

National Centre of Competence in Research
Financial Valuation and Risk Management

Working Paper No. 757

Investors' Attention and Stock Market Volatility

Daniel Andrei

Michael Hasler

First version: December 2010
Current version: June 2012

This research has been carried out within the NCCR FINRISK project on
“Macro Risk, Systemic Risks and International Finance”

Investors' Attention and Stock Market Volatility*

Daniel Andrei[†] Michael Hasler[‡]

June 20, 2012

Abstract

Investors' attention to the stock market is time-varying and strongly co-moves with stock market volatility. We build a theoretical model consistent with this observation. Our model features fluctuating attention to news, and implies a quadratic relationship between investors' attention and stock market volatility. We find empirical support for this relationship. Volatility and risk premium are counter-cyclical, and the relationship between them changes with the level of attention. Furthermore, the short-term asset—the claim to dividends in the near future—is volatile and commands a large equity premium during downturns, in line with recent empirical findings.

*We would like to thank Bernard Dumas and Julien Hugonnier for stimulating conversations and many insightful comments. We would also like to acknowledge comments from Tony Berrada, Jerome Detemple, Christian Dorion, Alexandre Jeanneret, Semyon Malamud, Michael Rockinger, and seminar participants at the Mathematical Finance Days Conference in Montréal and the Princeton-Lausanne Workshop on Quantitative Finance in Lausanne. Financial support from the Swiss Finance Institute and from NCCR FINRISK of the Swiss National Science Foundation is gratefully acknowledged.

[†]Swiss Finance Institute, University of Lausanne, dandrei@unil.ch, www.danielandrei.net

[‡]Swiss Finance Institute, Ecole Polytechnique Fédérale de Lausanne, michael.hasler@epfl.ch, <http://people.epfl.ch/michael.hasler>

1 Introduction

Let us imagine an economy in which investors have Gaussian priors on some unobservable variables at the beginning of history. Assume that all variables are normally distributed and, as new information becomes available, investors rationally update their estimates. In this economy, the conditional variance of investors' estimates—the *learning uncertainty*—is deterministic and quickly converges to its steady-state value. This result is unrealistic, as it does not allow uncertainty to fluctuate.

One way to overcome this result is to assume discreteness of states for the unobservable variable, approach advocated by David (1997) and Veronesi (1999). When the unobservable variable takes a finite number of values, uncertainty fluctuates and leads to time-varying return volatility.

In this paper, we propose an alternative framework featuring fluctuating uncertainty. Instead of assuming discreteness of states, our focus is on investors' *attention*, a variable strongly related with uncertainty. We assume that investors' attention is fluctuating. Naturally, a more attentive person has the tendency to learn better—to decrease uncertainty—hence fluctuating attention endogenously generates fluctuating uncertainty.

We build a general equilibrium model in which investors collect information on the unobservable state of the economy. They do so with a fluctuating attention: they are very attentive at times, and less attentive at other times. We characterize the market volatility implied by our model and show that it is driven by the fluctuating attention to news. A consequence of our model is that volatility is driven simultaneously by attention and uncertainty. First, attention increases volatility by incorporating more information into prices. Second, attention decreases volatility by reducing uncertainty. We show that for relatively high levels of attention the former effect prevails, whereas for low levels of attention the latter effect prevails. Consequently, these two competing effects create a quadratic relationship between attention and volatility. We find empirical support for this relationship.

Next, we measure the risk premium and show that it decreases with uncertainty. Equivalently, the risk premium increases with the attention. The latter result is new to our knowledge, whereas the former is in line with Veronesi (2000). We find a strong positive relationship between risk premium and volatility, but only for high levels of attention. For low levels of attention the relationship is ambiguous. As the relation between risk premium and volatility depends on the level of the attention, our work might explain why previous studies obtain mixed results about the nature of this relationship (Campbell, 1987; Glosten, Jagannathan, and Runkle, 1993).

Then, we build term structures of risk premia and volatilities and show that fluctuating attention increases both risk premia and volatilities in the short run. These

features, uncovered recently by [van Binsbergen, Brandt, and Kojen \(2010\)](#), challenge leading asset pricing models. Furthermore, we show that the term structure of forward equity yields fluctuates strongly over time, more for short maturities than for long maturities, consistent with [van Binsbergen, Hueskes, Kojen, and Vrugt \(2011\)](#).

Finally, we calibrate the model on US data and show that attention tends to be high in bad aggregate economic states. Thus, volatility and risk premium are counter-cyclical, during downturns the short-term asset commands a large equity premium, and the slopes of the term structures of forward equity yields and of risk premia are pro-cyclical. These results are consistent with [Mele \(2007, 2008\)](#), [van Binsbergen et al. \(2010\)](#), and [van Binsbergen et al. \(2011\)](#).

We focus on attention instead of uncertainty for two main reasons. The first reason is that uncertainty is inherently difficult to measure. [Massa and Simonov \(2005\)](#) and [Ozoguz \(2009\)](#) are two recent attempts, arguing that uncertainty is a priced risk factor, albeit many of the results from [Ozoguz \(2009\)](#) are only weakly significant. On the contrary, proxy measures for attention have been successfully built by [Da, Engelberg, and Gao \(2011\)](#), [Vlastakis and Markellos \(2012\)](#), [Dimpfl and Jank \(2011\)](#), and [Kita and Wang \(2012\)](#). These authors use Google search volumes on companies names or tickers and other economic terms to gauge investors' attention to publicly available sources of information. All these studies conclude that investors' attention is strongly time-varying and higher in periods of high volatility. Furthermore, [Vlastakis and Markellos \(2012\)](#) show that their attention index explains roughly 50% of the variability in the Market Volatility Index (VIX).

The second reason is that fluctuations in attention necessarily imply fluctuations in uncertainty. Because attention impacts the learning process of the investor, the Bayesian uncertainty resulting from learning has to move as well. Although we start by assuming fluctuating attention, both fluctuating attention and fluctuating uncertainty are present in our setup.

Inspired by [Da et al. \(2011\)](#), [Vlastakis and Markellos \(2012\)](#), [Dimpfl and Jank \(2011\)](#), [Kita and Wang \(2012\)](#), and to provide empirical support to the quadratic relationship between attention and volatility, we build an empirical measure of attention that we call "Focus on Economic News." 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.¹ The resulting index is depicted in [Figure 1](#), lower panel. It confirms that the attention is stochastic. In addition, the upper panel of [Figure 1](#) depicts the S&P500 volatility. A simultaneous analysis of both panels

¹More precisely, the index depicted in [Figure 1](#) 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." Other similar words in several combinations are used with almost identical results.

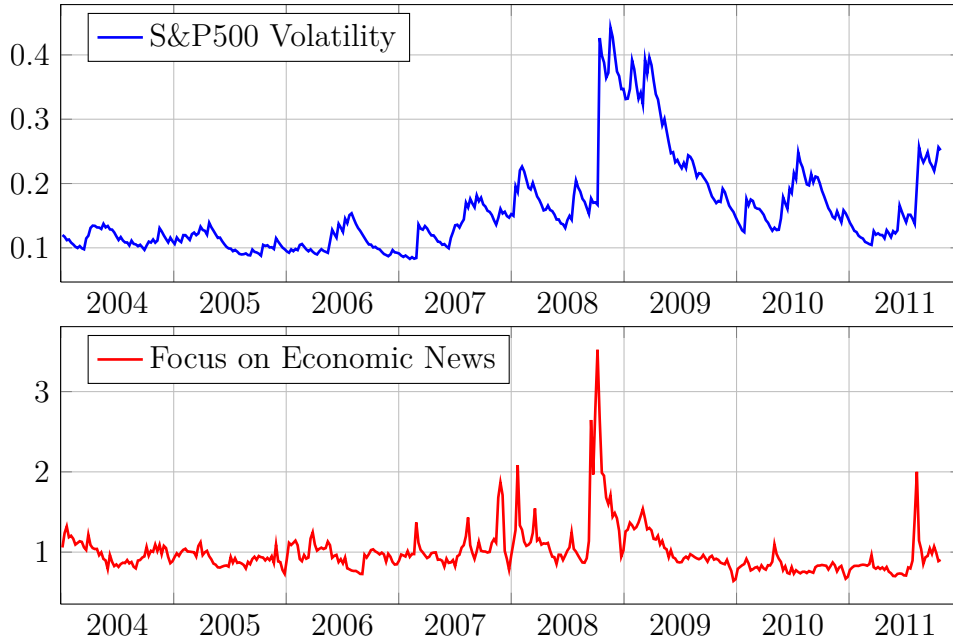


Figure 1: Focus on economic news and S&P500 volatility

The lower panel depicts a value weighted search index on financial and economic news from 2004 to 2011, at weekly frequency. The upper panel depicts the realized S&P500 annualized volatility from 2004 to 2011, resulted from a GARCH(1,1) estimation on weekly data.

suggest that there is a close connection between the “Focus on Economic News” index and the S&P500 volatility. We perform a quadratic fit of the 1-week ahead S&P500 volatility on the attention index and find that attention explains 11% of the variability in the future S&P500 volatility. Moreover, all coefficients are highly significant. This relationship is reported in Table 1.

	Estimate	t-stat	p-value	R^2
Constant	0.270	4.677	0	0.108
Attention	-0.071	-2.423	0.016	
Attention ²	0.011	3.658	0	

Table 1: Quadratic OLS fit of 1-week ahead S&P500 volatility on empirical attention

The table shows the results of a quadratic fit between between the S&P500 volatility and the empirical attention. The empirical attention corresponds to the “Focus on Economic News” index built using Google search volumes from 2004 to 2011. Standard errors are corrected using Newey and West (1987)’s estimator with 3 lags.

Our simple exercise, together with the work of Da et al. (2011) and other aforementioned empirical studies in the same vein, calls for a theoretical model featuring fluctuating attention, a task that we undertake in this paper.

2 Related Literature

Our work is related to several strands of the literature. First, the empirical work initiated by [Da et al. \(2011\)](#), pursued by [Vlastakis and Markellos \(2012\)](#), [Dimpfl and Jank \(2011\)](#), and [Kita and Wang \(2012\)](#), finding a strong positive relationship between attention and volatility, is the main motivation of our paper. We contribute to this strand of literature by providing a theoretical model able to reproduce the positive relationship between attention and volatility.

Second, our paper is related to the literature that studies learning and uncertainty in financial markets. The closest related papers in this literature are [Veronesi \(1999\)](#), [Veronesi \(2000\)](#), and [Brennan and Xia \(2001\)](#). [Veronesi \(1999, 2000\)](#) assume that the unobservable fundamental is driven by a continuous-time Markov chain. The assumed discreteness of states result in a stochastic filtered fundamental's volatility. [Brennan and Xia \(2001\)](#) assume that dividend and consumption are two distinct processes. The unobservable drift of the dividend is assumed to follow a mean-reverting process that needs to be filtered out. Linear filtering along with the distinction between dividend and consumption implies a market volatility that is constant, but higher than in an economy with complete information. Our contribution to this strand of literature is to obtain a stochastic filtered fundamental's volatility, without departing from a Gaussian setting. We also analyze simultaneously the impact of attention and uncertainty on asset returns.

Third, our work is related to the literature that studies the implications of attention for investment behavior, initiated by [Duffie and Sun \(1990\)](#). The closest related papers in this literature are [Detemple and Kihlstrom \(1987\)](#), [Peng and Xiong \(2006\)](#), [Huang and Liu \(2007\)](#), and [Hasler \(2012\)](#). [Detemple and Kihlstrom \(1987\)](#) is an attempt to solve for endogenous attention in a general equilibrium setting, yet the solution is only in implicit form. [Huang and Liu \(2007\)](#) and [Hasler \(2012\)](#) find solutions for optimal attention in partial equilibrium settings. We contribute to this strand of literature by considering a general equilibrium setting with fluctuating attention. In our model the attention is exogenous and driven by the state of the economy.

Finally, our work is related to the recent empirical literature that studies the term structure of risk premia and of volatilities. In this literature, [van Binsbergen et al. \(2010\)](#) show that the short-term asset (defined as the claim to dividends in the near future) is more volatile and bears a larger risk premium than the market. Additionally, [van Binsbergen et al. \(2011\)](#) show that the slope of the term structure of forward equity yields is pro-cyclical leading to strong fluctuations in short term yields. We contribute to this strand of research by building a general equilibrium model able to match qualitatively some of these findings.

3 A General Equilibrium Model with Fluctuating Attention to News

The novelty of our approach is to incorporate state-dependent attention in a continuous-time pure exchange economy (Lucas, 1978). The economy is characterized by a single output process (henceforth the *dividend*) having an unobservable drift (henceforth the *fundamental*). A single investor filters out the fundamental by observing the dividend and a signal. The signal has a particular feature: Its accuracy is time-varying and is related to the attention of the investor. Specifically, a higher attention translates into a higher accuracy, and vice versa.

The price of the single perishable consumption good is set to unity. 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 being the claim to a dividend process δ , whose dynamics are given by

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

The fundamental is assumed to be unobservable and to follow a mean reverting process

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

Since the fundamental is unobservable, the investor uses the information at hand to estimate it. The investor observes the current dividend δ and an informative signal s with dynamics

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

The vector $(Z^\delta, Z^f, Z^s)^\top$ is a 3-dimensional standard Brownian motion under the complete information filtration. The 3 Brownians are uncorrelated. The process Φ belongs to $[0, 1]$ and represents the positive correlation between the signal and the fundamental.

The above specification of the signal (adopted from Dumas, Kurshev, and Uppal, 2009) can be interpreted as follows. Assume that in this economy there is a very large number of public news sources (e.g., Wall Street Journal, Financial Times, CNBC, Internet, Bloomberg, Reuters). Each news source provides an unbiased noisy signal on the shock governing the fundamental, dZ_t^f . Because of the large number of news sources, the agent does not have the capacity to absorb each item of news (information overload). Consequently, at each point in time the agent collects an arbitrary number of news that depends on his attention. Since the collected news are of Gaussian type,

the agent can build a sufficient statistic by averaging them and properly adjusting the correlation Φ between this statistic and the fundamental. This results in the specification exposed in Equation (1).

Furthermore, the specification of the signal is comparable with a situation in which the investor observes a noisy signal of the fundamental $f_t dt$, as in Detemple and Kihlstrom (1987), Huang and Liu (2007), or Veronesi (2000).² We adopt the structure of the signal exposed in Equation (1) as it is better suited to illustrate the relationships between equilibrium variables and the correlation Φ . Indeed, in our case the correlation Φ belongs to a compact set, while in the above references the variance of the noisy signal belongs to the interval $[0, \infty[$.

In the spirit of Detemple and Kihlstrom (1987) and Huang and Liu (2007), Φ can be interpreted as the accuracy of the information flow, which in our setup is assumed to be time-varying. The investor can exert control on this accuracy: If she collects a large number of news, then Φ is close to 1 and the signal is very precise; if she collects a negligible amount of news, then Φ is close to 0 and the signal is pure noise. Since the investor exerts control on this accuracy (although the effort exerted by the investor is exogenous in our setup), we call the correlation process Φ *attention* to news and we interpret the signal s as the flow of news acquired by the investor.

Note that the main difference with respect to Detemple and Kihlstrom (1987), Peng and Xiong (2006), and Huang and Liu (2007) is that we do not model endogenous information acquisition. Instead, we simply assume that Φ is exogenously time-varying and determined by the current economic conditions, as it will be shown in the next Section. By adopting such a reduced form approach we are able to build a full-fledged general equilibrium model. Before going into the details of the equilibrium, it is necessary to characterize the dynamics of the attention Φ . This is a task that we undertake below.

3.1 Definition of Time-Varying Attention

In this section we characterize the dynamics of the attention, or the correlation Φ between the signal and the fundamental. For this, we construct a variable in such a way that it reflects the past performance of dividends and we call it *performance index*. This variable captures the recent development of the dividend. It is defined as follows

$$\phi_t \equiv \int_0^t e^{-\omega(t-u)} \frac{d\delta_u}{\delta_u} \quad (2)$$

²In the above references, the agent learns about the level of the fundamental, while in our case the agent learns about variations in the fundamental (we thank Jerome Detemple for providing this interpretation). Consequently, when the attention is close to 1, the uncertainty takes some time before converging toward zero.

where the parameter $\omega > 0$ represents the weight associated to the present relative to the past. If ω is large, the past dividend growth influences in an insignificant manner the performance index. On the other hand, if ω is small, the past dividend growth influences to a greater extent the current value of ϕ . [Kojien, 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.

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

$$d\phi_t = \omega \left(\frac{f_t}{\omega} - \phi_t \right) dt + \sigma_\delta dZ_t^\delta \quad (3)$$

It follows that the performance index fluctuates around the fundamental with a mean-reversion speed ω . The long term mean of the performance index is \bar{f}/ω .

We are now ready to introduce the link between current economic conditions and the attention Φ . The following definition is the core of our way to model time-varying attention.

Definition 1. *The attention Φ is defined as a function g of the performance index:*

$$\Phi_t = g(\phi_t) \equiv \frac{\Psi}{\Psi + (1 - \Psi) e^{\Lambda(\phi_t - \bar{f}/\omega)}} \quad (4)$$

where $\Lambda \in \mathbb{R}$ and $\Psi > 0$.

It follows from Equation (3) and Definition 1 that the attention Φ fluctuates around a long-run mean, the latter being given by Ψ . Moreover, the specification of the attention assumed in Equation (4) guarantees that the attention (correlation) Φ lies in $[0, 1]$, irrespective of the sign of the parameter Λ .

A particular case is obtained when the parameter ω is close to infinity, when, from the definition of the performance index in Equation (2), the agent simply looks at the current dividend growth. Put differently, if $\omega = \infty$ the performance index becomes a substitute of the current dividend growth. By assuming that $\omega \in [0, \infty[$ we let the investor decide how much of the history of past dividends to consider. The performance index then reflects the current and past dividend growths, with weights adjusted by the parameter ω .

According to the sign of the parameter Λ , the correlation Φ can either increase ($\Lambda < 0$) or decrease ($\Lambda > 0$) with the performance index ϕ . The calibration performed in Section 4 on US GDP data reveals that Λ is positive and significant, implying that our investor is more attentive and performs more accurate forecasts in bad aggregate eco-

conomic states (poor dividend performance) than in good aggregate economic states. This implication is in line with two pieces of empirical evidence. [Da, Gurun, and Warachka \(2011\)](#) show that analyst forecast errors are smaller when past 12 months return is negative than when it is positive, suggesting that information gathered by analysts in downturns is more accurate than in bullish phases.³ Additionally, [Garcia \(2012\)](#) documents that investors react strongly to good and bad news during recessions, whereas during expansions investors' sensitivity to information is much weaker.

3.2 Discussion on the Attention Process

Our model is based on the argument that investors are not constantly focused on the flow of information. Instead, periods when investors are relatively well focused are alternating with periods when they ignore the incoming news. These alternating periods are not predetermined, i.e., investors do not know today when they will be attentive in the future.

[Duffie and Sun \(1990\)](#) are the first to study theoretically the implications of attention on investment behavior. 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. [Chien, Cole, and Lustig \(2009\)](#) and [Bacchetta and Wincoop \(2010\)](#) adopt simplifying approaches by assuming that the periods of inattention are fixed. Other studies focusing on investment behavior are [Abel, Eberly, and Panageas \(2007\)](#), [Rossi \(2010\)](#), or [Duffie \(2010\)](#). Our approach is different in two respects. First, in our model investors trade and observe their wealth continuously (the aforementioned papers focus on investors' inattention to wealth, whereas we focus on investors' attention to financial news). Second, unlike [Duffie and Sun \(1990\)](#) the attention we consider is exogenous and depends on the dividend performance index only.

Endogenous fluctuating attention is hard to solve in general equilibrium settings. A notable attempt is offered by [Detemple and Kihlstrom \(1987\)](#), where the solution has only an implicit form. [Peng and Xiong \(2006\)](#), [Huang and Liu \(2007\)](#), and [Hasler \(2012\)](#) find solutions in partial equilibrium settings. Instead of solving for the endogenous attention, we choose to build a general equilibrium setting and to specify the attention process in a reduced form. The functional form that we specify in Equation (4) is quite general; the attention can either be positively or negatively correlated with the performance index (according to the sign of Λ), and the parameters Ψ , ω , and Λ can give rise to a large range of different dynamics. These degrees of freedom allow us to calibrate the model on US data, task that we undertake in Section 4. Below we proceed with a detailed discussion on the attention process.

³This claim holds under the assumption of continuous information, as in our case. Under discrete information, the reverse assertion is verified.

The unconditional distribution of the performance index is Gaussian with mean \bar{f}/ω and variance given by $\frac{\sigma_f^2}{2\lambda\omega(\lambda+\omega)} + \frac{\sigma_\delta^2}{2\omega}$ (see Appendix 10.A for a proof of these statements). We know from Equation (4) that, for $\Lambda > 0$, Φ is a strictly decreasing function of ϕ . This monotonicity allows to compute the density function of the attention Φ by a change of variable argument:

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

While the parameter Ψ dictates the location of the unconditional distribution of the attention Φ , two other important parameters govern the shape of this distribution. The first is Λ , the parameter which dictates the adjustment of the attention after changes in the performance index. The second is ω , the parameter which dictates how fast the performance index adjusts after changes in dividends. Figure 2 illustrates the probability density functions of the attention for different values of these two parameters. The black solid line corresponds to the calibration performed in Section 4 on US data. It shows that the attention is close to a regime-switching process. In other words, investors switch quickly from a period of very high attention to a period of very low attention, being moderately attentive only for brief periods. The two additional lines show that a decrease in the parameter Λ (dashed blue line) and respectively an increase in the parameter ω (dotted red line) have similar effects: both tend to bring the attention closer to its long-run mean.

Although the effects are similar, the parameters ω and Λ have different impacts on the process Φ . The parameter ω dictates the length of the history of dividends taken into account by the investor. If ω is large, the investor tends to focus more on recent dividend shocks, and the attention reverts quickly to its mean. Consequently, the unconditional distribution concentrates more around the long-term mean Ψ .

On the other hand, the parameter Λ controls the range the attention belongs to. A large parameter Λ would make the investor's attention be mainly in 2 states: either close to 0 or close to 1. The larger Λ is, the closer to a regime-switching process the attention becomes. The parameter Λ governs thus the amplitude of the attention movements.

Since in the present setup the attention Φ is observable, the setup remains conditionally Gaussian and the Kalman filter is applicable for the purpose of learning. The next Section defines the vector of filtered state variables.

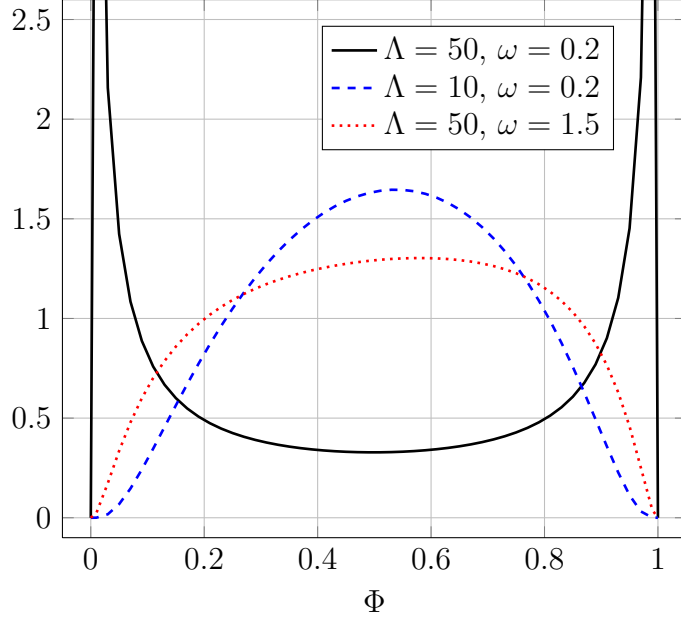


Figure 2: Probability density function of investor's attention

Probability density function of Φ for different values of Λ and ω . Other parameters are $\lambda = 0.86$, $\bar{f} = 0.026$, $\sigma_\delta = 0.015$, $\sigma_f = 0.057$, $\Psi = 0.52$. The black solid line illustrates the pdf for $\Lambda = 50$ and $\omega = 0.2$, the blue dashed line for $\Lambda = 10$ and $\omega = 0.2$, and the red dotted line for $\Lambda = 50$ and $\omega = 1.5$.

3.3 Filtering

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 ϕ 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—the attention Φ —is time-varying and is a function of the performance index, the assessed fundamental (filter) takes a non-standard form. The major change is that the conditional variance of investor's current assessment of f (simply referred to as the posterior variance, or Bayesian *uncertainty*) is time-varying. Intuitively, when the attention is high the uncertainty is low, whereas the opposite occurs when the attention is low. Following this reasoning, the vector of filtered state variables includes two additional terms: the performance index, which dictates the level of the attention, and the uncertainty that we denote by γ . Hence, the dynamics of the observed state vector

becomes

$$\begin{aligned}
d\zeta_t &= \left(\hat{f}_t - \frac{1}{2}\sigma_\delta^2 \right) dt + \begin{pmatrix} \sigma_\delta & 0 \end{pmatrix} dW_t \\
d\hat{f}_t &= \lambda (\bar{f} - \hat{f}_t) dt + \begin{pmatrix} \frac{\gamma_t}{\sigma_\delta} & \sigma_f \Phi_t \end{pmatrix} dW_t \\
d\phi_t &= \omega \left(\frac{\hat{f}_t}{\omega} - \phi_t \right) dt + \begin{pmatrix} \sigma_\delta & 0 \end{pmatrix} dW_t \\
d\gamma_t &= \left(\sigma_f^2 (1 - \Phi_t^2) - 2\lambda\gamma_t - \frac{\gamma_t^2}{\sigma_\delta^2} \right) dt
\end{aligned} \tag{5}$$

where $W \equiv (W^\delta, W^s)^\top$ is a 2-dimensional Brownian motion under the investor's observation filtration, and Φ is given by the functional form (4). The assessed fundamental is denoted by \hat{f} . The two Brownian motions governing this system are defined by

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

The proof of the above statements is provided in Appendix 10.B. 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 the uncertainty converged to its steady-state value. The deterministic nature of the uncertainty process obtained in the latter references makes this assumption plausible, as γ 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 the attention, which itself is time-varying, as shown in (5). Thus, uncertainty must be included in the state space. Although this increases considerably the complexity of the problem, we are still able to solve for the equilibrium by a linear-quadratic approximation.

A crucial implication arises from our modeling assumption of time-varying attention. The dynamics of the assessed fundamental \hat{f} depend on two diffusion components, the first 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 the attention changes, these weights move in opposite direction: A higher attention pushes the investor to give more weight to news, whereas a lower attention pushes the investor to give less weight to news and more weight to the dividend.

Consequently, the variance of the filtered fundamental, denoted henceforth by $\sigma^2(\hat{f}_t)$, is time-varying. It satisfies

$$\sigma^2(\hat{f}_t) = \frac{\gamma_t^2}{\sigma_\delta^2} + \sigma_f^2 \Phi_t^2 \tag{6}$$

Here is a verbal restatement of Equation (6). An increase in the attention has two opposing effects on the variance of the filtered fundamental. First, as attention increases, the investor assigns more weight to news and thus the variance of the filtered fundamental increases through the second term on the right hand side. Second, as attention increases, the investor assigns less weight to the dividend and thus the variance of the filtered fundamental decreases through the first term on the right hand side.

In other words, there are two forces driving the variance of the filtered fundamental. Fluctuating attention increases the variance of the filtered fundamental through better learning (a direct impact). Better learning, in turn, decreases uncertainty, thus dampening the initial effect (an indirect impact).

A note of caution is in order here. The deterministic dynamics of the uncertainty process outlined in the last Equation of (5) shows that there is no instantaneous correlation between attention and uncertainty. Indeed, there is no Brownian motion in the dynamics of γ . What these dynamics suggest is that uncertainty decreases deterministically when attention is high and increases deterministically when attention is low. Hence, the two competing effects on the variance of the filtered fundamental in Equation (6).

To summarize, we offer a framework to study the *simultaneous* impact of attention and uncertainty on asset prices. While in Veronesi (1999, 2000) attention is constant but uncertainty is fluctuating due to the assumed discreteness of states the fundamental can belong to, in our case attention drives uncertainty. Hence our equilibrium model permits to study the dynamic impact of both attention and uncertainty on asset returns. The computation of the equilibrium is exposed in what follows.

3.4 Equilibrium

Because our setup contains two observable Brownian motions and only one risky asset, markets are incomplete. The problem of the investor in this economy is to maximize expected utility from lifetime consumption subject to the lifetime budget constraint:

$$\begin{aligned} \sup_{c,n} \mathbb{E} \int_0^\infty e^{-\rho t} \frac{c_t^{1-\alpha}}{1-\alpha} dt \\ \text{s.t. } dV_t = [r_t V_t - c_t + n_t S_t (\mu_t - r_t)] dt + n_t S_t (\sigma_{1t} \sigma_{2t}) dW_t \end{aligned} \quad (7)$$

where V_t is investor's wealth, r_t is the risk-free rate, n_t is the number of shares of the risky asset, μ_t is the expected return of the risky asset, $(\sigma_{1t} \sigma_{2t})$ is the 2-dimensional diffusion vector of the risky asset, ρ is the subjective discount factor, and α is the coefficient of relative risk aversion.

Optimality and market clearing yields

$$\begin{aligned}
 c_t &= \delta_t \\
 n_t &= 1 \\
 \xi_t &= e^{-\rho t} \left(\frac{\delta_t}{\delta_0} \right)^{-\alpha}
 \end{aligned} \tag{8}$$

where ξ is the state price density. Because we are in a representative agent economy, the state price density is characterized as in the complete market setup of Karatzas, Lehoczky, and Shreve (1987) and Cox and Huang (1989).

The functional form of the state price density implies that an increase in expected dividend growth decreases the expected value of discount factors. Furthermore, precautionary savings imply that an increase in future dividend growth risk increases the expected value of discount factors. In our model future dividend growth risk is not constant, but depends crucially on the volatility of the expected growth rate $\sigma(\hat{f}_t)$. It is precisely the effect of fluctuating dividend growth risk which, through the discount factor channel, generates our results.

The price of the risky asset is computed as the expected sum of discounted future dividends:

$$S_t = \mathbb{E}_t \left[\int_t^\infty \frac{\xi_u}{\xi_t} \delta_u du \right] = \delta_t^\alpha \int_t^\infty e^{-\rho(u-t)} \mathbb{E}_t \left[e^{(1-\alpha)\zeta_u} \right] du \tag{9}$$

The dynamics of the risky security are of the form

$$\frac{dS_t}{S_t} = \left(\mu_t - \frac{\delta_t}{S_t} \right) dt + (\sigma_{1t} \sigma_{2t}) dW_t$$

where μ , σ_1 , and σ_2 are to be determined.

3.5 Transform Analysis

It can be easily seen that the dynamics of the state vector described in the system of Equations (5) are not affine. First, the performance index ϕ enters the dynamics of both \hat{f} and γ through the nonlinear functional form of the attention (Equation 4). Second, the uncertainty γ enters with a quadratic term in its own drift.

Consequently, the theory of affine processes (e.g., Duffie, 2008) cannot be used directly to solve for the conditional expectation in Equation (9). To overcome this difficulty, we first notice that \hat{f} and ϕ are mean reverting around their long term means \bar{f} and \bar{f}/ω respectively. Moreover, when Φ converges to its long term mean Ψ (or, equivalently, when the performance index ϕ is at its long term mean) we can solve for

the steady-state value of uncertainty, γ_{ss} . That is, γ_{ss} solves the following equation

$$\sigma_f^2(1 - \Psi^2) - 2\lambda\gamma_{ss} - \frac{\gamma_{ss}^2}{\sigma_\delta^2} = 0$$

This equation has two solutions, only one of them being positive.⁴ We have now a natural set of reference points (\bar{f} , \bar{f}/ω , and γ_{ss}) around which we can implement an approximation of the dynamics of the state vector. This task is undertaken below.

In order to obtain a sufficiently accurate approximation, we first augment the state space by adding ϕ^2 , \hat{f}^2 , γ^2 , $\hat{f}\phi$, $\gamma\phi$ and $\hat{f}\gamma$ in the vector of state variables. We define the 10-dimensional augmented state vector X by

$$X = \left(\zeta \quad \hat{f} \quad \phi \quad \gamma \quad \phi^2 \quad \hat{f}^2 \quad \gamma^2 \quad \hat{f}\phi \quad \gamma\phi \quad \hat{f}\gamma \right)^\top$$

We compute the dynamics of this augmented state vector by applying Itô's lemma. The drift and the variance-covariance matrix of the augmented state vector, $\mu(X)$ and $\sigma(X)\sigma(X)^\top$, are non-linear functions. Consequently, we perform a 3-dimensional second order Taylor approximation of the drift $\mu(X)$ and of the variance-covariance matrix $\sigma(X)\sigma(X)^\top$ with respect to the variables \hat{f} , ϕ and γ around the points \bar{f} , \bar{f}/ω and γ_{ss} respectively. More precisely, the approximation yields

$$\begin{aligned} \mu(X_t) &\approx K_0 + K_1 X_t \\ \sigma(X_t)\sigma(X_t)^\top &\approx H_0 + H_{11}\zeta_t + H_{12}\hat{f}_t + H_{13}\phi_t + H_{14}\gamma_t + H_{15}\hat{f}_t^2 + H_{16}\phi_t^2 + H_{17}\gamma_t^2 \\ &\quad + H_{18}\hat{f}_t\phi_t + H_{19}\gamma_t\phi_t + H_{110}\hat{f}_t\gamma_t \end{aligned}$$

where K_0 is a 10-dimensional vector and K_1 , H_0 , $H_{11} - H_{110}$ are 10-dimensional squared matrices which we do not expose here, but are available upon request.

This approximation can be performed at orders higher than 2 by adding state variables to the system and performing the same steps. As an exercise, we went up to the fourth order and obtained almost identical equilibrium quantities. This approximation allows us to work with an affine 10-dimensional vector of state variables, X . As the expectation term pertaining to Equation (9) is the moment-generating function of ζ , we can now apply the theory of affine processes to get a closed form expression for it. The moment-generating function satisfies

$$\mathbb{E}_t e^{(1-\alpha \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0)X_{t+\tau}} \approx e^{\bar{\alpha}(\tau) + \bar{\beta}(\tau)^\top X_t} \quad (10)$$

where $\bar{\alpha}(\tau)$ and $\bar{\beta}(\tau)$ solve an 11-dimensional system of Riccati equations with initial conditions $\bar{\alpha}(0) = 0$, $\bar{\beta}_1(0) = 1 - \alpha$, and $\bar{\beta}_i(0) = 0$, $\forall i > 1$. With slight abuse of

⁴The positive root of the equation is $\gamma_{ss} = -\lambda\sigma_\delta^2 + \sqrt{\sigma_\delta^2(\lambda^2\sigma_\delta^2 + \sigma_f^2(1 - \Psi^2))}$.

notation, this system of Riccati equations is written⁵

$$\begin{aligned}\bar{\beta}'(\tau) &= K_1^\top \bar{\beta}(\tau) + \frac{1}{2} \bar{\beta}(\tau)^\top H_1 \bar{\beta}(\tau) \\ \bar{\alpha}'(\tau) &= K_0^\top \bar{\beta}(\tau) + \frac{1}{2} \bar{\beta}(\tau)^\top H_0 \bar{\beta}(\tau)\end{aligned}\tag{11}$$

with K and H being defined above. Notice from (5) that the first state variable ζ does neither enter the variance-covariance matrix nor the drift. Hence we have $H_{11} = [0]_{10 \times 10}$ and $K_1(i, 1) = 0, \forall i$. This yields $\bar{\beta}_1(\tau) = 1 - \alpha$ and the system of Riccati equations reduces to a dimension of 10. We then solve this system numerically.

To address the concern of the accuracy of the above approximation, we implement the following procedure. Define the transform M by

$$M_t \equiv \mathbb{E}_t \left[e^{(1-\alpha)\zeta_u} \right].$$

Since M is a martingale, the associated partial differential equation is

$$\mathcal{L}^{\zeta, \hat{f}, \phi, \gamma} M + \frac{\partial}{\partial t} M = 0\tag{12}$$

where $\mathcal{L}^{\zeta, \hat{f}, \phi, \gamma}$ denotes the infinitesimal generator of $(\zeta, \hat{f}, \phi, \gamma)$. We substitute the right hand side of Equation (10), that is, the approximation of the moment generating function, into the left hand side of Equation (12). Then, we compute the residuals of the approximate solution for large ranges of values of (\hat{f}, ϕ, γ) . For this exercise, we set $T - t = 1$ and $\zeta = 0$, because the absolute values of the residuals seem to decrease with the time horizon and the log-consumption. We obtain residuals of order 10^{-8} at most. Our approximation scheme seems to provide very accurate results.

4 Calibration to the U.S. Economy

We proceed now to the calibration of our model to the U.S. economy. The investor is able to observe 2 processes: the dividend stream δ and a flow of information s . Hence, the investor uses δ and s to estimate the evolution of the non-observable variable f . Calibrating our model to observed data is challenging since we don't know which variable corresponds to the signal s . Although our theoretical model assumes that the flow of information s is observable, it is almost impossible to observe and quantify 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

⁵The matrix H_1 in the term $\frac{1}{2} \bar{\beta}(\tau)^\top H_1 \bar{\beta}(\tau)$ is 3-dimensional. A separate equation should be written for each β_i , but we avoid this here and prefer the form (11) for simplicity.

as a proxy for the output growth rate. Quarterly data from Q1:1969 to Q3:2010 are obtained from the Federal Reserve Bank of Philadelphia’s website.

Since we work with quarterly data, an immediate discretization of the stochastic differential equations exposed in (5) would provide biased estimators. Hence, we first solve this set of 4 stochastic differential equations. The solutions are provided in Appendix 10.C. We then approximate the continuous-time processes pertaining to those solutions using the following simple discretization scheme

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

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

where κ_1 and κ_2 are some arbitrary processes, $\Delta = t_2 - t_1 = \frac{1}{4}$, and $\epsilon_{t_1+\Delta} \sim N(0, \Delta)$. 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 δ and \hat{f} , instead of δ and s , also provides a well defined system.

4.1 Generalized Method of Moments Procedure

Our model is calibrated on the 2 time-series discussed above using Hansen (1982)’s Generalized Method of Moments (GMM) procedure. The vector of parameters is defined by $\Theta = (\sigma_\delta, \bar{f}, \lambda, \sigma_f, \omega, \Lambda, \Psi)^\top$. Consequently, we need 7 moment conditions to infer the vector of parameters Θ . For the sake of brevity, the moment conditions are exposed in Appendix 10.D.

The values, t-stats, and p-values of the vector Θ resulted from the GMM estimation are provided in Table 2. The sole parameter incurring a relatively small t-stat is the mean reversion speed λ of the fundamental. A test of the null hypothesis $H_0 : \lambda \leq 0.3$ is rejected at 95% confidence level. We want to point out that the value of λ is relatively far from what the long run risk literature assumes. In fact, papers dealing with long run uncertainty typically suppose that the mean reversion parameter is between 0 and 0.1. In Bansal and Yaron (2004) the AR(1) parameter of the fundamental is worth 0.979 at monthly frequency. This parameter would correspond to $\lambda = 0.25$. Barsky and De Long (1993) go even further by assuming that the fundamental is an integrated process. Although our dataset suggests that the hypothesis of Barsky and De Long (1993) and Bansal and Yaron (2004) have to be rejected, only the far future can potentially confirm that these authors’ hypothesis is sustainable. Indeed, 40 years of quarterly data are largely insufficient to estimate a parameter implying a half-life of roughly 3 years, i.e., the value proposed by Bansal and Yaron

(2004).

	Estimate	t-stat	p-value
σ_δ	0.015	11.18	0
\bar{f}	0.026	8.98	0
λ	0.861	2.97	0.003
σ_f	0.057	10.00	0
ω	0.201	11.35	0
Λ	50.065	7.13	0
Ψ	0.524	9.95	0

Table 2: Calibration to the U.S. economy (GMM estimation)

We obtain a large positive and significant value for the parameter Λ , suggesting that investors have the tendency to jump from very low attention states to high attention states. The parameter ω is positive and significant. Its value of 0.2, coupled with the large value of Λ , imply that the attention has the tendency to stay mainly in high or low states. The probability distribution function of the attention depicted in Figure 2 (black solid line) confirms this intuition. Moreover, since Λ is positive, the data confirms that attention is high in bad aggregate economic states ($\phi < \frac{\bar{f}}{\omega}$) and low in good aggregate economic states ($\phi > \frac{\bar{f}}{\omega}$). This can be interpreted as follows. When the economy is in a bullish period, the probability of a decrease in δ is relatively small. Thus, investors do not have the incentive to exert a strong learning effort. On the other hand, when the economy enters a recessionary phase, investors substantially worry about the fundamental driving the economy. In this situation, the probability of a decrease in future consumption is high, leading investors to estimate as accurately as possible the change in the fundamental.

A positive parameter Λ is consistent with empirical findings by [Da et al. \(2011\)](#) and [Garcia \(2012\)](#). On the theoretical side, [Hasler \(2012\)](#) finds (in a partial equilibrium setting) that forecast accuracy is decreasing with past returns which, again, is in line with a positive parameter Λ .

Finally, we obtain a low volatility of dividends σ_δ (which is equal to the volatility of consumption in our model). Additionally, we set the relative risk aversion to $\alpha = 3$ and the subjective discount rate to $\rho = 0.01$. We turn now to the analysis of our results.

5 Attention, Uncertainty, and Volatility

Attention and uncertainty are strongly related in our setup: High attention brings lower uncertainty, whereas low attention brings higher uncertainty. How is the volatility of asset returns driven by attention *and* uncertainty? We address this question below.

In the theoretical literature, spikes in volatility have been often related to spikes in

uncertainty (Veronesi, 1999; Timmermann, 1993, 2001; Ozoguz, 2009). In what follows we show that the primary effect of attention on volatility overcomes the secondary effect of the uncertainty. While we find a similar positive relationship between volatility and uncertainty, the effect of attention is clearly stronger. We show that high levels of investors' attention increase both volatility and risk premia.

The stock return diffusion vector follows from Equation (9) by applying Itô's lemma to the stock price S :

$$(\sigma_{1t} \ \sigma_{2t}) = \frac{1}{S_t} \frac{\partial S_t}{\partial x_t} \text{diff}(x_t) \quad (13)$$

where $\text{diff}(\cdot)$ is the diffusion operator. Denoting by S_f and S_ϕ the partial derivatives of the price with respect to the assessed fundamental \hat{f} and the performance index ϕ , the variance of stock returns is

$$\|\sigma_t\|^2 = \sigma_{2t}^2 + \sigma_{1t}^2 = \left(\frac{S_f}{S}\right)^2 \sigma_f^2 \Phi_t^2 + \left[\frac{S_f}{S} \frac{\gamma_t}{\sigma_\delta} + \sigma_\delta \left(1 + \frac{S_\phi}{S}\right)\right]^2 \quad (14)$$

Stock return volatility depends on a complex interaction between attention and uncertainty on the one hand, and investor's price valuations of those states (reflected in the price and its partial derivatives with respect to \hat{f} and ϕ) on the other hand. A similar form for the variance of stock returns is discussed by Veronesi (2000) and Brennan and Xia (2001). To make a parallel with Veronesi (2000), the term V_θ in his case is equivalent to $S_f \gamma_t / S$ in our case. If uncertainty is zero, V_θ is zero. If investors assign the same value to the asset for any value of \hat{f} (that would be the case for log-utility), then V_θ is zero.

The variance of stock returns expressed in Equation (14) has two terms. Both terms depend on the attention Φ , the first one directly, whereas the second one indirectly. The first term clearly shows a quadratic relationship between attention and return variance. The indirect effect in the second term is produced by the inverse relationship between attention and uncertainty. Naturally, and as explained in Section 3.3, as the agent learns better (when the attention is high), the uncertainty is diminished. Thus, our intuition is that the first term increases with Φ , while the second decreases.

As in Section 3.3, the effect of the attention on the stock return volatility can also be interpreted in terms of weights. First, as attention increases, 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.

We insist here on the fact that the direct effect of attention on return variance,

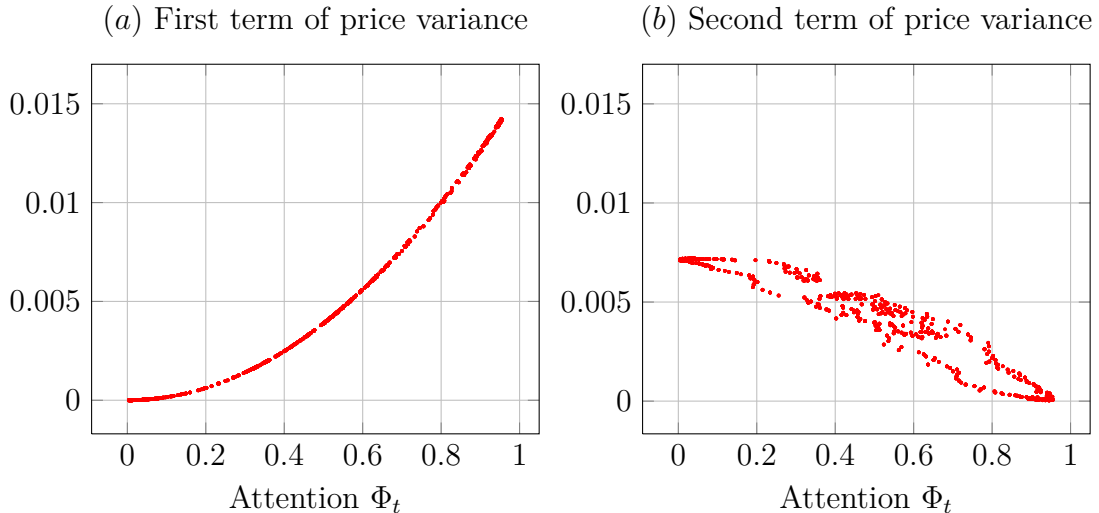


Figure 3: Decomposition of stock return variance

Panel (a) depicts the first term of (14), resulted from a simulations of 20 years of weekly data. Panel (b) depicts the second term of (14) resulted from the same simulation. The parameter values are shown in Table 2 and discussed in Section 4.

arising through the first term of Equation (14), is unambiguous. No matter the sign of S_f , i.e. the partial derivative of the price with respect to the assessed fundamental \hat{f} , the squared attention has a positive coefficient. The uncertainty effect is, however, dictated by the sign and magnitude of S_f . In a setup with power utility function, Veronesi (2000) shows that, for levels of risk aversion higher than 1, S_f is negative. Investors more risk-averse than log assign a lower relative value to the asset in high growth states—as they discount future dividends using their marginal utility of future consumption.

Partial derivatives of the price with respect to the assessed fundamental \hat{f} and the performance index ϕ are functions of the state vector. Hence, they change with the attention, preventing us from showing a unique relationship between attention and variance. Consequently, the simplest way to assess the magnitude and the effect of the attention on the two terms of the variance in Equation (14) is by simulations. We therefore simulate 20 years of weekly data (that is, 1040 data points) and we plot the two terms of the price variance as functions of the attention. The results are depicted in Figure 3.

Scales in both panels of Figure 3 are matched, with the aim to compare the magnitude of the two terms. Panel (a) confirms the quadratic relationship between the first term of the stock return variance and the attention. It is important to note that changes in the stock price and in its partial derivative with respect to the assessed fundamental, S_f , have little effect on the relationship, which remains clearly quadratic. Panel (b) confirms our initial intuition of negative relationship between the attention and the second term of the price variance. The relationship is, however, more sensitive

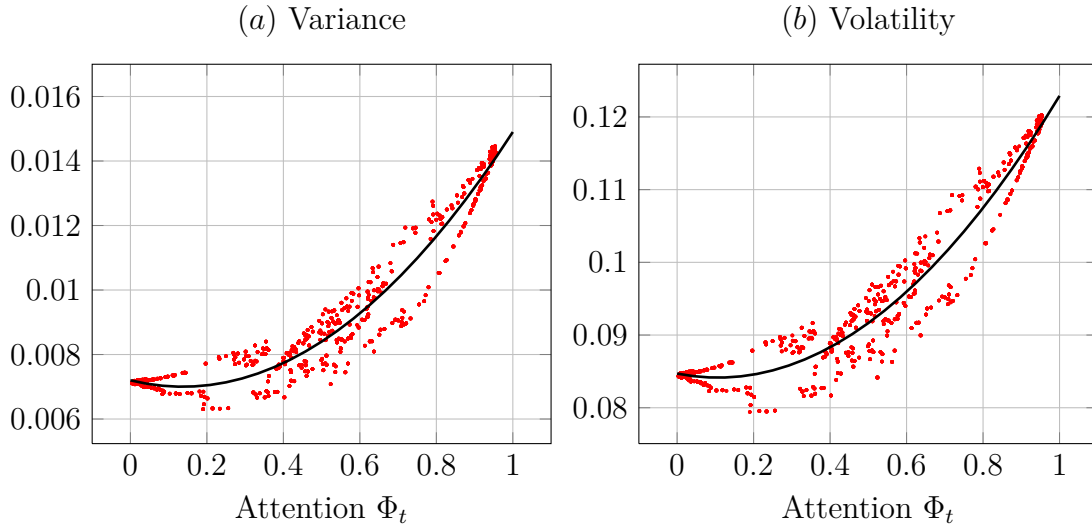


Figure 4: Investors' attention drives stock return variance and volatility

Panel (a) depicts the variance of asset returns, resulted from a simulations of 20 years of weekly data. Panel (b) depicts the volatility of asset returns, resulted from the same simulation. The black solid lines are quadratic fits. The parameter values are shown in Table 2 and discussed in Section 4.

to changes in the stock price and its partial derivatives S_f and S_ϕ . Furthermore, the relationship depicted in Panel (b) of Figure 3 approaches a linear form.

Inspection of both panels reveals that the direct effect is stronger than the indirect one when the attention is high and weaker when the attention is low. Thus, adding up the two terms of the stock return variance, one should expect a U-shaped (quadratic) relationship between attention and return variance. This is confirmed by Figure 4, in which we plot the variance and the volatility of asset returns, as functions of the attention Φ , resulted from the same simulation of 20 years of weekly data. Moreover, as Figure 3 suggests, we expect a positive coefficient for the quadratic term and a negative coefficient for the linear term.

Panel (a) of Figure 4 shows that the relationship between attention and stock return variance is indeed quadratic. It depicts the total return variance, i.e., the sum of both terms in Equation (14). To help envisioning the quadratic relationship, we added to the graph a quadratic fit of our simulation. In panel (b) of Figure 4 we plot the relationship between attention and volatility, which remains of a quadratic shape, as the added quadratic fit suggests.

To summarize, there are two opposing effects produced by fluctuations in attention. First, stock return variance increases quadratically with attention. Second, higher attention means better learning, which tends to decrease linearly the variance of stock returns. Overall, the relationship between price variance and attention is quadratic.

A natural question arising at this point is which of the two effects dominates. Is the attention increasing the volatility or reducing it? The scatter plots from Figure 4

suggest a U-shaped pattern. For low values of attention the second effect (decrease of uncertainty through better learning) is strong and can lead to lower variance/volatility. For high values of attention, however, the first effect (direct increase of the attention) clearly dominates and leads to higher variance/volatility.

We perform a quadratic fit of return volatility on attention. This fit is represented by the black solid line in panel (b) of Figure 4. The results of the estimation are shown in the last 3 columns of Table 3. For convenience, we show in the first three columns the results of the quadratic fit between the S&P500 volatility and the empirical attention, as exposed in the Introduction and in Table 1.

	Data			Model		
	Estimate	t-stat	R^2	Estimate	t-stat	R^2
α	0.270	4.677	0.108	0.085	376	0.952
β_1	-0.071	-2.423		-0.010	-8.47	
β_2	0.011	3.658		0.049	36.8	

Table 3: Quadratic OLS fit of volatility on attention (Data and Model)

The left panel of the table shows the results of a quadratic fit between between the S&P500 volatility and the empirical attention. The right panel shows the results of the quadratic fit performed with simulated data (20 years of weekly data). α is the intercept, β_1 is the first order coefficient, and β_2 the second order coefficient of the quadratic fit.

Although the coefficients obtained from our simulation have different magnitudes than the empirical ones, their sign is correct in all cases. Which confirms that, at least qualitatively, we have a correct relationship between attention and stock market volatility. Moreover, the sign of coefficients have an economic interpretation. First, higher attention means faster revelation of information into prices, which increases volatility (positive quadratic term). Second, higher attention means lower uncertainty, which decreases volatility (negative linear term). For high levels of attention, the former effect dominates the latter.

The contribution of our setup with respect to Veronesi (2000) is twofold. First, in our setup the attention (equivalent to the *quality* of information in Veronesi, 2000) is time-varying. This brings our model closer to recent empirical work on fluctuating attention. Second, in our setup uncertainty is time-varying due to fluctuating attention, which allows us to dynamically seize its impact, whereas Veronesi (2000) performs comparative statics analysis in terms of mean preserving spread of the distribution of the assessed fundamental.

One additional implication of our model is that, as attention increases, the price should be more strongly related to news. This can be easily seen from Equation (14). Is this implication supported by the data? In recent empirical work, Garcia (2012) shows that the predictability of stock returns using news' content is concentrated

during recessions, lending support to our implication. More precisely, [Garcia \(2012\)](#) finds that one standard deviation shock to news during recessions (i.e., during times of high attention) predicts a change in the conditional average return on the Dow Jones Industrial Average of twelve basis points over one day.

An important question could be raised at this point regarding the exogenous process of attention assumed in Equation (4). How would our results be affected if the attention itself depends on return volatility? In other words, what if investors observe high return volatility and become more attentive? Such a problem, comprising a feedback effect of volatility on attention, is very difficult to solve, as explained by [Detemple and Kihlstrom \(1987\)](#). We conjecture, however, that this feedback mechanism can only reinforce our results: Higher attention brings higher volatility which in turn brings higher attention and so on.

6 Attention, Uncertainty, and Risk Premium

By applying Itô's Lemma to the state price density (Equation 8) we obtain the risk free rate r and the market price of risk θ :

$$\frac{d\xi_t}{\xi_t} = -r_t dt - \theta_t dW_t$$

The first order conditions of the optimality problem (7) imply that the equilibrium excess return must satisfy

$$\mu_t - r_t = \alpha \sigma_{1t} \sigma_\delta = \alpha \text{Cov}_t \left(\frac{d\delta_t}{\delta_t}, \frac{dS_t}{S_t} \right) / dt \quad (15)$$

where σ_{1t} is the first component of the stock return diffusion defined in Equation (13).

Equation (15) states that the covariance between changes in dividend (which in equilibrium equals consumption) and the rate of return on the risky asset is proportional to the expected excess return on the risky asset, with coefficient of proportionality given by the risk aversion. It follows from Equation (15) that the equity risk premium equals

$$\mu_t - r_t = \alpha \left[\sigma_\delta^2 \left(1 + \frac{S_\phi}{S} \right) + \frac{S_f}{S} \gamma_t \right] \quad (16)$$

Here is a verbal restatement of Equation (16). Attention affects the equity premium only indirectly through the uncertainty channel. As attention gets higher, uncertainty is lower, and since S_f/S is negative, the risk premium increases. A similar result is obtained by [Veronesi \(2000\)](#): for a coefficient of risk aversion higher than 1, lower uncertainty increases the risk premium, whereas for α lower than 1 the opposite holds.

Since the risk premium depends on the covariance of consumption growth and stock returns, as attention increases the covariance of returns and consumption increases and so does the risk premium.

When the attention is high (uncertainty is low), the covariance between consumption growth and stock returns is higher than in the case where attention is low (uncertainty is high). Let us consider the latter case. A negative shock in dividends has two opposite effects on prices: (i) it tends to decrease the price through a direct dividend channel, and (ii) it tends to increase the price through an indirect discounting channel—discount rates decrease because the shock reduces the assessed fundamental. When attention is low (high uncertainty), the latter effect dominates. Indeed, low future consumption pushes the investor to demand more of the stock. This is the precautionary savings effect. Hence the covariance between dividend growth and stock return is negative and so is the risk premium. Conversely, when attention is high (low uncertainty), the former effect outweighs the discounting effect. This yields a positive covariance and a positive risk premium.

Figure 5 depicts the risk premium resulted from the simulation discussed previously. *Ceteris paribus*, a higher level of attention implies a higher risk premium. Taking into consideration the fact that for the same level of attention one can have several values of uncertainty, there are several levels of risk premia for the same level of attention, as shown in the scatter plot.

Veronesi (2000) uncovers the seemingly paradoxical result that uncertainty decreases the risk premium. We obtain the same result for a fixed level of attention: as uncertainty gets higher, the risk premium is lower. Letting the attention to vary has, however, a strong effect on the risk premium. As attention increases, uncertainty gets smaller, which in turn increases the risk premium. As suggested in Figure 5, the effect of the attention seems strong.

More to the point, Figure 5 shows that the relationship between attention and risk premium is not linear; risk premium increases more when the attention is high (bad aggregate economic states) than when the attention is low (good aggregate economic states). Risk-premia are therefore counter-cyclical, and lower much less during expansions than they increase during recessions, as in Mele (2007). With respect to Mele (2007) we have, nonetheless, a complementary explanation. While Mele (2007) suggests that the price-dividend ratio should be increasing and concave in the fundamental to generate this effect, our explanation is that, although the price dividend ratio is not increasing and concave, the relationship between risk premia and attention is increasing and convex.

It is worth pointing out that the risk premium is small and mostly negative in our setup. This result arises because we consider an investor whose elasticity of intertemporal substitution is below one and who consumes the aggregate dividend,

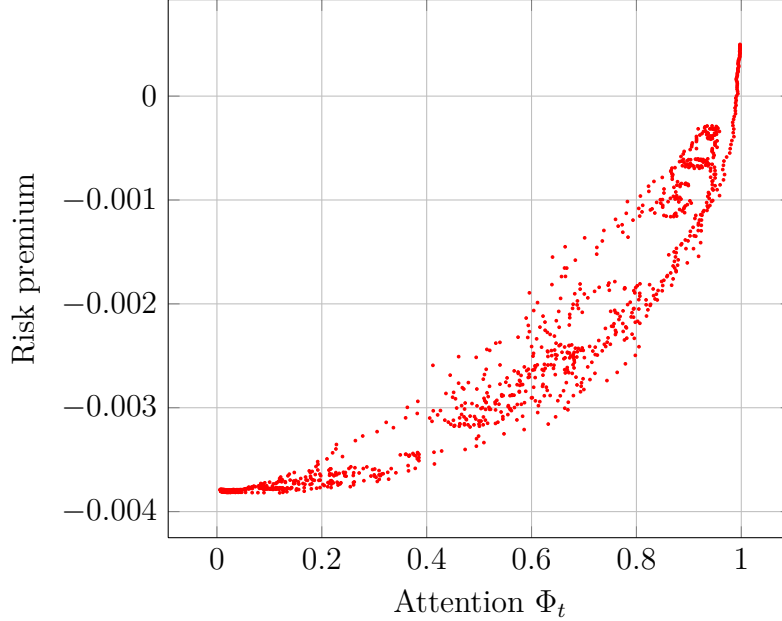


Figure 5: Investors’ attention drives risk premia

Equity risk premium, as a function of attention, resulted from the same simulation of 20 years of weekly data. The parameter values are shown in Table 2 and discussed in Section 4.

much as Veronesi (2000). In such a setup, bad news about dividends decreases current consumption but also future consumption, forcing the agent to demand more of the stock today in order to smooth consumption. This larger demand tends to increase the price. The result would be reversed if the elasticity of intertemporal substitution was larger than one, suggesting that Epstein-Zin preferences would increase the equity premium (Bansal and Yaron, 2004). Exponential utility might help as well to increase the equity premium (Veronesi, 1999). Finally, an economy where dividends and consumption follow separate processes (Brennan and Xia, 2001) would also generate a higher risk premium.

The results depicted in Figures 4 and 5 suggest that there is no clear relationship between volatility and risk premia. While higher attention unambiguously increases the risk premium, the relationship between attention and volatility is U-shaped. It is worth mentioning, nonetheless, that for reasonably large levels of attention (larger than 0.2), there is a clear positive relationship between volatility and risk premia. Volatility and risk premia are, therefore, counter-cyclical (Mele, 2007, 2008).

These results suggest that the nature of the relation between market risk premium and conditional market variance changes with the attention. While for low levels of attention there seems to be a negative relationship (Campbell, 1987; Glosten et al., 1993; Ozoguz, 2009), for high levels of attention the relationship turns positive. Hence, one implication of our model is that adding attention to the regression risk premium–

volatility could improve explanatory power. Indeed, as investors attention seems to explain anomalies in asset return predictability (Peng, Xiong, and Bollerslev, 2007; Da et al., 2011), it is a worthwhile endeavor to study these implications in future research.

7 Volatility and Risk Premium in the Short-Term

A recent paper by van Binsbergen et al. (2010) challenges leading asset pricing models by showing that risk premia and volatilities are concentrated mostly in the short-term. van Binsbergen et al. (2010) arrive to this conclusion by recovering prices of zero-coupon equity (dividend strips) on the aggregate stock market. Leading asset pricing models like Campbell and Cochrane (1999), Bansal and Yaron (2004), or Gabaix (2008) all predict the opposite, that is, higher risk premia and volatilities for long-term dividend strips. It is therefore an important task to build a general equilibrium model consistent with the findings of van Binsbergen et al. (2010).

The short-term asset computed by van Binsbergen et al. (2010) as the claim to first 2 years of dividends is very volatile and bears a high risk premium, especially during bad times. In what follows we show that fluctuating investors' attention increases both the volatility and the risk premium of the short-term asset exactly during bad times. Moreover, the increase of the volatility is proportionally higher for the short-term asset than for the market as a whole—meaning that the volatility increases more in the short-term than in the long term during bad times.

Similar to van Binsbergen et al. (2010), the short-term asset price at time t satisfies

$$ST A_t = \mathbb{E}_t \int_t^T \frac{\xi_u}{\xi_t} \delta_u du$$

where $T = 2$ is the short-term asset payout horizon. The relationship between the stock market price and the short-term asset price is

$$\begin{aligned} S_t &= ST A_t + \mathbb{E}_t \int_T^\infty \frac{\xi_u}{\xi_t} \delta_u du \\ &= ST A_t + LTA_t \end{aligned}$$

where LTA_t is the long-term asset, paying dividends from T to infinity.

Figure 6 depicts the volatility of the short term asset resulted from the simulation performed in the previous sections (Panel *a*) and one path of the short-term asset price (black solid line, panel *b*) and its corresponding dividend payout lagged 24 months (red dashed line, panel *b*). To increase readability of the graph, the lines from panel (*b*) correspond to the first 10 years of simulated data.

The relationship between the volatility of the short-term asset and attention, as

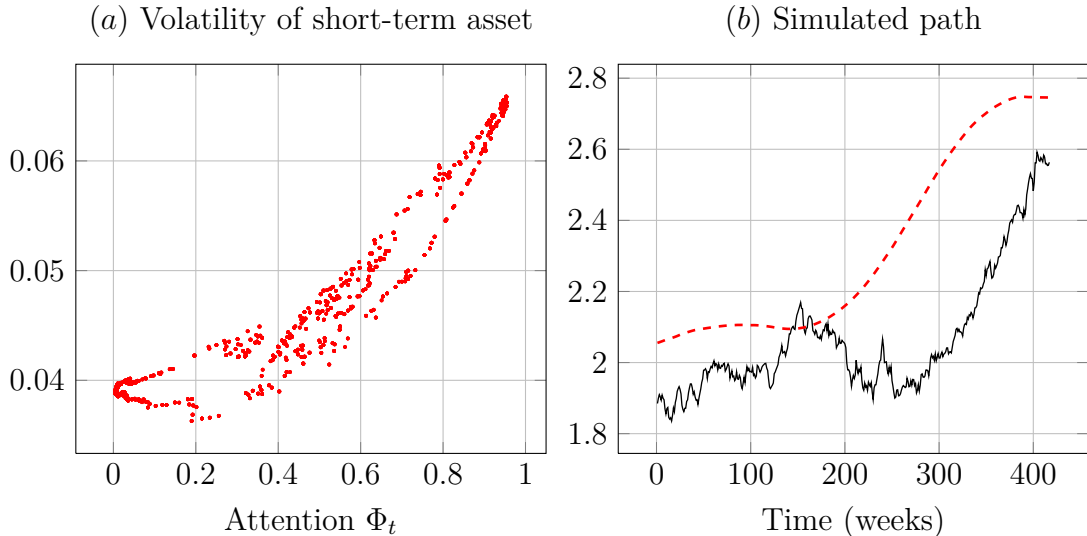


Figure 6: The volatility of the short-term asset increases with attention. Furthermore, the short-term asset is more volatile during downturns

Panel (a) depicts the volatility of the short-term asset resulted from a simulation of 20 years of weekly data. Panel (b) depicts the price of the short term asset, defined as the claim to the next 2 years of dividends (black solid line) and its lagged dividend realizations (red dashed line), resulted from the same simulation. The lines from panel (b) correspond to the first 10 years of simulated data. The parameter values are shown in Table 2 and discussed in Section 4.

shown in panel (a) of Figure 6, is similar with the one found for the market index. As attention increases, the short-term asset becomes more volatile, in a quadratic manner. All the intuition from Section 5 applies in this case as well.

Panel (b) of Figure 6 reveals that the short-term asset is more volatile than its associated dividend payout yielding excess volatility in the short-run. Next, the short-term asset volatility is larger in periods of poor dividend performance compared to periods of sustained growth. Figure 6 is thus in line with the findings of van Binsbergen et al. (2010).

When attention increases, the volatility of the short-term asset increases more than the volatility of the index itself. To illustrate this, we plot in Figure 7 the relationship between the short-term asset to market volatility ratio against the attention Φ . Obviously, the volatility of the short-term asset increases more with the attention than market volatility. The red scatter plot depicts the simulated values and the black line is the corresponding quadratic fit. There is a clear positive relationship between attention and the short-term asset to market volatility ratio.

Our model suggests that, in bad aggregate economic states, investors are more attentive to news. The increased attentiveness helps in reducing the uncertainty related to future expected growth rates which, in turn, tends to decrease the market volatility. However, higher attention speeds up information revelation, which increases volatility. The latter effect clearly dominates at high levels of attention. Finally, as attention

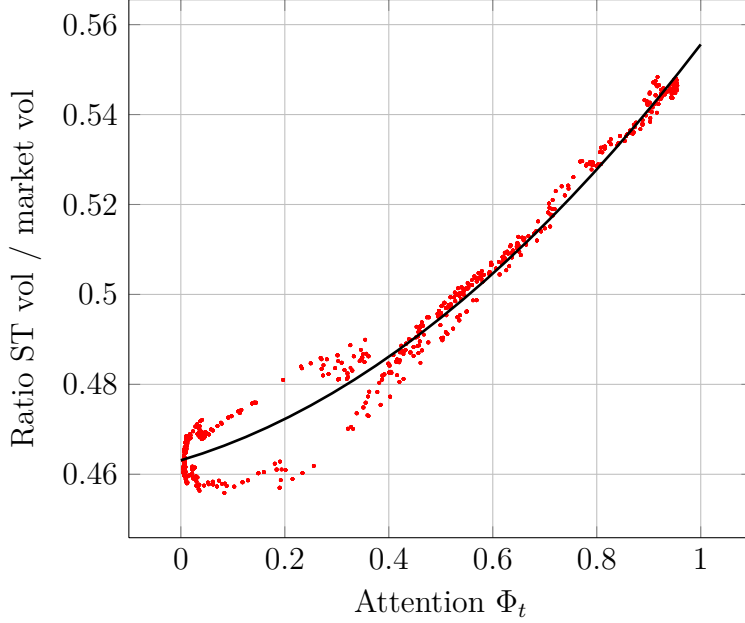


Figure 7: Short-term asset to market volatility ratio

Ratio of the short-term asset volatility to market volatility, as a function of attention. The black solid line represents a quadratic fit. The parameter values are shown in Table 2 and discussed in Section 4.

increases, expected dividend growth becomes more sensitive to news in the short-term than in the long-term. This implies a higher sensitivity to news for the short-term asset and, consequently, a higher short-term asset to market volatility ratio.

We now turn to the term structure of risk premia. The main questions that we address are whether there is more risk premium in the short term than in the long term and whether the attention can act upon the slope of the term structure of risk premia. The term structure can be computed by applying Itô's lemma to single dividend paying assets. Let $\delta_{t+\tau}$ denote the stochastic dividend paid out in τ years from today's date t , whose price today is denoted by $S_{t,\tau}$. We have

$$S_{t,\tau} = \mathbb{E}_t \left[\frac{\xi_{t,\tau}}{\xi_t} \delta_{t+\tau} \right]$$

$$\frac{dS_{t,\tau}}{S_{t,\tau}} = \mu_{t,\tau} dt + \sigma_{t,\tau} dW_t$$

The instantaneous risk premium of every dividend strip, $\mu_{t,\tau} - r_t$, can then be computed numerically. The risk premia for all maturities form a term structure of risk premia. As this term structure can be different according to state values, Figure 8 plots the term structure of risk premia in 3 situations. Panel (a) depicts a simulated path (10 years of weekly data) of the expected growth rate \hat{f} . There are 3 cases emphasized in the plot: case A (expected growth rate is at the maximum—good aggregate economic states), case B (expected growth rate is at the minimum—bad aggregate economic

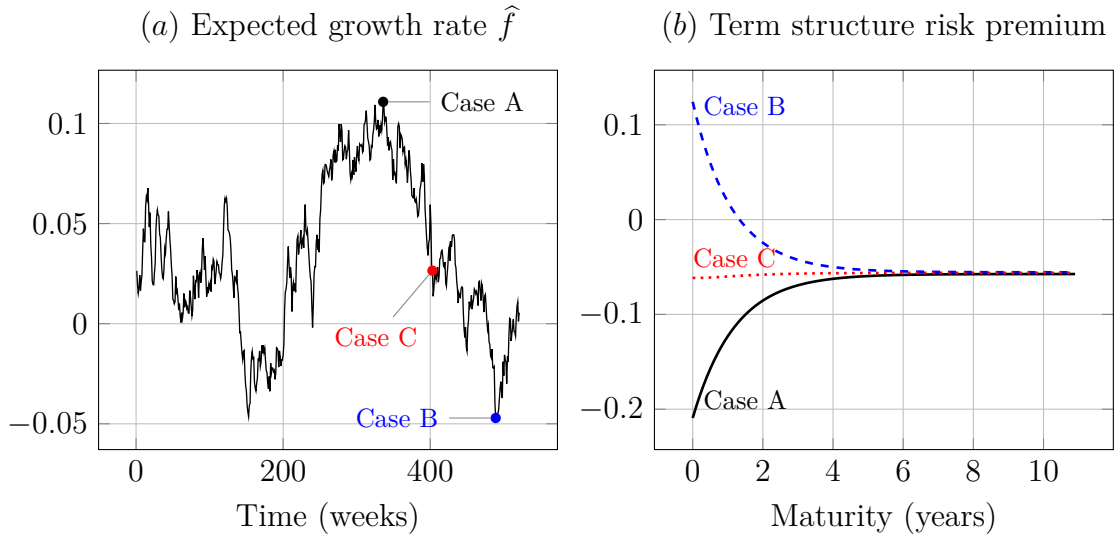


Figure 8: Term structure of risk premia

Panel (a) depicts a simulated path of the expected growth rate \hat{f} (10 years of weekly data). Panel (b) depicts the term structure of risk premia in the 3 cases outlined in panel (a): case A (expected growth rate is at the maximum—good aggregate economic states), case B (expected growth rate is at the minimum—bad aggregate economic states), and case C (expected growth rate is at its long-term mean \bar{f}). The parameter values are shown in Table 2 and discussed in Section 4.

states), and case C (expected growth rate is at its long-term mean \bar{f}). In case A the attention is at its lowest level, in case B the attention is at its highest level, while in case C the attention is near its long-term mean.

The resulting term structures of risk premia for each case are depicted in panel (b) of Figure 8. The term structure is upward-sloping in case A (good aggregate economic states), downward-sloping in case B (bad aggregate economic states), and almost neutral in case C (expected growth rate at long-term mean).

We uncover a downward slopping term structure of risk premia in bad aggregate economic states. Our result is therefore partly in line with van Binsbergen et al. (2010). Nonetheless, the relatively short data set (January 1996 to June 2009) used by van Binsbergen et al. (2010) does not completely rule out the possibility that the term structure of risk premia might be upward slopping in other periods. In that case, our theoretical model would be able to explain a time-varying slope of the term structure of risk premia.

Our general equilibrium model with fluctuating attention might therefore explain why van Binsbergen et al. (2010) found a higher risk premium and a higher volatility for the short-term asset.

8 The Term Structure of Forward Equity Yields

The term structure of forward equity yields obtained from dividend derivatives by van Binsbergen et al. (2011) is found to fluctuate strongly over time, more for short maturities than for long maturities. Moreover, van Binsbergen et al. (2011) uncover a pro-cyclical slope of the term structure of forward equity yields, whereas the slope of the term structure of expected dividend growth rates is counter-cyclical. Finally, forward equity yields strongly predict risk premia and dividend growth rates. For example, a high value of the forward equity yield implies that the expected dividend growth is low, and viceversa.

The present model is able to reproduce a term structure of forward yields consistent with van Binsbergen et al. (2011). As in the previous section, let $\delta_{t+\tau}$ denote the stochastic dividend paid out in τ years from t , whose price today is denoted by $S_{t,\tau}$. The equity yield at time t with maturity τ is defined as:

$$e_{t,\tau} = \frac{1}{\tau} \ln \left(\frac{\delta_t}{S_{t,\tau}} \right) = y_{t,\tau} + \theta_{t,\tau} - g_{t,\tau}$$

where $y_{t,n}$ is the nominal bond yield, $\theta_{t,\tau}$ is the risk premium required by investors to hold dividend risk of maturity τ , and $g_{t,\tau}$ is the per-period expected dividend growth rate. Note that $\theta_{t,\tau}$ is different from the instantaneous risk premium computed in the previous Section, as in that case we used the instantaneous risk free rate and not the bond yield.

Following van Binsbergen et al. (2011), the average per-period expected dividend growth rate over the next τ years is defined as:

$$\begin{aligned} g_{t,\tau} &= \frac{1}{\tau} \mathbb{E}_t \ln \left(\frac{\delta_{t+\tau}}{\delta_t} \right) \\ &= \hat{f}_t \frac{1 - e^{-\lambda\tau}}{\tau\lambda} + \bar{f} \left(1 - \frac{1 - e^{-\lambda\tau}}{\tau\lambda} \right) - \frac{\sigma_\delta^2}{2} \end{aligned} \quad (17)$$

As the maturity τ increases, the expected dividend growth approaches \bar{f} . On the contrary, as the maturity of the dividend goes to zero, the expected dividend growth approaches \hat{f}_t . The sum of the two factors multiplying \hat{f} and \bar{f} is 1, and describes the trade-off between using the posterior belief or the prior belief as maturity changes.

The forward equity yield is equal to the equity yield $e_{t,\tau}$ minus the nominal bond yield $y_{t,n}$:

$$e_{t,\tau}^f = e_{t,\tau} - y_{t,n} = \theta_{t,\tau} - g_{t,\tau} \quad (18)$$

Denoting by $B_{t,\tau}$ the price of a zero-coupon bond with maturity $t + \tau$, the resulting

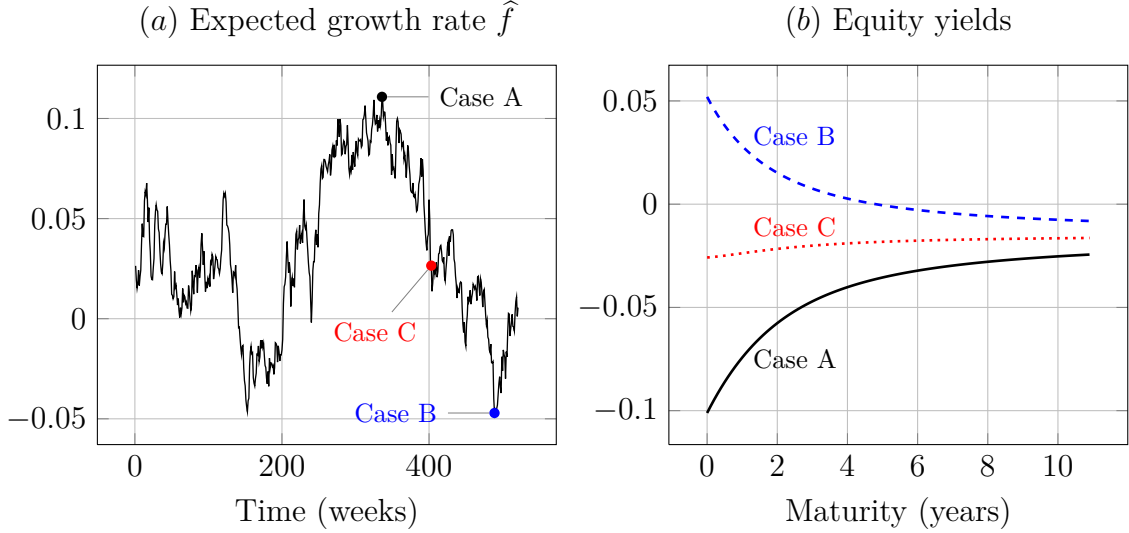


Figure 9: Term structure of forward equity yields

Panel (a) depicts a simulated path of the expected growth rate \hat{f} (10 years of weekly data). Panel (b) depicts the term structure of forward equity yields in the 3 cases outlined in panel (a): case A (expected growth rate is at the maximum—good aggregate economic states), case B (expected growth rate is at the minimum—bad aggregate economic states), and case C (expected growth rate is at its long-term mean \bar{f}). The parameter values are shown in Table 2 and discussed in Section 4.

nominal bond yield $y_{t,\tau}$ is equal to $(-1/\tau \ln B_{t,\tau})$. It is thus straightforward to compute forward equity yields in the context of the present model.

The term structure of forward equity yields is a function of the state variables. It is, therefore, time-varying. To seize its dynamics, we use the same technique as in the previous section and we plot in Figure 9 the term structure of forward equity yields for 3 different cases.

Panel (a) of Figure 9 depicts a simulated path (10 years of weekly data) of the expected growth rate \hat{f} . There are 3 cases emphasized in the plot: case A (expected growth rate is at the maximum—good aggregate economic states), case B (expected growth rate is at the minimum—bad aggregate economic states), and case C (expected growth rate is at its long-term mean \bar{f}).

In panel (b) of Figure 9 we build term structures of forward equity yields for each case. The term structure is upward-sloping in case A (good aggregate economic states), downward-sloping in case B (bad aggregate economic states), and almost neutral in case C (expected growth rate at long-term mean). In other words, the slope of the term structure of forward equity yields is pro-cyclical.

Can forward equity yields help predicting future dividend growth? Equation (18) implies that, by definition, forward equity yields must either predict risk premia or expected dividend growth, or both. High values of the forward equity yield imply that either risk premia are high or expected dividend growth are low, or both. In

the context of the present model, maturity risk premia $(\theta_{t,\tau})$ movements are small. Consequently, most of the variation in forward equity yields is driven by variation in expected dividend growth. Dividend growth rates are therefore strongly predictable, more so at shorter horizons, a result outlined in [van Binsbergen et al. \(2011\)](#).

Our model implies that forward equity yields are more volatile at shorter maturities. Equation (17) suggests that there are two effects on the volatility of expected dividend growth rates. First, a shorter maturity implies that the expected dividend growth approaches \hat{f}_t and thus is becoming more volatile. Second, for any maturity, as the current assessment of the growth rate \hat{f}_t is more volatile the expected growth rate becomes more volatile. The latter effect is particular to our model of fluctuating attention. That is, when attention is higher, the expected dividend growth rates are more volatile, more so in the short term.

And since variation in forward equity yields is driven by variation in expected dividend growth, it follows that short maturity equity yields are more volatile. Moreover, a higher attention strongly amplifies this effect.

Finally, since forward equity yields and expected dividend growth rates are negatively related, a pro-cyclical slope of the term structure of forward equity yields implies a counter-cyclical slope of the term structure of expected dividend growth rates. Figure 9 shows that, when the economy switches from good to bad times, equity yields increase because expected growth rates decline, which is in line with [van Binsbergen et al. \(2011\)](#). In our case, however, almost all the variation in forward equity yields is driven by variation in expected dividend growth rates, whereas [van Binsbergen et al. \(2011\)](#) found that dividend growth variation accounts for 62 – 81% of the variation in forward equity yields.

To make into a short statement the main points of this section, our model of fluctuating attention implies that forward equity yields with short maturities fluctuate strongly over time. Next, the slope of the term structure of forward equity yields is pro-cyclical, and the slope of the term structure of expected dividend growth rates is counter-cyclical. Finally, forward equity yields strongly predict dividend growth rates: a high value of the forward equity yield implies that the expected dividend growth is low, and viceversa.

9 Conclusion

We consider a continuous-time pure exchange economy where a single investor filters the unobservable fundamental by observing the output process and a signal representing the flow of information that she acquires. The accuracy of the signal is assumed to be stochastic and depends on a dividend performance index. We call this accuracy the investor’s attention to news.

The estimation performed on US data suggests that a period of poor dividend performance pushes the investor to acquire accurate information about the fundamental, whereas a period of high dividend performance leads to a lack of information acquisition. In other words, the investor willingness to gather accurate information is high in bearish periods and low in bullish periods.

By performing an affine approximation of the state vector, we solve for stock prices and focus, in particular, on the stock return volatility. We show that there is a quadratic relationship between attention and stock market volatility, relationship that we confirm empirically.

Moreover, because the attention is counter-cyclical, we show that volatility is counter-cyclical and the risk premium is counter-cyclical. During downturns, the short-term asset, computed as in [van Binsbergen et al. \(2010\)](#), is excessively volatile and commands a large equity premium. Moreover, the slope of forward equity yields is pro-cyclical, as shown empirically by [van Binsbergen et al. \(2011\)](#).

Several questions are now the subject of our ongoing research. First, this paper considers only one dimension of uncertainty, i.e., *learning uncertainty*. Another dimension of uncertainty comes from the way beliefs differ across investors—*dispersion of beliefs*. As [Massa and Simonov \(2005\)](#) show that both dimensions are priced, it is of interest to integrate both of them in the same setup. If both dimensions are priced, our conjecture is that if different investors learn from different sources of information, spikes in attention might contribute to polarization of beliefs. Thus, the effects produced by these two dimensions of risk might reinforce each other.

Next, dispersion of beliefs could come from different information or from different priors. We explore the path of different priors (model heterogeneity) in [Andrei and Hasler \(2012\)](#), although without fluctuating attention. In subsequent research we plan to see how fluctuating attention interacts with model heterogeneity.

Finally, other aspect worthwhile considering for future research is a general equilibrium model with endogenous attention, in which prices could also dictate the level of attention.

10 Appendix

10.A Unconditional moments of ϕ

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

$$B = \begin{bmatrix} \lambda & 0 \\ -1 & \omega \end{bmatrix}$$

and

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

The solution is found by applying Itô's lemma to

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

After integrating from 0 to t we obtain

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

Thus, the first moments of f and ϕ solve the following system of equations

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

It follows that the long term mean of f is \bar{f} and the long term mean of ϕ is $\frac{\bar{f}}{\omega}$. 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 $\frac{\sigma_f^2}{2\lambda}$. The long term variance of ϕ is found by replacing the solution for f in

$$\frac{e^{\lambda t} - e^{\omega t}}{\omega - \lambda} f_t + e^{\omega t} \phi_t - \phi_0 = \int_0^t \frac{\lambda \bar{f} (e^{\omega u} - e^{\lambda u})}{\lambda - \omega} du + \int_0^t \frac{\sigma_f (e^{\omega u} - e^{\lambda u})}{\lambda - \omega} dZ_u^f + \int_0^t \sigma_\delta e^{\omega u} dZ_u^\delta$$

and computing

$$\text{Var}[\phi_t] = \mathbb{E}[(\phi_t - \mathbb{E}[\phi_t])(\phi_t - \mathbb{E}[\phi_t])].$$

Finally, the long term variance of ϕ is

$$\lim_{t \rightarrow +\infty} \text{Var} [\phi_t] = \frac{\sigma_f^2}{2\lambda\omega(\lambda + \omega)} + \frac{\sigma_\delta^2}{2\omega}.$$

10.B Details on ζ , \hat{f} , ϕ , and γ

We have

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

or (as in [Liptser and Shiryaev, 2001](#))

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

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 γ represents the posterior variance. Notice that the dynamics of γ depend on ϕ 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 γ in the state space.

10.C Solutions for ζ , \hat{f} , ϕ , and γ

Since the dividend process δ 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 ϕ , 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

$$B = \begin{bmatrix} \lambda & 0 \\ -1 & \omega \end{bmatrix}$$

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

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

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

The dynamics of F obey

$$dF_t = \begin{bmatrix} \frac{e^{t\lambda} (\sigma_f \sigma_\delta \Phi_t dW_t^s + \gamma_t dW_t^\delta + dt \lambda \bar{f} \sigma_\delta)}{dW_t^\delta (e^{t\omega} ((\lambda - \omega) \sigma_\delta^2 + \gamma_t) - e^{t\lambda} \gamma_t) - (e^{t\lambda} - e^{t\omega}) \sigma_\delta (\sigma_f \Phi_t dW_t^s + dt \lambda \bar{f})} \\ \frac{\sigma_\delta}{(\lambda - \omega) \sigma_\delta} \end{bmatrix}.$$

After integrating from v to t and rearranging we obtain

$$\hat{f}_t = e^{-\lambda(t-v)} \hat{f}_v + \bar{f} (1 - e^{-\lambda(t-v)}) + \frac{1}{\sigma_\delta} \int_v^t e^{-\lambda(t-u)} \gamma_u dW_u^\delta + \sigma_f \int_v^t e^{-\lambda(t-u)} \Phi_u dW_u^s \quad (19)$$

$$\phi_t = \phi_v e^{-\omega(t-v)} + e^{-\omega t} \left[\frac{e^{\lambda v} - e^{\omega v}}{\omega - \lambda} \hat{f}_v - \frac{e^{\lambda t} - e^{\omega t}}{\omega - \lambda} \hat{f}_t + \frac{\lambda \bar{f}}{\omega - \lambda} \left(\frac{e^{\omega v} - e^{\omega t}}{\omega} + \frac{e^{\lambda t} - e^{\lambda v}}{\lambda} \right) \right. \\ \left. + \int_v^t \left(\sigma_\delta e^{\omega u} - \frac{\gamma_u}{\sigma_\delta (\lambda - \omega)} (e^{\lambda u} - e^{\omega u}) \right) dW_u^\delta + \frac{\sigma_f}{\omega - \lambda} \int_v^t \Phi_u (e^{\lambda u} - e^{\omega u}) dW_u^s \right]. \quad (20)$$

Substituting Equation (19) in Equation (20) yields the desired result. Finally, the dynamics of the posterior variance γ 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

$$\Delta = t - v$$

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

10.D Moment Conditions

Let's define the observable process d as

$$d_{t+\Delta} \equiv \log \frac{\delta_{t+\Delta}}{\delta_t} - \hat{f}_t \Delta = -\frac{1}{2} \sigma_\delta^2 \Delta + \sigma_\delta \epsilon_{t+\Delta}^\delta. \quad (21)$$

Equation (21) characterizes the first moment condition. That is,

$$\text{Var}(d_{t+\Delta}) = \sigma_\delta^2 \Delta. \quad (22)$$

The empirical counterpart of Equation (22) writes

$$\frac{1}{T} \sum_{i=1}^T [(d_{i\Delta} - \mu_d)^2 - \sigma_\delta^2 \Delta] = 0 \quad (23)$$

where $T = 165$ is the number of observations and $\mu_d = \frac{1}{T} \sum_{i=1}^T d_{i\Delta}$. The conditional expectation of the filtered fundamental \hat{f} satisfies

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

Its empirical counterpart is then

$$\frac{1}{T} \sum_{i=1}^T [\hat{f}_{i\Delta} - e^{-\lambda\Delta} \hat{f}_{(i-1)\Delta} - \bar{f} (1 - e^{-\lambda\Delta})] = 0. \quad (24)$$

Note that over $T = 165$ observations, the mean of $\hat{f}_{t+\Delta}$ is roughly equal to the mean of \hat{f}_t . Consequently, Equation (24) simplifies to

$$\frac{1}{T} \sum_{i=1}^T [\hat{f}_{i\Delta} - \bar{f}] = 0. \quad (25)$$

Unsurprisingly, the empirical average of an Ornstein-Uhlenbeck process has to be equated to its long term mean. The parameter λ is determined through the covariance between $\hat{f}_{t+\Delta}$ and \hat{f}_t . Indeed, simple computations yield the following equality

$$\text{Cov}(\hat{f}_{t+\Delta}, \hat{f}_t) = e^{-\lambda\Delta} \text{Var}(\hat{f}_t). \quad (26)$$

The empirical counterpart of Equation (26) is

$$\frac{1}{T} \sum_{i=1}^T [(\hat{f}_{i\Delta} - \mu_f)(\hat{f}_{(i-1)\Delta} - \mu_f) - e^{-\lambda\Delta} (\hat{f}_{(i-1)\Delta} - \mu_f)^2] = 0 \quad (27)$$

where $\mu_f = \frac{1}{T} \sum_{i=1}^T \hat{f}_{i\Delta}$. Equations (23), (25), and (27) directly determine the parameters σ_δ , \bar{f} , and λ . The parameters σ_f , ω , Λ , and Ψ have to be estimated jointly, because they all depend on each other. Since the steady-state value of the process γ is worth γ_{ss} , the empirical average of the posterior variance γ is set to γ_{ss} . That is,

$$\frac{1}{T} \sum_{i=1}^T [\gamma_{i\Delta} - \gamma_{ss}] = 0 \quad (28)$$

where γ_{ss} is defined in Section 3.5. Note that γ depends on the attention Φ , Φ depends on the dividend performance ϕ , and ϕ is driven by \hat{f} , ϵ^δ , and ϵ^s . Hence, the posterior variance

γ_t , the attention Φ_t , and the dividend performance index ϕ_t , for $t = 0, \Delta, \dots, T\Delta$, can be constructed recursively. Assuming that σ_δ , \bar{f} , and λ are already known, the vectors γ , Φ , and ϕ are T -dimensional functions of the parameters σ_f , ω , Λ , and Ψ . Equation (20) shows that the long term mean of the dividend performance ϕ is $\frac{\bar{f}}{\omega}$. Consequently, the fifth moment condition is

$$\frac{1}{T} \sum_{i=1}^T \left[\phi_{i\Delta} - \frac{\bar{f}}{\omega} \right] = 0. \quad (29)$$

The process Φ has a long term mean equal to Ψ . Therefore, the sixth moment condition is written

$$\frac{1}{T} \sum_{i=1}^T [\Phi_{i\Delta} - \Psi] = 0. \quad (30)$$

Equation (19) permits to characterize the conditional variance of \hat{f}

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

The empirical counterpart is given by

$$\begin{aligned} \frac{1}{T} \sum_{i=1}^T \left[\left(\hat{f}_{i\Delta} - e^{-\lambda\Delta} \hat{f}_{(i-1)\Delta} - \bar{f} (1 - e^{-\lambda\Delta}) \right)^2 \right. \\ \left. - \left(\frac{1}{\sigma_\delta^2} \gamma_{(i-1)\Delta}^2 + \sigma_f^2 e^{-2\lambda\Delta} \Phi_{(i-1)\Delta}^2 \right) e^{-2\lambda\Delta} \Delta \right] = 0 \end{aligned} \quad (31)$$

To summarize, Equations (23), (25), (27), (28), (29), (30), and (31) define a system of 7 moment conditions that needs to be solved to obtain the 7-dimensional vector of parameters Θ .

References

- [1] Abel, A., J. Eberly, and S. Panageas (2007). Optimal inattention to the stock market. *American Economic Review* 97, 244–249. [9]
- [2] Andrei, D. and M. Hasler (2012). Volatility clustering through persistent disagreement. *Working Paper, UNIL and EPFL*. [33]
- [3] Bacchetta, P. and E. v. Wincoop (2010). Infrequent portfolio decisions: A solution to the forward discount puzzle. *American Economic Review* 100(3), 870–904. [9]
- [4] 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. [17, 25, 26]
- [5] 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. [17]
- [6] Brennan, M. J. and Y. Xia (2001). Stock price volatility and equity premium. *Journal of Monetary Economics* 47(2), 249 – 283. [5, 19, 25]
- [7] Campbell, J. Y. (1987, June). Stock returns and the term structure. *Journal of Financial Economics* 18(2), 373–399. [2, 25]
- [8] Campbell, J. Y. and J. H. Cochrane (1999). By force of habit: A consumption-based explanation of aggregate stock market beha. *Journal of Political Economy* 107, 205–251. [26]
- [9] Chien, Y.-L., H. L. Cole, and H. Lustig (2009, September). Is the volatility of the market price of risk due to intermittent portfolio re-balancing? NBER Working Papers 15382, National Bureau of Economic Research, Inc. [9]
- [10] Cox, J. C. and C.-f. Huang (1989). Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory* 49(1), 33 – 83. [14]
- [11] Da, Z., J. Engelberg, and P. Gao (2011, October). In search of attention. *Journal of Finance* 66(5), 1461–1499. [3, 4, 5, 26]
- [12] Da, Z., U. G. Gurun, and M. Warachka (2011). Frog in the pan: Continuous information and momentum. *SSRN eLibrary*. [9, 18]
- [13] 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]
- [14] David, A. (2008). Heterogeneous beliefs, speculation and the equity premium. *Journal of Finance* 63, 41–83. [16]
- [15] Detemple, J. and R. Kihlstrom (1987). Acquisition d'information dans un modèle intertemporel en temps continu. *L'Actualité économique* 63, 118–137. [5, 7, 9, 23]
- [16] Dimpfl, T. and S. Jank (2011). Can internet search queries help to predict stock market volatility? University of Tübingen. [3, 5]
- [17] Duffie, D. (2008). *Financial Modeling with Affine Processes*. Preliminary and Incomplete Draft. [14]
- [18] Duffie, D. (2010, 08). Presidential address: Asset price dynamics with slow-moving capital. *Journal of Finance* 65(4), 1237–1267. [9]
- [19] 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. [5, 9]
- [20] 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. [6, 12]

- [21] Gabaix, X. (2008, January). Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. NBER Working Papers 13724, National Bureau of Economic Research, Inc. [26]
- [22] Garcia, D. (2012). Sentiment during recessions. *Working Paper*. [9, 18, 22, 23]
- [23] Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993, December). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48(5), 1779–1801. [2, 25]
- [24] Hansen, L. P. (1982, July). Large sample properties of generalized method of moments estimators. *Econometrica* 50(4), 1029–54. [17]
- [25] Hasler, M. (2012). Portfolio choice and costly dynamic information acquisition. *Working Paper, EPFL*. [5, 9, 18]
- [26] Huang, L. and H. Liu (2007, 08). Rational inattention and portfolio selection. *Journal of Finance* 62(4), 1999–2040. [5, 7, 9]
- [27] Karatzas, I., J. P. Lehoczky, and S. E. Shreve (1987). Optimal portfolio and consumption decisions for a “small investor” on a finite horizon. *SIAM Journal on Control and Optimization* 25(6), 1557–1586. [14]
- [28] Kita, A. and Q. Wang (2012). Investor attention and fx market volatility. Bangor Business School. [3, 5]
- [29] Koijen, R. S. J., J. C. Rodriguez, and A. Sbuelz (2009). Momentum and mean reversion in strategic asset allocation. *MANAGEMENT SCIENCE* 55(7), 1199–1213. [8]
- [30] Liptser, R. S. and A. N. Shiryaev (2001). *Statistics of Random Processes II*. Springer Verlag, New York. [35]
- [31] Lucas, Robert E., J. (1978). Asset prices in an exchange economy. *Econometrica* 46(6), pp. 1429–1445. [6]
- [32] Massa, M. and A. Simonov (2005, October). Is learning a dimension of risk? *Journal of Banking & Finance* 29(10), 2605–2632. [3, 33]
- [33] Mele, A. (2007, November). Asymmetric stock market volatility and the cyclical behavior of expected returns. *Journal of Financial Economics* 86(2), 446–478. [3, 24, 25]
- [34] Mele, A. (2008). Understanding stock market volatility: A business cycle perspective. *Working Paper*. [3, 25]
- [35] Newey, W. K. and K. D. West (1987, May). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55(3), 703–08. [4]
- [36] Ozoguz, A. (2009, November). Good times or bad times? investors’ uncertainty and stock returns. *Review of Financial Studies* 22(11), 4377–4422. [3, 19, 25]
- [37] Peng, L. and W. Xiong (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics* 80(3), 563 – 602. [5, 7, 9]
- [38] Peng, L., W. Xiong, and T. Bollerslev (2007). Investor attention and time-varying comovements. *European Financial Management* 13(3), 394–422. [26]
- [39] Rossi, A. (2010). Optimal inattention to the stock market, portfolio allocation, and consumption. *Working Paper*. [9]
- [40] Scheinkman, J. A. and W. Xiong (2003). Overconfidence and speculative bubbles. *The Journal of Political Economy* 111(6), 1183–1219. [12, 35]
- [41] Timmermann, A. (2001, July). Structural breaks, incomplete information, and stock prices. *Journal of Business & Economic Statistics* 19(3), 299–314. [19]

- [42] 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. [19]
- [43] van Binsbergen, J., M. Brandt, and R. Koijen (2010). On the timing and pricing of dividends. *Working Paper*. [3, 5, 26, 27, 29, 33]
- [44] van Binsbergen, J. H., W. Hueskes, R. Koijen, and E. B. Vrugt (2011, September). Equity yields. NBER Working Papers 17416, National Bureau of Economic Research, Inc. [3, 5, 30, 32, 33]
- [45] 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, 5, 13, 19, 25]
- [46] Veronesi, P. (2000, April). How does information quality affect stock returns? *Journal of Finance* 55(2), 807–837. [2, 5, 7, 13, 19, 20, 22, 23, 24, 25]
- [47] Vlastakis, N. and R. N. Markellos (2012). Information demand and stock market volatility. *Journal of Banking and Finance* 36, 1808–1821. [3, 5]