# Why Does Return Predictability Concentrate in Bad Times?*

Julien Cujean† Michael Hasler‡

September 22, 2015

**Abstract**

We build an equilibrium model to explain why stock return predictability is concentrated in bad times. The key ingredient is counter-cyclical investors' disagreement, which originates from heterogeneous models. As economic conditions deteriorate, difference in investors' learning speed increases, investors' opinions eventually polarize and disagreement spikes, causing returns to react to past news. The model explains several findings: the link between disagreement and future returns, time series momentum and how it crashes after sharp market rebounds. The model further predicts that time series momentum increases with disagreement and strengthens in bad times. We provide empirical support to these new predictions.

**Keywords.** Equilibrium Asset Pricing, Learning, Disagreement, Business Cycle, Predictability, Times Series Momentum, Momentum Crashes.

**JEL Classification.** D51, D83, G12, G14.

*We are particularly grateful to Kenneth Singleton and two anonymous referees for their insightful suggestions and comments. We would like to thank Alex Kostakis, Hui Chen, Peter Christoffersen, Pierre Collin-Dufresne, Michel Dubois, Darrell Duffie, Bernard Dumas, Laurent Frésard, Steve Heston, Julien Hugonnier, Alexandre Jeanneret, Scott Joslin, Andrew Karolyi, Leonid Kogan, Jan-Peter Kulak, Pete Kyle, Jeongmin Lee, Mark Loewenstein, Semyon Malamud, Erwan Morellec, Antonio Mele, Yoshio Nozawa, Chayawat Ornthanalai, Lubos Pastor, Lasse Pedersen, Rémy Praz, Marcel Rindisbacher, Alberto Rossi, René Stulz, Ngoc-Khanh Tran, Adrien Verdelhan, Pietro Veronesi, Jason Wei, Liyan Yang, and conference/seminar participants at the 4th Financial Risks International Forum, Collegio Carlo Alberto, EFA 2015, Eurofidai 2014, FIRS 2015, Goethe University Frankfurt, HEC Montréal, SFI-NCCR Workshop, UBC Winter Finance Conference 2015, University of California at San Diego, University of Geneva, University of Maryland, University of Neuchâtel, University of Toronto, and University of Virginia for their discussions and comments. Financial support from the University of Maryland, the University of Toronto, and the Connaught New Researcher Award is gratefully acknowledged. A previous version of this paper circulated under the title "Time Series Predictability, Investors' Disagreement, and Economic Conditions".

†University of Maryland, Robert H. Smith School of Business, 4466 Van Munching Hall, College Park, MD 20742, USA; +1 (301) 405 7707; jcujean@rhsmith.umd.edu; www.juliencujean.com

‡University of Toronto, Rotman School of Management, 105 St. George Street, Toronto, ON, M5S 3E6, Canada; +1 (416) 946 8494; michael.hasler@rotman.utoronto.ca; www.rotman.utoronto.ca/Hasler

# 1 Introduction

Stock return predictability concentrates in bad times.[1] For instance, Garcia (2013) shows that the content of news (Tetlock, 2007) better predicts future returns in recessions. Yet, the reason return predictability varies over the business cycle remains unclear.

In this paper, we provide a theoretical mechanism that causes stock return predictability to concentrate in bad times, a mechanism we validate empirically. We base our explanation on the empirical finding that disagreement among professional forecasters originates from model heterogeneity and moves *counter-cyclically* (Patton and Timmermann, 2010).[2] The main idea is that investors revise their expectations at different speeds. Some investors focus on sharp variations in the business cycle and revise their expectations rapidly, while others focus on persistent fluctuations and revise their expectations slowly. As economic conditions deteriorate, the difference in adjustment speeds increases, investors' opinions eventually polarize and disagreement spikes, causing returns to react to past news. Return predictability therefore concentrates in bad times.

We introduce model heterogeneity within a dynamic general equilibrium populated with two investors, $A$ and $B$. Investors trade one stock and a riskless bond and consume the dividends that the stock pays out. Agents do not observe the expected growth rate of dividends, which we call the fundamental, and have different views regarding the empirical process that governs it. Agent $A$ assumes that the fundamental transits smoothly from good to bad times. In contrast, Agent $B$ postulates that the fundamental may transit precipitously from good to bad times. While investors observe the same data, they disagree about the fundamental because they use different models and therefore interpret the data differently.

Model heterogeneity causes investors to revise their expectations at different speeds and to reassess fundamental uncertainty differently depending on economic conditions.[3] The pattern

[1] Rapach, Strauss, and Zhou (2010), Henkel, Martin, and Nardari (2011), and Dangl and Halling (2012) find that macro variables, such as the price-dividend ratio, have better predictive power in recessions. Similarly, Cen, Wei, and Yang (2014) and Loh and Stulz (2014) find that return predictability using investors' disagreement, as proxied by the dispersion of analysts' forecasts (Diether, Malloy, and Scherbina, 2002), is concentrated in recessions. We provide evidence that current excess returns and disagreement better predict aggregate future excess returns in recessions (Section 5).

[2] Similarly, Kandel and Pearson (1995) show that disagreement stems from model heterogeneity and Veronesi (1999), Carlin, Longstaff, and Matoba (2014), and Barinov (2014) show that it is counter-cyclical.

[3] Chalkley and Lee (1998), Veldkamp (2005), and Van Nieuwerburgh and Veldkamp (2006) obtain similar learning asymmetries. In their setup, however, the information flow fluctuates with economic conditions.

of investors' disagreement therefore changes over the business cycle. In good times, investors' expectations adjust at comparable speeds and disagreement exhibits little variation. The difference in investors' adjustment speeds becomes apparent in normal times. While investors have similar views, Agent $B$ adjusts her expectations significantly faster than Agent $A$, which exacerbates disagreement in the short term. As economic conditions deteriorate, the difference in investors' assessment of uncertainty strengthens. In the eyes of Agent $A$, Agent $B$ overstates bad economic outcomes. This polarization of opinions results in large disagreement spikes in bad times. We show that this counter-cyclical pattern of disagreement is strongly positively correlated to the observed dispersion of analysts' forecasts.

Counter-cyclical disagreement leads return predictability to concentrate in bad times. To see this, suppose there is a good news today. In bad times, the news polarizes opinions: while Agent $A$ revises her expectations upwards, Agent $B$ revises her expectations downwards. Agent $B$'s pessimism in turn induces returns to move opposite to the news, thereby producing short-term under-reaction. In normal times, the news exacerbates the difference in adjustment speeds: while the news corroborates both agents' expectations, it precipitates an upward revision in Agent $B$'s expectations. Agent $B$'s overoptimism relative to Agent $A$ in turn causes returns to over-react in the short term. In the long term, agents' opinions align and returns systematically revert. In good times, predictability vanishes because the news generates little disagreement and reversion to fundamentals therefore occurs instantly.

Return continuation in bad and normal times produces time series momentum in excess returns (Moskowitz, Ooi, and Pedersen, 2012). Consistent with empirical findings, returns exhibit momentum over a 1-year horizon and then revert over subsequent horizons. Moreover, the term structure of returns' serial correlation is hump shaped: momentum is strong at short horizons and decays at intermediate horizons. Importantly, at short horizons, time series momentum is strongest in bad times. The reason is that the polarization of opinions in bad times, as opposed to the difference in adjustment speeds in normal times, leads to a sharper spike in disagreement. By contrast, in good times, excess returns exhibit strong reversal at short horizons. An important consequence of this reversal spike is that a time series momentum strategy may crash after sharp market rebounds.

We provide empirical support to three new predictions of the model: *(1)* future excess returns are positively related to contemporaneous disagreement. Furthermore, time series momentum at short horizons *(2)* increases with disagreement and *(3)* is, therefore, strongest

in bad times. Using the dispersion of analysts' forecasts as an empirical proxy for disagreement to mirror our theoretical analysis, we first show that dispersion positively predicts future excess returns on the S&P 500. We then show that time series momentum increases significantly with dispersion at short horizons. We finally show that, over the last century, time series momentum at a one-month lag is significantly stronger during recessions.

This paper contributes to the vast literature on heterogeneous beliefs. While we borrow, to a large extent, the methodology of Dumas, Kurshev, and Uppal (2009) and the numerical method of David (2008), we introduce a new form of disagreement. In contrast with other sources of disagreement considered in the literature, investors in this paper disagree because they use heterogeneous models (an autoregressive process and a Markov chain).[4] The resulting pattern of disagreement differs from the dynamics that prevail when agents use different sets of parameters within the same parametric model. Specifically, when both agents use a Markov chain with different parameters (David, 2008), beliefs polarize in bad times, as in our model, but disagreement does not spike under our calibration—we need disagreement spikes to generate return predictability. When both agents use an autoregressive model with different parameters (e.g., Buraschi and Whelan (2013), Ehling, Gallmeyer, Heyerdahl-Larsen, and Illeditsch (2013) and Buraschi, Trojani, and Vedolin (2014)), disagreement concentrates in normal times—we need counter-cyclical disagreement to explain why predictability concentrates in bad times.

This paper is also related to the large literature on momentum. In particular, other studies link heterogeneous beliefs to momentum.[5] In a rational-expectations model, Banerjee, Kaniel, and Kremer (2009) show that heterogeneous beliefs only generate price drift in the presence of higher-order difference of opinions. In contrast, we obtain momentum, even

[4] In our model, Agent $A$'s model follows a mean-reverting process (Detemple (1986, 1991), Wang (1993), Brennan and Xia (2001), Scheinkman and Xiong (2003), Bansal and Yaron (2004), and Dumas et al. (2009)), while Agent $B$'s model follows a Markov switching process (David (1997, 2008), Veronesi (1999, 2000), Chen (2010), Bhamra, Kuehn, and Strebulaev (2010), and David and Veronesi (2013)). In the literature, the pattern of disagreement differs whether investors are overconfident (e.g., Dumas et al. (2009), Scheinkman and Xiong (2003), Xiong and Yan (2010)), whether they have different initial priors (e.g., Detemple and Murthy (1994), Zapatero (1998), Basak (2000)), or whether they have heterogeneous dogmatic beliefs (e.g., Kogan, Ross, Wang, and Westerfield (2006), Borovicka (2011), Chen, Joslin, and Tran (2012), Bhamra and Uppal (2013)). See Xiong (2014) for a literature review

[5] Other theories of momentum include Berk, Green, and Naik (1999), Holden and Subrahmanyam (2002), Johnson (2002), Sagi and Seasholes (2007), Makarov and Rytchkov (2012), Vayanos and Woolley (2013), Biais, Bossaerts, and Spatt (2010), Cespa and Vives (2012), Cujean (2013), Albuquerque and Miao (2014), and Andrei and Cujean (2014).

when heterogeneous prior beliefs are commonly known. Importantly, as in Ottaviani and Sorensen (2015), we also generate price under-reaction, short-term momentum and long-term reversal through the combination of heterogeneous beliefs and wealth effects. Unlike Ottaviani and Sorensen (2015), however, our paper offers an explanation for fluctuations in momentum over the business cycle. Specifically, counter-cyclical disagreement causes prices to under-react (Hong and Stein, 1999) in bad times and to over-react (Daniel, Hirshleifer, and Subrahmanyam, 1998) in normal times.[6] Our emphasis is that these two phenomena may alternate over the business cycle and that the former effect is stronger at short horizons.

The remainder of the paper is organized as follows. Section 2 presents and solves the model, Section 3 calibrates the model to the U.S. business cycle, Section 4 contains theoretical results, Section 5 tests the predictions of the model, and Section 6 concludes. Derivations and computational details are relegated to the Internet Appendix A.

# 2 The Model

We develop a dynamic general equilibrium in which investors use heterogeneous models to estimate the business cycle. We first describe the economy and the learning problem of investors. We then solve for the equilibrium stock price.

## 2.1 The Economy and Models of the Business Cycle

We consider an economy with an aggregate dividend that flows continuously over time. The market consists of two securities, a risky asset—the stock—in positive supply of one unit and a riskless asset—the bond—in zero net supply. The stock is a claim to the dividend process, $\delta$, which evolves according to

$$\mathrm{d}\delta_t = \delta_t f_t \mathrm{d}t + \delta_t \sigma_\delta \mathrm{d}W_t. \tag{1}$$

The random process $(W_t)_{t\geq 0}$ is a Brownian motion under the physical probability measure, which governs the empirical realizations of dividends. The expected dividend growth rate $f$—henceforth the *fundamental*—is unobservable.

[6] Barberis, Shleifer, and Vishny (1998) provide conditions to enforce under- or over-reaction, or both.

The economy is populated by two agents, $A$ and $B$, who consume the dividend and trade in the market. Agents understand that the fundamental affects the dividend they consume and the price of the assets they trade. Since the fundamental is unobservable, agents need to estimate it. To do so, they use the empirical realizations of dividends, the only source of information available. This information is, however, meaningless without a proper understanding of the data-generating process that governs dividends; agents need to have a model in mind.

Agent $A$ assumes that the fundamental follows a mean-reverting process. In particular, she has the following model in mind

$$\begin{aligned} \mathrm{d}\delta_t &= f_t^A \delta_t dt + \sigma_\delta \delta_t \mathrm{d}W_t^A \\ \mathrm{d}f_t^A &= \kappa \left(\overline{f} - f_t^A\right) \mathrm{d}t + \sigma_f \mathrm{d}W_t^f, \end{aligned} \tag{2}$$

where $W^A$ and $W^f$ are two independent Brownian motions under Agent $A$'s probability measure $\mathbb{P}^A$, which reflects her views about the data-generating process. Under Agent $A$'s representation of the economy, the fundamental $f^A$ transits smoothly from good states to bad states of the business cycle, reverting to a long-term mean $\overline{f}$ at speed $\kappa$.

Agent $B$, instead, believes that the fundamental follows a 2-state continuous-time Markov chain and therefore uses the following model

$$\begin{aligned} &\mathrm{d}\delta_t = f_t^B \delta_t dt + \sigma_\delta \delta_t \mathrm{d}W_t^B \\ &f_t^B \in \{f^h, f^l\} \text{ with generator matrix } \Lambda = \begin{pmatrix} -\lambda & \lambda \\ \psi & -\psi \end{pmatrix}, \end{aligned} \tag{3}$$

where $W^B$ is a Brownian motion under $B$'s probability measure $\mathbb{P}^B$. Under Agent $B$'s model, the fundamental $f^B$ is either high $f^h$ or low $f^l$. The economy transits from the high to the low state with intensity $\lambda > 0$ and from the low to the high state with intensity $\psi > 0$.

Each model has a natural interpretation. While Agent $A$'s model is typically used to fit consumption growth in the long-run risk literature (Bansal and Yaron (2004)), the perspective of Agent $B$ is closer to models used in the economics literature to forecast business cycle

turning points.[7] The financial economics literature has focused, to a large extent, on the two types of model presented in Equations (2) and (3) to forecast the growth rate of dividends. These models, however, have always been considered separately. In this paper, we consider these models within a unified framework and show that the resulting disagreement among investors can explain several empirical facts related to return predictability.

## 2.2 Bayesian Learning and Disagreement

Agents learn about the fundamental by observing realizations of the dividend growth rate and, given the model they have in mind, update their expectations accordingly. Doing so, they come up with an estimate of the fundamental, which, from now on, we call the *filter*. We present the dynamics of the filter of Agents $A$ and $B$ in Proposition 1 below.

**Proposition 1.**

1. *The filter,* $\widehat{f}_t^A = \mathbb{E}_t^{\mathbb{P}^A}\left[f_t^A\right]$*, of Agent* $A$ *evolves according to the dynamics*

$$d\widehat{f}_t^A = \kappa\left(\overline{f} - \widehat{f}_t^A\right)dt + \frac{\gamma}{\sigma_\delta} d\widehat{W}_t^A, \tag{4}$$

   *where* $\gamma = \sqrt{\sigma_\delta^2\left(\sigma_\delta^2\kappa^2 + \sigma_f^2\right)} - \kappa\sigma_\delta^2$ *denotes the steady-state posterior variance and where* $\widehat{W}^A$ *is a Brownian motion under Agent A's probability measure* $\mathbb{P}^A$*.*

2. *The filter,* $\widehat{f}_t^B = \mathbb{E}_t^{\mathbb{P}^B}\left[f_t^B\right] = \mathbb{P}_t^B\left(f_t^B = f^h\right)f^h + \left(1 - \mathbb{P}_t^B\left(f_t^B = f^h\right)\right)f^l$ *, of Agent* $B$ *satisfies the dynamics*

$$d\widehat{f}_t^B = (\lambda + \psi)\left(f_\infty - \widehat{f}_t^B\right)dt + \frac{1}{\sigma_\delta}\left(\widehat{f}_t^B - f^l\right)\left(f^h - \widehat{f}_t^B\right)d\widehat{W}_t^B, \tag{5}$$

   *where* $f_\infty = \lim_{t\to\infty}\mathbb{E}\left[f_t^B\right] = f^l + \frac{\psi}{\lambda+\psi}\left(f^h - f^l\right)$ *denotes the unconditional mean of the filter and where* $\widehat{W}^B$ *is a Brownian motion under Agent B's probability measure* $\mathbb{P}^B$*.*

[7] The model in Equation (2) is adopted by Detemple (1986), Brennan and Xia (2001), Scheinkman and Xiong (2003), Dumas et al. (2009), Xiong and Yan (2010), Colacito and Croce (2013), and Bansal, Kiku, Shaliastovich, and Yaron (2014) among others. The model in Equation (3) is based on the work of David (1997) and Veronesi (1999) and extensions thereof. In the Economics literature, time series models of the business cycles include Threshold AutoRegressive (TAR) models, Markov-Switching AutoRegressive (MSAR) models, and Smooth Transition AutoRegressive (STAR) models. See Hamilton (1994) and Milas, Rothman, and van Dijk (2006) for further details.

**Proof**. See Appendix A.1 for Part *1.* and Lipster and Shiryaev (2001a) for Part *2.* ■

The mechanism through which Agent $A$ and $B$ update their views differs. First, both agents' expectations mean-revert, yet at a different speed $\lambda + \psi \neq \kappa$ and towards a different long-term mean $f_\infty \neq \overline{f}$, as shown in Equations (4) and (5). In particular, unlike Agent $A$'s filter that mechanically inherits the mean-reversion of her initial model, the mean-reversion in Agent $B$'s filter reflects her effort to continuously reassess the likelihood of each state using the information that continuously flows from dividends. Second, unlike the volatility of Agent $A$'s filter, which is constant, the volatility of Agent $B$'s filter is stochastic. In Section 3.2, we discuss how these learning differences affect the dynamics of disagreement among the two agents.

We choose to work under Agent $A$'s probability measure, a choice we clarify in Section 3.1. We convert Agent $B$'s views into those of Agent $A$ through the change of measure:

$$\left.\frac{\mathrm{d}\mathbb{P}^B}{\mathrm{d}\mathbb{P}^A}\right|_{\mathscr{F}_t} \equiv \eta_t = \exp\left[-\frac{1}{2}\int_0^t \frac{g_u^2}{\sigma_\delta^2}\mathrm{d}u - \int_0^t \frac{g_u}{\sigma_\delta}\mathrm{d}\widehat{W}_u^A\right], \tag{6}$$

where $g \equiv \widehat{f}^A - \widehat{f}^B$ represents agents' disagreement about the estimated fundamental. Importantly, the change of measure in (6) implies that Agents $A$ and $B$ have equivalent perceptions of the world, a result we establish in Proposition 2.

**Proposition 2.** *The probability measures $\mathbb{P}^A$ and $\mathbb{P}^B$ restricted to the filtration $\mathcal{F}_t$ are equivalent, for all $t \in \mathbb{R}_+$. The Brownian motions $\widehat{W}^A$ and $\widehat{W}^B$ therefore satisfy the relation*

$$d\widehat{W}_t^B = d\widehat{W}_t^A + \frac{g_t}{\sigma_\delta}dt.$$

**Proof**. See Appendix A.2. ■

Proposition 2 ensures that we can translate the perception of Agent $B$ into that of Agent $A$. To illustrate the economic relevance of this result, suppose that Agent $A$'s model is the true data-generating process. Then, observing a finite history of data, Agent $B$ cannot falsify her own model. That is, if agents could sell a claim contingent on which model is correct, this claim would never pay off (except perhaps at an infinite horizon).

Because the economy runs forever, a conceptual issue—although immaterial for the equilibrium construction—is that agents asymptotically observe an infinite history of data.

Hence, at least one agent asymptotically falsifies her model. Which model is wrong requires a statement about the true data-generating process, $f$. While we, even as modelers, do not observe the true fundamental, $f$, we think about it in the spirit of Hong, Stein, and Yu (2007). If Agent $A$'s and $B$'s models are both reasonable representations of the world, then the true data-generating process should be some combination, $C$, of the two and noise, $\epsilon$:

$$f_t = C(f_t^A, f_t^B, \epsilon).$$

In Section 3.1, we will show that the combination, $C$, of agents' models is such that Agent $A$'s model is historically closer to the true data-generating process.

## 2.3 Equilibrium

Agents choose their portfolios and consumption plans to maximize their expected lifetime utility of consumption. They have power utility preferences defined by

$$U(c,t) \equiv e^{-\rho t}\frac{c^{1-\alpha}}{1-\alpha},$$

where $\alpha > 0$ is the coefficient of relative risk aversion and $\rho > 0$ the subjective discount rate.

Since markets are complete, we solve the consumption-portfolio problem of both agents using the martingale approach of Karatzas, Lehoczky, and Shreve (1987) and Cox and Huang (1989). We present the equilibrium risk-free rate, $r^f$, and the market price of risk, $\theta$, in Proposition 3 below.

**Proposition 3.** *The market price of risk, $\theta$, as perceived by Agent A and the risk-free rate, $r^f$, satisfy*

$$\theta_t = \alpha\sigma_\delta + \frac{(1-\omega(\eta_t))}{\sigma_\delta}g_t \tag{7}$$

$$r_t^f = \rho + \alpha\widehat{f}_t^A - \frac{1}{2}\alpha(\alpha+1)\sigma_\delta^2 + (1-\omega(\eta_t))\,g_t\left(\frac{1}{2}\frac{\alpha-1}{\alpha\sigma_\delta^2}\omega(\eta_t)g_t - \alpha\right), \tag{8}$$

*where*

$$\omega(\eta_t) = \left(\frac{1}{\phi_A}\right)^{\frac{1}{\alpha}} \left(\left(\frac{\eta_t}{\phi_B}\right)^{\frac{1}{\alpha}} + \left(\frac{1}{\phi_A}\right)^{\frac{1}{\alpha}}\right)^{-1} \tag{9}$$

*denotes the consumption share of Agent $A$ and where $\phi_A$ and $\phi_B$ are the Lagrange multipliers associated with Agent $A$'s and Agent $B$'s budget constraints, respectively.*

**Proof**. See Appendix A.3. ■

The market price of risk is the product of the diffusion of Agent $A$'s consumption growth

$$\sqrt{\frac{1}{\mathrm{d}t}\mathrm{var}_t^{\mathbb{P}^A}\left(\frac{\mathrm{d}c_{At}}{c_{At}}\right)} = \sigma_\delta + \frac{1-\omega(\eta_t)}{\alpha\sigma_\delta} g_t \tag{10}$$

and her coefficient of risk aversion, $\alpha$. Agent $A$ wants to be compensated for holding assets that co-vary positively with her consumption growth and is willing to pay a premium for holding assets that co-vary negatively with her consumption growth. In particular, her consumption varies either when dividends fluctuate or when her consumption share fluctuates. Hence, the market price of risk is proportional to the diffusion of dividend growth and the diffusion of the growth of her consumption share. The market price of risk then scales with risk aversion, reflecting that the more risk averse Agent $A$ is, the more she wants to hedge.

The consumption share of Agent $B$, $1-\omega$, determines the extent to which disagreement affects the market price of risk. To see this, suppose disagreement is positive today (Agent $A$ is more optimistic than Agent $B$). A positive dividend shock tomorrow then indicates that Agent $A$'s beliefs were more accurate than Agent $B$'s—the likelihood, $\eta$, of Agent $B$'s model relative to Agent $A$'s decreases and the consumption share of Agent $A$ increases through (9). In this case, Agent $A$ wants to be rewarded for holding assets that co-vary positively with the dividend shock and thus with her consumption share. Moreover, as the consumption share of Agent $B$ decreases, the risk of Agent $A$'s consumption share growth decreases: were Agent $A$ the only agent populating the economy, disagreement would become irrelevant and Agent $A$ would be compensated for holding assets that are positively correlated with dividend growth only. This mechanism applies symmetrically when disagreement is negative.

The risk-free rate reacts to the fundamental through one main channel—the income effect. Suppose that only one agent populates the economy (disagreement is irrelevant) and that the

fundamental increases today. Equation (8) then indicates that the risk-free rate increases. The reason is that a fundamental increase today leads to a consumption increase tomorrow. Hence, the agent consumes more today, saves less and the bond price drops (the risk-free rate increases). This *income effect* dominates irrespective of the value of risk aversion.

Disagreement affects the risk-free rate through two opposite channels—the income and the substitution effects—and constitutes a key driver of the risk-free rate because agents use the bond to hedge their consumption growth risk. To see this, notice that the risk of Agent $A$'s consumption growth can be alternatively written as

$$\frac{1}{\mathrm{d}t}\mathrm{var}_t^{\mathbb{P}^A}\left(\frac{\mathrm{d}c_{At}}{c_{At}}\right)=\frac{1}{\mathrm{d}t}\left[\mathrm{var}_t^{\mathbb{P}^A}\left(\frac{\mathrm{d}\delta_t}{\delta_t}\right)+\mathrm{var}_t^{\mathbb{P}^A}\left(\frac{\mathrm{d}\omega(\eta_t)}{\omega(\eta_t)}\right)+2\mathrm{cov}_t^{\mathbb{P}^A}\left(\frac{\mathrm{d}\delta_t}{\delta_t},\frac{\mathrm{d}\omega(\eta_t)}{\omega(\eta_t)}\right)\right].$$

Three sources of risk drive Agent $A$'s consumption growth: the risk of dividend growth, the risk of consumption share growth and the co-movement between the two. When agents' coefficient of risk aversion is less than one, the risk-free rate decreases with all three sources of risk, because consumption growth risk induces agents to save more today to consume tomorrow. That is, the income effect dominates. When the risk aversion coefficient is larger than one, instead, the risk-free rate increases with the risk of consumption share growth. The reason is that, as agents become more risk averse, consumption growth risk induces agents to lock in consumption by saving less and consuming more today. This *substitution effect* decreases the bond price and hence increases the risk-free rate.

We conclude the equilibrium description by providing in Proposition 4 the equilibrium stock price $S$, the derivation of which follows the methodology in Dumas et al. (2009).

**Proposition 4.** *Assuming that the coefficient of relative risk aversion $\alpha$ is an integer, the equilibrium stock price satisfies*

$$\frac{S_t}{\delta_t}=\omega(\eta_t)^\alpha\frac{S_t}{\delta_t}\Big|_{O.U.}+(1-\omega(\eta_t))^\alpha\frac{S_t}{\delta_t}\Big|_{M.C.}+\omega\left(\eta_t\right)^\alpha\sum_{j=1}^{\alpha-1}\left(\begin{array}{c}\alpha\\ j\end{array}\right)\left(\frac{1-\omega\left(\eta_t\right)}{\omega\left(\eta_t\right)}\right)^j F^j\left(\widehat{f}_t^A,g_t\right) \tag{11}$$

*where $\frac{S}{\delta}\big|_{O.U.}$ and $\frac{S}{\delta}\big|_{M.C.}$ denote the prices that prevail in a representative-agent economy populated by Agent $A$ and $B$, respectively. The functions $F^j\left(\widehat{f}^A,g\right)$ represent price adjustment for investors' disagreement. We derive these equilibrium functions in Appendix A.4.*

**Proof**. See Appendix A.4. ■

The price in Equation (11) has three terms. When agents have logarithmic utilities ($\alpha = 1$), only the first two terms are relevant and the price is just an average of the prices that obtain in a representative-agent economy populated by Agent $A$ and $B$, respectively. When agents are more risk averse than a logarithmic agent ($\alpha > 1$), the third term becomes relevant. This term isolates how the joint interaction of fundamental and disagreement operates on the price through the function $F$.

# 3 Calibration and Dynamics of Disagreement

In this section we calibrate the models of Agent $A$ and $B$ and discuss their relative fit to the U.S. business cycle (Section 3.1). We then study the resulting term structure of disagreement and show that disagreement is counter-cyclical and persists over long horizons (Section 3.2).

## 3.1 Calibration and Model Fit to the U.S. Economy

In our model, agents use a single source of information—the time series of dividends $\delta$—to update their expectations. As a proxy for the dividend stream, we use the S&P 500 dividend time series recorded at a monthly frequency from January 1871 to November 2013, which we obtain from Robert Shiller's website. Looking back to the 19th century allows us to cover a large number of business cycle turning points, but obviously adds strong seasonality effects (Bollerslev and Hodrick (1992)). To reduce these effects, we apply the filter developed by Hodrick and Prescott (1997) to the time series of dividends.

We assume that both agents observe the S&P 500 dividend time series after seasonalities have been smoothed out. Since this data is available monthly, agents need to first estimate a discretized version of their model by Maximum Likelihood.[8] Agents then map the parameters they estimated into their continuous-time model. Doing so, agents obtain the parameter values presented in Table 1. All discussions and results that follow are based on these parameter values. For convenience, we discuss the methodological details in Appendix A.5.

[8]See Hamilton (1994) for the likelihood function of each model. We present the estimated parameters, their standard errors, and their statistical significance in Table 5 in Appendix A.5.

| Parameter | Symbol | Value |
|---|---|---|
| Volatility of Dividend Growth | $\sigma_\delta$ | $0.0225^{***}$ $(3.94 \times 10^{-4})$ |
| Mean-Reversion Speed of $f^A$ | $\kappa$ | $0.1911^{***}$ $(0.0264)$ |
| Long-Term Mean of $f^A$ | $\bar{f}$ | $0.0630^{***}$ $(0.0083)$ |
| Volatility of $f^A$ | $\sigma_f$ | $0.0056^{***}$ $(2.34 \times 10^{-4})$ |
| High State of $f^B$ | $f^h$ | $0.0794^{***}$ $(0.0032)$ |
| Low State of $f^B$ | $f^l$ | $-0.0711^{***}$ $(0.0038)$ |
| Intensity of $f^B$: High to Low | $\lambda$ | $0.3022$ $(1.1294)$ |
| Intensity of $f^B$: Low to High | $\psi$ | $0.3951$ $(1.1689)$ |
| Relative Risk Aversion | $\alpha$ | 2 |
| Subjective Discount Rate | $\rho$ | 0.01 |

**Table 1: Parameter Calibration.**

This table reports the estimated parameters of the continuous-time model of Agent $A$ and Agent $B$. Standard errors (computed using the Delta Method) are reported in brackets and statistical significance at the 10%, 5%, and 1% levels is labeled with * , **, and ***, respectively. The last two rows report our choice of preference parameters.

The parameters of Table 1 show that both models produce distinct interpretations of the data. First, Agent $A$ finds a low reversion speed $\kappa$ and therefore concludes that the fundamental is highly persistent, a conclusion largely supported by the long-run risk literature (e.g., Bansal and Yaron (2004)). Second, Agent $B$ finds that the high state $f^h$ and the low state $f^l$ identify an expansion and a recession state, respectively. Importantly, Agent $B$ finds an asymmetry between the two states—expansions are more persistent than recessions ($\psi > \lambda$). The sum of $\lambda$ and $\psi$ further imply that Agent $B$'s filter reverts about 3.5 times faster than that of Agent $A$. In addition, the long-term mean of Agent $B$'s filter ($f_\infty \approx 0.014$) is lower than that of Agent $A$ ($\overline{f} = 0.063$). Finally, in contrast to the volatility of Agent $B$'s filter, that of Agent $A$ is constant. Overall, the two models differ with respect to persistence, long-term means, and volatilities.

Which model is best is obviously specific to an agent's needs. We can, however, determine which model better fits historical data by applying a model selection method such as the *Akaike* Information Criterion (AIC). Because an Ornstein-Uhlenbeck process is more versatile and relies on fewer parameters than a 2-state Markov chain, the information criterion

favors Agent $A$'s model.[9] Hence, over the last century, Agent $A$'s probability measure was closer to the physical probability measure *in-sample*. This fact ultimately justifies our choice of computing and analyzing the equilibrium under Agent $A$'s probability measure $\mathbb{P}^A$. It does, however, not necessarily question Agent $B$'s rationality—a model that performs better in-sample does not necessarily perform better *out-of-sample*.[10] In particular, for agents who process the data in real time, it is difficult to assess the relative fit of their model.

## 3.2 Dynamics of Disagreement

We first show that model heterogeneity leads to countercyclical disagreement (Section 3.2.1). We then identify which particular aspect of model heterogeneity is key to generate this result relative to alternative specifications (Section 3.2.2).

### 3.2.1 Differences in Adjustment Speed and Polarization of Beliefs

Our previous discussion suggests that model heterogeneity causes agents to disagree about three aspects of the fundamental: its persistence, its long-term mean, and its volatility. We now show that these three aspects together give rise to *counter-cyclical* disagreement that persists over long horizons, consistent with the observed pattern of disagreement among forecasters (e.g., Kandel and Pearson (1995) and Patton and Timmermann (2010)).

We first define three regimes of the economy on which we focus our study. To do so, we adopt the terminology used to describe the different phases of the business cycle: we say that the economy is going through *good times* when agents' expectations are above $\bar{f}$, while the economy is going through *bad times* when agents' expectations are below

$$f_m \equiv \frac{1}{2} f^l + \frac{1}{2} f^h,$$

the point at which Agent $B$ assigns equal probabilities to the expansion and the recession states. Otherwise, we say the economy is in *normal times* when filters lie between $f_m$

[9] The Akaike Information Criterion is defined as follows: $AIC = 2K - 2\log(L)$, where $K$ is the number of parameters estimated and $L$ is the Likelihood function. Hence, the smaller the criterion is, the better is the model. We find that the AIC for Agent $A$ and $B$'s model are $AIC_{\text{Agent A}} = -1.7229 \times 10^4$ and $AIC_{\text{Agent B}} = -1.2154 \times 10^4$.

[10] For instance, Welch and Goyal (2008) show that, while any predictive model performs better than the historical mean in-sample, the latter tends to have better predictive power out-of-sample.

| Good Times | Normal Times | Bad Times |
|---|---|---|
| $(\widehat{f}^A, \widehat{f}^B) \in [\bar{f}, f^h], \quad g = 0$ | $(\widehat{f}^A, \widehat{f}^B) \in [f_m, \bar{f}], \quad g = 0$ | $(\widehat{f}^A, \widehat{f}^B) \in [f^l, f_m], \quad g = 0$ |

**Table 2: Definition of Regimes.**
This table describes the 3 different regimes of the economy—good, normal, and bad times.

and $\bar{f}$. We postpone an extensive discussion on the relevance of $\bar{f}$ and $f_m$ as the thresholds defining each regime to Section 3.2.2. Furthermore, to emphasize that future disagreement—as opposed to current disagreement—drives our result, we set current disagreement to zero in each case. Finally, as a convention, we assume that agents' filters always start in the middle of each interval.[11] Table 2 summarizes the definition of the three regimes.

To illustrate how agents update their expectations in each regime, we plot in Figure 1 Agent $A$ and $B$'s average filter over time. Each panel corresponds to a different regime. In good times (the left panel), both agents adjust their views at comparable speeds and disagreement exhibits little variation. In normal times (the middle panel), the difference in adjustment speed between Agent $A$ and $B$ becomes apparent. Although both agents expect economic conditions to improve—both filters move upwards on average—Agent $B$ adjusts her expectations significantly faster than Agent $A$. Due to this difference in adjustment speed, disagreement increases in the short term. Disagreement does not evaporate over subsequent horizons, but persists in the long term.

The learning pattern differs in bad times, as illustrated in the right-hand panel of Figure 1. Two forces are at play. First, there is a *polarization* of opinions in the short term: agents disagree about the direction in which the economy is heading and average filters diverge. In particular, Agent $A$ is optimistic and expects the economy to recover, while Agent $B$ is pessimistic and expects economic conditions to further deteriorate. This polarization of opinions creates a spike in disagreement in the short term. Second, the difference in adjustment speed reappears in the long run. As Agent $B$ realizes that the economy is recovering, she rapidly catches up with Agent $A$ and eventually becomes more optimistic than Agent $A$. As a result, disagreement evaporates in the medium term and regenerates in the long term through the difference in adjustment speeds.

[11] The dynamics of disagreement are consistent within each region, irrespective of the starting point within each region, except in two knife-edge intervals, $(0.063, 0.067)$ and $(-0.056, -0.0711)$, around $\bar{f}$ and $f^l$, respectively. In Appendix A.9.2, we explain why disagreement dynamics differ, but show that our results remain unaffected, within these intervals.

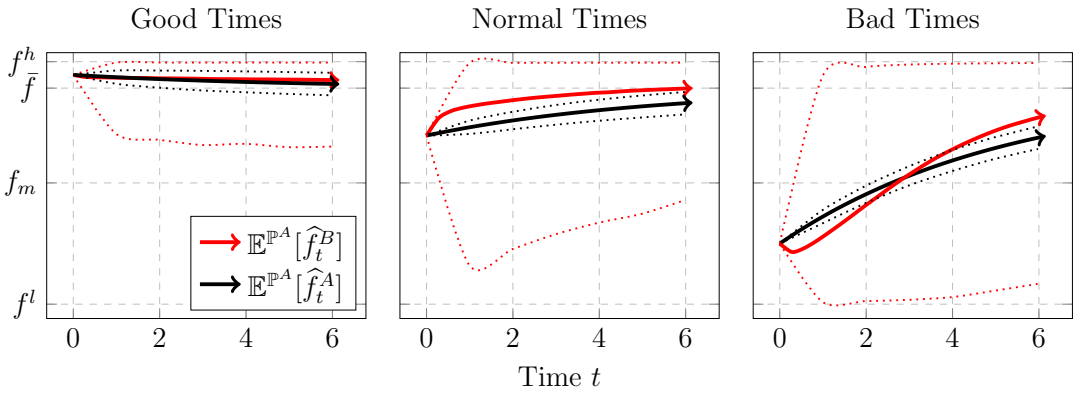


**Figure 1: Filtered Dynamics in 3 States of the Economy.**

The solid black and solid red lines represent Agent $A$ and $B$'s average filters ($\mathbb{E}^{\mathbb{P}^A}[\widehat{f}_t^A]$ and $\mathbb{E}^{\mathbb{P}^A}[\widehat{f}_t^B]$), respectively. The dotted lines depict the corresponding 90% confidence intervals. Filters are plotted against time. Each Panel corresponds to a specific state of the economy: good, normal, and bad times.

Finally, while Agent $A$ postulates constant uncertainty throughout the business cycle, the variance of Agent $B$'s filter greatly varies over the business cycle, as the confidence interval in each panel of Figure 1 demonstrates. Specifically, the variance of Agent $B$'s filter increases as economic conditions deteriorate. Hence, the difference in agents' assessment of economic uncertainty is greater in bad times. We now show analytically that this effect precisely generates the polarization of beliefs in bad times that we just described.

### 3.2.2 Discussion of the Main Mechanism and Alternative Specifications

In our framework, model heterogeneity makes disagreement counter-cyclical: starting in good times, there is little disagreement because agents have similar views. Disagreement then increases in normal times through the difference in agents' adjustment speeds and spikes in bad times through the polarization of agents' beliefs. Short-term disagreement is largest in bad times because opposite views generate more disagreement than a same view adjusting at different speeds. In this section, we investigate which aspect of model heterogeneity—difference in persistence, long-term mean or volatility—specifically generates *counter-cyclical* disagreement. We explain how disagreement differs under alternative specifications in which agents use different sets of parameters within the same parametric model.

To guide our discussion, we introduce a formal definition of "adjustment speed" and "polarization of beliefs", which we provide in Definition 1.

**Definition 1.** *The speed, $\Sigma_t^i$, at which Agent $i$ updates her expectations conditional on the initial value of her filter is the rate of change of her average filter, $\mathbb{E}_0^{\mathbb{P}^A}[\widehat{f}_t^i|\widehat{f}_0^i = x_0]$, over time*

$$\Sigma_t^i := \frac{d}{dt}\mathbb{E}_0^{\mathbb{P}^A}[\widehat{f}_t^i|\widehat{f}_0^i = x_0], \quad i = A, B.$$

*Furthermore, a polarization of beliefs occurs at time $t$ when $sign(\Sigma_t^A) \neq sign(\Sigma_t^B)$.*

An important observation is that the volatility of Agent $B$'s filter introduces a nonlinear term in the filter's drift under $\mathbb{P}^A$,

$$\frac{1}{\mathrm{d}t}\mathbb{E}_t^{\mathbb{P}^A}[\mathrm{d}\widehat{f}_t^B] = (\lambda + \psi)\left(f_\infty - \widehat{f}_t^B\right) + \underbrace{\frac{1}{\sigma_\delta^2}\left(\widehat{f}_t^A - \widehat{f}_t^B\right)\left(\widehat{f}_t^B - f^l\right)\left(f^h - \widehat{f}_t^B\right)}_{\text{nonlinear change of measure}}, \tag{12}$$

through the change of measure of Proposition 2. To understand how this term affects Agent $B$'s expectations, we first linearize the drift in (12), which boils down to assuming that both agents use Agent $A$'s model with different parameters. Denoting by $\widetilde{f}^B$ the filter resulting from this linearization, we obtain an approximation for the adjustment speed of Agent $B$, which we present in Proposition 5 along with the adjustment speed of Agent $A$.

**Proposition 5.** *The adjustment speeds of the filter $\widehat{f}^A$ and $\widetilde{f}^B$, which prevail under the calibration in Table 1 and in a neighborhood of $t \approx 0$, are approximately given by*

$$\Sigma_t^A = \frac{d}{dt}\mathbb{E}_0^{\mathbb{P}^A}\left[\widehat{f}_t^A|\widehat{f}_0^A = x_0\right] \approx \kappa(\bar{f} - x_0) \tag{13}$$

$$\Sigma_t^B = \frac{d}{dt}\mathbb{E}_0^{\mathbb{P}^A}\left[\widetilde{f}_t^B|\widetilde{f}_0^B = x_0\right] \approx (\lambda + \psi)(f_\infty - x_0) + \frac{(f^l - x_0)(x_0 + f^l)(\overline{f}\kappa + x_0(\lambda + \psi - \kappa))}{\sigma_\delta^2}t. \tag{14}$$

**Proof**. See Appendix A.6. ■

The first term in (13) and (14) shows that Agent $A$'s and $B$'s expectations revert towards their long-term mean $\bar{f}$ and $f_\infty$ at speed $\kappa$ and $\lambda + \psi$, respectively. Clearly, differences in persistence or long-term mean—even within the same parametric model—generate differences in adjustment speeds across agents. Most importantly, the two linear filters can only generate polarization of beliefs *in normal times*, i.e., agents' beliefs diverge in the interval

$x_0 \in (f_\infty, \overline{f})$. This observation in turn implies that, under the alternative specification in which both agents use Agent $A$'s model with different parameters (e.g., Buraschi and Whelan (2013), Ehling et al. (2013) and Buraschi et al. (2014)), polarization of beliefs only arises within the range spanned by the long-term means (i.e., in normal times, but not in bad times). Hence, under this alternative specification, short-term disagreement concentrates in normal times. Moreover, in this range, the second term in (14) reveals that both agents' expectations eventually align.[12]

The polarization of beliefs in bad times, which arises in our model, must therefore come from the nonlinearity of the change of measure in (12). This second-order effect makes Agent $B$'s expectations centrifugal under $\mathbb{P}^A$. To show this, we now perform a second-order approximation of the drift of Agent $B$'s filter. Denoting by $\check{f}^B$ the filter resulting from this approximation, we obtain a new expression for the adjustment speed of Agent $B$, which we highlight in Proposition 6.

**Proposition 6.** *The adjustment speed of the filter $\check{f}^B$, which prevails under the calibration in Table 1 and in a neighborhood of $t \approx 0$, is approximately given by*

$$\begin{aligned}\Sigma_t^B &= \frac{d}{dt}\mathbb{E}_0^{\mathbb{P}^A}\left[\check{f}_t^B | \check{f}_0^B = x_0\right] \\ &\approx (\lambda+\psi)(f_\infty - x_0) + \sigma_\delta^{-4}\left(f^l - x_0\right)\left(f^l + x_0\right)\left(2x_0(f^l)^2 + \overline{f}\kappa\sigma_\delta^2 + x_0\sigma_\delta^2(-\kappa+\lambda+\psi) - 2x_0^3\right)t.\end{aligned} \tag{15}$$

**Proof**. See Appendix A.6. ■

While a second-order approximation does not modify Agent $A$'s adjustment speed (because her model is linear), it introduces an additional term

$$\varphi \times t \equiv \frac{1}{\sigma_\delta^4}\left(f^l - x_0\right)\left(f^l + x_0\right)\left(2x_0(f^l)^2 - 2x_0^3\right)t \tag{16}$$

in Agent $B$'s. The coefficient $\varphi$ is of order $o(\sigma_\delta^{-4})$ and our calibration implies that, for any $t > \sigma_\delta$, the expression in (16) drives the speed of learning in (15). The coefficient $\varphi$ changes sign in the neighborhood of $x_0 = f^l$, $x_0 = f^h$ and $x_0 = f_m \approx 0$. Specifically, as $\check{f}^B$ approaches $f^l$, the adjustment speed becomes positive to reflect the process upwards and as

[12] Our calibration implies that $\lambda + \psi > \kappa$ and $f^l < 0$ and thus the term multiplying $t$ in (14) is positive.

$\check{f}^B$ approaches $f^h$, it becomes negative to reflect the process downwards. Most important is the behavior of $\varphi$ around $f_m$. As $\check{f}^B$ rises above $f_m$, the term $2x_0((f^l)^2 - x_0^2) > 0$ makes the adjustment speed positive. Instead, when $\check{f}^B$ drops below $f_m$, this term makes the adjustment speed negative. That is, under $\mathbb{P}^A$, the nonlinear change of measure in (12) causes Agent $B$'s expectations to become centrifugal outward $f_m$ (see Figure 14 in Appendix A.6). The overall effect is that the polarization of beliefs—agents' expectations diverge—now only occurs when expectations drop below $f_m$ (i.e., in bad times).

We draw four conclusions from the discussion above. First, to generate polarization of beliefs in bad times there must be at least one agent who does not postulate constant uncertainty throughout the business cycle (i.e., her filter exhibits stochastic volatility). Second, to generate differences in adjustment speed agents' model must feature either differences in persistence or long-term mean, or both. Third, the relevant business cycle turning point for Agent $A$ is $\bar{f}$, while that for Agent $B$ is $f_m$, thus justifying the specification of regimes in Table 2. Finally, the behavior of adjustment speeds implies that, relative to Agent $A$, Agent $B$ revises her beliefs rapidly in both normal and good times and slowly in bad times.

Our discussion also suggests that polarization of beliefs in bad times occurs when both agents use Agent $B$'s model with different parameters (David, 2008). We conclude this section by showing that this specification, however, does not generate a spike in disagreement *under our calibration*. To see this, suppose that Agent $A$ estimates a Markov chain using Agent $B$'s state values $f^l$ and $f^h$, but with intensities $\lambda = \frac{\kappa(f^h - \bar{f})}{f^h - f^l}$ and $\psi = \frac{\kappa(\bar{f} - f^l)}{f^h - f^l}$ that match the persistence and long-term mean of her initial model.[13] This specification also produces polarization of beliefs in bad times and differences in adjustment speed in normal times, as in our framework. Under this specification, however, the dispersion around average disagreement is significantly different. To illustrate this, we plot plot in Figure 2 the distribution of disagreement under our specification (upper panel) and that under the specification of David (2008) using our calibration (lower panel).[14]

Although disagreement moves in the same direction with both specifications, the distribution of disagreement shifts immediately with our specification, reflecting short-term spikes in disagreement. We emphasize, however, that the specification of David (2008) can generate spikes in disagreement under alternative calibrations.

[13] Assuming that both agents use the same state values is consistent with the calibration of David (2008).
[14] We relegate computational details to Appendix A.7.

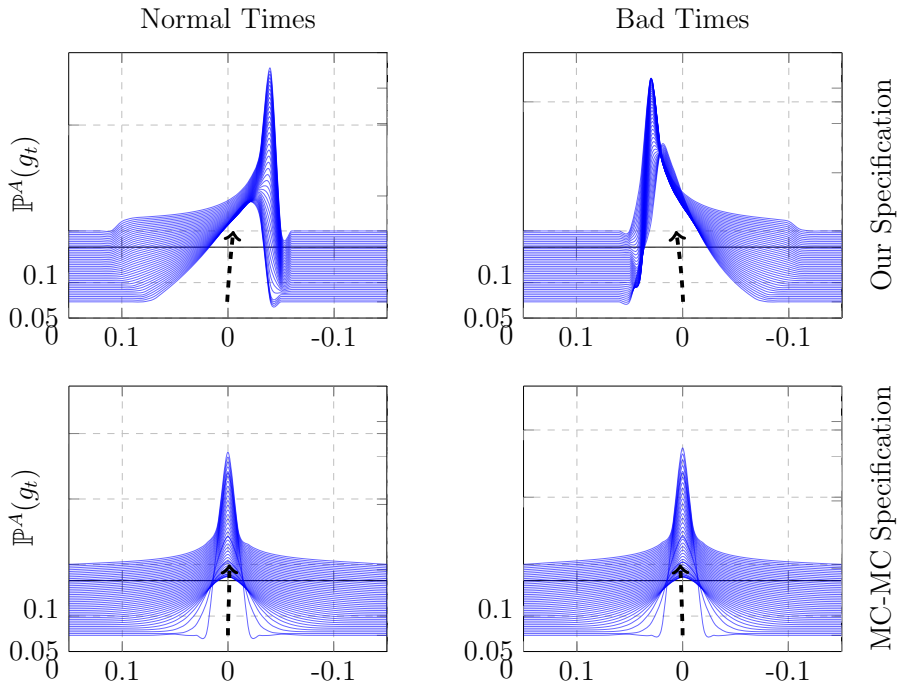


**Figure 2: Distribution of Disagreement: OU-MC and MC-MC Specifications.** This figure plots the distribution of disagreement along with average disagreement ($\mathbb{E}^{\mathbb{P}^A}[g_t]$ (the solid red line) over a 1-month horizon under Agent $A$'s probability measure ($\mathbb{P}^A(\widehat{f}_t^B)$). The upper panels correspond to our specification, while the lower panels correspond to the specification in David (2008) under our calibration. Each column corresponds to a specific state of the economy: Normal and Bad Times.

# 4 Disagreement driving Stock Return Predictability

Recent empirical studies find that stock return predictability concentrates in bad times. In this section, our objective is to determine the mechanism that causes return predictability to vary over the business cycle. We show that counter-cyclical disagreement leads to return continuation, which strengthens as economic conditions deteriorate.

To understand how disagreement, current and future excess returns are related, and thus how disagreement and current returns predict future returns, it is insightful to first describe agents' trading strategy. Following Dumas et al. (2009), we decompose Agent $A$'s strategy into two components, which we present in Proposition 7.

**Proposition 7.** *The number of shares, $Q$, that Agent $A$ holds can be decomposed into a*

*myopic portfolio, $M$, and a hedging portfolio, $H$, according to*

$$Q_t = M_t + H_t = \frac{\mu_t - r_t^f}{\alpha \sigma_t^2} \frac{V_t}{S_t} + \frac{\alpha - 1}{\alpha \sigma_t S_t} E_t^{\mathbb{P}^A} \left[ \int_t^\infty \frac{\xi_s}{\xi_t} c_{As} \left( \frac{\mathcal{D}_t \xi_s}{\xi_s} - \frac{\mathcal{D}_t \xi_t}{\xi_t} \right) ds \right] \tag{17}$$

*where $\xi$ denotes the state-price density, which is given by*

$$\xi_t = e^{-\rho t} \delta_t^{-\alpha} \left[ \left( \frac{\eta_t}{\phi_B} \right)^{\frac{1}{\alpha}} + \left( \frac{1}{\phi_A} \right)^{\frac{1}{\alpha}} \right]^{\alpha} \tag{18}$$

*and where $V$ denotes Agent $A$'s wealth, which we provide in Appendix A.8.*

**Proof**. See Appendix A.8. ■

Agent $A$'s portfolio in (17) tells us how she trades on return predictability. While the first part, $M$, is a myopic demand through which Agent $A$ seeks to extract the immediate Sharpe ratio, the second term, $H$, is an hedging demand through which she exploits return predictability. To see this, notice that the hedging demand in (17) incorporates Agent $A$'s outlook on future returns through the future response of state-price densities to a shock occurring today, $\mathcal{D}_t \xi_s$. This response represents Agent $A$'s attempt to predict future returns.

We first analyze the relation between disagreement and contemporaneous excess returns, which are contained in Agent $A$'s myopic demand (Section 4.1). We then study how Agent $A$ attempts to predict future returns through her hedging demand and explain how current disagreement and future excess returns are related (Section 4.2). We finally show that disagreement creates a positive relation between current and future excess returns in the short term, a phenomenon known as "time series momentum" (Section 4.3).

## 4.1 Disagreement and Contemporaneous Excess Returns

In our model, disagreement and contemporaneous excess returns are negatively related. To show this, we observe that contemporaneous (expected) excess returns, $\mu - r^f = \sigma\theta$, are the product of the stock return diffusion, $\sigma$, and the market price of risk, $\theta$, which we discussed in Section 2.3. To infer the properties of contemporaneous excess returns, we discuss the

diffusion of stock returns, which satisfies

$$\sigma = \frac{1}{\sigma_\delta S}\left[\gamma \underbrace{\frac{\partial}{\partial \widehat{f}^A} S}_{<0} + \left(\gamma - \left(f^h + g - \widehat{f}^A\right)\left(\widehat{f}^A - g - f^l\right)\right)\underbrace{\frac{\partial}{\partial g} S}_{>0} - g\eta \underbrace{\frac{\partial}{\partial \eta} S}_{>0}\right] + \sigma_\delta. \quad (19)$$

We first explain the sign of the price sensitivities in (19). The price decreases with the fundamental through the income effect (when risk aversion is larger than one (Veronesi, 2000)): anticipating that a fundamental increase today leads to higher consumption tomorrow, Agent $A$ decreases her savings, which decreases the price. In contrast, the price increases with disagreement through the substitution effect: an increase in disagreement implies that Agent $A$ is optimistic relative to Agent $B$ and therefore increases her stock holdings, which leads to a price increase. Finally, an increase in the likelihood of Agent $B$'s model implies an increase in the risk of Agent $A$'s consumption growth through (9) and (10), to which she responds by increasing her hedging demand, leading to a price increase.

We now identify the main term driving the diffusion of stock returns in (19). Although Equation (19) indicates a nontrivial relation between the stock return diffusion and the state variables, Figure 3 shows that most of the variations in the diffusion is implied by the third diffusion component, $-\frac{\eta g}{\sigma_\delta S}\frac{\partial}{\partial \eta} S$, irrespective of the relative likelihood, $\eta$ (and the fundamental (see Figure 15 in Appendix A.9.1)). Because the price sensitivity, $\frac{\partial}{\partial \eta} S > 0$, is positive, the diffusion of stock returns reacts negatively to an increase in disagreement—it is decreasing in disagreement and changes sign around zero. As a result, the diffusion of stock returns and disagreement mostly have opposite signs: as shown in Figure 3, the diffusion of stock returns is positive when disagreement is negative and vice-versa.

The market price of risk in (7), instead, is increasing in disagreement: it is positive when disagreement is positive and mostly negative when disagreement is negative. Hence, contemporaneous excess returns—the product between the diffusion of stock returns and the market of risk—have an inverse U-shaped relation with disagreement around zero. Specifically, contemporaneous excess returns are low when disagreement is strong and high when disagreement is weak. The main conclusion is that contemporaneous excess returns are decreasing with absolute disagreement.

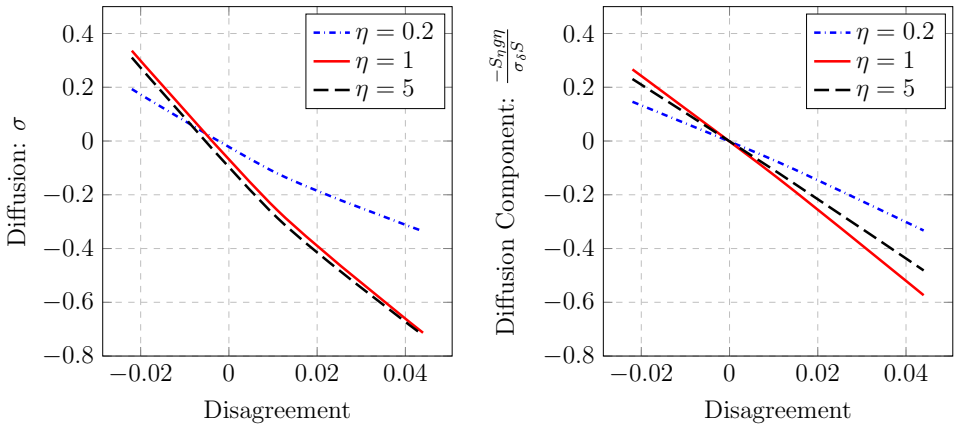


**Figure 3: Stock Return Diffusion vs. Disagreement.**

The left and the right panels depict the stock return diffusion and the third diffusion component, $\frac{-S_\eta g \eta}{\sigma_\delta S}$, respectively, both as a function of disagreement for different values of the change of measure. The other state variables are set to $\delta = \eta = 1$ and $\widehat{f}^A = \bar{f}$. The values for the change of measure $\eta = 0.2$, $\eta = 1$, and $\eta = 5$ yield consumption shares that are worth $\omega(0.2) \approx 0.7$, $\omega(1) = 0.5$, and $\omega(5) \approx 0.3$, respectively. The range for the disagreement is the 90% confidence interval obtained from 1,000 simulations of the economy over a 100-year horizon.

## 4.2 Disagreement, Future Excess Returns and Reaction to News

We now show that future excess returns react to contemporaneous news in a way that is predictable if disagreement spikes in the short term. In bad times, news polarizes opinions, causing returns to persistently under-react. In normal times, news exacerbates the difference in adjustment speed, causing returns to persistently over-react. In good times, the news provokes an immediate return adjustment. As a result, the content of news (Tetlock, 2007) better predicts future returns in bad times (Garcia, 2013).

Agent $A$'s hedging demand in (17) shows that forecasting future returns in our model involves computing the response of future state-price densities to a shock occurring today. Hence, as emphasized in Dumas et al. (2009), the concept of future returns that is relevant for portfolio choice is the future response, $\mathcal{D}_t \xi_s / \xi_s$, of state-price densities, as opposed to the usual multiperiod rate of return, $\xi_t / \xi_s$. We will accordingly refer to "future returns" and the "the response of state-price densities" indifferently and denote future returns cumulated up to an horizon $s > t$ by $R(t,s) = \mathcal{D}_t \xi_s / \xi_s$. This concept can be mapped into the empirical measures of Tetlock (2007) and Garcia (2013): its average, $\mathbb{E}_t^{\mathbb{P}^A}[R(t,s)]$, is the regression coefficient of the news arriving at time $t$ on cumulative stock returns from time $t$ to $s$.

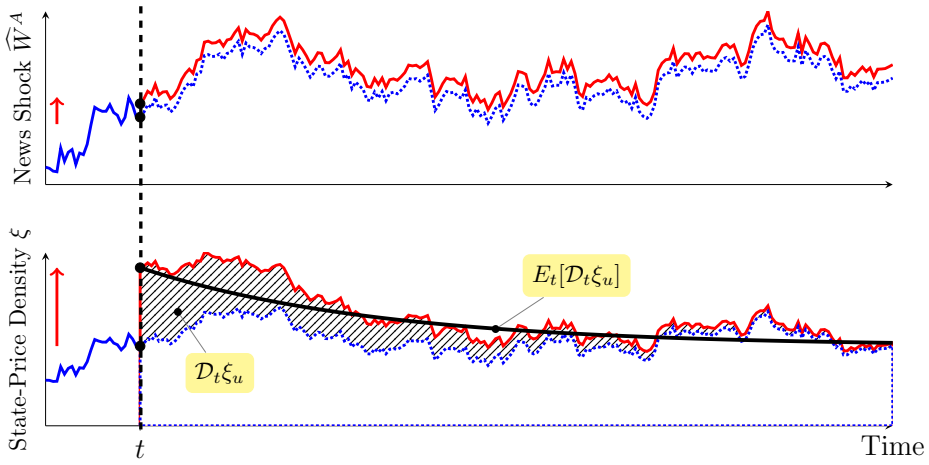


**Figure 4: Illustration of the State-Price Density's Response to a News Shock.** The upper panel shows a trajectory of news before (dashed blue line) and after (solid red line) a news surprise. The lower panel shows the associated trajectory of the state-price density before (dashed blue line) and after (solid red line) a news surprise. The shaded area represents the reaction $\mathcal{D}\xi$ to a news shock and the solid black line represents the average reaction $E[\mathcal{D}\xi]$ of the state-price density.

We start with an illustration of the concept of future response. To compute a response to a news shock, Agent $A$ first considers a trajectory of the Brownian innovation $\widehat{W}^A$, the only source of news in our model. She then contemplates a news surprise today (an initial perturbation of the news trajectory), while keeping the news trajectory otherwise unchanged, as illustrated in the upper panel of Figure 4. To each news trajectory, both perturbed and unperturbed, corresponds a trajectory of the state-price density, as illustrated in the lower panel. The difference, $\mathcal{D}\xi$, between these trajectories (the shaded area) precisely captures the reaction of the state-price density to the news surprise relative to the state-price density that would have prevailed, had there been no news surprise. Because Agent $A$ is interested in all possible trajectories, she computes an average reaction $E[\mathcal{D}\xi]$ (the solid black line).

When the news surprise is "small", the reaction, $\mathcal{D}\xi$, of the state-price density becomes a well-defined mathematical object known as a Malliavin derivative.[15] It has the economic meaning of an impulse-response function following a shock in initial values (in our case, a news shock). However, unlike a standard impulse-response function, a Malliavin derivative takes future uncertainty into account: Agent $A$ does not assume that the world becomes

[15]See, e.g., Detemple and Zapatero (1991), Detemple, Garcia, and Rindisbacher (2003, 2005), Berrada (2006), and Dumas et al. (2009) for applications of Malliavin calculus in Financial Economics.

deterministic after the shock has occurred and therefore computes an average response.

To obtain an analytical expression for average future returns in our model, notice that the state-price density, $\xi(\delta, \eta)$, in (18) depends on two state variables, dividends and the likelihood of Agent $B$'s model. As a result, the way future state-price densities react to a news depends, first, on how they react to a change in these state variables and, second, on how these state variables themselves react to the news. Applying chain rule to (18), the average response of the state-price density satisfies

$$\mathbb{E}_t^{\mathbb{P}^A}[R(t,s)] = \underbrace{-\alpha \frac{\mathcal{D}_t \delta_s}{\delta_s}}_{\text{fundamental channel}} + \underbrace{\mathbb{E}_t^{\mathbb{P}^A}\left[(1-\omega(\eta_s))\frac{\mathcal{D}_t \eta_s}{\eta_s}\right]}_{\text{disagreement channel}}. \tag{20}$$

We now decompose this chain reaction and relegate mathematical details to Appendix 7.

As the news hits, dividends react in a perfectly predictable fashion,

$$\frac{\mathcal{D}_t \delta_s}{\delta_s} = \sigma_\delta + \frac{\gamma}{\kappa \sigma_\delta}\left(1 - e^{-\kappa(s-t)}\right) > 0, \quad \forall s > t \in \mathbb{R}_+ \tag{21}$$

because the fundamental mean reverts with constant uncertainty, $\frac{\gamma}{\sigma_\delta}$, under Agent $A$'s measure. In particular, Equation (21) indicates that dividends respond positively to a good news and that their response gradually weakens at Agent $A$'s learning speed, $\kappa$. In contrast, the likelihood of Agent $B$'s model reacts ambiguously to the news on average

$$\mathbb{E}_t^{\mathbb{P}^A}\left[\frac{\mathcal{D}_t \eta_s}{\eta_s}\right] = -\frac{1}{\sigma_\delta^2}\mathbb{E}_t^{\mathbb{P}^A}\left[\int_t^s g_u \mathcal{D}_t g_u \mathrm{d}u\right] \tag{22}$$

as the news may either corroborate or invalidate Agent $B$'s model *depending on economic conditions*. Specifically, when the reaction of disagreement, $g \times \mathcal{D}g$, to the news is persistently negative, Agent $B$ is overly optimistic relative to Agent $A$ and a good news therefore corroborates her model, as Equation (22) indicates. If, instead, the reaction of disagreement is persistently positive, Agent $B$ is pessimistic and a good news invalidates her model.

As dividends and the likelihood of Agent $B$'s model move, the state-price density in turn respond. It decreases after a dividend increase, $\frac{\partial}{\partial \delta}\xi(\delta,\eta) = -\alpha\xi(\delta,\eta)/\delta < 0$, and increases following an increase in the likelihood of Agent $B$'s model, $\frac{\partial}{\partial \eta}\xi(\delta,\eta) = (1-\omega(\eta))\xi(\delta,\eta)/\eta > 0$. The reason is that an increase in dividends decreases marginal utility. An increase in the

likelihood, $\eta$, instead, increases the risk of Agent $A$'s consumption share growth, which increases her hedging demand and thus her marginal utility, $\phi_A \xi$.

Overall, a news shock moves state-price densities through two channels, the fundamental and disagreement. The fundamental channel in (20) triggers a negative state-price density reaction following a good news, as it hints future dividend abundance. The disagreement channel conveys an ambiguous reaction depending on whether the news corroborates Agent $B$'s model, thereby making Agent $A$'s consumption share growth riskier. Our calibration indicates that the magnitude of the former channel ranges from $-0.045$ to $-0.058$, while that of the latter ranges from $-0.3$ to $0.3$. Clearly, disagreement is the main channel driving future returns in our model. As a result, average cumulative returns move alongside the average reaction of the likelihood, $\eta$, and therefore against the average reaction of disagreement:

$$\text{sign}(\mathbb{E}_t^{\mathbb{P}^A}[R(t,s)]) \approx \text{sign}(\mathbb{E}_t^{\mathbb{P}^A}[\mathcal{D}_t \eta_s]) \equiv -\text{sign}\left(\mathbb{E}_t^{\mathbb{P}^A}\left[\int_t^s g_u \mathcal{D}_t g_u \mathrm{d}u\right]\right). \tag{23}$$

We plot the 3-year average cumulative return in the left panel of Figure 5 along with the average reaction of disagreement, $E_t[g_u \mathcal{D}_t g_u]$, in the right panel. When interpreting average cumulative returns, based on (23), a negative sign implies that returns move opposite to the news, i.e., they under-react. A positive sign either means that returns adjust to the news or over-react, in which case returns subsequently revert, i.e., average cumulative returns change direction. Finally, when interpreting the average reaction of disagreement, a negative sign implies that Agent $B$ is overly optimistic and a positive sign means that she is pessimistic.

In bad times (the black dashed line), the news exacerbates disagreement by polarizing agents' opinions: Agent $A$ interprets the news positively, whereas Agent $B$ interprets it negatively and disagreement spikes (the right panel). Since the news invalidates Agent $B$'s model, returns move against the news, i.e., a good news is followed by negative returns in the short term (the left panel). Hence, Agent $B$'s pessimism slows down returns' adjustment to the news (decreasing reaction). A similar under-reaction phenomenon arises in Hong and Stein (1999) and Ottaviani and Sorensen (2015).

In normal times (the solid red line), the news shock precipitates an upward adjustment in Agent $B$'s expectations, sharpening the difference in adjustment speeds across agents. While agents interpret the news in similar ways, Agent $B$ overreacts to the news and becomes overly optimistic relative to Agent $A$ (the right panel). Because the news corroborates Agent $B$'s

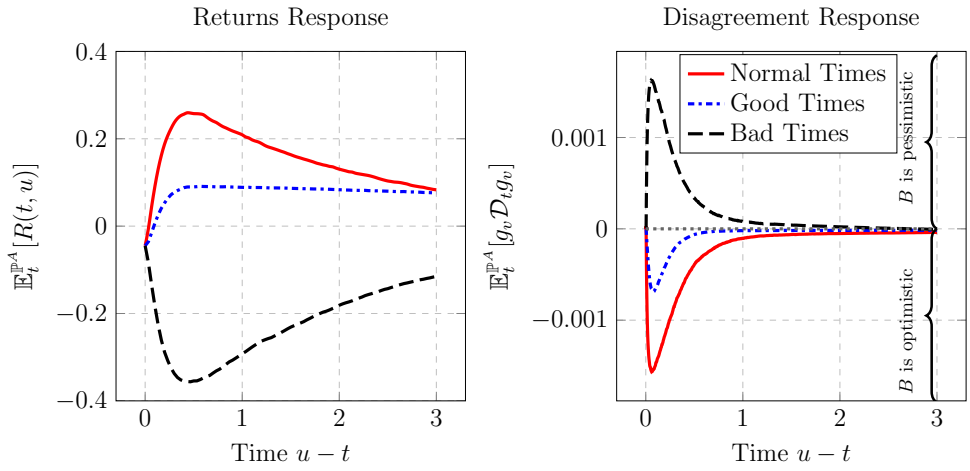


**Figure 5: Impulse Response of Excess Returns to a News Shock.**
The left panel plots the total (sum of fundamental and disagreement channels) impulse response of stock returns. The right panel plots the response of disagreement to a news shock $\mathbb{E}_t^{\mathbb{P}^A}[g_u \mathcal{D}_t g_u]$. Each line corresponds to a different regime.

overoptimism, returns over-react to the news in the short term (the left panel). That is, Agent $B$'s overreaction accelerates return's adjustment to the news (increasing reaction), similar to the over-reaction phenomenon of Daniel et al. (1998).

In good times (the dash-dotted blue line), the reaction of disagreement to the news is weak and exhibits little persistence (the right panel). Because the news is consistent with Agent $B$'s model and induces little disagreement, returns adjust immediately (the left panel). Hence, return continuation only arises when disagreement spikes in the short term.

In the long term (after about 6 months), excess returns systematically revert, i.e., cumulative returns change direction. This later phase of correction arises through the long-term behavior of Agent $B$'s consumption share, $1-\omega(\eta)$, which multiplies the reaction of disagreement to the news shock in (20). Since the drift of Agent $B$'s consumption share,

$$-\frac{1}{\mathrm{d}t}\mathbb{E}_t^{\mathbb{P}^A}[\mathrm{d}\omega_t] = \frac{g_t^2(2\omega_t - 1 - \alpha)(1-\omega_t)\omega_t}{2\alpha^2\sigma_\delta^2} < 0,$$

is negative and decreases with absolute disagreement, strong disagreement today leads to a strong downward trend in Agent $B$'s future consumption share. Agent $B$'s consumption share therefore gradually dampens the disagreement response in (22) (the right panel of Figure 5), which generates a reversion in excess returns in the long term.

We draw three conclusions from our discussion. First, in our model, continuation arises through a spike in disagreement whereby future returns react to past news. Reversal, instead, arises through a correction in the likelihood, $\eta$, of Agent $B$'s model. Second, whether a spike in disagreement results from the polarization of opinions in bad times or the difference in adjustment speeds in normal times dictates whether returns under- or over-react. Hence, unlike Ottaviani and Sorensen (2015), continuing under- and over-reaction alternate over the business cycle. Third, in our model, the content of news predicts future returns (Tetlock, 2007) and this effect concentrates in bad times (Garcia, 2013).

## 4.3 Time Series Momentum and Momentum Crashes

One of the most pervasive facts in finance is momentum (Jegadeesh and Titman, 1993). Recently, Moskowitz et al. (2012) uncover a similar pattern in aggregate returns, which they coin "time series momentum". In our model, counter-cyclical disagreement creates a positive relation between contemporaneous and future excess returns in the short term. We now show that this relation explains time series momentum, its term structure, how it crashes following sharp market rebounds and why it is stronger in bad times at short horizons.

To relate contemporaneous and future excess returns, which we have discussed separately, we use Agent $A$'s strategy in (17). Forcing Agent $A$ to hold the entire supply, $Q_t \equiv 1$, of the stock, market clearing implies that she consumes, $c_{A,t} = \delta_t$, aggregate dividends and that her wealth, $V_t = S_t$, coincides with the stock price. We then substitute these identities in (17) along with the instantaneous response of the state-price density, which is just a negative Sharpe ratio, $\frac{\mathcal{D}_t \xi_t}{\xi_t} = -\theta_t$. We finally obtain a CAPM-like relation between contemporaneous (expected) excess returns and the concept of future excess returns, $R$, that matters to investors in our model:

$$\mu_t - r_t^f = \sigma_t \left( \sigma_t + \frac{1-\alpha}{\alpha} \frac{1}{S_t} \mathbb{E}_t^{\mathbb{P}^A} \left[ \int_t^\infty \frac{\xi_s}{\xi_t} \delta_s R(t,s) \mathrm{d}s \right] \right). \tag{24}$$

The relation in (24) indicates that contemporaneous excess returns only predict future excess returns when agents are nonmyopic. Specifically, if $\alpha = 1$, investors myopically reap the immediate Sharpe ratio (the first term in (17)) and contemporaneous excess return, $\mu_t - r_t^f = \sigma_t^2$, do not provide information regarding future returns. When $\alpha > 1$, instead, return predictability arises through the inter-temporal hedge (the second term in (17)) that

reflects the prospects on future returns. Hence, contemporaneous and future returns relate through agents' hedging demands. Since contemporaneous excess returns are negatively related to disagreement (Section 4.1) and future excess returns are negatively related to the reaction of disagreement (Section 4.2), contemporaneous and future excess returns are positively related—returns exhibit time series momentum.

To analyze the pattern of serial correlation of returns in a way that is consistent with empirical studies, we compute monthly excess returns as

$$r_t^e \equiv \int_{t-\Delta}^{t} \left( \frac{\mathrm{d}S_u + \delta_u \mathrm{d}u}{S_u} - r_u^f \mathrm{d}u \right) = \int_{t-\Delta}^{t} \sigma_u (\theta_u \mathrm{d}u + \mathrm{d}\widehat{W}_u^A),$$

with $\Delta = 1/12 = 1$ month. Following Banerjee et al. (2009), we then measure time series momentum at different lags $h$ according to

$$\rho(h) = \frac{\mathrm{cov}^{\mathbb{P}^A}\left(r_{t+h\Delta}^e, r_t^e\right)}{\mathrm{var}^{\mathbb{P}^A}\left(r_t^e\right)}.$$

The coefficient, $\rho(h)$, determines whether returns exhibit momentum ($\rho(h) > 0$) or reversal ($\rho(h) < 0$). In Figure 6, we plot the $t$-statistics of this coefficient for lags ranging from 1 month to 3 years. For robustness, we also report the coefficient value in Appendix A.9.3.

The serial correlation of returns has a similar term structure in normal times (the center panel) and in bad times (the right panel): returns exhibit time series momentum over an horizon of 10 to 18 months, followed by reversal over subsequent horizons. Moreover, the magnitude of momentum varies according to the horizon considered: momentum is large at short horizons (up to four months) and then decays over longer horizons. In Appendix A.9.4 we show that these patterns are also robust unconditionally. Both the magnitude and the timing of time series momentum are in line with empirical evidence.[16] Indeed, Moskowitz et al. (2012) find significant time series momentum at short horizons (1 to 6 months), weaker momentum at intermediate horizons (7 to 15 months), and reversal at longer horizons.

Time series momentum results from a spike in disagreement, whereby returns continue to react to past news (Section 4.2). Hence, although similar, the term structures of momentum in normal times and in bad times differ in one important dimension. Because the polarization

[16] In Appendix A.9.8 we confirm empirically that time series momentum persists up to one year and is followed by long-term reversal in both NBER expansions and NBER recessions.

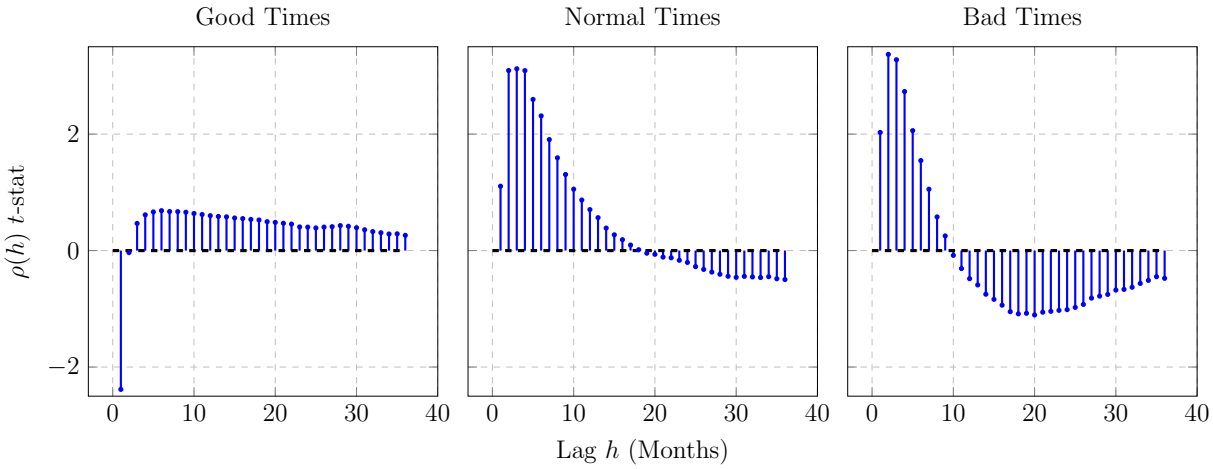


**Figure 6: Time Series Momentum of Excess Returns.**

This figure plots the $t$-statistics of the coefficient $\rho(h)$ for lags $h$ ranging from 1 month to 3 years. Each panel corresponds to a different state of the economy. Standard errors are adjusted using Newey and West (1987) procedure. The values reported above are obtained from 10,000 simulations of the economy over a 20-year horizon.

of opinions in bad times induces a sharper spike in disagreement relative to the difference in adjustment speeds in normal times, time series momentum at short horizons is significantly larger in bad times. That is, more disagreement leads to more time series momentum. This implication finds strong empirical support (see Section 5.2).

Excess returns have different time series properties in good times (the left panel): returns exhibit strong reversal in the very short term and weak momentum thereafter. The reason is that the main force driving returns in good times is reversion to fundamentals, as opposed to disagreement (see Figure 5). Suppose there is a bad news today. Investors, who want to smooth consumption over time, save by increasing their holdings in the stock, thus causing an immediate price increase. As the fundamental reverts, they unwind part of their stock holdings, leading to a price decrease and thus to reversal.

An important consequence of the initial reversal spike, which occurs in good times, is that a time series momentum strategy may crash if the market rises sharply. To see this, suppose we start implementing a momentum strategy in bad times. The right panel then indicates that this strategy remains profitable as long as the economy does not transit suddenly from bad to good times. If, instead, the market sharply rebounds, the trend suddenly reverts and the momentum strategy crashes. Sharp trend reversals typically occur at the end of financial

crises. For instance, a momentum strategy incurred large losses at the end of the Global Financial Crisis in March, April and May of 2009 (Moskowitz et al., 2012).[17] Furthermore, we show in Appendix A.9.7 that the model implies stronger time series momentum in extreme markets, consistent with Moskowitz et al. (2012).

# 5 Testable Predictions and Empirical Evidence

Our model offers three novel predictions: *(1)* future excess returns are positively related to contemporaneous absolute disagreement. Furthermore, time series momentum at short horizons *(2)* increases with absolute disagreement and *(3)* is, therefore, strongest in bad times. In this section, we first define several measures that quantify these qualitative predictions to obtain testable predictions (Section 5.1). These measures have immediate empirical counterparts that allow us to guide our empirical analysis. Mirroring our theoretical analysis, we then provide empirical support to the predictions of the model (Section 5.2).

## 5.1 Testable Predictions

Section 5.1.1 quantifies the relation between absolute disagreement and future excess returns. Section 5.1.2 measures the relation between momentum, disagreement and economic conditions.

### 5.1.1 Dispersion and Future Excess Returns

Empirical evidence suggests that high dispersion of analysts' forecasts predicts low future returns in the cross section of firms (Diether et al., 2002). In contrast to this *negative* cross-sectional relation, our model predicts a *positive* relation between the aggregate dispersion of forecasts and aggregate future returns. To quantify this relation, we start by defining several explanatory variables. In our model, three state variables drive excess returns: disagreement, the fundamental, and the likelihood, $\eta$. In the empirical literature, disagreement is usually proxied by the dispersion of analyst forecast (Diether et al., 2002), i.e., the absolute value of disagreement, $|g|$. We accordingly define monthly dispersion $G$, the monthly fundamental

[17] Daniel and Moskowitz (2013) and Barroso and Santa-Clara (2015) document a similar phenomenon in the cross-section of stocks.

$F$, and the monthly change of measure $E$ as

$$G_t = \int_{t-\Delta}^{t} |g_u| \mathrm{d}u, \qquad F_t = \int_{t-\Delta}^{t} \widehat{f}_u^A \mathrm{d}u, \qquad E_t = \int_{t-\Delta}^{t} \eta_u \mathrm{d}u. \tag{25}$$

To measure the relation between excess returns and dispersion, we run the regression

$$r^e_{t+h\Delta} = \alpha(h) + \beta_G(h) G_t + \beta_F(h) F_t + \beta_E(h) E_t + \epsilon_{t+h\Delta}, \tag{26}$$

in which we let excess returns be the dependent variable and lagged dispersion, fundamental and change of measure be the explanatory variables at different lags $h$. While we control for the effect of the fundamental and the change of measure in (26), we focus our discussion on the coefficient $\beta_G(h)$, which precisely quantifies the relation between excess returns and dispersion at different lags $h$. Since this model-implied coefficient does not have an explicit solution, we compute it through simulations. We plot the $t-$statistics of this coefficient in Figure 7 for different lags ranging from 0 month to 1 year. The left panel controls for the fundamental only, while the right panel controls for both the fundamental and the change of measure. For robustness, we also report the value of the coefficient $\beta_G(h)$ in Appendix A.9.5.

As explained in Section 4.1, dispersion is negatively related to contemporaneous excess returns (the lower-lags bars in each panel in Figure 7). It is, however, positively related to future excess returns (the higher-lags bars). To see this, suppose that dispersion is high today. Dispersion then decreases in the future because it is counter-cyclical.[18] Since dispersion and excess returns are contemporaneously negatively related, future excess returns are high, creating a positive relation between dispersion and future excess returns. This observation further implies that, in our model, contemporaneous excess returns are pro-cyclical, while future excess returns are countercyclical. Finally, comparing the left and right panels of Figure 7 shows that controlling for the change of measure has little impact: in both cases, dispersion positively predicts future excess return at lags larger than three months, although statistical significance is stronger when controlling for the fundamental only. Hence, control-

[18] To confirm that dispersion is counter-cyclical in our model, we regress dispersion $G$ on the fundamental $F$ using 1,000 simulations over a 100-year horizon. We find that the fundamental and dispersion are, indeed, negatively related with a regression coefficient of $-1.0051$. The $R^2$ of the regression is 0.1832 and, using Newey and West (1987) standard errors, the regression coefficient is significant at the 1% level.

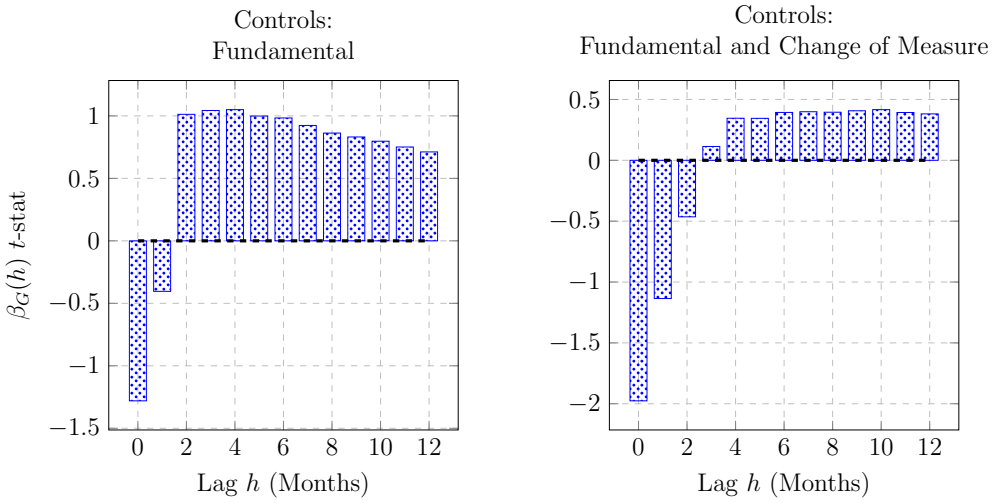


**Figure 7: Excess Returns vs. Dispersion.**

This figure plots the $t$-statistics of the response of excess returns to dispersion $\beta_G(h)$ for lags $h$ ranging from 0 month to 1 year. The left panel controls for the fundamental only (i.e. $\beta_E(h) = 0$), whereas the right panel controls for both the fundamental and the change of measure. Standard errors are adjusted using Newey and West (1987) procedure. The values reported above are obtained from 1,000 simulations of the economy over a 100-year horizon.

ling for the change of measure only alters the results quantitatively but not qualitatively, an observation on which we rely to guide our empirical analysis in Section 5.2.

### 5.1.2 Time Series Momentum

In Section 4.3, we showed that short-term time series momentum increases with dispersion. We consider two approaches to measure this effect. In a first approach, we capture periods of high dispersion with a dummy variable, $Y_{G,t}(p)$, that takes value 1 when monthly dispersion, $G$, is larger than its $p-$th percentile. We then run the following regression

$$r^e_{t+\Delta} = \alpha_G(p) + \beta_{1,G}(p) r^e_t + \beta_{2,G}(p) r^e_t Y_{G,t}(p) + \epsilon_{t+\Delta}. \tag{27}$$

The coefficient $\beta_{2,G}(p)$ measures time series momentum in excess of that present in periods of low dispersion. In an alternative approach, we compute time series momentum at a 1-month lag, $\beta_M(1)_j$, by running the following regression over $36-$month rolling windows:

$$r^e_{t+\Delta} = \alpha_M(1)_j + \beta_M(1)_j r^e_t + \epsilon_{t+\Delta}, \quad t \in (j\Delta, j\Delta + 36\Delta),$$

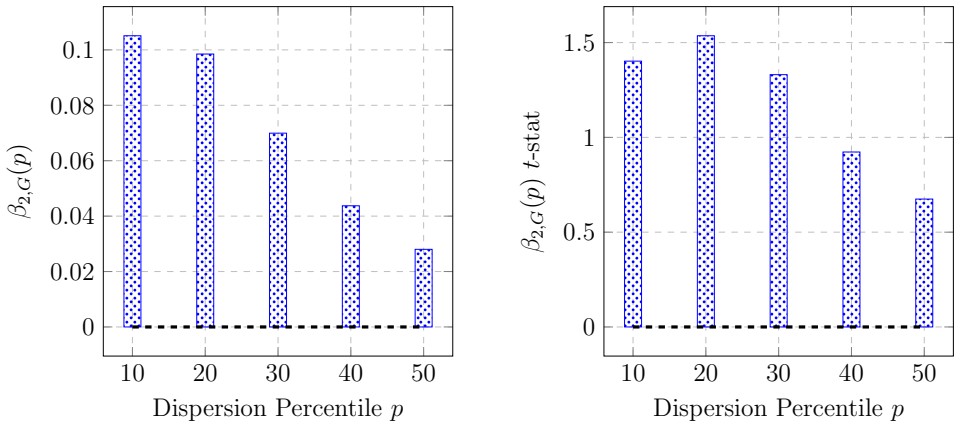


**Figure 8: Excess 1-Month Time Series Momentum in High Dispersion Periods.** This figure plots the excess 1-month time series momentum $\beta_{2,G}(p)$ and its $t$-statistics when dispersion is larger than its $p-$th percentile. Standard errors are adjusted using Newey and West (1987) procedure. The values reported above are obtained from 1,000 simulations of the economy over a 100-year horizon.

where $j = 0, \ldots, N-1$ is the index of each 36-month rolling window and $N$ is the total number of windows. We then regress the $t$-statistics of $\beta_M(1)_j$ on the aggregate dispersion, $AG_j = \sum_{t=j\Delta}^{j\Delta+36\Delta} G_t$, computed over each 36-month window:

$$\beta_M(1)\ t\text{-stat}_j = \alpha + \beta AG_j + \epsilon_j. \tag{28}$$

The coefficient $\beta$ measures the sensitivity of time series momentum (at a 1-month lag) to a change in aggregate dispersion.

We compute the model-implied statistics under the two approaches through simulations. Figure 8 plots the statistics, $\beta_{2,G}(p)$, of the first approach as a function of the $p-$th percentile considered. Consistent with our qualitative analysis, additional disagreement implies additional time series momentum at a 1-month lag. Moreover, this relation is most significant for the 20$-$th percentile of dispersion. Similarly, the model-implied statistics, $\beta = 22.8481$, of the second approach is significant at the 5% confidence level, which shows that momentum at a 1-month lag increases with dispersion.[19]

Finally, our model predicts strongest momentum in bad times at short horizons, a result

[19]We describe the statistics $\beta_{2,G}$ and $\beta$ for momentum at a 1-month lag only. A positive relation between time series momentum and dispersion also holds at a 2-, 3-, and 4-month lag. The coefficient $\beta$ is computed using 1,000 simulations over a 100-year horizon. The $R^2$ of the regression is 0.0345 and, using Newey and West (1987) standard errors, the slope coefficient is significant at the 5% level.

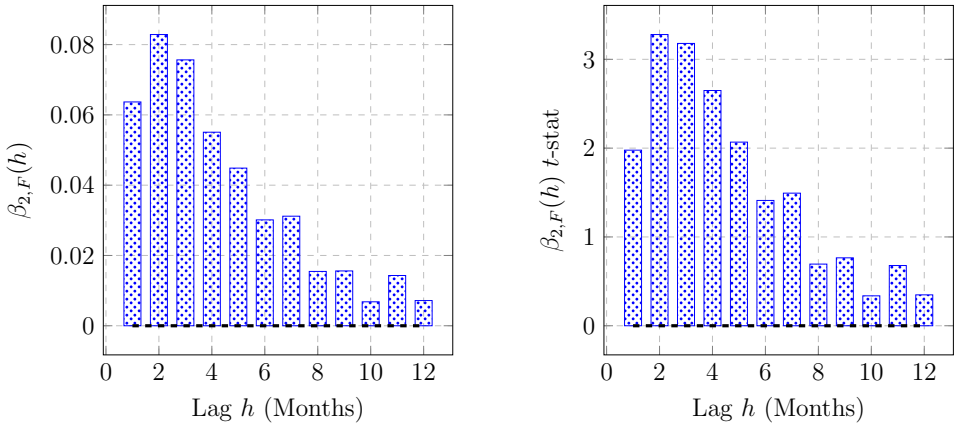


**Figure 9: Excess Time Series Momentum in Recessions.**
This figure plots the excess time series momentum coefficient $\beta_{2,F}(h)$ and its $t$-statistics for lags $h$ ranging from 1 month to 1 year. Standard errors are adjusted using Newey and West (1987) procedure. The values reported above are obtained from 1,000 simulations of the economy over a 100-year horizon.

we quantify as follows. We first identify bad times in our model in a way that can be mapped into our empirical analysis (Section 5.2) in which we use NBER recession dates. Since NBER recessions have historically accounted for 30% of the business cycle, we capture recessions with a dummy variable $Y_F$ that takes value 1 if the monthly fundamental, $F$, is below its 30−th percentile.[20] We then run the following regression:

$$r^e_{t+h\Delta} = \alpha_F(h) + \beta_{1,F}(h) r^e_t + \beta_{2,F}(h) r^e_t Y_{F,t} + \epsilon_{t+h\Delta}. \tag{29}$$

The coefficient $\beta_{2,F}(h)$ measures time-series momentum in excess of that present in expansions at a $h-$months lags. We plot this coefficient in Figure 9 as a function the lag considered. Consistent with our qualitative analysis, time series momentum is significantly larger in recessions than in expansions (up to the 5-month lag).

## 5.2 Empirical Evidence

We now mirror the quantitative analysis of Section 5.1 by constructing empirical counterparts to the dispersion, $G$, the fundamental, $F$, and the change of measure, $E$, defined in (25). As

[20] Note that our results are robust to any other threshold lying between the 10−th and the 50−th percentile.

is customary in the literature, we proxy for the dispersion and the fundamental using data on analyst forecast of firms' earnings per share. Unfortunately, a model-free proxy for the change of measure is unavailable. Our model, however, indicates that this variable has no qualitative impact on the timing of predictability—dispersion positively predicts future excess returns, whether or not we control for the change of measure in the predictive regression (Section 5.1.1). Moreover, the two other predictions of the model relate exclusively to time series momentum. Since we showed in Section 4.2 that time series momentum is a consequence of disagreement, as opposed to the change of measure, the unavailability of a model-free proxy for the change of measure does not affect our empirical analysis altogether.

To construct an empirical proxy for dispersion, we follow Diether et al. (2002). We first obtain monthly data on analyst forecast from I/B/E/S for the time period ranging from February 1976 to November 2014. Denoting by $f^{i,j}$ the forecast of analyst $j$ regarding firm $i$'s one fiscal year earnings per share, we define the dispersion, $D^i$, of analyst forecasts regarding firm $i$'s earnings as

$$D_t^i = \frac{\underset{j}{\text{std}}(f_t^{i,j})}{|\underset{j}{\text{mean}}(f_t^{i,j})|},$$

where $\underset{j}{\text{mean}}(.)$ and $\underset{j}{\text{std}}(.)$ denote the mean and the standard deviation of forecasts computed across analysts, respectively. We then compute the aggregate dispersion, $D$, across firms as

$$D_t = \frac{1}{N_t} \sum_{i=1}^{N_t} D_t^i,$$

where $N_t$ is the number of firms in the S&P 500 that have at least two analyst forecasts and a mean forecast different from zero at time $t$.[21] Finally, regressing aggregate dispersion, $D_t$, on time $t$ shows that dispersion has a strong linear downward trend (see left panel of Table 3). To obtain a de-trended proxy for dispersion, $Disp$, we use the residuals of this regression.

In Section 3, we estimated the two sets of beliefs, $\widehat{f}^A$ and $\widehat{f}^B$, using monthly dividends on the S&P 500. These calibrated beliefs in turn provide a model-implied time series of

[21] Over the time period 02/1976 to 11/2014, a mean of 490 firms meet these two conditions each month.

| Dep. Var. | $D_t$ |
|---|---|
| Const. | $0.2122^{***}$ |
| | (0.0135) |
| Time $t$ | $-0.0033^{***}$ |
| | (0.0005) |
| Adj. $R^2$ | 0.1470 |
| Obs. | 466 |

| Dep. Var. | $GMF_t$ |
|---|---|
| Const. | $-0.0300^{***}$ |
| | (0.0074) |
| $Y_t^{12}$ | $0.1241^{*}$ |
| | (0.0643) |
| Adj. $R^2$ | 0.0310 |
| Obs. | 466 |

**Table 3: Linear Trend in Dispersion and 12-Month Seasonality in Fundamental.** The left and right panels report the outputs obtained by regressing dispersion $D_t$ on time $t$ and the fundamental $GMF$ on the 12-month seasonality dummy variable $Y^{12}$, respectively. Standard errors are reported in brackets and statistical significance at the 10%, 5%, and 1% levels is labeled with * , **, and ***, respectively. Standard errors are adjusted using Newey and West (1987) procedure. Data are at the monthly frequency from 02/1976 to 11/2014.

dispersion, $Disp^{MI}$:

$$Disp^{MI} = \left| \widehat{f}^{A} - \widehat{f}^{B} \right|.$$

If the two filtering processes, $\widehat{f}^{A}$ and $\widehat{f}^{B}$, offer a reasonable approximation of agents' disagreement, then the model-implied dispersion, $Disp^{MI}$, should be correlated with the actual dispersion among forecasters, $Disp$. To investigate this matter, we regress the actual dispersion of forecasts, $Disp$, on the model-implied dispersion, $Disp^{MI}$. The right panel of Table 4 shows that there is a remarkably strong positive relation between the two, thus lending support to our assumption that agents use heterogeneous forecasting models. Moreover, conditioning on NBER recessions with a dummy variable, $Rec$, we find that the actual dispersion is higher in recessions than in expansions (left panel of Table 4). The actual dispersion is therefore counter-cyclical as our model predicts.

Finally, we build a proxy for the fundamental using the aggregate mean forecast, $MF = \frac{1}{N_t} \sum_{i=1}^{N_t} \underset{j}{\text{mean}}(f_t^{i,j})$, from which we compute its growth rate, $GMF$, as

$$GMF_t = \frac{MF_t - MF_{t-\Delta}}{MF_{t-\Delta}},$$

where $\Delta = 1/12 = 1$ month. We choose not to work with the log growth rate because the aggregate mean forecast is sometimes negative. We also remove three outliers from the time series of growth rates, despite their minor influence on the results. We finally eliminate

| Dep. Var. | $Disp_t$ |
|---|---|
| Const. | $-0.0070$ |
| | (0.0043) |
| $Rec_t$ | $0.0534^{***}$ |
| | (0.0137) |
| Adj. $R^2$ | 0.0383 |
| Obs. | 466 |

| Dep. Var. | $Disp_t$ |
|---|---|
| Const. | $-0.0320^{***}$ |
| | (0.0081) |
| $Disp_t^{MI}$ | $9.9362^{***}$ |
| | (2.3770) |
| Adj. $R^2$ | 0.0597 |
| Obs. | 454 |

**Table 4: Dispersion in Recessions and Actual vs. Model-Implied Dispersion.** The left panel reports the outputs obtained by regressing dispersion *Disp* on a dummy variable *Rec* that equals 1 during NBER recessions. The right panel reports the outputs obtained by regressing dispersion *Disp* on the model-implied dispersion $Disp^{MI}$ estimated in Section 3. Standard errors are reported in brackets and statistical significance at the 10%, 5%, and 1% levels is labeled with * , **, and ***, respectively. Standard errors are adjusted using Newey and West (1987) procedure. The left and right panels consider monthly data from 02/1976 to 11/2014 and from 02/1976 to 11/2013, respectively.

seasonalities from the resulting time series of growth rates by running the following regression

$$GMF_t = \alpha + \beta Y_t^{12} + Fund_t,$$

where $Y^{12}$ is a 12-month seasonality dummy variable (see right panel of Table 3). We use the residual, $Fund$, of this regression as a proxy for the fundamental.

### 5.2.1 Dispersion and Future Excess Returns

We start by testing our model's first prediction that future excess returns are positively related to current dispersion. To do so, we run the empirical equivalent to the regression in Equation (26). Specifically, we measure the relation, $\beta_G(h)$, between excess returns and dispersion at different lags $h$ by running the regression in (27) in which we substitute the fundamental, $F$, and the dispersion, $G$, by their empirical proxy, $Fund$ and $Disp$, respectively, and take excess returns, $r^e$, to be the monthly excess return on the S&P 500. We report our results in Figure 10 for different lags ranging from 1 month to 1 year.

Except for the 1-month lag, comparing Figures 7 and 10 shows that the model-implied pattern and the empirical pattern are in line with each other, particularly so at the 2-month horizon. Hence, consistent with the model's prediction, dispersion positively predicts future excess returns. Our finding, however, contrasts with the negative relation between dispersion

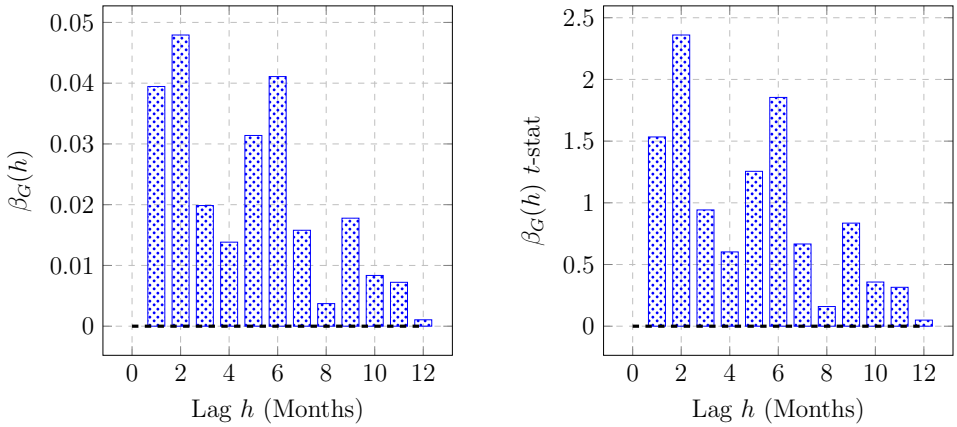


**Figure 10: Future Excess Returns vs. Dispersion.**

This figure plots the response of excess returns to dispersion $\beta_G(h)$ and its $t$-statistics for lags $h$ ranging from 1 month to 1 year. Standard errors are adjusted using Newey and West (1987) procedure. Monthly data range from 02/1976 to 11/2014.

and future returns, which is observed in the cross-section of firms (Diether et al., 2002). A possible explanation is that short-sale constraints become relevant in the cross-section of firms, as they may create a negative relation between dispersion and returns (Miller, 1977). Confirming this conjecture, Anderson, Ghysels, and Juergens (2005), Boehme, Danielsen, Kumar, and Sorescu (2009), and Carlin et al. (2014) find a positive cross-sectional relation in markets in which short-sale constraints are absent.

### 5.2.2 Dispersion and Time Series Momentum

We now test our model's second prediction that short-term time series momentum increases with dispersion. To that purpose, we use the two measures defined in Section 5.1.2. Specifically, we first measure excess 1-month time series momentum in high dispersion periods, $\beta_{2,G}(p)$, by running the regression in (27) in which we substitute dispersion $G$, by its empirical proxy, $Disp$. Second, we compute the sensitivity, $\beta$, of 1-month time series momentum to a change in aggregate dispersion: we estimate time series momentum over 36-month rolling windows through (28), record its $t-$statistics, and regress these $t-$statistics on the mean monthly dispersion, $MG$, which replaces our theoretical measure of aggregate dispersion, $AG$, in (28). We plot the coefficient $\beta_{2,G}(p)$ and its $t$-statistics for percentiles ranging from 10 to 50 in Figure 11 and report in Figure 12 the regression outputs for the coefficient $\beta$.

Figures 11 and 12 indicate the presence of time series momentum in excess of that present

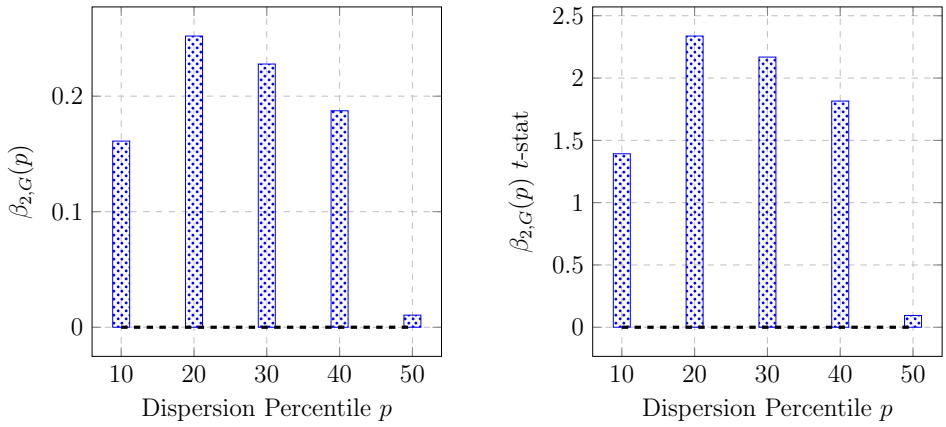


**Figure 11: Excess 1-Month Time Series Momentum in High Dispersion Periods.** This figure plots the excess 1-month time series momentum $\beta_{2,G}(p)$ and its $t$-statistics when dispersion is larger than its $p-$th percentile. Standard errors are adjusted using Newey and West (1987) procedure. Monthly data range from 02/1976 to 11/2014.

in periods of low dispersion, lending support to our model's second prediction. Specifically, both measures, $\beta_{2,G}(p)$ and $\beta$, have values that are consistent with their theoretical counterparts (discussed in Section 5.1.2). First, comparing Figures 8 and 11, we observe that, both in the model and in the data, the relation between disagreement and momentum is particularly strong when the dispersion threshold is the $20-$th percentile and weakens when the threshold increases from the $20-$th to the $50-$th percentile. Second, Figure 12 shows that 1-month time series momentum increases with dispersion. The empirical coefficient $\beta$ is, as its theoretical counterpart (see Section 5.1.2), significant at the 5% confidence level. That is, 1-month time series momentum and dispersion are both counter-cyclical, as predicted by the model. Furthermore, the data also lend support to the prediction that there is short-term time series reversal in low dispersion periods (see Appendix A.9.6).

### 5.2.3 Time Series Momentum over the Business Cycle

We finally test our model's third prediction that time series momentum at short horizons is strongest in recessions. To cover sufficiently many business cycle turning points, we extend the sample period and focus on monthly S&P 500 excess returns, $r^e$, from January 1871 to November 2013, which we obtain from Robert Shiller's website. Looking back to the 19th century allows us to account for a large number of recessions—NBER reports 29 recessions since the beginning of our sample. We then measure excess time series momentum

| Dep. Var. | $\beta_M(1)$ $t$-stat |
|---|---|
| Const. | $-0.2132^{*}$ |
| | (0.1241) |
| $MG$ | $4.3784^{**}$ |
| | (2.0397) |
| Adj. $R^2$ | 0.0441 |
| Obs. | 429 |

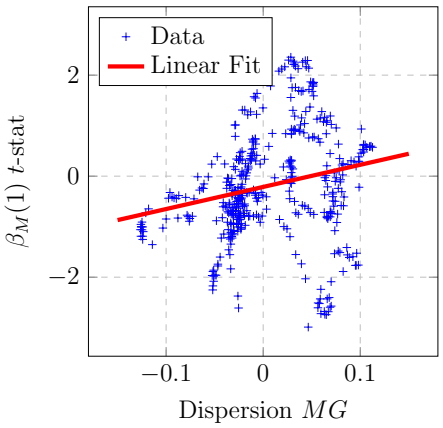


**Figure 12: 1-Month Time Series Momentum vs. Dispersion.**
The table reports the outputs of the following regression: $t$-statistics of 1-month time series momentum coefficient computed over 36-month rolling windows on the mean dispersion computed over the same 36-month rolling windows. The figure plots the data and the linear fit. Standard errors are reported in brackets and statistical significance at the 10%, 5%, and 1% levels is labeled with * , **, and ***, respectively. Standard errors are adjusted using Newey and West (1987) procedure. Monthly data range from 02/1976 to 11/2014.

in recessions, $\beta_2(h)$, by running the empirical equivalent to the regression in (29):

$$r^e_{t+h\Delta} = \alpha + \beta_1(h) r^e_t + \beta_2(h) Y_t r^e_t + \epsilon_{t+h\Delta}, \tag{30}$$

where $Y_t$ is a dummy variable that takes value 1 if $t$ belongs to an NBER recession. The coefficient $\beta_1(h)$ in Equation (30) captures the time series momentum effect of Moskowitz et al. (2012) at a $h-$month lag. The second coefficient $\beta_2(h)$, absent in Moskowitz et al. (2012), measures time series momentum in excess of that present in expansions.

Figure 13 reports the coefficient $\beta_2(h)$ and its $t$-statistics for lags $h$ ranging from 1 month to 1 year. Our model predicts that time series momentum is significantly stronger in recessions up to the 4-month lag (see Figure 9). Figure 13 supports this prediction at a 1-month lag, but indicates that excess time series momentum in recessions vanishes over subsequent lags. Our finding that time series momentum in the short term is stronger in recessions contrasts with evidence reported in the cross-section of returns. In particular, Chordia and Shivakumar (2002) and Cooper, Gutierrez, and Hameed (2004) find that cross-sectional momentum is weaker in down-markets than in up-markets. While Moskowitz et al. (2012) show that time series and cross-sectional momentum are strongly related, our result suggests that

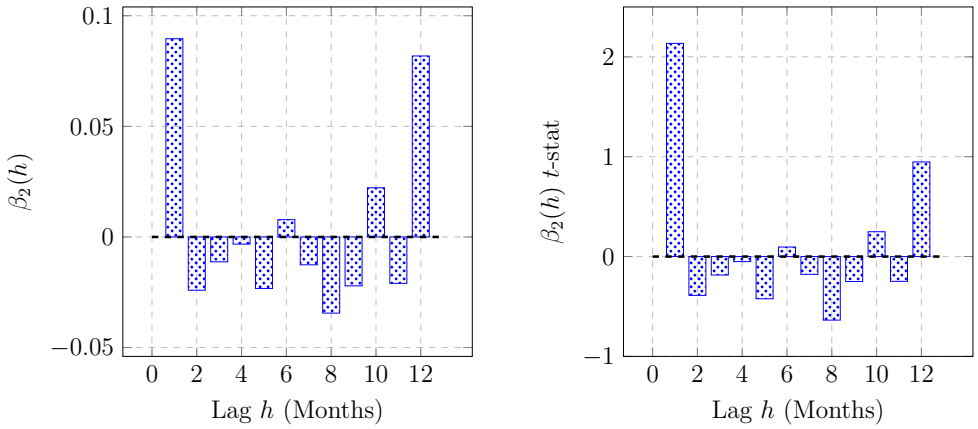


**Figure 13: Excess Time Series Momentum in Recessions.**
This figure plots the excess time series momentum in recessions $\beta_2(h)$ and its $t$-statistics for lags $h$ ranging from 1 month to 1 year. Standard errors are adjusted using Newey and West (1987) procedure. Monthly data range from 01/1871 to 11/2013.

their relation varies over the business cycle.

# 6 Conclusion

This paper suggests at least two interesting avenues for future research. First, our analysis describes how a single stock—an index—reacts to news shocks, but individual stocks composing the index may react differently. In particular, the performance of one stock relative to another may vary over the business cycle. Some may be "losers", others may be "winners", and this relation may persist or revert depending on economic conditions. Extending our framework to an economy with two trees would allow us to study this cross-sectional relation. Second, in our framework, investors estimate heterogeneous models, both of which measures a different aspect of the business cycle. In a representative-agent economy, we would like to study how the agent picks dynamically one model over the other depending on economic conditions. We believe that such endogenous "paradigm shifts" can explain several empirical facts regarding the dynamics of stock return volatility.

# References


Albuquerque, R. and J. Miao (2014). Advance information and asset prices. *Journal of Economic Theory 149*(0), 236 – 275.

Anderson, E. W., E. Ghysels, and J. L. Juergens (2005). Do heterogeneous beliefs matter for asset pricing? *Review of Financial Studies 18*(3), 875–924.

Andrei, D. and J. Cujean (2014). Information percolation, momentum, and reversal. *Working Paper*.

Banerjee, S., R. Kaniel, and I. Kremer (2009). Price drift as an outcome of differences in higher-order beliefs. *Review of Financial Studies 22*(9), 3707–3734.

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron (2014). Volatility, the macroeconomy, and asset prices. *The Journal of Finance 69*(6), 2471–2511.

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.

Barberis, N., A. Shleifer, and R. Vishny (1998). A model of investor sentiment. *Journal of Financial Economics 49*(3), 307 – 343.

Barinov, A. (2014, 3). Analyst disagreement and aggregate volatility risk. *Journal of Financial and Quantitative Analysis FirstView*, 1–43.

Barroso, P. and P. Santa-Clara (2015). Momentum has its moments. *Journal of Financial Economics 116*(1), 111 – 120.

Basak, S. (2000, January). A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk. *Journal of Economic Dynamics and Control 24*(1), 63–95.

Berk, J. B., R. C. Green, and V. Naik (1999, October). Optimal investment, growth options, and security returns. *Journal of Finance 54*(5), 1553–1607.

Berrada, T. (2006). Incomplete information, heterogeneity, and asset pricing. *Journal of Financial Econometrics 4*(1), 136–160.

Bhamra, H. S., L.-A. Kuehn, and I. A. Strebulaev (2010). The levered equity risk premium and credit spreads: A unified framework. *Review of Financial Studies 23*(2), 645–703.

Bhamra, H. S. and R. Uppal (2013). Asset prices with heterogeneity in preferences and beliefs. *Review of Financial Studies*.

Biais, B., P. Bossaerts, and C. Spatt (2010). Equilibrium asset pricing and portfolio choice under asymmetric information. *Review of Financial Studies 23*(4), 1503–1543.

Boehme, R. D., B. R. Danielsen, P. Kumar, and S. M. Sorescu (2009). Idiosyncratic risk and the cross-section of stock returns: Merton (1987) meets miller (1977). *Journal of Financial Markets 12*(3), 438 – 468.

Bollerslev, T. and R. Hodrick (1992). *Financial Market Efficiency Tests*. Number no. 4108. National Bureau of Economic Research.

Borovicka, J. (2011). Survival and long-run dynamics with heterogeneous beliefs under recursive preferences. *Working Paper* (WP-2011-06).

Brennan, M. J. and Y. Xia (2001). Stock price volatility and equity premium. *Journal of Monetary Economics 47*(2), 249 – 283.

Buraschi, A., F. Trojani, and A. Vedolin (2014). When uncertainty blows in the orchard: Comovement and equilibrium volatility risk premia. *The Journal of Finance 69*(1), 101–137.

Buraschi, A. and P. Whelan (2013). Term structure models and differences in beliefs. *Working Paper*.

Carlin, B. I., F. A. Longstaff, and K. Matoba (2014). Disagreement and asset prices. *Journal of Financial Economics 114*(2), 226 – 238.

Cen, L., K. C. J. Wei, and L. Yang (2014). Disagreement, underreaction, and stock returns. *Working Paper*.

Cespa, G. and X. Vives (2012). Dynamic trading and asset prices: Keynes vs. Hayek. *Review of Economic Studies 79*(2), 539–580.

Chalkley, M. and I. H. Lee (1998, July). Learning and Asymmetric Business Cycles. *Review of Economic Dynamics 1*(3), 623–645.

Chen, H. (2010). Macroeconomic conditions and the puzzles of credit spreads and capital structure. *The Journal of Finance 65*(6), 2171–2212.

Chen, H., S. Joslin, and N.-K. Tran (2012). Rare disasters and risk sharing with heterogeneous beliefs. *Review of Financial Studies 25*(7), 2189–2224.

Chordia, T. and L. Shivakumar (2002, 04). Momentum, business cycle, and time-varying expected returns. *Journal of Finance 57*(2), 985–1019.

Colacito, R. and M. M. Croce (2013). International asset pricing with recursive preferences. *The Journal of Finance 68*(6), 2651–2686.

Cooper, M. J., R. C. Gutierrez, and A. Hameed (2004, 06). Market states and momentum. *Journal of Finance 59*(3), 1345–1365.

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.

Cujean, J. (2013). The social dynamics of performance. *Working Paper*.

Dangl, T. and M. Halling (2012). Predictive regressions with time-varying coefficients. *Journal of Financial Economics 106*(1), 157–181.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam (1998, December). Investor psychology and security market under- and overreactions. *Journal of Finance 53*(6), 1839–1885.

Daniel, K. D. and T. J. Moskowitz (2013). Momentum crashes. *Working Paper*.

David, A. (1997, December). Fluctuating confidence in stock markets: Implications for returns and volatility. *Journal of Financial and Quantitative Analysis 32*(04), 427–462.

David, A. (2008). Heterogeneous beliefs, speculation and the equity premium. *Journal of Finance 63*, 41–83.

David, A. and P. Veronesi (2013). What ties return volatilities to price valuations and fundamentals? *Journal of Political Economy 121*(4), pp. 682–746.

Detemple, J., R. Garcia, and M. Rindisbacher (2005). Representation formulas for malliavin derivatives of diffusion processes. *Finance and Stochastics 9*, 349–367.

Detemple, J. and S. Murthy (1994, April). Intertemporal asset pricing with heterogeneous beliefs. *Journal of Economic Theory 62*(2), 294–320.

Detemple, J. B. (1986, June). Asset pricing in a production economy with incomplete information. *Journal of Finance 41*(2), 383–91.

Detemple, J. B. (1991, July). Further results on asset pricing with incomplete information. *Journal of Economic Dynamics and Control 15*(3), 425–453.

Detemple, J. B., R. Garcia, and M. Rindisbacher (2003). A monte carlo method for optimal portfolios. *The Journal of Finance 58*(1), pp. 401–446.

Detemple, J. B. and F. Zapatero (1991). Asset prices in an exchange economy with habit formation. *Econometrica 59*(6), pp. 1633–1657.

Diether, K. B., C. J. Malloy, and A. Scherbina (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance 57*(5), pp. 2113–2141.

Duffie, D., J. Pan, and K. Singleton (2000, November). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica 68*(6), 1343–1376.

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.

Ehling, P., M. Gallmeyer, C. Heyerdahl-Larsen, and P. Illeditsch (2013). Disagreement about inflation and the yield curve. *Working Paper*.

Garcia, D. (2013, 06). Sentiment during recessions. *Journal of Finance 68*(3), 1267–1300.

Hajek, B. (1985). Mean stochastic comparison of diffusions. *Wahrscheinlichkeitstheorie verw. Gebiete 68*, 315–329.

Hamilton, J. D. (1994). *Time Series Analysis*.

Henkel, S. J., J. S. Martin, and F. Nardari (2011, March). Time-varying short-horizon predictability. *Journal of Financial Economics 99*(3), 560–580.

Heston, S. L., M. Loewenstein, and G. Willard (2007). Options and bubbles. *Review of Financial Studies 20*, 359–390.

Hodrick, R. J. and E. C. Prescott (1997, February). Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking 29*(1), 1–16.

Holden, C. W. and A. Subrahmanyam (2002, January). News events, information acquisition, and serial correlation. *Journal of Business 75*(1), 1–32.

Hong, H. and J. C. Stein (1999, December). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance 54*(6), 2143–2184.

Hong, H., J. C. Stein, and J. Yu (2007, 06). Simple forecasts and paradigm shifts. *Journal of Finance 62*(3), 1207–1242.

Jegadeesh, N. and S. Titman (1993, March). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance 48*(1), 65–91.

Johnson, T. C. (2002, 04). Rational momentum effects. *Journal of Finance 57*(2), 585–608.

Judd, K. L. (1998). *Numerical Methods in Economics*. MIT Press.

Kandel, E. and N. D. Pearson (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy 103*(4), pp. 831–872.

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.

Karatzas, I. and S. Shreve (1988). *Brownian Motion and Stochastic Calculus*. Springer, Graduate Texte in Mathematics.

Kogan, L., S. A. Ross, J. Wang, and M. M. Westerfield (2006). The price impact and survival of irrational traders. *The Journal of Finance 61*(1), pp. 195–229.

Lipster, R. S. and A. N. Shiryaev (2001a). *Statistics of Random Processes II*. Springer Verlag, New York.

Lipster, R. S. and A. N. Shiryaev (2001b). *Statistics of Random Processes, Volume 1*. Springer-Verlag.

Loh, R. K. and R. M. Stulz (2014). Is sell-side research more valuable in bad times? *Working Paper*.

Lucas, Robert E., J. (1978). Asset prices in an exchange economy. *Econometrica 46*(6), pp. 1429–1445.

Makarov, I. and O. Rytchkov (2012). Forecasting the forecasts of others: Implications for asset pricing. *Journal of Economic Theory 147*(3), 941 – 966.

Milas, C., P. Rothman, and D. van Dijk (2006). *Nonlinear time series analysis of business cycles*, Volume 276. Emerald Group Publishing.

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion. *The Journal of Finance 32*(4), pp. 1151–1168.

Moskowitz, T. J., Y. H. Ooi, and L. H. Pedersen (2012). Time series momentum. *Journal of Financial Economics 104*(2), 228–250.

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.

Ottaviani, M. and P. N. Sorensen (2015). Price reaction to information with heterogeneous beliefs and wealth effects: Underreaction, momentum, and reversal. *American Economic Review 105*(1), 1–34.

Patton, A. and A. Timmermann (2010). Why do forecasters disagree? lessons from the term structure of cross-sectional dispersion. *Journal of Monetary Economics 7*, 803–820.

Rapach, D. E., J. K. Strauss, and G. Zhou (2010, February). Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy. *Review of Financial Studies 23*(2), 821–862.

Revuz, D. and M. Yor (1999). *Continuous Martingales and Brownian Motion*. Springer-Verlag Berlin Heidelberg.

Sagi, J. S. and M. S. Seasholes (2007, May). Firm-specific attributes and the cross-section of momentum. *Journal of Financial Economics 84*(2), 389–434.

Scheinkman, J. A. and W. Xiong (2003). Overconfidence and speculative bubbles. *The Journal of Political Economy 111*(6), 1183–1219.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance 62*(3), 1139–1168.

Van Nieuwerburgh, S. and L. Veldkamp (2006, May). Learning asymmetries in real business cycles. *Journal of Monetary Economics 53*(4), 753–772.

Vayanos, D. and P. Woolley (2013). An institutional theory of momentum and reversal. *Review of Financial Studies 26*(5), 1087–1145.

Veldkamp, L. L. (2005, October). Slow boom, sudden crash. *Journal of Economic Theory 124*(2), 230–257.

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.

Veronesi, P. (2000). How does information quality affect stock returns? *Journal of Finance 55, 2*, pages 807–837.

Wang, J. (1993, April). A model of intertemporal asset prices under asymmetric information. *Review of Economic Studies 60*(2), 249–82.

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies 21*(4), 1455–1508.

Xiong, W. (2014). *Bubbles, Crises, and Heterogeneous Beliefs, published in Handbook for Systemic Risks.*

Xiong, W. and H. Yan (2010, April). Heterogeneous expectations and bond markets. *Review of Financial Studies 23*(4), 1433–1466.

Yamada, T. and S. Watanabe (1971). On the uniqueness of solutions of stochastic differential equations. *Journal of Mathematics of Kyoto University 11*, 155–167.

Zapatero, F. (1998, April). Effects of financial innovations on market volatility when beliefs are heterogeneous. *Journal of Economic Dynamics and Control 22*(4), 597–626.

# A Internet Appendix

## A.1 Proof of Proposition 1

We follow the notations in Lipster and Shiryaev (2001a) and write the observable process as

$$\frac{\mathrm{d}\delta_t}{\delta_t} = \left(A_0 + A_1 f_t^A\right)\mathrm{d}t + A_1 \mathrm{d}W_t^f + A_2 \mathrm{d}W_t^A$$

and the unobservable process as

$$\mathrm{d}f_t^A = \left(a_0 + a_1 f_t^A\right)\mathrm{d}t + b_1 \mathrm{d}W_t^f + b_2 \mathrm{d}W_t^A.$$

Using the SDEs in (1) and (2), we write $A \circ A = \sigma_f^\delta$, $b \circ b = \sigma_f^2$ and $b \circ A = 0$. Applying Theorem 12.7 in Lipster and Shiryaev (2001a), the dynamics of the filter satisfy

$$\mathrm{d}\widehat{f}_t^A = \left(a_0 + a_1 \widehat{f}_t^A\right)\mathrm{d}t + \left(A \circ A + \gamma A_1^\top\right)(A \circ A)^{-1}\left(\frac{\mathrm{d}\delta_t}{\delta_t} - \left(A_0 + A_1 \widehat{f}_t^A\right)\mathrm{d}t\right),$$

where the steady-state posterior variance $\gamma$ solves the algebraic equation

$$a_1\gamma + \gamma a_1^\top + b \circ b - \left(b \circ A + \gamma A_1^\top\right)(A \circ A)^{-1}\left(b \circ A + \gamma A_1^\top\right)^\top = 0.$$

Substituting the coefficients yields Equation (4), the steady-state posterior variance $\gamma$, and

$$\mathrm{d}\widehat{W}_t^A = \frac{1}{\sigma_\delta}\left(\frac{\mathrm{d}\delta_t}{\delta_t} - \widehat{f}_t^A \mathrm{d}t\right).$$

□

## A.2 Proof of Proposition 2

We want to demonstrate the equivalence of the probability measures $\widehat{\mathbb{P}}^A$ and $\widehat{\mathbb{P}}^B$. To do so, we start with the following definition.

**Definition 2.** *The probability measures $\widehat{\mathbb{P}}^A$ and $\widehat{\mathbb{P}}^B$ are equivalent if and only if they are absolutely continuous with respect to each other under $\mathcal{F}_t$ $\forall t \in \mathbb{R}_+$.*

From Girsanov Theorem (Theorem 5.1, Karatzas and Shreve (1988)), the probability measures $\widehat{\mathbb{P}}^A$ and $\widehat{\mathbb{P}}^B$ are absolutely continuous with respect to each other if and only if the local martingale in (6) is a strictly positive martingale, i.e., $E^{\widehat{\mathbb{P}}^A}[\eta_t] = 1$ for all $t \in \mathbb{R}_+$. Hence, to prove the equivalence of $\widehat{\mathbb{P}}^A$ and $\widehat{\mathbb{P}}^B$, we must show that (6) is a martingale. To do so, we write the dynamics of agents'

disagreement under $\mathbb{P}^A$ as

$$\begin{aligned} \mathrm{d}g_t = &\left[\kappa\left(\bar{f} - \widehat{f}_t^A\right) - (\lambda+\psi)\left(f_\infty + g_t - \widehat{f}_t^A\right) - \frac{g_t}{\sigma_\delta^2}\left(\widehat{f}_t^A - g_t - f^l\right)\left(f^h + g_t - \widehat{f}_t^A\right)\right] dt \\ &+ \frac{\gamma - \left(\widehat{f}_t^A - g_t - f^l\right)\left(f^h + g_t - \widehat{f}_t^A\right)}{\sigma_\delta} \mathrm{d}\widehat{W}_t^A. \end{aligned} \tag{31}$$

and use the following result, which we formulate in Theorem 1.

**Theorem 1.** *The process $\eta$ defined in (6) is a true martingale (as opposed to a local martingale) if and only if the process $g$ defined in (31) has a unique nonexplosive strong solution under $\widehat{\mathbb{P}}^A$ and $\widehat{\mathbb{P}}^B$.*

**Proof**. The proof follows as a special case of Theorem A.1 in Heston, Loewenstein, and Willard (2007). See also Exercise 2.10 in Revuz and Yor (1999) and Theorem 7.19 in Lipster and Shiryaev (2001b) for related results. ■

We now show that the stochastic differential equation in (31) has a unique nonexplosive strong solution under $\widehat{\mathbb{P}}^A$ and $\widehat{\mathbb{P}}^B$. Rewrite the process $g_t$ in (31) as

$$\mathrm{d}g_t = \left(\mu(\widehat{f}_t^B) - \lambda(\widehat{f}_t^B) g_t\right) \mathrm{d}t + \sigma(\widehat{f}_t^B) \mathrm{d}\widehat{W}_t^A \tag{32}$$

under $\widehat{\mathbb{P}}^A$ and as

$$\mathrm{d}g_t = \left(\mu(\widehat{f}_t^B) - \left(\kappa + \frac{\gamma}{\sigma_\delta^2}\right) g_t\right) \mathrm{d}t + \sigma(\widehat{f}_t^B) \mathrm{d}\widehat{W}_t^B \tag{33}$$

under $\widehat{\mathbb{P}}^B$, where the functions

$$\begin{aligned} \mu &: (f^l, f^h) \to \left(\kappa(\bar{f} - f^h) - (\lambda+\psi)(f_\infty - f^l), \kappa(\bar{f} - f^l) - (\lambda+\psi)(f_\infty - f^h)\right) \\ \sigma &: (f^l, f^h) \to \left(\frac{1}{\sigma_\delta}(\gamma - \frac{1}{4}(f^h - f^l)^2), \frac{\gamma}{\sigma_\delta}\right) \\ \lambda &: (f^l, f^h) \to \left(\kappa, \kappa + \frac{1}{4\sigma_\delta^2}(f^h - f^l)^2\right) \end{aligned}$$

are defined as

$$\begin{aligned} \mu(x) &:= \kappa(\bar{f} - x) - (\lambda+\psi)(f_\infty - x) \\ \sigma(x) &:= \frac{1}{\sigma_\delta}(\gamma - (x - f^l)(f^h - x)) \\ \lambda(x) &:= \kappa + \frac{1}{\sigma_\delta^2}(x - f^l)(f^h - x). \end{aligned} \tag{34}$$

We then have the following result, which we highlight in Lemma 1.

**Lemma 1.** *The processes in (32) and (33) have a unique strong solution.*

**Proof**. We first prove the result under $\widehat{\mathbb{P}}^A$ and then show that the result under $\widehat{\mathbb{P}}^B$ follows as a special case. To prove strong existence, we construct a sequence of successive approximations to $g_t$ in (32) by setting

$$g_t^{(k+1)} := g_t^{(0)} + \int_0^t \left(\mu(\widehat{f}_s^{(k)}) - \lambda(\widehat{f}_s^{(k)})g_s^{(k)}\right) \mathrm{d}s + \int_0^t \sigma(\widehat{f}_s^{(k)}) \mathrm{d}\widehat{W}_s^A \tag{35}$$

for $k \geq 0$, where $\widehat{f}^{(k)}$ denotes some approximation of $\widehat{f}^B$. From (35), we can write $g_t^{(k+1)} - g_t^{(k)} = B_t + M_t$ where

$$B_t := \int_0^t \left(\mu(\widehat{f}_s^{(k)}) - \mu(\widehat{f}_s^{(k-1)}) - \lambda(\widehat{f}_s^{(k)})g_s^{(k)} + \lambda(\widehat{f}_s^{(k-1)})g_s^{(k-1)}\right) \mathrm{d}s$$

and

$$M_t := \int_0^t (\sigma(\widehat{f}_s^{(k)}) - \sigma(\widehat{f}_s^{(k-1)})) \mathrm{d}\widehat{W}_s^A.$$

Now observe that the functions in (34) are locally Lipschitz continuous. In particular, for all $x, y \in (f^l, f^h)$, we have

$$|\sigma(x) - \sigma(y)| = \frac{1}{\sigma_\delta}|x-y||f^h + f^l - (x+y)| \leq \frac{1}{\sigma_\delta} \max\left\{|f^h - f^l|, |f^l - f^h|\right\} |x-y| \tag{36}$$

and

$$|\mu(x) - \mu(y)| = |\lambda + \psi - \kappa||x-y| \leq (\lambda + \psi + \kappa)|x-y|.$$

Lipschitz continuity for $\lambda(\cdot)$ directly follows from (36). Furthermore, $\lambda(\cdot)$ is bounded. We can therefore let $f_t^{(k)} \equiv \widehat{f}_t^B$ for all $k \geq 0$ and write accordingly

$$g_t^{(k+1)} - g_t^{(k)} = -\int_0^t \lambda(\widehat{f}_s^B) \left(g_s^{(k)} - g_s^{(k-1)}\right) \mathrm{d}s. \tag{37}$$

Taking the absolute value of both sides of (37) and observing that $\lambda(x) > 0$ for all $x \in (f^l, f^h)$, we have

$$\left|g_t^{(k+1)} - g_t^{(k)}\right| \leq \int_0^t \lambda(\widehat{f}_s^B) \left|g_s^{(k)} - g_s^{(k-1)}\right| \mathrm{d}s \leq \overline{\lambda} \int_0^t \left|g_s^{(k)} - g_s^{(k-1)}\right| \mathrm{d}s \tag{38}$$

where

$$\overline{\lambda} := \sup_{x \in (f^l, f^h)} \lambda(x) = \lambda\left(\frac{f^h + f^l}{2}\right).$$

Iterating over (38), we further get

$$\left|g_t^{(k+1)} - g_t^{(k)}\right| \leq \left|g_t^{(1)} - g_t^{(0)}\right| \frac{(\overline{\lambda}t)^k}{k!}.$$

Strong existence then directly follows from the last part of the proof of Theorem 2.9, Karatzas and Shreve (1988). The result under $\widehat{\mathbb{P}}^B$ follows as a special case by setting $\lambda(x) \equiv \kappa + \frac{\gamma}{\sigma_\delta^2}$.

To prove uniqueness, we adapt the proof of Yamada and Watanabe (1971). Suppose that there are two strong solutions $g^1$ and $g^2$ to (32) with $g_0^1 = g_0^2$, $\widehat{\mathbb{P}}^A-$a.s. It is then sufficient to show that $g^1$ and $g^2$ are indistinguishable. Using (32), we can write

$$\mathrm{d}(g_t^1 - g_t^2) = \lambda(\widehat{f}_t^B)(g_t^1 - g_t^2)\mathrm{d}t.$$

Integrating and taking the absolute value, we obtain

$$|g_t^1 - g_t^2| \leq \int_0^t \lambda(\widehat{f}_s^B)|g_s^1 - g_s^2|\mathrm{d}s \leq \overline{\lambda}\int_0^t |g_s^1 - g_s^2|\mathrm{d}s \leq 0$$

where the last inequality follows from Gronwall inequality (Problem 2.7, Karatzas and Shreve (1988)). Similarly, strong uniqueness under $\widehat{\mathbb{P}}^B$ follows as a special case when $\lambda(x) \equiv \kappa + \frac{\gamma}{\sigma_\delta^2}$. ■

It now remains to show that the disagreement process does not explode both $\widehat{\mathbb{P}}^A-$ and $\widehat{\mathbb{P}}^B-$almost surely, as we do in Lemma 2. Our result is actually stronger: we bound the cumulative distribution of $g$ by a scaled Gaussian cumulative distribution function, which guarantees uniform integrability of $g$.

**Lemma 2.** *At any time $t \in \mathbb{R}_+$, the process $g_t$ defined in (31) is finite $\widehat{\mathbb{P}}^A-$ and $\widehat{\mathbb{P}}^B-$almost surely, i.e.,*

$$\lim_{c\to\infty} \widehat{\mathbb{P}}^A(|g_t| \geq c) = \lim_{c\to\infty} \widehat{\mathbb{P}}^B(|g_t| \geq c) = 0, \quad \forall t \in \mathbb{R}_+.$$

**Proof**. We prove the result under $\widehat{\mathbb{P}}^A$. The result under $\widehat{\mathbb{P}}^B$ follows as a special case when $\lambda(x) \equiv \kappa + \frac{\gamma}{\sigma_\delta^2}$. Applying Ito's lemma, let $A_t := e^{\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s} g_t$ satisfy

$$\mathrm{d}A_t = \mu(\widehat{f}_t^B)e^{\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}\mathrm{d}t + e^{\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}\sigma(\widehat{f}_t^B)\mathrm{d}\widehat{W}_t^A, \quad A_0 = g_0 \tag{39}$$

and let $A^i$, $i = 1, 2$ have dynamics

$$\mathrm{d}A_t^i = (-1)^i m_t^A \mathrm{d}t + e^{\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}\sigma(\widehat{f}_t^B)\mathrm{d}\widehat{W}_t^A, \quad (-1)^i A_0^i \geq (-1)^i A_0 \tag{40}$$

under $\widehat{\mathbb{P}}^A$. Combining (39) and (40), we then obtain

$$A_t^i - A_t = A_0^i - A_0 + \int_0^t \left((-1)^i m_s^A - \mu(\widehat{f}_s^B)e^{\int_0^s \lambda(\widehat{f}_u^B)\mathrm{d}u}\right)\mathrm{d}s, \quad i =, 1, 2. \tag{41}$$

Now set

$$m_t^A := \sup_{x,y\in(f^l,f^h)} e^{\int_0^t \lambda(y_s)\mathrm{d}s}|\mu(x)| = e^{\overline{\lambda}t} \sup_{x\in(f^l,f^h)} |\mu(x)| = \exp\left(\overline{\lambda}t\right) \max\left\{|\mu(f^l)|, |\mu(f^h)|\right\}$$

and observe that

$$(-1)^i e^{\int_0^t \lambda(\widehat{f}_u^B)\mathrm{d}u}\mu(\widehat{f}_t^B) \leq e^{\int_0^t \lambda(\widehat{f}_u^B)\mathrm{d}u}|\mu(\widehat{f}_t^B)| \leq m_t^A, \quad i = 1, 2 \tag{42}$$

for all $t \in \mathbb{R}_+$. Inequalities in (42) and the expressions in (41) together imply that

$$A_t^1 \leq A_t \leq A_t^2, \quad \widehat{\mathbb{P}}^A - a.s. \tag{43}$$

for all $t \in \mathbb{R}_+$. Furthermore, rewriting $g$ as

$$g_t = e^{-\int_0^t \lambda(\widehat{f}_u^B)\mathrm{d}u} A_t = e^{-\int_0^t \lambda(\widehat{f}_u^B)\mathrm{d}u}(A_t^+ - A_t^-)$$

it then follows that

$$|g_t| = e^{-\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}(A_t^+ + A_t^-) \leq A_t^+ + A_t^-, \quad \widehat{\mathbb{P}}^A - a.s. \tag{44}$$

where the second inequality follows from that $e^{-\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s} \in (0, 1)$ for all $t \in \mathbb{R}_+$, since

$$\lambda(x) > 0, \quad \forall x \in (f^l, f^h).$$

Combining (43) and (44), we obtain

$$|g_t| \leq \sum_{i=1,2} ((-1)^i A_t^i)^+ \leq \max\{(-A_t^1)^+, (A_t^2)^+\}, \quad \widehat{\mathbb{P}}^A - a.s. \tag{45}$$

Using (45), we can write that, for any positive constant $c \geq 0$,

$$\mathbf{1}_{|g_t|\geq c} \leq \sum_{i=1,2} \mathbf{1}_{((-1)^i A_t^i)^+ \geq c} \equiv \sum_{i=1,2} \mathbf{1}_{(-1)^i A_t^i \geq c} \tag{46}$$

where the last equality follows from that $\mathbf{1}_{X^+\geq c} = \mathbf{1}_{X\geq c}$ for any $c$ positive. Taking expectations of (46) under $\widehat{\mathbb{P}}^A$, we get

$$\widehat{\mathbb{P}}^A(|g_t| \geq c) \leq \sum_{i=1,2} \widehat{\mathbb{P}}^A\left((-1)^i A_t^i \geq c\right), \quad \forall t \in \mathbb{R}_+. \tag{47}$$

Finally, adapting the proof of Theorem 1.4 in Hajek (1985), let $\widehat{A}^i$, $i = 1, 2$ have dynamics

$$\mathrm{d}\widehat{A}_t^i = (-1)^i m_t^A \mathrm{d}t + \sigma_t^A \mathrm{d}\widehat{W}_t^A, \quad (-1)^i \widehat{A}_0^i \geq (-1)^i A_0^i \tag{48}$$

under $\widehat{\mathbb{P}}^A$ and set

$$\sigma_t^A := \sup_{x,y\in(f^l,f^h)} e^{\int_0^t \lambda(y_s)\mathrm{d}s}|\sigma(x)| \equiv \exp\left(\overline{\lambda}t\right)\frac{1}{\sigma_\delta}\max\left\{\gamma, \left|\gamma - \frac{1}{4}(f^h - f^l)^2\right|\right\}.$$

Furthermore, assume without loss of generality that there exists a Brownian motion $\widehat{W}$ on the same probability space as $\widehat{W}^A$, which is independent of $(A^1, A^2, \widehat{f}^B, \widehat{W}^A)$. Let $a^{i,j}$, $i,j = 1,2$ be defined by

$$\begin{aligned} a_t^{i,j} &= \widehat{A}_0^i + (-1)^i \int_0^t m_s^A \mathrm{d}s \\ &+ \left[\int_0^t e^{\int_0^s \lambda(\widehat{f}_u^B)\mathrm{d}u}\sigma(\widehat{f}_s^B)\mathrm{d}\widehat{W}_s^A + (-1)^j \int_0^t \left((\sigma_s^A)^2 - e^{\int_0^s \lambda(\widehat{f}_u^B)\mathrm{d}u}\sigma(\widehat{f}_s^B)^2\right)^{\frac{1}{2}}\mathrm{d}\widehat{W}_s\right]. \end{aligned}$$

First, observe that for each $j = 1,2$, the process in the square bracket is a continuous martingale with quadratic variation equal to $\int_0^t (\sigma_s^A)^2\mathrm{d}s$. As a result, each $a^{i,j}$, $j = 1,2$ has the same distribution as $\widehat{A}^i$, i.e.,

$$a^{i,j} \sim \widehat{A}^i, \quad i,j = 1,2. \tag{49}$$

Second, define the process $\bar{a}_t^i := \frac{1}{2}\sum_{j=1,2} a_t^{i,j}$, $i = 1,2$, which, applying Ito's lemma, satisfies

$$\mathrm{d}\bar{a}_t^i = (-1)^i m_t^A \mathrm{d}t + e^{\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}\sigma(\widehat{f}_t^B)\mathrm{d}\widehat{W}_t^A. \tag{50}$$

Combining (40) and (50), we obtain from the initial conditions that

$$(-1)^i \bar{a}_t^i \geq (-1)^i A_t^i, \quad \widehat{\mathbb{P}}^A - a.s.$$

for $i = 1,2$. Since $(-1)^i\bar{a}_t^i \leq \max\{(-1)^i a_t^{i,1}, (-1)^i a_t^{i,2}\}$, $i = 1,2$, we further have that

$$\mathbf{1}_{(-1)^i A_t^i \geq c} \leq \mathbf{1}_{(-1)^i a_t^{i,1}\geq c} + \mathbf{1}_{(-1)^i a_t^{i,2}\geq c}.$$

Taking expectations under $\widehat{\mathbb{P}}^A$ and using (49), we obtain that, for any $c \in \mathbb{R}$, the processes $A^i$ and $\widehat{A}^i$, $i = 1,2$ satisfy

$$\mathbb{P}^A((-1)^i A_t^i \geq c) \leq 2\mathbb{P}^A((-1)^i \widehat{A}_t^i \geq c), \quad \forall t \in \mathbb{R}_+. \tag{51}$$

Combining the inequalities in (47) and (51), we obtain that

$$\widehat{\mathbb{P}}^A(|g_t| \geq c) \leq 2\sum_{i=1,2}\widehat{\mathbb{P}}^A\left((-1)^i\widehat{A}_t^i \geq c\right), \quad \forall t \in \mathbb{R}_+. \tag{52}$$

Observing that each $\widehat{A}^i$, $i = 1,2$ in (48) is a Gaussian process, the probabilities on the right-hand

side of (52) are explicitly given by

$$\mathbb{P}^A\left((-1)^i\widehat{A}_t^i \geq c\right) = \Phi\left(\frac{(-1)^i\widehat{A}_0^i + \int_0^t m_s^A \mathrm{d}s - c}{\sqrt{\int_0^t (\sigma_s^A)^2 \mathrm{d}s}}\right). \tag{53}$$

Taking limits on both sides of (52) and using (53) yields

$$\lim_{c\to\infty} \mathbb{P}^A(|g_t| \geq c) \leq 2 \lim_{c\to\infty} \sum_{i=1,2} \Phi\left(\frac{(-1)^i\widehat{A}_0^i + \int_0^t m_s^A \mathrm{d}s - c}{\sqrt{\int_0^t (\sigma_s^A)^2 \mathrm{d}s}}\right) = 0, \quad \forall t \in \mathbb{R}_+$$

as desired. ■

We have shown that the process $\eta$ in (6) is a martingale and therefore $E^{\widehat{\mathbb{P}}^A}[\eta_t] = 1$ for all $t \in \mathbb{R}_+$. As a result, $\widehat{\mathbb{P}}^A$ is absolutely continuous with respect to $\widehat{\mathbb{P}}^B$ under $\mathcal{F}_t \ \ \forall t \in \mathbb{R}_+$ and the claim follows from Girsanov Theorem (Theorem 5.1, Karatzas and Shreve (1988)).

□

## A.3 Maximization Problems

We write Agent $A$'s problem as follows

$$\max_{c_A} \mathbb{E}^{\mathbb{P}^A}\left[\int_0^\infty e^{-\rho t}\frac{c_{At}^{1-\alpha}}{1-\alpha}\mathrm{d}t\right] + \phi_A\left(X_{A,0} - \mathbb{E}^{\mathbb{P}^A}\left[\int_0^\infty \xi_t c_{At}\mathrm{d}t\right]\right),$$

where $\phi_A$ denotes the Lagrange multiplier of Agent $A$'s static budget constraint and $\xi$ is the state-price density perceived by Agent $A$. Agent $B$ solves an analogous problem but under her own probability measure $\mathbb{P}^B$. Rewriting Agent $B$'s problem under Agent $A$'s probability measure $\mathbb{P}^A$ yields

$$\max_{c_B} \mathbb{E}^{\mathbb{P}^A}\left[\int_0^\infty \eta_t e^{-\rho t}\frac{c_{Bt}^{1-\alpha}}{1-\alpha}\mathrm{d}t\right] + \phi_B\left(X_{B,0} - \mathbb{E}^{\mathbb{P}^A}\left[\int_0^\infty \xi_t c_{Bt}\mathrm{d}t\right]\right).$$

The first-order conditions lead to the following optimal consumption plans

$$c_{At} = \left(\phi_A e^{\rho t}\xi_t\right)^{-\frac{1}{\alpha}} \qquad c_{Bt} = \left(\frac{\phi_B e^{\rho t}\xi_t}{\eta_t}\right)^{-\frac{1}{\alpha}}. \tag{54}$$

Clearing the market yields the following characterization of the state-price density $\xi$:

$$\xi_t = e^{-\rho t}\delta_t^{-\alpha}\left[\left(\frac{\eta_t}{\phi_B}\right)^{\frac{1}{\alpha}} + \left(\frac{1}{\phi_A}\right)^{\frac{1}{\alpha}}\right]^\alpha. \tag{55}$$

Substituting Equation (55) into Equation (54) gives the consumption share of Agent $A$, $\omega(\eta)$, which satisfies:

$$\omega\left(\eta_t\right)=\frac{\left(\frac{1}{\phi_A}\right)^{\frac{1}{\alpha}}}{\left(\frac{\eta_t}{\phi_B}\right)^{\frac{1}{\alpha}}+\left(\frac{1}{\phi_A}\right)^{\frac{1}{\alpha}}}.$$

The consumption share of Agent $A$ is a function of the likelihood $\eta$. In particular, an increase in $\eta$ raises the likelihood of Agent $B$'s model relative to Agent $A$'s model. The consumption share of Agent $A$ is therefore decreasing in $\eta$. That is, the more likely Agent $B$'s model becomes, the less Agent $A$ can consume. This result applies symmetrically to the consumption share, $1-\omega(\eta)$, of Agent $B$.

The dynamics of the state-price density $\xi$ satisfy

$$\frac{\mathrm{d}\xi_t}{\xi_t}=-r_t^f\mathrm{d}t-\theta_t\mathrm{d}\widehat{W}_t^A.$$

Therefore, applying Itô's lemma to the state-price density defined in (55) determines the risk-free rate $r^f$ and the market price of risk $\theta$ provided in Proposition 3.

□

## A.4 Proof of Proposition 4

Following Dumas et al. (2009), we assume that the coefficient of relative risk aversion $\alpha$ is an integer. This assumption allows us to obtain the following convenient expression for the equilibrium stock price:[22]

$$\begin{aligned}\frac{S_t}{\delta_t}&=\mathbb{E}_t^{\mathbb{P}^A}\left[\int_t^\infty\frac{\xi_u\delta_u}{\xi_t\delta_t}du\right]\\&=\omega\left(\eta_t\right)^\alpha\sum_{j=0}^{\alpha}\left(\begin{array}{c}\alpha\\j\end{array}\right)\left(\frac{1-\omega\left(\eta_t\right)}{\omega\left(\eta_t\right)}\right)^j\mathbb{E}_t^{\mathbb{P}^A}\left[\int_t^\infty e^{-\rho(u-t)}\left(\frac{\eta_u}{\eta_t}\right)^{\frac{j}{\alpha}}\left(\frac{\delta_u}{\delta_t}\right)^{1-\alpha}du\right].\end{aligned}\tag{56}$$

We start by computing the first and the last term of the sum in (56). These terms correspond to the prices of a Lucas (1978) economy in which the representative agent assumes that fundamental follows an Ornstein-Uhlenbeck process and a 2-state Markov chain, respectively. These prices have (semi) closed-form solutions, which we present in Proposition 8.

**Proposition 8.** *Suppose the economy is populated by a single agent.*

1. *If the agent's filter follows the Ornstein-Uhlenbeck process described in Equation (4), the*

[22] We refer the reader to Dumas et al. (2009) for the details of the derivation.

*equilibrium price-dividend ratio satisfies*

$$\left.\frac{S_t}{\delta_t}\right|_{O.U.} = \int_0^\infty e^{-\rho\tau+\alpha(\tau)+\beta_2(\tau)\widehat{f}_t^A}\, d\tau,$$

*where the functions $\alpha(\tau)$ and $\beta_2(\tau)$ are the solutions to a set of Ricatti equations.*

2. *If the agent's filter follows the filtered 2-state Markov chain process described in Equation (5), the equilibrium price-dividend ratio satisfies*

$$\begin{aligned}\left.\frac{S_t}{\delta_t}\right|_{M.C.} &= \pi_t H_1 + (1-\pi_t)H_2 = \frac{\widehat{f}_t^B - f^l}{f^h - f^l}H_1 + \left(1 - \frac{\widehat{f}_t^B - f^l}{f^h - f^l}\right)H_2 \\ &= \frac{\widehat{f}_t^A - g_t - f^l}{f^h - f^l}H_1 + \left(1 - \frac{\widehat{f}_t^A - g_t - f^l}{f^h - f^l}\right)H_2,\end{aligned}$$

*where*

$$\begin{aligned}H &= \begin{pmatrix} H_1 & H_2 \end{pmatrix}^\top = A^{-1}\mathbf{1}_2 \\ A &= -\Omega - (1-\alpha)\begin{pmatrix} f^h & 0 \\ 0 & f^l \end{pmatrix} + (\rho + \frac{1}{2}\alpha(1-\alpha)\sigma_\delta^2)\mathbf{Id}_2.\end{aligned}$$

$\mathbf{Id}_2$ *is a 2-by-2 identity matrix and* $\mathbf{1}_2$ *is a 2-dimensional vector of ones.*

**Proof**.

1. Following Duffie, Pan, and Singleton (2000), the functions $\alpha(\tau)$ and $\beta(\tau) \equiv (\beta_1(\tau), \beta_2(\tau))$ solve the following system of Ricatti equations

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

with boundary conditions $\beta(0) = (1-\alpha, 0)$ and $\alpha(0) = 0$. The $H$ and $K$ matrices satisfy

$$\begin{aligned}K_0 &= \begin{pmatrix} -\frac{1}{2}\sigma_\delta^2 \\ \kappa\bar{f} \end{pmatrix} & K_1 &= \begin{pmatrix} 0 & 1 \\ 0 & -\kappa \end{pmatrix} \\ H_0 &= \begin{pmatrix} \sigma_\delta^2 & \gamma \\ \gamma & \left(\frac{\gamma}{\sigma_\delta}\right)^2 \end{pmatrix} & H_1 &= 0_2 \otimes 0_2.\end{aligned}$$

The solutions to this system are

$$\begin{aligned}\beta_1(\tau) =& 1-\alpha \\ \beta_2(\tau) =& -\frac{(\alpha-1)e^{-\kappa\tau}\left(e^{\kappa\tau}-1\right)}{\kappa} \\ \alpha(\tau) =& \frac{(\alpha-1)^2\gamma^2 e^{-2\kappa\tau}\left(e^{2\kappa\tau}(2\kappa\tau-3)+4e^{\kappa\tau}-1\right)}{4\kappa^3\sigma_\delta^2} \\ & -\frac{(\alpha-1)e^{-2\kappa\tau}\left(4\kappa\sigma_\delta^2 e^{\kappa\tau}\left(e^{\kappa\tau}(\kappa\tau-1)+1\right)\left(\kappa\bar{f}-\alpha\gamma+\gamma\right)+2\alpha\kappa^3\tau\sigma_\delta^4 e^{2\kappa\tau}\right)}{4\kappa^3\sigma_\delta^2}.\end{aligned}$$

□

2. See Veronesi (2000). □

■

We now rewrite the price in (56) as

$$\frac{S_t}{\delta_t} = \omega(\eta_t)^\alpha \frac{S_t}{\delta_t}\bigg|_{O.U.} + \omega(\eta_t)^\alpha \sum_{j=1}^{\alpha-1}\begin{pmatrix}\alpha\\ j\end{pmatrix}\left(\frac{1-\omega(\eta_t)}{\omega(\eta_t)}\right)^j F^j\left(\widehat{f}_t^A, g_t\right) + (1-\omega(\eta_t))^\alpha \frac{S_t}{\delta_t}\bigg|_{M.C.}. \tag{57}$$

The last step consists in computing the intermediate terms $F^j$ in (57), which relate to heterogeneous beliefs. Each term solves a differential equation, which we present in Proposition 9.

**Proposition 9.** *The function $F^j$, defined as*

$$F^j\left(\widehat{f}_t^A, g_t\right) \equiv \mathbb{E}_t^{\mathbb{P}^A}\left[\int_t^\infty e^{-\rho(u-t)}\left(\frac{\eta_u}{\eta_t}\right)^{\frac{j}{\alpha}}\left(\frac{\delta_u}{\delta_t}\right)^{1-\alpha} du\right], \tag{58}$$

*solves the following partial differential equation*

$$\widetilde{\mathscr{L}}^{\widehat{f}^A, g} F^j + X^j F^j + 1 = 0, \tag{59}$$

*where $\widetilde{\mathscr{L}}$ denotes the infinitesimal generator of $\left(\widehat{f}^A, g\right)$ under the probability measure $\widetilde{\mathbb{P}}^A$.*

**Proof.**
We introduce two sequential changes of probability measure, one from $\mathbb{P}^A$ to a probability measure $\overline{\mathbb{P}}$ according to

$$\frac{\mathrm{d}\overline{\mathbb{P}}}{\mathrm{d}\mathbb{P}^A}\bigg|_{\mathscr{F}_t} \equiv e^{-\frac{1}{2}\int_0^t\left(\frac{j}{\alpha}\frac{g_s}{\sigma_\delta}\right)^2 \mathrm{d}s - \int_0^t \frac{j}{\alpha}\frac{g_s}{\sigma_\delta}\mathrm{d}\widehat{W}_s^A}$$

and one from $\overline{\mathbb{P}}$ to a probability measure $\widetilde{\mathbb{P}}$ according to

$$\left.\frac{\mathrm{d}\widetilde{\mathbb{P}}}{\mathrm{d}\overline{\mathbb{P}}}\right|_{\mathscr{F}_t} \equiv e^{-\frac{1}{2}\int_0^t (1-\alpha)^2\sigma_\delta^2 \mathrm{d}s + \int_0^t (1-\alpha)\sigma_\delta \mathrm{d}\overline{W}(s)},$$

where, by Girsanov's Theorem, $\overline{W}$ is a $\overline{\mathbb{P}}-$Brownian motion satisfying

$$\overline{W}_t = \widehat{W}_t^A + \int_0^t \frac{j}{\alpha}\frac{g_s}{\sigma_\delta}\mathrm{d}s, \tag{60}$$

and $\widetilde{W}$ is a $\widetilde{\mathbb{P}}-$Brownian motion satisfying

$$\widetilde{W}_t = \overline{W}_t - \int_0^t (1-\alpha)\,\sigma_\delta \mathrm{d}t. \tag{61}$$

Implementing sequentially the changes of probability measures in Equations (60) and (61) allows us to rewrite the interior expectation in (58) as

$$F^j\left(\widehat{f}_t, g_t\right) = \mathbb{E}^{\widetilde{P}}\left[\int_t^\infty e^{\int_t^u X_s^j \mathrm{d}s}\mathrm{d}u \middle| \mathscr{F}_t\right], \tag{62}$$

where

$$X_t^j = -\left(\rho + \frac{1}{2}(1-\alpha)\,\alpha\sigma_\delta^2\right) + \frac{1}{2}\frac{j}{\alpha}\left(\frac{j}{\alpha}-1\right)\frac{g_t^2}{\sigma_\delta^2} - (1-\alpha)\,\frac{j}{\alpha}g_t + (1-\alpha)\,\widehat{f}_t^A.$$

To obtain the partial differential equation that the function $F^j$ has to satisfy, we use the fact that Equation (62) can be rewritten as follows

$$\begin{aligned}
F^j\left(\widehat{f}_t, g_t\right) &= e^{-\int_0^t X_s^j \mathrm{d}s}\mathbb{E}^{\widetilde{P}}\left[\int_t^\infty e^{\int_0^u X_s^j \mathrm{d}s}\mathrm{d}u \middle| \mathscr{F}_t\right] \\
&= e^{-\int_0^t X_s^j \mathrm{d}s}\left(-\int_0^t e^{\int_0^u X_s^j \mathrm{d}s}\mathrm{d}u + \mathbb{E}^{\widetilde{P}}\left[\int_0^\infty e^{\int_0^u X_s^j \mathrm{d}s}\mathrm{d}u \middle| \mathscr{F}_t\right]\right) \\
&\equiv e^{-\int_0^t X_s^j \mathrm{d}s}\left(-\int_0^t e^{\int_0^u X_s^j \mathrm{d}s}\mathrm{d}u + \widetilde{M}_t\right),
\end{aligned}$$

where $\widetilde{M}$ is a $\widetilde{\mathbb{P}}-$Martingale. An application of Itô's lemma along with the Martingale Representation Theorem then gives the partial differential equation in (59).[23] □
■

We numerically solve Equation (59) for each term $j$ through Chebyshev collocation. In partic-

[23] As proved in David (2008), the boundary conditions are absorbing in both the $\widehat{f}^A-$ and the $g-$dimension.

ular, we approximate the functions $F^j(\widehat{f}^A, g)$ for $j = 1, \ldots, \alpha - 1$ as follows:

$$P^j\left(\widehat{f}^A, g\right) = \sum_{i=0}^{n}\sum_{k=0}^{m} a_{i,k}^j T_i\left(\widehat{f}^A\right) T_k\left(g\right) \approx F^j(\widehat{f}^A, g),$$

where $T_i$ is the Chebyshev polynomial of order $i$. Following Judd (1998), we mesh the roots of the Chebyshev polynomial of order $n$ with those of the Chebyshev polynomial of order $m$ to obtain the interpolation nodes. We then substitute $P^j(\widehat{f}^A, g)$ and its derivatives in Equation (59), and we evaluate this expression at the interpolation nodes. Since all the boundary conditions are absorbing, this approach directly produces a system of $(n+1) \times (m+1)$ equations with $(n+1) \times (m+1)$ unknowns that we solve numerically. □

## A.5 Derivation of the Parameter Values Provided in Table 1

Agents first estimate a discretized version of their model and then map the parameters they estimated into their continuous-time model. In particular, Agent $A$ estimates the following discrete-time model

$$\log\left(\frac{\delta_{t+1}}{\delta_t}\right) = f_t^A + \sqrt{v^\delta}\epsilon_{1,t+1} \tag{63}$$

$$f_{t+1}^A = m^{f^A} + a^{f^A} f_t^A + \sqrt{v^{f^A}}\epsilon_{2,t+1}, \tag{64}$$

while Agent $B$ estimates the following discrete-time model

$$\log\left(\frac{\delta_{t+1}}{\delta_t}\right) = f_t^B + \sqrt{v^\delta}\epsilon_{3,t+1} \tag{65}$$

$$f_t^B \in \{s^h, s^l\} \quad \text{with transition matrix} \quad P = \begin{pmatrix} p^{hh} & 1-p^{hh} \\ 1-p^{ll} & p^{ll} \end{pmatrix}.$$

$\epsilon_1$, $\epsilon_2$, and $\epsilon_3$ are normally distributed with zero-mean and unit-variance, and $\epsilon_1$ and $\epsilon_2$ are independent. The transition matrix $P$ contains the probabilities of staying in the high and the low state over the following month. We estimate the discrete-time models in Equations (63) and (65) by Maximum Likelihood. We report the estimated parameters, their standard errors, and their statistical significance in Table 5.

We then map the parameters of Table 5 into the associated continuous-time models. Straightforward applications of Itô's lemma show that the dividend stream, $\delta$, and the fundamental perceived

| Parameter | Symbol | Estimate |
|---|---|---|
| Variance Dividend Growth | $v^{\delta}$ | $4.23 \times 10^{-5***}$ $(1.48 \times 10^{-6})$ |
| Persistence Growth Rate $f^A$ | $a^{f^A}$ | $0.9842^{***}$ $(0.0022)$ |
| Mean Growth Rate $f^A$ | $m^{f^A}$ | $9.96 \times 10^{-4***}$ $(9.39 \times 10^{-5})$ |
| Variance Growth Rate $f^A$ | $v^{f^A}$ | $2.53 \times 10^{-6***}$ $(2.12 \times 10^{-7})$ |
| High State of $f^B$ | $s^h$ | $0.0066^{***}$ $(2.7 \times 10^{-4})$ |
| Low State of $f^B$ | $s^l$ | $-0.0059^{***}$ $(3.17 \times 10^{-4})$ |
| Prob. of Staying in High State | $p^{hh}$ | $0.9755^{***}$ $(0.0912)$ |
| Prob. of Staying in Low State | $p^{ll}$ | $0.9680^{***}$ $(0.0940)$ |

**Table 5: Output of the Maximum-Likelihood Estimation.**
Parameter values resulting from a discrete-time Bayesian-Learning Maximum-Likelihood estimation. The estimation is performed on monthly S&P500 dividend data from 01/1871 to 11/2013. Standard errors are reported in brackets and statistical significance at the 10%, 5%, and 1% levels is labeled with *, **, and ***, respectively.

by Agent $A$, $f^A$, satisfy

$$\begin{aligned}
\log\left(\frac{\delta_{t+\Delta}}{\delta_t}\right) &= \int_t^{t+\Delta}\left(f_u^A - \frac{1}{2}\sigma_\delta^2\right)\mathrm{d}u + \sigma_\delta\left(W_{t+\Delta}^A - W_t^A\right) \\
&= \int_t^{t+\Delta}\left(f_u^B - \frac{1}{2}\sigma_\delta^2\right)\mathrm{d}u + \sigma_\delta\left(W_{t+\Delta}^B - W_t^B\right) \\
&\approx \left(f_t^A - \frac{1}{2}\sigma_\delta^2\right)\Delta + \sigma_\delta\left(W_{t+\Delta}^A - W_t^A\right) \qquad (66) \\
&\approx \left(f_t^B - \frac{1}{2}\sigma_\delta^2\right)\Delta + \sigma_\delta\left(W_{t+\Delta}^B - W_t^B\right) \qquad (67)
\end{aligned}$$

$$f_{t+\Delta}^A = e^{-\kappa\Delta}f_t^A + \bar{f}\left(1 - e^{-\kappa\Delta}\right) + \sigma_f\int_t^{t+\Delta} e^{-\kappa(t+\Delta-u)}\mathrm{d}W_u^f. \qquad (68)$$

The relationship between the transition matrix $P$ and the generator matrix $\Lambda$ is written as follows

$$\begin{aligned} P &= \begin{pmatrix} p^{hh} & 1-p^{hh} \\ 1-p^{ll} & p^{ll} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\psi}{\lambda+\psi} + \frac{\lambda}{\lambda+\psi}e^{-(\lambda+\psi)\Delta} & \frac{\lambda}{\lambda+\psi} - \frac{\lambda}{\lambda+\psi}e^{-(\lambda+\psi)\Delta} \\ \frac{\psi}{\lambda+\psi} - \frac{\psi}{\lambda+\psi}e^{-(\lambda+\psi)\Delta} & \frac{\lambda}{\lambda+\psi} + \frac{\psi}{\lambda+\psi}e^{-(\lambda+\psi)\Delta} \end{pmatrix}. \end{aligned} \tag{69}$$

We perform the Maximum-Likelihood estimation on monthly data and accordingly set $\Delta = 1/12 =$ 1 month.

Matching Equation (63) to Equation (66) and Equation (64) to Equation (68) yields the following system of equation for $\kappa$, $\bar{f}$, $\sigma_\delta$, and $\sigma_f$

$$\begin{aligned} a^{f^A} &= e^{-\kappa\Delta} \\ m^{f^A} &= \bar{f}\left(1-e^{-\kappa\Delta}\right) - \frac{1}{2}\sigma_\delta^2\Delta \\ v^{\delta} &= \sigma_\delta^2\Delta \\ v^{f^A} &= \frac{\sigma_f^2}{2\kappa}\left(1-e^{-2\kappa\Delta}\right), \end{aligned} \tag{70}$$

where the last equation relates the variance of the Ornstein-Uhlenbeck process to its empirical counterpart. Matching Equation (65) to Equation (67) yields the following system of equations for $f^l$ and $f^h$

$$\begin{aligned} s^h &= \left(f^h - \frac{1}{2}\sigma_\delta^2\right)\Delta \\ s^l &= \left(f^l - \frac{1}{2}\sigma_\delta^2\right)\Delta. \end{aligned} \tag{71}$$

Solving the system comprised of Equations (69), (70), and (71) yields the parameters presented in Table 1.

□

## A.6 Approximation of the Filter's Adjustment Speed

In this appendix, we derive an approximation for the adjustment speed of agents' filter, as defined in Definition 1. Define the vector $X_t := (\widehat{f}_t^A, \widehat{f}_t^B)^\top$ with dynamics

$$\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}\widehat{W}_t^A$$

under $\widehat{\mathbb{P}}^A$ where

$$\mu(X) := \begin{pmatrix} \kappa\overline{f} \\ (\lambda+\psi)f_\infty \end{pmatrix} + \begin{pmatrix} -\kappa & 0 \\ -\frac{f^l f^h}{\sigma_\delta^2} & -(\lambda+\psi-\frac{f^l f^h}{\sigma_\delta^2}) \end{pmatrix} X_t + \phi(X_t) \tag{72}$$

and

$$\sigma(X) := \begin{pmatrix} \frac{\gamma}{\sigma_\delta} \\ \frac{1}{\sigma_\delta}(\widehat{f}_t^B - f^l)(f^h - \widehat{f}_t^B) \end{pmatrix}.$$

Observe that the change of measure from $\widehat{\mathbb{P}}^B$ to $\widehat{\mathbb{P}}^A$ introduces a nonlinear component

$$\phi(X) = \begin{pmatrix} 0 \\ \frac{1}{\sigma_\delta^2}((f^h+f^l)\widehat{f}_t^B - (\widehat{f}_t^B)^2)(\widehat{f}_t^A - \widehat{f}_t^B) \end{pmatrix}$$

in the otherwise affine drift in (72). We perform now a first-order Taylor expansion of $\mu(\cdot)$ around the initial point $X_0$

$$\widetilde{\mu}(X_t) = \mu(X_0) + \nabla_X \mu(X)|_{X_0}(X - X_0) = \Lambda + \Omega(X_t - X_0) \tag{73}$$

where $\Lambda = \begin{pmatrix} \kappa\overline{f} & (\lambda+\psi)f_\infty \end{pmatrix}^\top$ and

$$\Omega = \begin{pmatrix} -\kappa & 0 \\ \frac{(f^h-x_0)(x_0-f^l)}{\sigma_\delta^2} & -\frac{(f^h-x_0)(x_0-f^l)}{\sigma_\delta^2} - (\lambda+\psi) \end{pmatrix}.$$

Denote by $\widetilde{X}_t$ the process associated with the linearized drift in (73). Observe that its drift is strictly affine and that its conditional expectation therefore satisfies

$$E_0^A[\widetilde{X}_t] = -\Omega^{-1}\Lambda + \exp(\Omega t)(X_0 + \Omega^{-1}\Lambda). \tag{74}$$

Furthermore, for $t$ small enough, the expression in (74) is approximately given by

$$\frac{\mathrm{d}}{\mathrm{d}t} E_0^A[\widetilde{X}_t]\Big|_{t=\epsilon} \approx \Omega(I + \Omega\epsilon)(X_0 + \Omega^{-1}\Lambda). \tag{75}$$

Finally, observe that our calibration implies that $f^l \approx -f^l$, $f_\infty \approx 0$ and that, for $\epsilon$ small, only terms of order $o(\frac{1}{\sigma_\delta^2})\epsilon$ dominate. Therefore, under our calibration, (75) is approximately given by the following simpler expression

$$\frac{\mathrm{d}}{\mathrm{d}t} E_0^A[\widetilde{X}_t]\Big|_{t=\epsilon} \approx \begin{pmatrix} \kappa(\overline{f} - x_0) \\ (\lambda+\psi)(f_\infty - x_0) \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{(f^l-x_0)(x_0+f^l)(\overline{f}\kappa + x_0(\lambda+\psi-\kappa))}{\sigma_\delta^2} \end{pmatrix} \epsilon,$$

which we provide in Proposition 5.

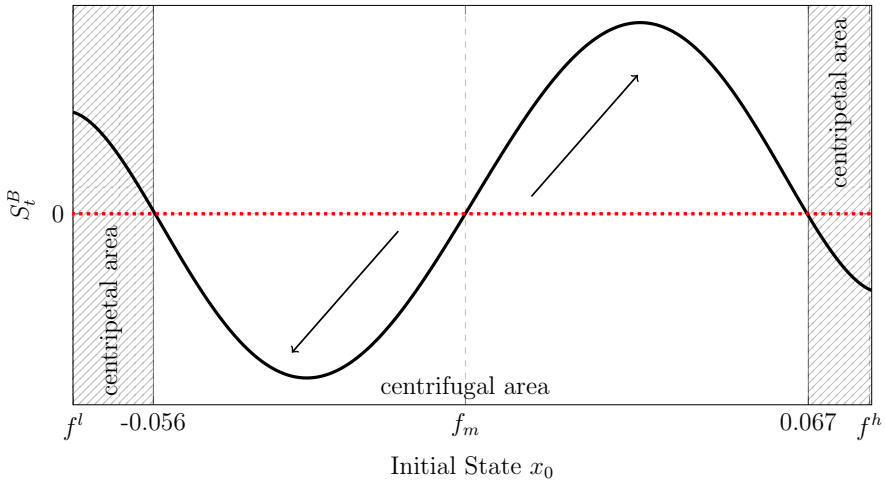


**Figure 14: Agent $B$'s Adjustment Speed under $\mathbb{P}^A$.**
This figure plots the second-order approximation of Agent $B$'s adjustment speed under $\mathbb{P}^A$ as a function of the state of the economy. The dashed areas represent the states in which Agent $B$'s expectations are attracted towards $f_m$, while the central area represents the states in which Agent $B$'s expectations are repelled outward $f_m$.

We now augment the vector $X_t$ with the quadratic term $(\widehat{f}^B)^2$ and accordingly defining a new vector $Y_t := (\widehat{f}_t^A, \widehat{f}_t^B, (\widehat{f}_t^B)^2)^\top$. An application of Ito's lemma shows that the drift of this process satisfies

$$\mu(Y) := \begin{pmatrix} \kappa\overline{f} \\ (\lambda+\psi)f_\infty \\ \frac{(f^h)^2(f^l)^2}{\sigma_\delta^2} \end{pmatrix} + \begin{pmatrix} -\kappa & 0 & 0 \\ -\frac{f^l f^h}{\sigma_\delta^2} & -(\lambda+\psi) - \frac{f^l f^h}{\sigma_\delta^2} & -\frac{f^h+f^l}{\sigma_\delta^2} \\ 0 & -2\frac{f^h f^l (f^h+f^l)}{\sigma_\delta^2} + 2f_\infty(\lambda+\psi) & \frac{(f^h)^2+6f^h f^l+(f^l)^2}{\sigma_\delta^2} - 2(\lambda+\psi) \end{pmatrix} Y_t$$

$$+ o((\widehat{f}_t^A)^2, (\widehat{f}_t^B)^3)$$

Now performing the first-order Taylor expansion of (73) around the initial point $Y_0$, the vector $\Lambda$ becomes

$$\Lambda = \begin{pmatrix} \overline{f}\kappa \\ f_\infty(\lambda+\psi) - \frac{(f^l+f^h-2x_0)x_0^2}{\sigma_\delta^2} \\ \frac{2f^h x_0^2 f^l + (f^h)^2(f^l)^3 + 9x_0^4 - 6(f^l+f^h)x_0^3}{\sigma_\delta^2} \end{pmatrix}$$

and, using that $f_\infty \approx 0$, the matrix $\Omega$ becomes

$$\Omega = \begin{pmatrix} -\kappa & 0 & 0 \\ \frac{(f^h-x_0)(x_0-f^l)}{\sigma_\delta^2} & \frac{f^h f^l-(\lambda+\psi)\sigma_\delta^2+(f^l+f^h-3x_0)x_0}{\sigma_\delta^2} & -\frac{f^l+f^h-2x_0}{\sigma_\delta^2} \\ \frac{2(f^h-x_0)x_0(x_0-f^l)}{\sigma_\delta^2} & -\frac{2(f^h(f^l+f^h)f^l+f^h x_0 f^l+9x_0^3-6(f^l+f^h)x_0^2)}{\sigma_\delta^2} & \frac{6f^h f^l+f^{2l}+(f^h)^2+12x_0^2-2(\lambda+\psi)\sigma_\delta^2-10(f^l+f^h)x_0}{\sigma_\delta^2} \end{pmatrix}.$$

Furthermore, using the definition of adjustment speed of Definition (1) and considering that our calibration implies that $f^l \approx -f^l$ and that, for $\epsilon$ small, only terms of order $o(\frac{1}{\sigma_\delta^2})\epsilon$ dominate, we can finally write

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}E_0^A[\widetilde{X}_t]\right|_{t=\epsilon} \approx \begin{pmatrix} \kappa(\overline{f}-x_0) \\ (\lambda+\psi)(f_\infty-x_0) \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{(f^l-x_0)(f^l+x_0)(2x_0(f^l)^2+\overline{f}\kappa\sigma_\delta^2+x_0\sigma_\delta^2(-\kappa+\lambda+\psi)-2x_0^3)}{\sigma_\delta^4} \end{pmatrix}\epsilon, \tag{76}$$

which yields the expressions in Proposition 6.

The nonlinearity of the change of measure introduces an additional term in the approximate expression for Agent $B$'s speed of learning in Equation (76). This term changes changes sign in the neighborhood of $f^l$, $f^m$ and $f^h$ in a way that makes Agent $B$'s expectations centrifugal outward $f^m$ under $\mathbb{P}^A$, as illustrated in Figure 14.

## A.7 Conditional Distribution of the Filter

In this appendix, we compute the distribution of disagreement in our model and that implied by the model specification in David (2008).

We first compute the conditional distribution of Agent $B$'s filter under $\mathbb{P}^A$, $\mathbb{P}^A(\widehat{f}_t^B|\widehat{f}_t^A = y, \widehat{f}_0^B = x_0)$, by solving the Fokker-Planck equation

$$\begin{aligned}\frac{\partial}{\partial t}p(x,y,t) = &-\frac{\partial}{\partial x}\left(p(x,y,t)\left((\lambda+\psi)(f_\infty-x)+\frac{1}{\sigma_\delta^2}(f^h-x)(x-f^l)(y-x)\right)\right) \\ &+\frac{1}{2}\frac{\partial^2}{\partial x^2}\left(p(x,y,t)\frac{1}{\sigma_\delta^2}(f^h-x)^2(x-f^l)^2\right)\end{aligned}$$

with boundary conditions $\lim_{x\nearrow f^h} p(x,y,t;x_0) = \lim_{x\searrow f^l} p(x,y,t;x_0) = 0$ and initial condition $p(x,y,0;x_0) = \widehat{\delta}_{x=x_0}$, where $\widehat{\delta}$ is the Dirac delta function. Then, observing that the

marginal distribution of Agent $A$'s filter, $\mathbb{P}^A(\widehat{f}_t^A|\widehat{f}_0^A = x_0)$ is explicitly given by

$$h(y,t) = \frac{\exp\left(-\frac{1}{2}\left(\frac{\gamma^2}{2\kappa\sigma_\delta^2}\left(1-\exp(-2\kappa t)\right)\right)^{-1}\left(y-\left(x_0\exp(-\kappa t)+\overline{f}(1-\exp(-\kappa t))\right)\right)\right)}{\sqrt{2\pi\frac{\gamma^2}{2\kappa\sigma_\delta^2}\left(1-\exp(-2\kappa t)\right)}},$$

we obtain the marginal density of disagreement, $g$, as

$$f(z,t) = \int_{f^l}^{f^h} p(x, x+z, t)h(x+z)\mathrm{d}x.$$

Second, we compute the distribution of disagreement in the framework of David (2008). In this case, $\widehat{f}^A$ follows a filtered Markov chain and we further need to obtain its marginal distribution by solving the Fokker-Planck equation

$$\frac{\partial}{\partial t}h(y,t) = -\frac{\partial}{\partial y}\left(h(y,t)\kappa(\overline{f}-y)\right) + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(p(y,t)\frac{1}{\sigma_\delta^2}(f^h-y)^2(y-f^l)^2\right)$$

with boundary conditions $\lim_{y\nearrow f^h} p(y,t) = \lim_{y\searrow f^l} p(y,t) = 0$ and initial condition $p(y,0) = \widehat{\delta}_{y=y_0}$. We then obtain the marginal density of $g$ as

$$f(z,t) = \int_{(f^l+z)\vee f^l}^{(f^h+z)\wedge f^h} h(y,t)p(y-z,y,t)\mathrm{d}y.$$

## A.8 Proof of Proposition 7

Following the methodology in Dumas et al. (2009), Agent $A$'s wealth, $V$, satisfies

$$\begin{aligned} V_t &= \delta_t\omega(\eta_t)^\alpha\sum_{j=0}^{\alpha-1}\binom{\alpha-1}{j}\left(\frac{1-\omega(\eta_t)}{\omega(\eta_t)}\right)^j\mathbb{E}_t^{\mathbb{P}^A}\left[\int_t^\infty e^{-\rho(u-t)}\left(\frac{\eta_u}{\eta_t}\right)^{\frac{j}{\alpha}}\left(\frac{\delta_u}{\delta_t}\right)^{1-\alpha}\mathrm{d}u\right] \\ &\overset{(\alpha=2)}{=} \delta_t\omega(\eta_t)^2\left.\frac{S_t}{\delta_t}\right|_{O.U.} + \delta_t\omega(\eta_t)(1-\omega(\eta_t))F\left(\widehat{f}_t^A, g_t\right). \end{aligned} \tag{77}$$

To derive the myopic and hedging components of Agent $A$'s strategy, $Q$, observe Agent $A$'s wealth, $V$ , satisfies the dynamics

$$dV_t = r_t^f V_t\mathrm{d}t + \left(\mu_t - r_t^f\right)Q_tS_t\mathrm{d}t - c_{At}\mathrm{d}t + \sigma_tQ_tS_t\mathrm{d}\widehat{W}_t^A, \tag{78}$$

where $r^f$ is the risk-free rate defined in Equation (8), $\mu - r^f$ is the risk premium on the stock, $Q$ is the number of shares held by Agent $A$, and $\sigma$ is the diffusion of stock returns. Applying Ito's lemma to Agent $A$'s discounted wealth using (77), we obtain the following martingale

$$\begin{aligned}
\mathrm{d}\left(\xi_t V_t + \int_0^t \xi_s c_{As} ds\right) &= \phi_t \mathrm{d}\widehat{W}_t^A \\
&= \mathbb{E}_t^{\mathbb{P}^A}\left(\int_t^\infty \mathcal{D}_t\left(\xi_s c_{As}\right)\mathrm{d}s\right)\mathrm{d}\widehat{W}_t^A \\
&= \left(\xi_t \sigma_t Q_t S_t - V_t \theta_t \xi_t\right)\mathrm{d}\widehat{W}_t^A,
\end{aligned}$$

where the first and second equalities follow from the Martingale Representation Theorem and the Clark-Ocone Theorem, respectively. Matching the diffusion terms in (78) and the expression above, the number of shares $Q$ satisfies

$$Q_t = \frac{\mu_t - r_t^f}{\sigma_t^2}\frac{V_t}{S_t} + \frac{1}{\xi_t \sigma_t S_t}\mathbb{E}_t^{\mathbb{P}^A}\left(\int_t^\infty \mathcal{D}_t\left(\xi_s c_{As}\right)\mathrm{d}s\right). \tag{79}$$

Finally, using the fact that

$$\begin{aligned}
\xi_s c_{As} &= \left(\phi_A e^{\rho s}\right)^{-1/\alpha}\xi_s^{\frac{\alpha-1}{\alpha}} \text{ and thus} \\
\mathcal{D}_t\left(\xi_s c_{As}\right) &= \frac{\alpha-1}{\alpha}\mathcal{D}_t(\xi_s)c_{As},
\end{aligned}$$

we can rewrite Equation (79) as follows

$$\begin{aligned}
Q_t &= \frac{\mu_t - r_t^f}{\sigma_t^2}\frac{V_t}{S_t} + \frac{\alpha-1}{\alpha\sigma_t S_t}\mathbb{E}_t^{\mathbb{P}^A}\left(\int_t^\infty \frac{\xi_s c_{As}}{\xi_t}\frac{\mathcal{D}_t\xi_s}{\xi_s}\mathrm{d}s\right) \\
&= \frac{\mu_t - r_t^f}{\sigma_t^2}\frac{V_t}{S_t} + \frac{\alpha-1}{\alpha\sigma_t S_t}\mathbb{E}_t^{\mathbb{P}^A}\left(\int_t^\infty \frac{\xi_s c_{As}}{\xi_t}\left(\frac{\mathcal{D}_t\xi_s}{\xi_s} - \frac{\mathcal{D}_t\xi_t}{\xi_t} + \frac{\mathcal{D}_t\xi_t}{\xi_t}\right)\mathrm{d}s\right) \\
&= \frac{\mu_t - r_t^f}{\sigma_t^2}\frac{V_t}{S_t} + \frac{\alpha-1}{\alpha\sigma_t S_t}\mathbb{E}_t^{\mathbb{P}^A}\left(\int_t^\infty \frac{\xi_s c_{As}}{\xi_t}\left(\frac{\mathcal{D}_t\xi_s}{\xi_s} - \frac{\mathcal{D}_t\xi_t}{\xi_t}\right)\mathrm{d}s\right) - \frac{(\alpha-1)\theta_t V_t}{\alpha\sigma_t S_t} \\
&= \frac{\mu_t - r_t^f}{\alpha\sigma_t^2}\frac{V_t}{S_t} + \frac{\alpha-1}{\alpha\sigma_t S_t}\mathbb{E}_t^{\mathbb{P}^A}\left(\int_t^\infty \frac{\xi_s c_{As}}{\xi_t}\left(\frac{\mathcal{D}_t\xi_s}{\xi_s} - \frac{\mathcal{D}_t\xi_t}{\xi_t}\right)\mathrm{d}s\right) \\
&\equiv M_t + H_t.
\end{aligned}$$

The expression for the state-price density in (18) is derived in Appendix A.3.

To obtain an explicit expression for the average reaction of the state-price density to a

Brownian shock today

$$\mathbb{E}_t^{\mathbb{P}^A}\left(R(t,s)\right) := \mathbb{E}_t^{\mathbb{P}^A}\left(\frac{\mathcal{D}_t\xi_s}{\xi_t}\right), \tag{80}$$

as given in (20), we decompose the Malliavin derivative of the stochastic discount factor as

$$\mathcal{D}_t\xi_s = -\alpha\frac{\xi_s}{\delta_s}\mathcal{D}_t\delta_s + \frac{\xi_s(1-\omega(\eta_s))}{\eta_s}\mathcal{D}_t\eta_s \tag{81}$$

with

$$\mathcal{D}_t\delta_s = \delta_s\left(\sigma_\delta + \int_t^s \mathcal{D}_t\widehat{f}_v^A \mathrm{d}v\right) \tag{82}$$

$$\mathcal{D}_t\eta_s = -\frac{\eta_s}{\sigma_\delta}\left(g_t + \int_t^s \mathcal{D}_t g_v \mathrm{d}\widehat{W}_v^A + \frac{1}{\sigma_\delta}\int_t^s g_v\mathcal{D}_t g_v \mathrm{d}v\right) \tag{83}$$

$$\mathcal{D}_t\widehat{f}_s^A = \frac{\gamma}{\sigma_\delta}e^{-\kappa(s-t)} \tag{84}$$

and where

$$\mathrm{d}\mathcal{D}_t g_v = \nabla\mu_g(\widehat{f}_v^A, g_v)^\top \begin{pmatrix} \frac{\gamma}{\sigma_\delta}e^{-\kappa(v-t)} \\ \mathcal{D}_t g_v \end{pmatrix}\mathrm{d}v + \nabla\sigma_g(\widehat{f}_v^A, g_v)^\top \begin{pmatrix} \frac{\gamma}{\sigma_\delta}e^{-\kappa(v-t)} \\ \mathcal{D}_t g_v \end{pmatrix}\mathrm{d}\widehat{W}_v^A,$$

with initial condition $\mathcal{D}_t g_t = \sigma_g(\widehat{f}_t^A, g_t)$. The coefficients $\mu_g$ and $\sigma_g$ represent the drift and the diffusion of disagreement in (31), respectively, and the operator $\nabla$ stands for the gradient. Substituting Eq. (84) in Eq. (82) yields

$$\frac{\mathcal{D}_t\delta_s}{\delta_s} = \sigma_\delta + \frac{\gamma}{\kappa\sigma_\delta}\left(1 - e^{-\kappa(s-t)}\right),$$

while substituting Eqs. (82), (83), and (84) in Eq. (81) yields

$$\frac{\mathcal{D}_t\xi_s}{\xi_s} = -\alpha\left(\sigma_\delta + \frac{\gamma}{\kappa\sigma_\delta}\left(1 - e^{-\kappa(s-t)}\right)\right) - \frac{1-\omega(\eta_s)}{\sigma_\delta}\left(g_t + \int_t^s \mathcal{D}_t g_v \mathrm{d}\widehat{W}_v^A + \frac{1}{\sigma_\delta}\int_t^s g_v\mathcal{D}_t g_v \mathrm{d}v\right).$$

Taking conditional expectations at time $t$ and setting $g_t = 0$ implies that (80) satisfies

$$\mathbb{E}_t^{\mathbb{P}^A}\left(R(t,s)\right) = -\alpha\left(\sigma_\delta + \frac{\gamma}{\kappa\sigma_\delta}\left(1 - e^{-\kappa(s-t)}\right)\right) - \frac{1}{\sigma_\delta^2}\mathbb{E}_t^{\mathbb{P}^A}\left((1-\omega(\eta_s))\int_t^s g_v\mathcal{D}_t g_v \mathrm{d}v\right).$$

□

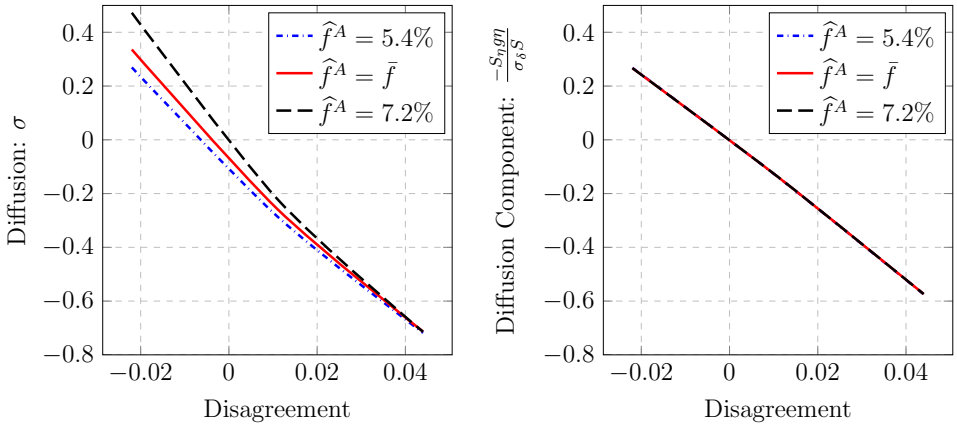


**Figure 15: Stock Return Diffusion vs. Disagreement.**

The left panel depicts the stock return diffusion against disagreement for different values of the fundamental. The right panel depicts the third diffusion component, $\frac{-S_\eta g \eta}{\sigma_\delta S}$, against disagreement for different values of the change fundamental. Note that $\delta = \eta = 1$. The values for the fundamental satisfy: $\widehat{f}^A_{min} = \bar{f} - 2\frac{\gamma}{\sigma_\delta\sqrt{2\kappa}}$ and $\widehat{f}^A_{max} = \bar{f} + 2\frac{\gamma}{\sigma_\delta\sqrt{2\kappa}}$. The range for the disagreement is the 90% confidence interval obtained from 1,000 simulations of the economy over a 100-year horizon.

## A.9 Additional Figures

### A.9.1 Robustness of the relation between the diffusion and disagreement to the fundamental (Section 4.1)

The left panel of Figure 15 depicts the relation between the stock return diffusion and disagreement for different values of the fundamental. The right panel depicts the main driving component, $-\frac{\eta g}{\sigma_\delta S}\frac{\partial}{\partial \eta}S$, of the diffusion. Figure 15 shows that the conclusions of Section 4.1 are robust when varying the fundamental. In particular, the term, $-\frac{\eta g}{\sigma_\delta S}\frac{\partial}{\partial \eta}S$, remains the main driver of the diffusion of stock returns and the diffusion is negatively related to disagreement, irrespective of the value of the fundamental.

### A.9.2 Robustness of the main result to knife-edge cases (Section 4.2)

The dynamics of disagreement are consistent, irrespective of the starting point within each region defined in Table 2, except in two knife-edge cases around $\bar{f}$ and in a close neighborhood of the recession state $f^l$. To see this, notice that we show in Section 3.2.2 that Agent $B$'s expectations are centrifugal outward $f_m$. Agent $B$'s expectations, however, cannot be strictly centrifugal over the entire domain $[f^l, f^h]$, as they would exit the domain otherwise. Hence, in the neighborhood of the recession state, $f^l$, and the expansion state, $f^h$, Agent $B$'s expectations become centripetal to reflect the process inside the domain, as we illustrate in

Figure 14 (Appendix A.6). Figure 14 provides the numerical values of the points at which Agent $B$'s expectations become centripetal. The main consequence of these two centripetal areas is that opinions stop polarizing in bad times when Agent $B$'s expectations are between $-0.056$ and $f^l = -0.0711$, while agents expectations polarize in good times when Agent $B$'s expectations are between $\bar{f} = 0.063$ and 0.067.

To show that our main result is robust within these two intervals, we repeat the analysis of Section 4.2 and plot in Figure 16 the response of future state-price densities in the two knife-edge cases. Comparing the response that prevails in good times (the blue dash-dotted line in the left panel of Figure 5) to the response that prevails in the knife-edge case whereby opinions polarize in good times (the blue dash-dotted line in Figure 16), we observe that in both cases returns adjust immediately to the news (the two responses have the same sign and the same shape). Similarly, comparing the response that prevails in bad times (the black dashed line in the left panel of Figure 5) to the response that prevails in the knife-edge case whereby opinions stop polarizing in bad times (the black dashed line in the left panel of Figure 16), we observe that in both cases returns under-react and then revert.

The reason our results are insensitive to these two knife-edge cases is as follows. While Agent $A$ postulates constant uncertainty throughout the business cycle, Agent $B$'s reassesses uncertainty in a way that greatly varies over the business cycle, as the confidence interval depicted in Figure 1 demonstrates. Specifically, the left panel shows that the variance of Agent $B$'s filter is small in good times. As a result, while opinions polarize in good times in the first knife-edge case, the polarization of beliefs is so small that it does not generate a spike in disagreement. The right panel of Figure 1, instead, shows that the variance of Agent $B$'s filter increases tremendously in bad times. As a result, while opinions stop polarizing in bad times in the second knife-edge case, the variance of Agent $B$'s filter is so large that her expectations almost instantly exit the knife-edge region. Hence, the shape of the impulse response remains unaffected in both cases.

### A.9.3 Coefficient values of time series momentum (Section 4.3)

Figure 17 depicts the time series momentum coefficient $\rho(h)$ for lags $h$ ranging from 1 month to 3 years. Each panel corresponds to a different state of the economy. The $t$-statistics are provided in Section 4.3.

### A.9.4 Unconditional time series momentum pattern

Our analysis of time series momentum in Section 4.3 is conditional on the state of the economy (good, normal and bad times, respectively). Since excess returns are negatively serially correlated at short horizons in good times, the unconditional pattern of serial correlation could be potentially inconsistent with Moskowitz et al. (2012). The left panel of Figure 18 above, which plots the unconditional pattern of time series momentum in our model (com-

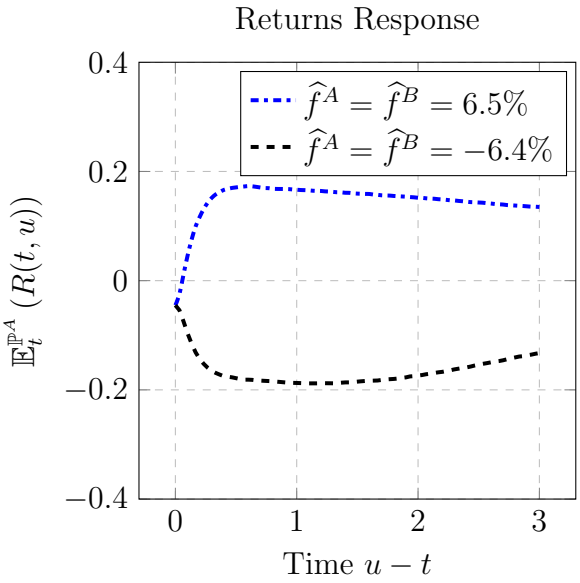


**Figure 16: Impulse Response of Excess Returns to a News Shock in the two Knife-Edge Cases.**

The top knife-edge region is such that $\widehat{f}^A = \widehat{f}^B = 6.5\%$, whereas the bottom knife-edge region is such that $\widehat{f}^A = \widehat{f}^B = -6.4\%$.

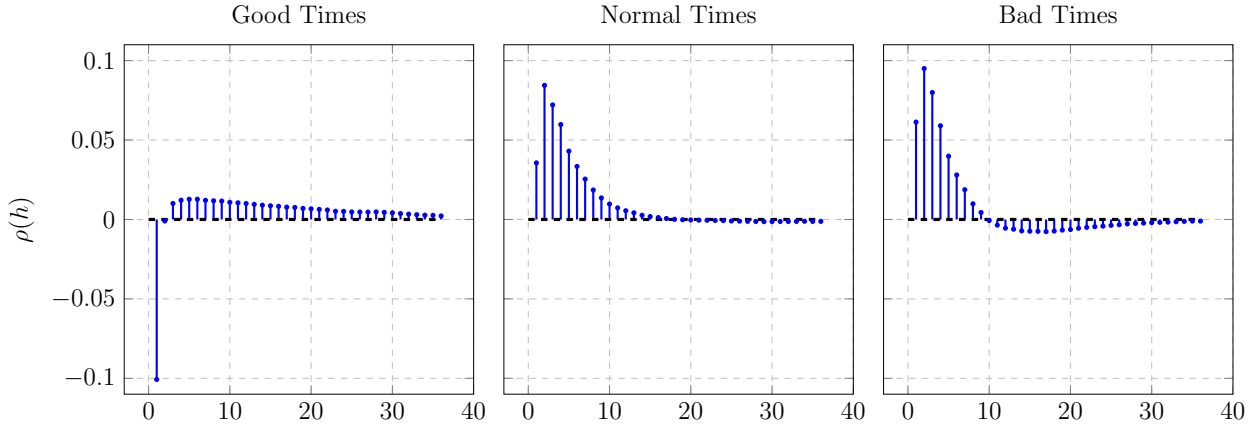


**Figure 17: Time Series Momentum of Excess Returns.**

This figure plots the time series momentum coefficient $\rho(h)$ for lags $h$ ranging from 1 month to 3 years. Each panel corresponds to a different state of the economy. Standard errors are adjusted using Newey and West (1987) procedure. The values reported above are obtained from 10,000 simulations of the economy over a 20-year horizon.

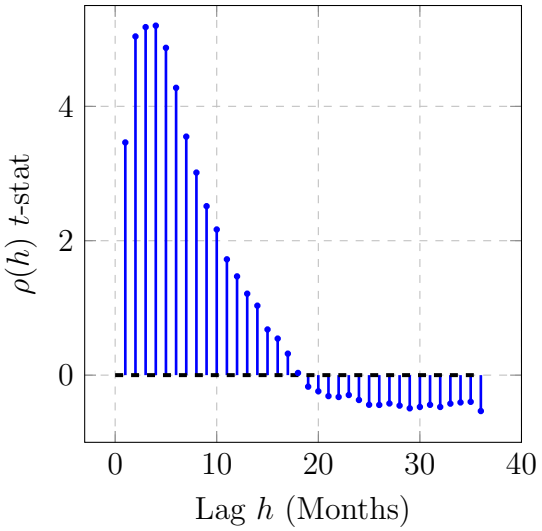


**Figure 18: Unconditional Time Series Momentum of Excess Returns.**
This figure plots the $t$-statistics of the coefficient $\rho(h)$ for lags $h$ ranging from 1 month to 3 years. Standard errors are adjusted using Newey and West (1987) procedure. The values reported above are obtained from 1,000 simulations of the economy over a 100-year horizon.

puted over a 100-year horizon), shows, however, that this is not the case. Specifically, it shows that there is time series momentum up to an 18-month horizon followed by long-term reversal for larger horizons, consistent with the empirical findings of Moskowitz et al. (2012).

### A.9.5 Coefficient values of $\beta_G$ (Section 5.1.1)

Figure 19 depicts the response of excess returns to dispersion $\beta_G(h)$ for lags $h$ ranging from 0 month to 1 year. The left panel controls for the fundamental only, whereas the right panel controls for both the fundamental and the change of measure. The $t$-statistics are provided in Section 5.1.1.

### A.9.6 Empirical pattern of time series momentum in periods of low dispersion (Section 5.2.2)

To verify that we observe, as predicted by the model, short-term time series reversal in low dispersion periods, we run the following regression:

$$r^e_{t+\Delta} = \alpha(p) + \beta_1(p) r^e_t + \beta_2(p) r^e_t Y_t(p) + \epsilon_{t+\Delta},$$

where $\Delta = 1/12 = 1$ month, $r^e$ denotes the monthly excess returns on the S&P 500 and $Y_t(p)$ is a dummy variable that takes value 1 when the monthly dispersion, $Disp$, defined in Section

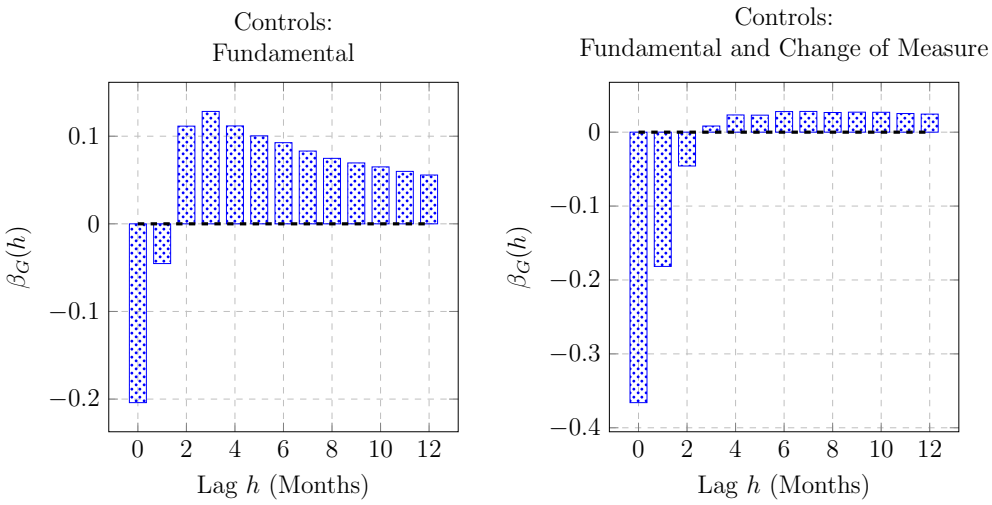


**Figure 19: Excess Returns vs. Dispersion.**

This figure plots the response of excess returns to dispersion $\beta_G(h)$ for lags $h$ ranging from 0 month to 1 year. The left panel controls for the fundamental only (i.e. $\beta_E(h) = 0$), whereas the right panel controls for both the fundamental and the change of measure. Standard errors are adjusted using Newey and West (1987) procedure. The values reported above are obtained from 1,000 simulations of the economy over a 100-year horizon.

5.2, is smaller than its $p-$th percentile. The coefficient $\beta_1(p)$ measures 1-month time series momentum, while the coefficient $\beta_2(p)$ measures additional time series momentum present in low dispersion period; accordingly, the sum $\beta_1(p) + \beta_2(p)$ measures 1-month time series momentum in low dispersion periods. Figure 20 below shows that the data lend support to the prediction of the model: we observe significant 1-month reversal during low dispersion periods.

### A.9.7 Additional Prediction: U-shaped relation between excess returns and time series momentum

In our model, excess returns only become extreme in bad times, when time series momentum is strongest. As a result, our model delivers strongest time series momentum in extreme markets, consistent with Moskowitz et al. (2012). More generally, the U-shaped relation between excess returns and momentum in Moskowitz et al. (2012) is not related to the sign of market returns, but rather to the volatility of market returns: the authors regress the returns on time series momentum against market returns and squared market returns and show that, while the relation between time series momentum and market returns is not significant, the relation between time series momentum and squared market returns is significantly positive. Hence, time series momentum is particularly strong during turbulent

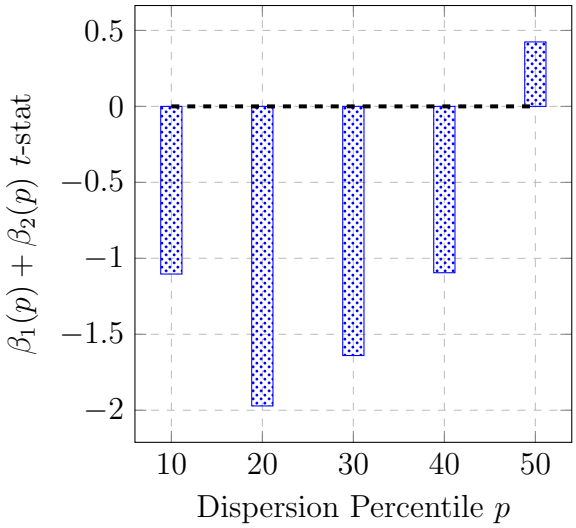


**Figure 20: Time Series Momentum of Excess Returns in Low Disagreement Periods.**

This figure plots the $t$-statistics of 1-month time series momentum $\beta_1(p) + \beta_2(p)$ when dispersion is smaller than its $p$-th percentile. Standard errors are adjusted using Newey and West (1987) procedure. Data are at monthly frequency from 02/1976 to 11/2014.

times (i.e. periods of high volatility).

To demonstrate that our model is consistent with this finding, we run the following regression:

$$r^{TM}_{t+\Delta} = \alpha + \beta_1 \bar{\sigma}_t + \beta_2 \bar{\sigma}^2_t + \epsilon_{t+\Delta},$$

where $\Delta = 1/12 = 1$ month, $\bar{\sigma}_t = \int_{t-\Delta}^{t} \sigma_u du$ is the monthly diffusion of stock returns, and $r^{TM}_{t+\Delta} = sign\left(r^e_t\right) r^e_{t+\Delta}$ is the returns on time series momentum, as defined in Moskowitz et al. (2012). If the sign of excess returns, and hence the sign of the diffusion, does not matter, then the first coefficient, $\beta_1$, should be insignificant. If its magnitude matters, however, the second coefficient, $\beta_2$, should be significantly positive. The $t$-statistics of the intercept, linear coefficient, $\beta_1$, and quadratic coefficient, $\beta_2$, are 0.6693, $-0.7672$, and 1.7402, respectively. Consistent with Moskowitz et al. (2012), only the quadratic coefficient is significant (at the 10% confidence level). As a result, we obtain a U-shaped relation between time series momentum and the diffusion of stock returns, as illustrated in Figure 21. It confirms the intuition that the returns on time series momentum are high during extreme markets.

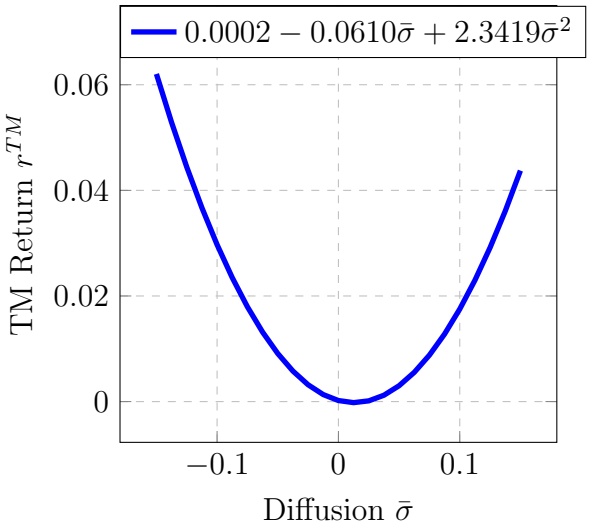


**Figure 21: Time Series Momentum Returns in Extreme Markets.**
This figure plots the relation between time series momentum returns and the stock return diffusion in our model. The relation is obtained from 1,000 simulations of the economy over a 100-year horizon.

### A.9.8 Persistence of time series momentum

To verify that we observe, as predicted by the model, persistent time series momentum in both expansions and recessions, we run the following regression:

$$r^e_{t+h\Delta} = \alpha(h) + \beta_1(h) r^e_t + \beta_2(h) r^e_t Y_t + \epsilon_{t+h\Delta},$$

where $\Delta = 1/12 = 1$ month, $r^e$ denotes the monthly excess returns on the S&P 500 and $Y$ is a dummy variable that takes value 1 in NBER recessions. The first coefficient, $\beta_1(h)$, captures time series momentum in NBER expansions (what we refer to as normal and good times in our model), which we plot in the right panel of Figure 22. Second, the sum of the coefficients, $\beta_1(h) + \beta_2(h)$, captures time series momentum in NBER recessions (what we refer to as bad times in our model). Both panels show that there is on average time series momentum up to the 12-month lag, followed by reversal on average over subsequent horizons.

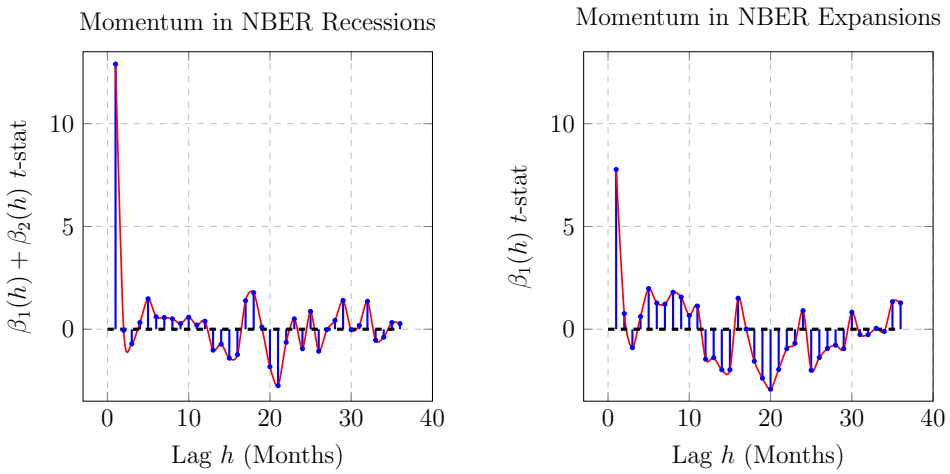


**Figure 22: Time Series Momentum in NBER Expansions and Recessions.** The left and right panels plot the $t$-statistics of the coefficient $\beta_1(h) + \beta_2(h)$ and $\beta_1(h)$, respectively, for lags $h$ ranging from 1 month to 3 years. Standard errors are adjusted using Newey and West (1987) procedure. Monthly data range from 01/1871 to 11/2013.