veronesi, 2006). Ultimately, the U.S. Federal Reserve increased interest rates several times, the economy began to lose speed, Nasdaq’s profitability plummeted—triggering probably a spike in investors’ attention—and the dot-com bubble burst in March 2000.

A related paper trying to rationalize fluctuations of uncertainty is Timmermann (2001). In Timmermann (2001)’s model, observable structural breaks generate a spike in uncertainty and push agents to filter out the new (unknown) value of the fundamental. As new data comes in, uncertainty is gradually reduced. Timmermann (2001)’s model bears a similar flavor with ours, in that learning gradually reduces uncertainty. In our case, a period of high attention gradually reduces uncertainty. We also provide the flip side of this argument: periods of low attention propagate accumulation of uncertainty, a precursor of panic states. Our model of fluctuating attention is therefore the origin of rich dynamics of uncertainty, volatilities, and risk premia.

Bacchetta, Tille, and van Wincoop (2012) develop a model of time-varying risk in the fundamentals. They show that, if the asset price depends negatively on its future variance, any fundamental variable which affects this future variance may generate a circular relationship and thus a self-fulfilling shift in risk. This fundamental variable becomes a focal point of fear in the market during a panic. Similarly, in our model news become a focal point during periods of high attention. The key aspect of our model is, however, different from theirs. While in Bacchetta et al. (2012) the time-varying risk is self-fulfilling and appears in *sunspot equilibria*, in our model the time-varying risk is determined by fluctuations in investor attention and appears in a pure exchange economy equilibrium.

## 6 Extensions and Other Comments

This Section is dividend in two parts. First, we consider the extension of our model to the case when attention depends on stock return surprises. Second, we perform two tests

that further sustain counter-cyclical attention and tend to reject the potential assumption that noise in the signals is counter-cyclical.

## 6.1 The Case When Attention Depends on Return Surprises

The performance index depends now on past return surprises

$$\phi_t = \int_0^t e^{-\omega(t-u)} \left( \frac{dS_u + \delta_u du}{S_u} - \mu_u du \right). \quad (16)$$

Applying Itô's lemma to Equation (16) yields the following dynamics

$$d\phi_t = -\omega\phi_t dt + \sigma_t dW_t,$$

where  $\sigma \equiv (\sigma_{1t} \ \sigma_{2t})$  is the stock return diffusion vector.

The relation between attention and the performance index is provided in Definition 1. As before we approximate the price-dividend ratio by the exponential form exposed in Equation (10). Applying Itô's lemma to that equation yields the following functional form for the stock return diffusion

$$\sigma_t \equiv (\sigma_{1t} \ \sigma_{2t}) = \left( \sigma_\delta + \beta_{11} \frac{\gamma_t}{\sigma_\delta} + \beta_{12} \sigma_{1t} \quad \beta_{11} \sigma_f \Phi_t + \beta_{12} \sigma_{2t} \right).$$

Solving for  $\sigma_1$  and  $\sigma_2$  yields

$$\sigma_t = \left( \frac{\beta_{11}\gamma_t + \sigma_\delta^2}{(1-\beta_{12})\sigma_\delta} \quad \frac{\beta_{11}\sigma_f\Phi_t}{1-\beta_{12}} \right).$$

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 2.5.1 yields the price-dividend ratio.

When attention depends on return surprises, the price-dividend ratio satisfies

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

where

$$\begin{aligned} \beta_0 &= 3.6156 & \beta_{11} &= 1.1207 \\ \beta_{12} &= -0.0006 & \beta_{13} &= -11.3618. \end{aligned} \quad (17)$$

The coefficients exposed in (11) and (17) show that the price-dividend ratio doesn't depend on whether attention is driven by dividend or return surprises. Moreover, the diffusion of the price-dividend ratio is also very similar in both cases because  $\beta_{12}$  is small. Consequently, the levels of the price-dividend ratio, risk free rate, stock return volatility,

Parameter	Symbol	(a) Estimate based on dividend surprises	(b) Estimate based on return surprises
Mean price-dividend	$E(I)$	38.24	38.26
Volatility price-dividend	$\text{Vol}\left(\frac{dI}{I}\right)$	2.73%	2.74%
Mean risk premium	$E(\mu - r)$	2.56%	2.46%
Volatility risk premium	$\text{Vol}(d(\mu - r))$	1.27%	2.52%
Mean risk free rate	$E(r)$	2.93%	2.96%
Volatility risk free rate	$\text{Vol}(dr)$	1.77%	2.24%
Mean return volatility	$E( \sigma )$	3.98%	3.92%
Volatility of return volatility	$\text{Vol}(d \sigma )$	0.83%	1.37%

**Table 4: Unconditional Moments**

Columns (a) and (b) show the unconditional asset pricing moments when attention depends on dividend surprises and return surprises, respectively. 20'000 simulations of daily data are performed over a 100 years horizon. First, averages are computed over each time-series to provide 20'000 estimates. Then, the median of the 20'000 estimates is reported above.

market prices of risks, and risk premia are very close to each other in both cases.

Table 4 exposes the unconditional asset pricing moments resulting from 20'000 simulations of daily data considered over an horizon of 100 years. This table confirms 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.

Because the performance index is more volatile when it depends on return surprises, attention becomes also more volatile in that case. Consequently, the variables that are significantly driven by attention inherit a larger volatility when the performance index depends on return surprises. As shown in Proposition 2 and Equations (14) and (15), these variables are the market prices of risk, the risk free rate, the stock return volatility, and the risk premium. Table 4 confirms that the volatility of risk premia almost doubles when we move from dividend to return surprises. While Chien et al. (2012) show that the presence of “infrequent traders” helps explain the observed volatility of risk premia, our complementary explanation resides in the fact that investors become more attentive to news when return surprises are negative, and vice versa. Moreover, the volatility of volatility increases by a factor of 1.65, whereas the volatility of the risk free rate rises by a factor of 1.27 only.

## 6.2 Attention versus Noise in the Signals

Our estimation implies that attention is higher in bad economic times than in good economic times ( $\Lambda > 0$ ). The question is: do we really extract from the data counter-

	(1)	(2)	(3)
	All Sample	NBER Rec.	NBER Exp.
Nb. Obs	176	29	147
Intercept	0.002 (0.50)	-0.015* (-1.88)	0.015*** (3.23)
Slope	0.974*** (7.52)	0.859** (2.22)	0.634*** (4.31)
R-Squared	0.245	0.155	0.113
Wald p-value	0.821	0.093	0.003

**Table 5: Test of forecast optimality: NBER recessions vs. NBER expansions**

The Mincer and Zarnowitz (1969) regression specification is  $Y_{t+1} = a_0 + a_1 \hat{Y}_{t+1,t} + u_{t+1}$ , where  $Y_{t+1}$  is the predicted variable (real GDP growth in our case) and  $\hat{Y}_{t+1,t}$  is the forecast based on information available at time  $t$  ( $\hat{f}_t$  in our case). The t-statistics are in parenthesis below the coefficients. The Wald test specification is the joint restriction  $\{a_0 = 0, a_1 = 1\}$ . Statistical significance at 10%, 5%, and 1% is labeled respectively with \*/\*\*/\*\*\*.

cyclical attention or rather counter-cyclical noise<sup>21</sup> in the signals? Intuitively, more noise in the signals generates more uncertainty, and consequently larger risk premia and return volatility as in our fluctuating attention model (see Sections 3 and 4).

In order to differentiate counter-cyclical attention and counter-cyclical noise in the signals, we perform two tests. The first consists in investigating whether analysts forecasts are more accurate in bad times than in good times. If attention is counter-cyclical, then forecasts should be more accurate in recessions than in expansions, whereas the reverse assertion should hold if the noise in the signals is counter-cyclical.

The first test is performed as follows. We use Mincer and Zarnowitz (1969) regressions of realized real GDP growth on a constant and the corresponding analyst forecast. The testable implication of these regressions is that the associated coefficients should be 0 and 1, i.e., that the forecast error is conditionally (and unconditionally) unbiased.<sup>22</sup>

Accordingly, we use NBER monthly recession indicators to divide our sample in two parts: recessionary and expansionary periods. Given that the Philadelphia Fed sends the questionnaires and the advance report from Bureau of Economic Analysis to panelists at the end of the first month of each quarter, we use the NBER indicator exactly at that date in order to decide if we are in a recession or an expansion. Then, we perform Mincer and Zarnowitz (1969) regressions for the whole sample and the two resulting subsamples. In each case, we perform Wald tests of the joint restriction on the intercept and the slope. The results are presented in Table 5.

<sup>21</sup>Note that the signal would be defined by  $ds_t = f_t dt + \sigma_t^s dZ_t^s$ , with  $\sigma^s$  counter-cyclical. Bansal and Shaliastovich (2010) consider this type of signal specification and fit the variance of  $s$  to observed analyst forecast dispersion using a jump-diffusion model.

<sup>22</sup>See Patton and Timmermann (2007) for a discussion and an application of Mincer and Zarnowitz (1969) regressions.



When we use the entire sample (first column of Table 5), the data tells us that predictions are quite accurate. The Wald joint test of the coefficients cannot be rejected. When we split the data in NBER recessions (column 2) and NBER expansions (column 3), we observe a higher R-squared for the recession sample. More important, the Wald test cannot be rejected at 5% confidence level in the recession sample, while it is rejected in the expansion sample. As such, these regressions should be viewed as evidence that we observe better forecasts in bad times, offering support to a positive parameter  $\Lambda$  and thus to counter-cyclical attention.<sup>23</sup> Also, this first test rejects the assumption that signals are more noisy in recessions than in expansions.

The second test consists in looking at the interdependence between returns and news shocks in recessions and in expansions. If attention is counter-cyclical, then returns should be more correlated to news shocks in recessions than in expansions. In contrast, if the noise in the signals is counter-cyclical, then returns should be less correlated to news shocks in recessions than in expansions.

Figure 7 depicts the model implied slope  $\beta^{RN}$  and correlation coefficient  $\rho^{RN}$  obtained by regressing returns on news shocks. These coefficients satisfy

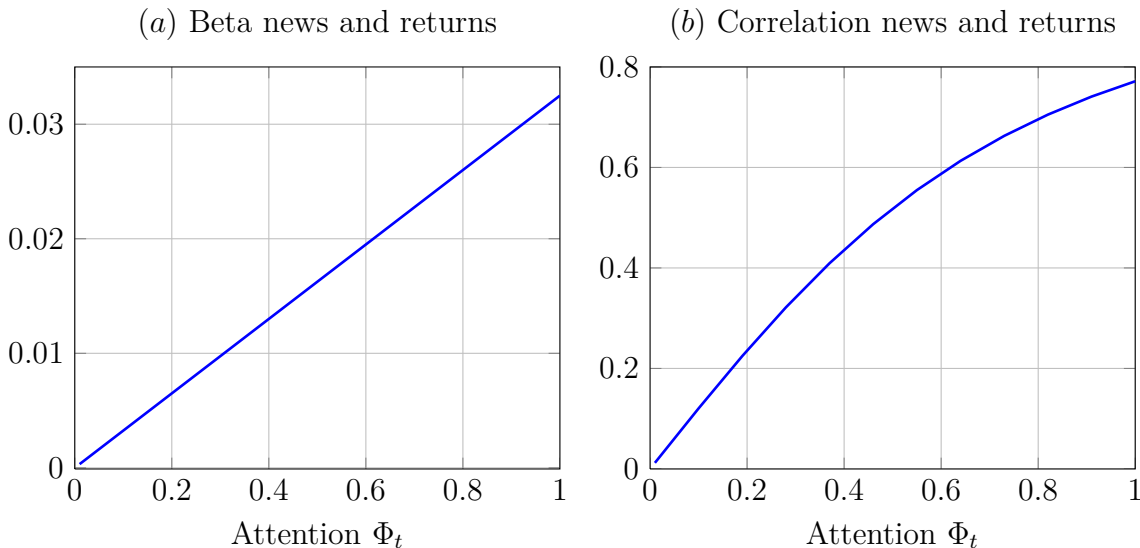
$$\beta_t^{RN} \equiv \frac{\text{Cov}_t \left( \frac{dS_t + \delta_t dt}{S_t}, dW_t^s \right)}{\text{Var}_t (dW_t^s)} \quad (18)$$

$$\rho_t^{RN} \equiv \frac{\text{Cov}_t \left( \frac{dS_t + \delta_t dt}{S_t}, dW_t^s \right)}{\text{Vol}_t (dW_t^s) \text{Vol}_t \left( \frac{dS_t + \delta_t dt}{S_t} \right)}. \quad (19)$$

Figure 7 confirms that both the slope and the correlation increase with attention. Hence the model implied interconnection between news and returns is strong in bad times and weak in good times, consistent with the empirical findings in Garcia (2013). In contrast, counter-cyclical noise in the signals implies that the weight assigned by the investor to news shocks declines in bad times, making returns less sensitive to news in recessions than in expansion. The latter prediction is inconsistent with Garcia (2013), who shows that Dow Jones Industrial Average returns are significantly more connected to Wall Street Journal news in recessions than in expansions. To summarize, our two tests lend strong support to counter-cyclical attention and reject counter-cyclical noise in the signals.

---

<sup>23</sup>As an additional exercise, we also divide our sample in periods when the Google attention index is higher than its mean and periods when it is lower than its mean. The adjusted R-squared nearly doubles (from 0.16 to 0.30) when moving from the low to the high attention state. Also, note that Van Nieuwerburgh and Veldkamp (2006) find that the absolute forecast error on the level of the nominal GDP tends to be larger in bad times than in good times. This finding does absolutely not contradict our results given that we focus on the growth rate of real GDP.



**Figure 7: Connection between news shocks and returns**

Panel (a) depicts the slope coefficient  $\beta^{RN}$  obtained by regressing returns on news shocks (see definition in Equation 18). Panel (b) depicts the correlation,  $\rho^{RN}$ , between returns and news shocks (see definition in Equation 19). The parameter values used for these plots are presented in Table 1.

## 7 Concluding Remarks

We have developed a simple setup to show how investors’ attention and learning uncertainty affect *simultaneously* the dynamics of asset returns. The model predicts that volatilities and risk premia increase quadratically with attention and uncertainty. These predictions are supported by our empirical analysis. Furthermore, investors’ learning generates a lead-lag relation between attention and uncertainty. We have shown that this feature can yield “panic states,” when stock prices are volatile and investors demand a high risk premia. This connects our model to events observed in recent years, such as the turmoil in the Fall of 2008.

We hope this paper makes a useful step forward in the important task of understanding the effect of learning on asset prices. No doubt that more complicated extensions of the present model deserve to be addressed in future research. Here are a few ideas.

For instance, and as we mentioned earlier, investors’ processing of information encompasses several dimensions, including dispersion of beliefs. Massa and Simonov (2005) show that both learning uncertainty and dispersion of beliefs are priced. It is therefore of interest to integrate both of them in the same setup. 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. We could observe states of high uncertainty associated with wildly divergent views. Recent results from Carlin, Longstaff, and Matoba (2013) indicate that this might be a fruitful research avenue. They find that increased disagreement is associated with higher expected re-

turns, higher return volatility, and larger trading volume. Moreover, disagreement is time varying and increases during periods of extreme uncertainty.

Learning can also be biased. For example, investors could be overconfident, as in [Dumas et al. \(2009\)](#). Moreover, high uncertainty probably feeds biased expectations. This could further increase the volatility of asset prices and, when expectations turn out to be wrong, could generate major corrections of the stock market. One such episode is the Nasdaq “bubble,” when investors overlooked traditional valuation metrics such as the P/E ratio in favor of (over)confidence in technological advancements.

We believe that a richer setup with leveraged firms and market liquidity can provide additional insights. Extended periods of low attention and low uncertainty, or “great moderations,” may cause firms to be less concerned about liquidity and to hold less capital. This, in turn, encourages increased debt levels and enables a period of financial instability. When uncertainty and attention start rising again, leverage and market liquidity collapse, and the economy enters into a financial panic.

Finally, [Karlsson et al. \(2009\)](#) suggest that, because retail investors temporarily ignore their portfolios, liquidity dries up exactly in major market downturns. Examples provided are the Asian crisis of 1997, the Russian debt default in 1998, and the credit crunch of 2008. If these investors are believed to be “unsophisticated,” then their failure to rebalance their portfolios can help to explain the counter-cyclical volatility of aggregate risk (see [Chien et al. 2012](#), [Duffie 2010](#), and other papers in the literature on slow moving capital). Fluctuations in investors’ attention to news can only reinforce these effects, providing an interesting connection between attention to wealth and attention to news.

# References

- [1] Abel, A., J. Eberly, and S. Panageas (2007). Optimal inattention to the stock market. *American Economic Review* 97, 244–249. [5, 8]
- [2] Ai, H. (2010, 08). Information quality and long-run risk: Asset pricing implications. *Journal of Finance* 65(4), 1333–1367. [3, 28]
- [3] Bacchetta, P., C. Tille, and E. van Wincoop (2012, December). Self-fulfilling risk panics. *American Economic Review* 102(7), 3674–3700. [32]
- [4] Bacchetta, P. and E. v. Wincoop (2010). Infrequent portfolio decisions: A solution to the forward discount puzzle. *American Economic Review* 100(3), 870–904. [8]
- [5] Bansal, R. and I. Shaliastovich (2010, September). Confidence risk and asset prices. *American Economic Review* 100(2), 537–41. [35]
- [6] Bansal, R. and I. Shaliastovich (2011). Learning and asset-price jumps. *Review of Financial Studies* 24(8), 2738–2780. [5, 16]
- [7] Bansal, R. and A. Yaron (2004, 08). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59(4), 1481–1509. [3, 10, 13, 14, 30]
- [8] Barsky, R. B. and J. B. De Long (1993, May). Why does the stock market fluctuate? *The Quarterly Journal of Economics* 108(2), 291–311. [14]
- [9] Beeler, J. and J. Y. Campbell (2012). The long-run risks model and aggregate asset prices: An empirical assessment. *Critical Finance Review* 1, 141–182. [14, 24]
- [10] Benzoni, L., P. Collin-Dufresne, and R. S. Goldstein (2011). Explaining asset pricing puzzles associated with the 1987 market crash. *Journal of Financial Economics* 101(3), 552 – 573. [18, 19, 20]
- [11] Bloom, N. (2009, 05). The impact of uncertainty shocks. *Econometrica* 77(3), 623–685. [21, 30]
- [12] Brennan, M. J. and Y. Xia (2001). Stock price volatility and equity premium. *Journal of Monetary Economics* 47(2), 249 – 283. [2, 3, 26, 28]
- [13] Carlin, B. I., F. A. Longstaff, and K. Matoba (2013). Disagreement and asset prices. *Working Paper*. [37]
- [14] Chauvet, M., S. Gabriel, and C. Lutz (2012). Fear an loathing in the housing market: Evidence from search query data. *Working Paper*. [4]
- [15] Chien, Y., H. Cole, and H. Lustig (2012, September). Is the volatility of the market price of risk due to intermittent portfolio rebalancing? *American Economic Review* 102(6), 2859–96. [8, 34, 38]
- [16] Cox, J. C., J. Ingersoll, Jonathan E, and S. A. Ross (1985, March). A theory of the term structure of interest rates. *Econometrica* 53(2), 385–407. [16]
- [17] Da, Z., J. Engelberg, and P. Gao (2011, October). In search of attention. *Journal of Finance* 66(5), 1461–1499. [1, 4, 5, 13]
- [18] Da, Z., U. G. Gurun, and M. Warachka (2011). Frog in the pan: Continuous information and momentum. *SSRN eLibrary*. [15]
- [19] David, A. (1997, December). Fluctuating confidence in stock markets: Implications for returns and volatility. *Journal of Financial and Quantitative Analysis* 32(04), 427–462. [2, 3, 26, 29]
- [20] David, A. (2008). Heterogeneous beliefs, speculation and the equity premium. *Journal of Finance* 63, 41–83. [12]
- [21] Detemple, J. and R. Kihlstrom (1987). Acquisition dinformation dans un modèle intertemporel en temps continu. *L’Actualité économique* 63, 118–137. [5, 16]

- [22] Detemple, J. B. (1986, June). Asset pricing in a production economy with incomplete information. *Journal of Finance* 41(2), 383–91. [2, 3, 5]
- [23] Detemple, J. B. (1991, July). Further results on asset pricing with incomplete information. *Journal of Economic Dynamics and Control* 15(3), 425–453. [3]
- [24] Dimpfl, T. and S. Jank (2011). Can internet search queries help to predict stock market volatility? University of Tübingen. [1, 4]
- [25] Dothan, M. U. and D. Feldman (1986, June). Equilibrium interest rates and multiperiod bonds in a partially observable economy. *Journal of Finance* 41(2), 369–82. [3, 5]
- [26] Duffie, D. (2010, 08). Presidential address: Asset price dynamics with slow-moving capital. *Journal of Finance* 65(4), 1237–1267. [8, 17, 38]
- [27] Duffie, D. and L. G. Epstein (1992a). Asset pricing with stochastic differential utility. *Review of Financial Studies* 5(3), 411–36. [18]
- [28] Duffie, D. and L. G. Epstein (1992b, March). Stochastic differential utility. *Econometrica* 60(2), 353–94. [18]
- [29] Duffie, D. and T.-s. Sun (1990, February). Transactions costs and portfolio choice in a discrete-continuous-time setting. *Journal of Economic Dynamics and Control* 14(1), 35–51. [8]
- [30] Dumas, B., A. Kurshev, and R. Uppal (2009). Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility. *The Journal of Finance* 64(2), 579–629. [2, 5, 10, 38]
- [31] Epstein, L. G. and S. E. Zin (1989, July). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica* 57(4), 937–69. [17]
- [32] Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance* 68(3), 1267–1300. [15, 23, 36]
- [33] Gennotte, G. (1986, July). Optimal portfolio choice under incomplete information. *Journal of Finance* 41(3), 733–46. [3, 5]
- [34] Graham, J. R. and C. R. Harvey (2013). The equity risk premium in 2013. *Work*. [28, 30, 31]
- [35] Hansen, L. P. (1982, July). Large sample properties of generalized method of moments estimators. *Econometrica* 50(4), 1029–54. [4, 13]
- [36] Hasler, M. (2012). Portfolio choice and costly dynamic information acquisition. *Working Paper, EPFL*. [5, 16]
- [37] Huang, L. and H. Liu (2007, 08). Rational inattention and portfolio selection. *Journal of Finance* 62(4), 1999–2040. [5]
- [38] Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp (2009, October). Rational attention allocation over the business cycle. Working Paper 15450, National Bureau of Economic Research. [5, 16]
- [39] Karlsson, N., G. Loewenstein, and D. Seppi (2009, April). The ostrich effect: Selective attention to information. *Journal of Risk and Uncertainty* 38(2), 95–115. [15, 16, 38]
- [40] Kita, A. and Q. Wang (2012). Investor attention and fx market volatility. Bangor Business School. [1, 4]
- [41] Kojien, R. S. J., J. C. Rodriguez, and A. Sbuelz (2009). Momentum and mean reversion in strategic asset allocation. *MANAGEMENT SCIENCE* 55(7), 1199–1213. [6]
- [42] Kreps, D. M. and E. L. Porteus (1978, January). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica* 46(1), 185–200. [17]
- [43] Laster, D., P. Bennett, and I. S. Geoum (1999, February). Rational bias in macroeconomic forecasts. *The Quarterly Journal of Economics* 114(1), 293–318. [23]

- [44] Liptser, R. S. and A. N. Shiryaev (2001). *Statistics of Random Processes II*. Springer Verlag, New York. [44]
- [45] Lucas, Robert E., J. (1978). Asset prices in an exchange economy. *Econometrica* 46(6), pp. 1429–1445. [4]
- [46] Massa, M. and A. Simonov (2005, October). Is learning a dimension of risk? *Journal of Banking & Finance* 29(10), 2605–2632. [1, 29, 37]
- [47] Mincer, J. A. and V. Zarnowitz (1969, September). The evaluation of economic forecasts. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, NBER Chapters, pp. 1–46. National Bureau of Economic Research, Inc. [35]
- [48] Ozoguz, A. (2009, November). Good times or bad times? investors’ uncertainty and stock returns. *Review of Financial Studies* 22(11), 4377–4422. [1, 29]
- [49] Pastor, L. and P. Veronesi (2006, July). Was there a nasdaq bubble in the late 1990s? *Journal of Financial Economics* 81(1), 61–100. [32]
- [50] Pastor, L. and P. Veronesi (2009, November). Learning in financial markets. *Annual Review of Financial Economics* 1(1), 361–381. [1]
- [51] Pastor, L. and P. Veronesi (2012, 08). Uncertainty about government policy and stock prices. *Journal of Finance* 67(4), 1219–1264. [30]
- [52] Patton, A. J. and A. G. Timmermann (2007). Testing forecast optimality under unknown loss. *Journal of American Statistical Association* 102, 1172–1184. [35]
- [53] Patton, A. J. and A. G. Timmermann (2008). The resolution of macroeconomic uncertainty: Evidence from survey forecasts. *Working Paper*. [15]
- [54] Peng, L. and W. Xiong (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics* 80(3), 563 – 602. [5, 16]
- [55] Rossi, A. (2010). Optimal inattention to the stock market, portfolio allocation, and consumption. *Working Paper*. [8]
- [56] Scheinkman, J. A. and W. Xiong (2003). Overconfidence and speculative bubbles. *The Journal of Political Economy* 111(6), 1183–1219. [2, 5, 10, 44]
- [57] Schmidt, D. (2013). Investors’ attention and stock covariation: Evidence from google sport searches. *Working Paper*. [4]
- [58] Timmermann, A. (2001, July). Structural breaks, incomplete information, and stock prices. *Journal of Business & Economic Statistics* 19(3), 299–314. [21, 32]
- [59] Timmermann, A. G. (1993, November). How learning in financial markets generates excess volatility and predictability in stock prices. *The Quarterly Journal of Economics* 108(4), 1135–45. [2, 21]
- [60] Van Nieuwerburgh, S. and L. Veldkamp (2006). Learning asymmetries in real business cycles. *Journal of Monetary Economics* 53(4), 753–772. [36]
- [61] Van Nieuwerburgh, S. and L. Veldkamp (2010). Information acquisition and under-diversification. *Review of Economic Studies* 77(2), 779–805. [5]
- [62] Veldkamp, L. (2006a). Media frenzies in markets for financial information. *American Economic Review* 96(3), 577–601. [5]
- [63] Veldkamp, L. L. (2006b). Information markets and the comovement of asset prices. *Review of Economic Studies* 73(3), 823–845. [5, 16]
- [64] Veronesi, P. (1999). Stock market overreaction to bad news in good times: A rational expectations equilibrium model. *Review of Financial Studies* 12(5), 975–1007. [2, 3, 21, 26, 29]

- [65] Veronesi, P. (2000, April). How does information quality affect stock returns? *Journal of Finance* 55(2), 807–837. [2, 3, 26, 27, 28]
- [66] Vlastakis, N. and R. N. Markellos (2012). Information demand and stock market volatility. *Journal of Banking and Finance* 36, 1808–1821. [1, 4, 5]
- [67] Weil, P. (1989, November). The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics* 24(3), 401–421. [17]
- [68] Xiong, W. and H. Yan (2010, April). Heterogeneous expectations and bond markets. *Review of Financial Studies* 23(4), 1433–1466. [5]



# A Appendix

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

Consider

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

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

$$F_t - F_0 = \begin{bmatrix} \int_0^t \lambda \bar{f} e^{\lambda u} du + \int_0^t \sigma_f e^{\lambda u} dZ_u^f \\ \int_0^t \sigma_\delta e^{\omega u} dZ_u^\delta \end{bmatrix}.$$

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

$$\begin{aligned} \text{Var}[f_t] &= \mathbb{E}[(f_t - \mathbb{E}[f_t])(f_t - \mathbb{E}[f_t])] \\ &= \mathbb{E}\left[\left(\int_0^t \sigma_f e^{\lambda u} dZ_u^f\right)^2\right] \\ &= \frac{\sigma_f^2 (1 - e^{-2\lambda t})}{2\lambda}. \end{aligned}$$

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

$$f_\Phi(\Phi_t) = \left| \frac{1}{g'(g^{-1}(\Phi_t))} \right| f_\phi(g^{-1}(\Phi_t)) = \frac{\exp\left(-\frac{\omega \log^2\left(\frac{\bar{\Phi}(\Phi_t-1)}{(\Phi-1)\Phi_t}\right)}{\Lambda^2 \sigma_\delta^2}\right)}{\sqrt{\pi} (\Lambda \Phi_t - \Lambda \Phi_t^2) \sqrt{\frac{\sigma_\delta^2}{\omega}}}$$

## A.2 Details on $\zeta$ , $\hat{f}$ , $\phi$ , and $\gamma$

We have

$$df_t = \left( \lambda \bar{f} + (-\lambda) f_t \right) dt + \sigma_f dZ_t^f + \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} dZ_t^\delta \\ dZ_t^s \end{bmatrix}$$

or (as in Liptser and Shiryaev (2001))

$$d\vartheta_t = [a_0(t, \vartheta) + a_1(t, \vartheta) f_t] dt + b_1(t, \vartheta) dZ_t^f + b_2(t, \vartheta) \begin{bmatrix} dZ_t^\delta \\ dZ_t^s \end{bmatrix}.$$

Moreover, the observable process is given by

$$d\vartheta_t = \left( \begin{bmatrix} -\frac{1}{2}\sigma_\delta^2 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} f_t \right) dt + \begin{bmatrix} 0 \\ \Phi_t \end{bmatrix} dZ_t^f + \begin{bmatrix} \sigma_\delta & 0 \\ 0 & \sqrt{1 - \Phi_t^2} \end{bmatrix} \begin{bmatrix} dZ_t^\delta \\ dZ_t^s \end{bmatrix}$$

or

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

Using Liptser and Shiryaev (2001)'s notations, we get

$$\begin{aligned} b \circ b &= b_1 b_1' + b_2 b_2' = \sigma_f^2 \\ B \circ B &= B_1 B_1' + B_2 B_2' = \begin{bmatrix} \sigma_\delta^2 & 0 \\ 0 & 1 \end{bmatrix} \\ b \circ B &= \begin{bmatrix} 0 & \sigma_f \Phi_t \end{bmatrix}. \end{aligned}$$

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

$$\begin{aligned} d\hat{f}_t &= [a_0 + a_1 \hat{f}_t] dt + [(b \circ B) + \gamma_t A_1'] (B \circ B)^{-1} [d\vartheta_t - (A_0 + A_1 \hat{f}_t) dt] \\ \dot{\gamma}_t &= a_1 \gamma_t + \gamma_t a_1' + (b \circ b) + [(b \circ B) + \gamma_t A_1'] (B \circ B)^{-1} [(b \circ B) + \gamma_t A_1']' \end{aligned}$$

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 including 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

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

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

$$Y_t = \begin{bmatrix} \hat{f}_t \\ \phi_t \end{bmatrix}, \quad dY_t = (A - BY_t) dt + C \begin{bmatrix} dW_t^\delta \\ dW_t^s \end{bmatrix}$$

with

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

$$C = \begin{bmatrix} \frac{\gamma_t}{\sigma_\delta} & \sigma_f \Phi_t \\ \sigma_\delta & 0 \end{bmatrix}$$

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

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

The dynamics of  $F$  obey

$$dF_t = \begin{bmatrix} \frac{e^{t\lambda}(\sigma_f \sigma_\delta \Phi_t dW_t^s + \gamma_t dW_t^\delta + dt \lambda \bar{f} \sigma_\delta)}{e^{t\omega} \sigma_\delta dW_t^\delta} \end{bmatrix}.$$

After integrating from  $v$  to  $t$  and rearranging we obtain

$$\begin{aligned} \hat{f}_t &= e^{-\lambda(t-v)} \hat{f}_v + \bar{f} (1 - e^{-\lambda(t-v)}) + \frac{1}{\sigma_\delta} \int_v^t e^{-\lambda(t-u)} \gamma_u dW_u^\delta + \sigma_f \int_v^t e^{-\lambda(t-u)} \Phi_u dW_u^s \\ \phi_t &= e^{-\omega(t-v)} \phi_v + \int_v^t \sigma_\delta e^{-\omega(t-u)} dW_u^\delta. \end{aligned}$$

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

$$\frac{\partial}{\partial t} \begin{bmatrix} G_t & F_t \end{bmatrix} = \begin{bmatrix} G_t & F_t \end{bmatrix} \begin{bmatrix} -2\lambda & \frac{1}{\sigma_\delta^2} \\ \sigma_f^2 (1 - \Phi_t^2) & 0 \end{bmatrix}$$

where  $\gamma_t = \frac{G_t}{F_t}$ . The solution is obtained through exponentiation and is given by

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

where

$$\begin{aligned} \Delta &= t - v \\ i_{v,t} &= \sigma_f^2 \int_v^t (1 - \Phi_u^2) du. \end{aligned}$$

## 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  $t = 0, \Delta, \dots, T\Delta$ , 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 pertaining to the solution of the system of equations (4) using the following simple discretization scheme

$$\begin{aligned} \int_{t_1}^{t_2} \kappa_{1,u} du &\approx \kappa_{1,t_1} \Delta \\ \int_{t_1}^{t_2} \kappa_{2,u} dW_u &\approx \kappa_{2,t_1} \epsilon_{t_1+\Delta} \end{aligned}$$

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

$$\begin{aligned} \hat{f}_{t+\Delta} = & e^{-\lambda\Delta} \hat{f}_t + \bar{f} (1 - e^{-\lambda\Delta}) + \\ & \frac{1}{\sigma_\delta} \int_t^{t+\Delta} e^{-\lambda(t+\Delta-u)} \gamma_u dW_u^\delta + \sigma_f \int_t^{t+\Delta} e^{-\lambda(t+\Delta-u)} \Phi_u dW_u^s \end{aligned} \quad (20)$$

The following moment condition must hold

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

or, its empirical counterpart

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

The Federal Reserve Bank of Philadelphia not only provides the 1-quarter ahead forecast (which in our case is denoted by  $\hat{f}_t$ ), but also 2-quarters ahead and 3-quarters 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:

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

and

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

or, their empirical counterpart

$$\begin{aligned} 0 &= \frac{1}{T} \sum_{i=1}^t [g_{1,i\Delta} - e^{-\lambda\Delta} \hat{f}_{i\Delta} - \bar{f} (1 - e^{-\lambda\Delta})] \\ 0 &= \frac{1}{T} \sum_{i=1}^t [g_{2,i\Delta} - e^{-\lambda\Delta} g_{1,i\Delta} - \bar{f} (1 - e^{-\lambda\Delta})] \end{aligned}$$

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

The following moment condition must hold

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

or, its empirical counterpart

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

where  $\mu_{(\cdot)}$  represents the sample average.

### A.4.3 Unconditional autocovariance of $\phi_t$

The following moment condition must hold

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

or, its empirical counterpart

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

### A.4.4 Unconditional variance of $\phi_t$

The following moment condition must hold

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

or, its empirical counterpart

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

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

Take the first diffusion term in Equation (20):

$$\begin{aligned} \text{Var}_t \left( \frac{1}{\sigma_\delta} \int_t^{t+\Delta} e^{-\lambda(t+\Delta-u)} \gamma_u dW_u^\delta \right) &= \frac{1}{\sigma_\delta^2} e^{-2\lambda(t+\Delta)} \mathbb{E}_t \left[ \left( \int_t^{t+\Delta} e^{\lambda u} \gamma_u dW_u^\delta \right)^2 \right] \\ &= \frac{1}{\sigma_\delta^2} e^{-2\lambda(t+\Delta)} \mathbb{E}_t \left[ \int_t^{t+\Delta} e^{2\lambda u} \gamma_u^2 du \right] \\ &\approx \frac{1}{2\lambda\sigma_\delta^2} \gamma_t^2 (1 - e^{-2\lambda\Delta}) \end{aligned} \quad (21)$$

The second equality in Equation (21) 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 Equation (20) is obtained similarly. The conditional variance of  $\hat{f}_{t+\Delta}$  is then

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

which represents a moment condition. Its empirical counterpart is

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

### 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

$$\begin{aligned}
\text{Var}_t(\phi_{t+\Delta}) &= \sigma_\delta^2 e^{-2\omega(t+\Delta)} \mathbb{E}_t \left[ \left( \int_t^{t+\Delta} e^{\omega u} dW_u^\delta \right)^2 \right] \\
&= \sigma_\delta^2 e^{-2\omega(t+\Delta)} \mathbb{E}_t \left[ \int_t^{t+\Delta} e^{2\omega u} du \right] \\
&= \frac{\sigma_\delta^2}{2\omega} (1 - e^{-2\omega\Delta})
\end{aligned} \tag{22}$$

or

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

The second equality in Equation (22) 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

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

or, its empirical counterpart

$$0 = \frac{1}{T} \sum_{i=1}^T [\Phi_{i\Delta} - \bar{\Phi}]$$

#### A.4.8 Unconditional variance of $\Phi_t$

The unconditional variance of the 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)

$$\text{Var}[\Phi_t] = 0.054$$

or, its empirical counterpart

$$0 = \frac{1}{T} \sum_{i=1}^T [(\Phi_{i\Delta} - \mu_{\Phi,1:T})^2 - 0.054]$$

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

Let's define the observable process  $d$  as

$$d_{t+\Delta} \equiv \log \frac{\delta_{t+\Delta}}{\delta_t} - \hat{f}_t \Delta = -\frac{1}{2} \sigma_\delta^2 \Delta + \sigma_\delta \epsilon_{t+\Delta}^\delta$$

We can thus write

$$\text{Var}(d_{t+\Delta}) = \sigma_\delta^2 \Delta$$

or, its empirical counterpart

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

To summarize, there are 11 boxed equations which define a system of 11 moment conditions that needs to be solved to obtain the 7-dimensional vector of parameters  $\Theta = (\lambda, \bar{f}, \omega, \bar{\Phi}, \Lambda, \sigma_f, \sigma_\delta)^\top$ .