# Why Does Return Predictability Concentrate in Bad Times?*

Julien Cujean† Michael Hasler‡

May 26, 2016

**Abstract**

We build an equilibrium model to explain why stock return predictability concentrates in bad times. The key feature is that investors assess uncertainty through different models. As economic conditions deteriorate, uncertainty rises and investors' opinions polarize. Disagreement thus spikes in bad times, causing returns to react to past news. This phenomenon creates a positive relation between disagreement and future returns; it also generates time series momentum, which strengthens in bad times, increases with disagreement and crashes after sharp market rebounds. 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 Francesco d'Acunto, Pat Akey, Maria Cecilia Bustamante, Wen Chen, Hui Chen, Peter Christoffersen, Pierre Collin-Dufresne, Michel Dubois, Darrell Duffie, Bernard Dumas, Laurent Frésard, Steve Heston, Julien Hugonnier, Alexandre Jeanneret, Yoontae Jeon, Scott Joslin, Andrew Karolyi, Leonid Kogan, Alex Kostakis, Jan-Peter Kulak, Pete Kyle, Jeongmin Lee, Danmo Li, 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 in recessions the content of news (Tetlock, 2007) better predicts future returns. Yet, the reason return predictability varies over the business cycle remains unclear.

We provide a theoretical mechanism that explains why stock return predictability concentrates in bad times, a mechanism we validate empirically. We base our explanation on the empirical fact that disagreement among forecasters exhibits *counter-cyclical spikes*, a consequence of their use of different models to build their forecasts (Patton and Timmermann, 2010).[2] The main idea of our paper is that investors forecast fundamentals and assess uncertainty using different models; they thus interpret the same news differently depending on economic conditions. An investor who revises her assessment of uncertainty throughout the business cycle concludes that it is counter-cyclical (Veronesi, 1999) and thus gives little weight to news in bad times. As a result, after a long streak of bad news, good news is never "good enough" for this investor to abandon her pessimistic views. In contrast, an investor who assumes that uncertainty does not vary always interprets good news favorably. The opinions of these two investors thus polarize in bad times, causing disagreement to spike and returns to react to past news. Return predictability therefore concentrates in bad times.

We develop a dynamic general equilibrium model populated with two agents, $A$ and $B$, who differ in terms of their assessment of uncertainty. Agents trade a stock and a riskless bond and consume the dividends the stock pays out. They do not observe the expected growth rate of dividends, which we call the fundamental, and use different models for the empirical process that governs it. Agent $A$ assumes that uncertainty does not vary and uses a continuous-state model whereby the fundamental oscillates through good and bad times. In contrast, Agent $B$ uses a discrete-state model whereby the fundamental alternates between a high and a low state. Unlike Agent $A$, she must continuously reassess the uncertainty of

[1] Rapach, Strauss, and Zhou (2010), Henkel, Martin, and Nardari (2011), Dangl and Halling (2012), and Piatti and Trojani (2015) 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 Carlin, Longstaff, and Matoba (2014) and Barinov (2014) show that it is counter-cyclical.

each discrete state as information flows continuously from dividends.

Because agents assess uncertainty differently, they revise their expectations at different speeds and in different directions depending on the state of the economy.[3] A necessary condition for this effect to generate spikes in disagreement is that the difference in agents' assessment of uncertainty be sufficiently large. In particular, dividend volatility must be low so that Agent $A$ faces low uncertainty, while the distance between the high and the low state of Agent $B$'s model must be large so that Agent $B$ faces high overall uncertainty. Agent $B$'s views then become self-reinforcing relative to Agent $A$'s—she over-reacts to good news after a long streak of good news and under-reacts to good news after a long streak of bad news. As a result, good news in good states causes Agent $B$ to revise her expectations upwards significantly faster than Agent $A$. In contrast, good news in bad states always corroborates Agent $A$'s view that the economy is recovering, but is never "good enough" to convince Agent $B$ that economic conditions are improving. Both the difference of adjustment speeds in good states and the polarization of beliefs in bad states generate disagreement spikes.

For disagreement spikes to be counter-cyclical, Agent $B$ must face counter-cyclical uncertainty. Good states must be more persistent than bad states so that the high and the low state of Agent $B$'s model are close to and far from the mean of Agent $A$'s model, respectively. This asymmetry accentuates the polarization of beliefs in bad states and dampens the difference in adjustment speeds in good states. As a result, disagreement exhibits little variations in good times as both agents face low uncertainty and adjust their beliefs alongside. In normal times uncertainty rises and agents' expectations adjust at different speeds, which exacerbates disagreement. In bad times agents' opinions polarize, causing disagreement to spike. Finally, these counter-cyclical spikes in disagreement must arise through an economically plausible calibration, not through an arbitrary choice of parameters. We estimate both agents' models and show that the data directly implies this pattern of disagreement.

In our model counter-cyclical spikes in disagreement lead return predictability to concentrate in bad times. Because variations in disagreement—positive or negative—command a risk premium, contemporaneous returns decrease with the *square* of disagreement. The sign of disagreement only determines the direction in which returns move in the future. News in bad times polarizes opinions—Agent $B$ under-reacts relative to Agent $A$—inducing future

[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.

returns to move opposite to news and to persistently under-react. News in normal times precipitates a revision in Agent $B$'s expectations—she over-reacts relative to Agent $A$—causing future returns to persistently over-react. In good times beliefs move alongside and returns adjust immediately. Since disagreement spikes raise the risk premium persistently, both under- and over-reaction create positive serial correlation in returns at short horizons, a phenomenon known as "time series momentum" (Moskowitz, Ooi, and Pedersen, 2012). That disagreement spikes are counter-cyclical makes this phenomenon stronger in bad times.

The model can explain several features associated with time series momentum. Excess returns exhibit momentum over a 1-year horizon and then revert over subsequent horizons. This later phase of reversal arises through the long-run behavior of agents' consumption shares, which adjust to gradually dampen the effect of short-term disagreement spikes. Moreover, the serial correlation of returns has a hump-shaped term structure: momentum is strong at short horizons and decays at intermediate horizons. 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 predictions of the model. We construct an empirical proxy for the square of disagreement using the dispersion of analysts' forecasts. Based on this proxy the model predicts that *(1)* future excess returns are positively related to dispersion, *(2)* time series momentum at short horizons increases with dispersion and *(3)* is, therefore, strongest in bad times. We test these predictions and find that dispersion positively predicts future excess returns on the S&P 500 and that time series momentum increases significantly with dispersion at short horizons. We show that, over the last century, time series momentum at a 1-month lag is significantly stronger during NBER recessions.

This paper contributes to the vast literature on heterogeneous beliefs. While we borrow the methodologies of David (2008) and Dumas, Kurshev, and Uppal (2009), we introduce a new form of disagreement. Investors in this paper disagree because they assess uncertainty through different models, an Ornstein-Uhlenbeck process (Agent $A$) and a 2-state Markov chain (Agent $B$).[4] The resulting pattern of disagreement exhibits counter-cyclical spikes, a

[4]Agent $A$'s model is based on Detemple (1986, 1991), Brennan and Xia (2001), Scheinkman and Xiong (2003) and Dumas et al. (2009), while Agent $B$'s is based on David (1997, 2008) and Veronesi (1999, 2000). The pattern of disagreement differs whether investors are overconfident (e.g., Scheinkman and Xiong (2003)), whether they have different initial priors (e.g., Detemple and Murthy (1994)), or whether they have dogmatic beliefs (e.g., Kogan, Ross, Wang, and Westerfield (2006)). See Xiong (2014) for a literature review.

pattern that does not arise when both agents use one of the two models with different, but economically plausible parameters. Specifically, when both agents use Agent $A$'s model (e.g., Buraschi and Whelan (2013), Ehling, Gallmeyer, Heyerdahl-Larsen, and Illeditsch (2013), Buraschi, Trojani, and Vedolin (2014)), their assessment of uncertainty does not vary, which makes it challenging to generate spikes in disagreement and thus momentum. When both agents use Agent $B$'s model (David, 2008), there exists combinations of parameters that give rise to counter-cyclical disagreement spikes, as in our model. However, the parameter values estimated in David (2008) imply that disagreement never spikes and returns do not exhibit momentum. Intuitively, fitting the same model to the same data cannot produce sufficient heterogeneity in parameters across agents, and thus sufficiently different assessments of uncertainty, to generate momentum.

This paper is also related to the literature that studies the link between heterogeneous beliefs and momentum.[5] In a rational-expectations model, Banerjee, Kaniel, and Kremer (2009) show that heterogeneous beliefs generate price drift in the presence of higher-order differences of opinions. In contrast, we obtain momentum, even when heterogeneous prior beliefs are commonly known. Importantly, Ottaviani and Sorensen (2015) also obtain short-term momentum and long-term reversal through the combination of heterogeneous beliefs and wealth effects. Both in this paper and theirs, prices under-react because someone—Agent $B$ in this paper and the marginal trader in theirs—reacts opposite to news. However, the mechanism in Ottaviani and Sorensen (2015) does not explain why this effect concentrates in bad times, the main focus of our paper. In contrast, prices in our model under-react in bad times exclusively, while they over-react in normal times and counter-cyclical spikes in disagreement make the former effect stronger.[6]

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 on return predictability; Section 5 tests the predictions of the model; and Section 6 concludes. Derivations and computational details are relegated to the Internet Appendix A.

[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), Albuquerque and Miao (2014), and Andrei and Cujean (2014).

[6]Similarly, Hong and Stein (1999) generate price under-reaction, while (Daniel, Hirshleifer, and Subrahmanyam, 1998) generate price over-reaction. Barberis, Shleifer, and Vishny (1998) provide conditions to enforce under- or over-reaction, or both.

# 2 The Model

We develop a dynamic general equilibrium in which investors use different models to estimate the business cycle. In this section, we describe the economy, we solve the learning and optimization problems of investors, and characterize 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.

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 using 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, the uncertainty of which does not vary. 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 $f^A$ evolves continuously over the business

cycle, reverting to a long-term mean $\overline{f}$ at speed $\kappa$ with constant uncertainty $\sigma_f$.

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 $\psi > 0$.

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. Agent $A$'s model serves a canonical model of dividend growth in the heterogeneous beliefs literature (e.g., Scheinkman and Xiong (2003) and Dumas et al. (2009)). The perspective of Agent $B$ is based on the work of David (1997) and Veronesi (1999) and is closer to models used in the economics literature to forecast business cycle turning points.[7] These models, however, have always been considered separately. When considered jointly these models lead to counter-cyclical spikes in disagreement among agents, the key feature of this framework.

## 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*

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

[7] See Hamilton (1994) and Milas, Rothman, and van Dijk (2006) for further details.

*where* $\gamma = \sqrt{\sigma_\delta^2\left(\sigma_\delta^2\kappa^2 + \sigma_f^2\right)} - \kappa\sigma_\delta^2$ *denotes Agent A's 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,* $\widehat{f}_t^B = \mathbb{E}_t^{\mathbb{P}^B}\left[f_t^B\right]$, *of Agent B evolves according to the dynamics*

$$d\widehat{f}_t^B = (\lambda + \psi)\left(f_\infty - \widehat{f}_t^B\right)dt + v(\widehat{f}_t^B)d\widehat{W}_t^B, \tag{5}$$

*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, where* $v(\widehat{f}^B) = \frac{1}{\sigma_\delta}\left(\widehat{f}^B - f^l\right)\left(f^h - \widehat{f}^B\right)$ *defines Agent B's posterior variance and where* $\widehat{W}^B$ *is a Brownian motion under Agent B's probability measure* $\mathbb{P}^B$.

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

Importantly, the two filters in Equations (4) and (5) differ in terms of uncertainty. While Agent $A$'s uncertainty $\gamma/\sigma_\delta$ does not fluctuate, Agent $B$'s uncertainty $v(\cdot)$ does, reflecting her effort to continuously reassess the likelihood of each state using dividends that flow continuously. We analyze how this difference affects agents' disagreement in Section 3.

We choose to work under Agent $A$'s probability measure, a choice we clarify in Section 3.2. 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 Equation (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, $\epsilon_t$:

$$f_t = C(f_t^A, f_t^B, \epsilon_t). \tag{7}$$

In Section 3.2, we 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 standard martingale approach (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 under $\mathbb{P}^A$, $\theta$, and the risk-free rate, $r^f$, satisfy*

$$\theta_t = \alpha\sigma_\delta + \frac{(1-\omega_t)}{\sigma_\delta}g_t \tag{8}$$

$$r_t^f = \rho + \alpha\widehat{f}_t^A - \frac{1}{2}\alpha(\alpha+1)\sigma_\delta^2 + (1-\omega_t)\,g_t\left(\frac{1}{2}\frac{\alpha-1}{\alpha\sigma_\delta^2}\omega_t g_t - \alpha\right), \tag{9}$$

*where*

$$\omega_t = \frac{(1/\phi_A)^{1/\alpha}}{(1/\phi_A)^{1/\alpha} + (\eta_t/\phi_B)^{1/\alpha}} \tag{10}$$

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

**Proof**. See Appendix A.3. ■

We focus our discussion on the market price of risk in Equation (8), which is a key determinant of return predictability. The risk-free rate in Equation (9) is discussed in details in David (2008) and Buraschi and Whelan (2013). The market price of risk is the product of the diffusion of Agent $A$'s consumption growth

$$\sqrt{\frac{1}{\mathrm{d}t}\mathrm{var}_t^{\mathbb{P}^A}\left(\frac{\mathrm{d}c_{At}}{c_{At}}\right)} = \sigma_\delta + \frac{1-\omega_t}{\alpha\sigma_\delta}g_t \tag{11}$$

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. 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 (10). 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.

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, there exists an equilibrium in which the stock price satisfies*

$$\frac{S_t}{\delta_t} = \omega_t^\alpha \frac{S_t}{\delta_t}\bigg|_{O.U.} + (1-\omega_t)^\alpha \frac{S_t}{\delta_t}\bigg|_{M.C.} + \omega_t^\alpha \sum_{j=1}^{\alpha-1} \begin{pmatrix} \alpha \\ j \end{pmatrix} \left(\frac{1-\omega_t}{\omega_t}\right)^j F^j\left(\widehat{f}_t^A, g_t\right), \tag{12}$$

*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 disagreement. Under the calibration of Section 3.2, this equilibrium is unique.*

**Proof**. See Proposition 3 in David (2008) for existence and Appendix A.4. ■

The price in Equation (12) 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. It describes the joint effect of fundamental and disagreement on the price.

# 3 Dynamics of Disagreement and Calibration

In this section, we analyze the dynamics of disagreement that arise when one agent uses a discrete-state model and the other a continuous-state model. In Section 3.1, we provide restrictions on model parameters under which this difference of models creates counter-cyclical spikes in disagreement, the key feature driving our results. We then calibrate both models and let the data decide whether these parameter restrictions are satisfied (Section 3.2). We illustrate the resulting pattern of disagreement in Section 3.3: agents' expectations polarize in bad times, adjust at different speeds in normal times and move alongside in good times, precisely leading to counter-cyclical spikes in disagreement.

## 3.1 A Mechanism for Counter-Cyclical Spikes in Disagreement

In our framework, agents use different models and thus learn differently, as apparent from Equations (4) and (5). First, Agent $B$'s discrete-state model implies state-dependent uncertainty $v(\widehat{f}^B)$, whereas Agent $A$'s continuous-state model implies constant uncertainty $\gamma/\sigma_\delta$. We show that this difference in agents' assessment of uncertainty causes Agent $A$ to perceive Agent $B$'s views as self-reinforcing (centrifugal), a mechanism that can lead disagreement among agents to spike. Second, agents' views evolve at different speeds ($\lambda+\psi$ and $\kappa$) and have different long-term means ($f_\infty$ and $\overline{f}$). This difference in persistence and long-term mean has an asymmetric effect on disagreement, which can lead disagreement spikes to concentrate in bad states of the economy.

To guide the analysis, we introduce the concepts 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 $sign(\Sigma_t^A) \neq sign(\Sigma_t^B)$.*

The change of measure of Proposition 2 modifies the speed at which Agent $B$ learns. Expressing Agent $B$'s expectations under $\mathbb{P}^A$ introduces a nonlinear term in Agent $B$'s adjustment speed through the state-dependent uncertainty $v(\widehat{f}^B)$ of her filter:

$$\Sigma_t^B = \mathbb{E}_0^{\mathbb{P}^A}\left[(\lambda+\psi)\left(f_\infty-\widehat{f}_t^B\right)+\underbrace{\frac{1}{\sigma_\delta}\left(\widehat{f}_t^A-\widehat{f}_t^B\right)v(\widehat{f}_t^B)}_{\text{change of measure}}\right]. \tag{13}$$

To understand how the change of measure affects the speed at which Agent $B$ learns, we make two simplifying assumptions, which are consistent with our calibration in Section 3.2.

**Assumption 1.** *Agent $B$'s model has symmetric states $f^h=-f^l\equiv f$ around 0.*

**Assumption 2.** *The volatility of dividends $\sigma_\delta$ is such that terms of order $o(\sigma_\delta^4)$ are negligible.*

We then perform a second-order approximation of the change of measure in (13) and denote by $\widetilde{\Sigma}_t^B$ the resulting adjustment speed for Agent $B$, which we present in Proposition 5.

**Proposition 5.** *Under Assumptions 1 and 2, the second-order approximation of Agent $B$'s adjustment speed in a neighborhood of $t \approx 0$ satisfies*

$$\widetilde{\Sigma}_t^B = \underbrace{(1-\pi(x_0,t))\overbrace{(\lambda+\psi)(f_\infty - x_0)}^{\equiv \Sigma_t^B \text{ under } \mathbb{P}^B} + \pi(x_0,t)\overbrace{\kappa(\overline{f}-x_0)}^{\equiv \Sigma_t^A}}_{\text{weighted adjustment speed}} + \underbrace{2\frac{\pi(x_0,t)^2}{t}x_0}_{\text{centrifugal effect} \equiv C(x_0)} \tag{14}$$

*where $\pi(x_0,t) = v(x_0)t/\sigma_\delta$.*

**Proof**. See Appendix A.5. ■

The nonlinear change of measure in (13) has two effects on Agent $B$'s adjustment speed. First, it adjusts the dynamics of Agent $B$'s expectations to make them consistent with those of Agent $A$. Specifically, the first term in (14) is a weighted average of agents' adjustment speeds under their respective probability. Agent $B$'s uncertainty drives the weight $\pi(x_0,t)$, which reaches its maximum, $\pi(f_m,t) = (f/\sigma_\delta)^2 t$, when Agent $B$'s filter is equal to

$$f_m = \frac{1}{2}f^l + \frac{1}{2}f^h (\equiv 0 \text{ under Assumption 1}),$$

the point at which Agent $B$ is most uncertain about the state of the economy. For ease of interpretation, assume that $t = (\sigma_\delta/f)^2$ so that $\pi(x_0,(\sigma_\delta/f)^2) \in [0,1]$ in all states. It follows that the first term in Equation (14) causes agents' disagreement to increase when Agent $B$ is confident that the economy is in a bad or a good state.

Second, the change of measure causes Agent $B$'s expectations to become self-reinforcing (centrifugal) outward $f_m$ through the second term, $C(x_0)$, in Equation (14). To illustrate this effect, we plot this term as a function of the initial value of Agent $B$'s filter in Figure 1. As Agent $B$'s expectations rise above $f_m$, they increase at an accelerated rate, which generates a *difference in adjustment speeds* across the two agents in good states. As Agent $B$'s expectations drop below $f_m$, they decrease at an accelerated rate, which creates a *polarization of beliefs* across the two agents in bad states. In other words, good news in good states causes Agent $B$ to revise her expectations upwards significantly faster than Agent $A$. In contrast,

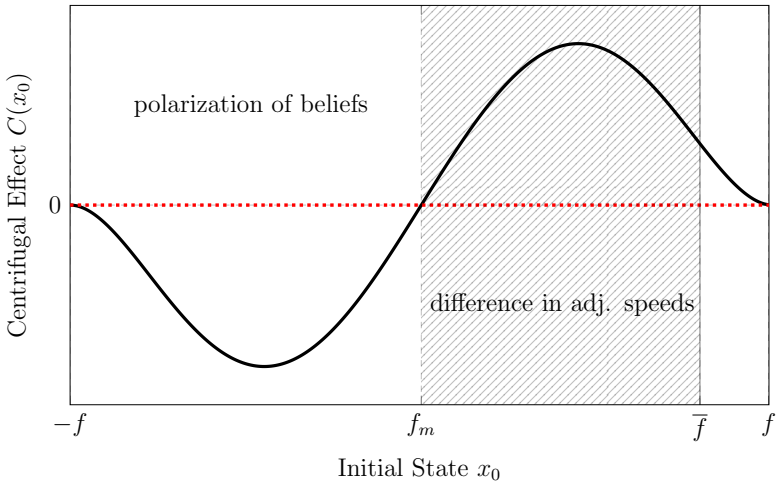


**Figure 1: Centrifugal Effect on the Speed at which Agent $B$ Learns.**
This figure illustrates the centrifugal effect in Agent $B$'s adjustment speed (see Proposition 5) as a function of the initial value of Agent $B$'s filter.

good news in bad states always corroborates Agent $A$'s view that the economy is recovering, but is never "good enough" to convince Agent $B$ that economic conditions are improving.

By amplifying disagreement in good and bad states, the centrifugal effect can produce *spikes* in short-term disagreement. Specifically, the maximum magnitude of this effect is

$$\max_{x_0\in[-f,f]} |C(x_0)| = \frac{8}{5^{3/2}}\frac{f}{\sigma_\delta}v(f/\sqrt{5}), \text{ around time } t \approx \left(\frac{\sigma_\delta}{f}\right)^2.$$

It follows that the larger the ratio of the distance between Agent $B$'s states to the volatility of dividends is, the larger and the sharper changes in disagreement are. For instance, a ratio of $f/\sigma_\delta \equiv 10$ implies that disagreement can change by up to sevenfold Agent $B$'s uncertainty, $v$, in a matter of days ($t \approx 0.01$). The economic relevance of such sudden and large changes in disagreement is that they can generate stock return momentum in our model, a key result we present in Section 4. In that respect, we reserve the terminology "disagreement spikes" for changes in disagreement that are sufficiently large and sharp to produce momentum. Since the ratio $f/\sigma_\delta$ determines the magnitude of changes in disagreement, the existence of disagreement spikes in the model places an implicit restriction on this ratio.

**Condition 1.** Under Assumptions 1 and 2, a necessary condition for disagreement spikes to arise in the model is that the ratio $f/\sigma_\delta$ be sufficiently large.

Importantly, without state-dependent uncertainty the centrifugal effect in Equation (14) is absent. For instance, consider the specification under which both agents use Agent $A$'s linear continuous-state model (e.g., Buraschi and Whelan (2013), Ehling et al. (2013) and Buraschi et al. (2014)). In this case the change of measure in Equation (13) only has a linear effect on Agent $B$'s adjustment speed, making it challenging to generate sharp increases in disagreement. A discrete-state model instead implies state-dependent uncertainty, which can lead disagreement to spike.

Since the centrifugal effect acts symmetrically on the magnitude of disagreement spikes across good and bad states, any asymmetry—pro-cyclicality or counter-cyclicality—must arise through the weighted average in Equation (14). Notice first that the speed at which disagreement increases is the difference between agents' adjustment speeds, which satisfies

$$\Sigma_t^A - \widetilde{\Sigma}_t^B = (1 - \pi(x_0, t)) \left( \underbrace{(\lambda - \psi) f + \kappa \overline{f}}_{\text{asymmetric effect}} - (\kappa - \psi - \lambda) x_0 \right) - C(x_0). \tag{15}$$

Given that the centrifugal effect makes disagreement spikes positive in bad states and negative in good states, a necessary condition for disagreement spikes to be stronger in bad states is that the asymmetric effect in Equation (15) be positive. When this effect is positive, it accentuates the polarization of beliefs in bad states and dampens the difference in adjustment speeds in good states. This condition for counter-cyclical spikes imposes a restriction on the persistence of good states relative to bad states, which we provide in Condition 2.

**Condition 2.** Under Assumptions 1 and 2, a necessary condition for disagreement spikes to be counter-cyclical is that $\psi - \lambda < \kappa \frac{\overline{f}}{f}$.

Conditions 1 and 2 jointly define parameter restrictions under which disagreement exhibits counter-cyclical spikes in our framework. The mechanism for disagreement spikes we highlight is not unique to our framework. Specifically, a similar mechanism applies when both agents use Agent $B$'s model with different parameters (David, 2008). Under both model specifications, there exists combinations of parameters that give rise to disagreement spikes and counter-cyclicality. However, these two effects must arise through an economically plausible mechanism, not through an arbitrary choice of parameters. We therefore let the data decide whether the parameter restrictions of Conditions 1 and 2 are satisfied and whether these estimated effects distinctly identify our specification, a matter we now investigate.

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

In our model, agents use a single source of information—the time series of dividends—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.6.

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 persistent. Second, Agent $B$ finds that the high state $f^h$ and the low state $f^l$ are symmetric around 0, consistent with Assumption 1, but that transition intensities between the two states are asymmetric—expansions are more persistent than recessions ($\psi > \lambda$). The value of the transition intensities $\psi$ and $\lambda$ further imply that Agent $B$'s filter reverts about 3.5 times faster than that of Agent $A$ and that the long-term mean of Agent $B$'s filter ($f_\infty \approx 0.014$) is significantly lower than that of Agent $A$ ($\overline{f} = 0.063$).

This calibration allows us to make a statement about the true data-generating process in (7). Specifically, we can determine which model better fits historical data by applying a model selection method such as the *Akaike* Information Criterion (AIC).[9] Because an Ornstein-Uhlenbeck process is more versatile than a 2-state Markov chain, the information criterion favors Agent $A$'s model.[10] Hence, over the last century, Agent $A$'s probability measure was closer to the physical probability measure *in-sample*. This fact ultimately

[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.6.

[9] The Akaike Information Criterion is defined as $AIC = 2K - 2\log(L)$, where $K$ is the number of parameters estimated and $L$ the likelihood function. The smaller the criterion is, the better the model does.

[10] The AIC for Agent $A$ and $B$'s models are $AIC_{\text{Agent A}} = -1.7229 \times 10^4$ and $AIC_{\text{Agent B}} = -1.2154 \times 10^4$.

| 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 |
| Lagrange Multipliers | $\phi_A = \phi_B$ | 1 |

**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 three rows report our choice of preference parameters and initial consumption shares ($\phi_A = \phi_B \Leftrightarrow \omega_0 = 1 - \omega_0 = 0.5$).

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*.[11] For agents who process the data in real time, it is difficult to assess the relative fit of their model.

Importantly, this calibration satisfies Conditions 1 and 2, thus allowing for counter-cyclical spikes in disagreement. Notice first that the low volatility of dividends in Table 1 is consistent with Assumption 2 and implies that the ratio of Condition 1 can produce disagreement changes of up to fivefold Agent $B$'s uncertainty within days. We show in Section 4 that these disagreement spikes are sufficiently large to produce momentum in stock returns (Condition 1). To emphasize the particularity of this result, consider the specification under which both agents use Agent $B$'s model with different parameters (David, 2008). The parameter values estimated in David (2008) then imply that disagreement never spikes and

[11] 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.

stock returns do not exhibit momentum under this specification. In other words, fitting the same discrete-state model to the data does not produce sufficient heterogeneity in parameters to generate momentum in the model, hence the relevance of our specification. Second, the relative persistence of good and bad states meets Condition 2. We now show that this condition is sufficient to make disagreement spikes counter-cyclical under our calibration.

## 3.3 Estimated Dynamics of Disagreement

In this section we illustrate how the mechanism of Section 3.1 and the calibration of Section 3.2 combine to produce counter-cyclical spikes in disagreement, consistent with observed patterns among forecasters (e.g., Kandel and Pearson (1995) and Patton and Timmermann (2010)). We first define three regimes of the economy, which we use to describe the different phases of the business cycle throughout the analysis. Based on the results of Section 3.1, the relevant business cycle turning point for Agent $A$ is her long-term mean, $\bar{f}$, while that for Agent $B$ is the point of maximum uncertainty $f_m$; accordingly, 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 $f_m$. Otherwise, we say the economy is in *normal times* when expectations lie between $f_m$ and $\bar{f}$. To emphasize that future disagreement—as opposed to current disagreement—drives our result, we set current disagreement to zero in each case. As a convention, we assume that agents' filters start in the middle of each interval.[12] Table 2 summarizes the definition of the three regimes.

To illustrate the behavior of disagreement over the business cycle, we plot Agent $A$ and $B$'s average filter over time in Figure 2 and highlight each regime in a separate panel. In good times (the left panel), both agents adjust their views at comparable speeds and

[12]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.11.1, we explain why disagreement dynamics differ, but show that our results remain unaffected, within these intervals.

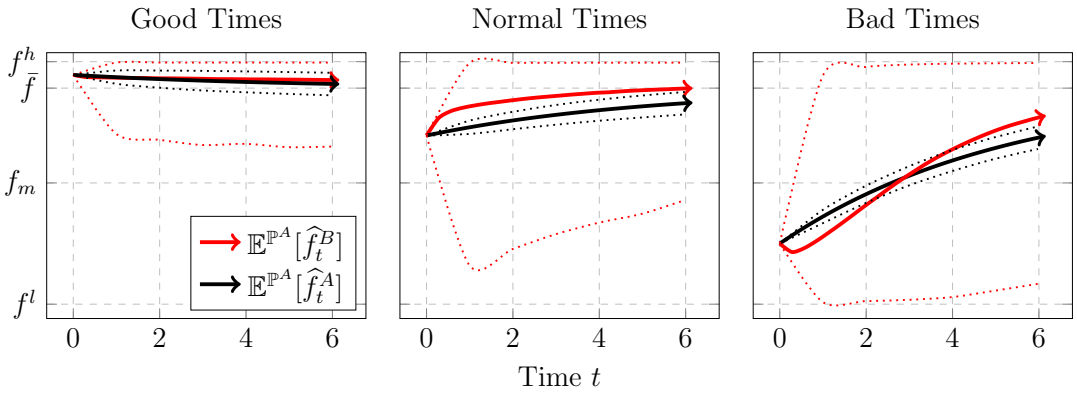


**Figure 2: Filtered Dynamics in 3 States of the Economy.**

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.

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 the centrifugal effect in Equation (14). This difference in adjustment speeds causes disagreement to spike in the short term.

In bad times (the right panel), beliefs polarize in the short term through the centrifugal effect in Equation (14). Good news in bad times always corroborates Agent $A$'s optimistic views, but is never good enough to invalidate Agent $B$'s pessimistic views, reinforcing her beliefs that economic conditions are deteriorating. This polarization of opinions creates disagreement spikes that are strongest in bad times through the asymmetric effect in Equation (15). This asymmetry arises because the high state of Agent $B$'s model is more persistent than its low state. As a result, Agent $B$'s uncertainty increases as economic conditions deteriorate (see the confidence interval in each panel of Figure 2), making the effect of the change of measure in Equation (13) stronger in bad times. In the long run Agent $B$ eventually realizes that the economy is recovering and rapidly catches up with Agent $A$, thus becoming more optimistic than Agent $A$. Disagreement therefore evaporates in the medium term and regenerates in the long term through the difference in adjustment speeds.

# 4 Disagreement driving Stock Return Predictability

In this section we show that counter-cyclical spikes in disagreement cause stock return predictability to concentrate in bad times. In our model fluctuations in disagreement command a risk premium: contemporaneous excess returns decrease with the *square* of disagreement (Section 4.1). The sign of disagreement only determines the direction in which excess returns adjust in the future (Section 4.2). In bad times Agent $B$ reacts opposite to news, beliefs polarize and future excess returns under-react. In normal times Agent $B$ over-reacts to news, beliefs adjust at different speeds and future excess returns over-react. In good times beliefs move alongside, disagreement is nearly constant and future excess returns adjust immediately. Since spikes in disagreement—positive *or* negative—persistently raise the risk premium, both under- and over-reaction create positive serial correlation in excess returns in the short term, a phenomenon known as "time series momentum" (Section 4.3). That disagreement spikes are counter-cyclical makes this phenomenon stronger in bad times.

To understand how disagreement, contemporaneous excess returns, and future excess returns are related, it is insightful to describe trading strategies. Following Dumas et al. (2009), Proposition 6 decomposes Agent $A$'s strategy into two components.

**Proposition 6.** *The number of shares, $Q$, that Agent $A$ holds can be decomposed into a myopic portfolio, $M$, and a hedging portfolio, $H$, according to*

$$Q_t = M_t + H_t = \frac{V_t}{\alpha\sigma_t S_t}\theta_t + \frac{\alpha-1}{\alpha\sigma_t S_t} E_t^{\mathbb{P}^A}\left[\int_t^\infty \frac{\xi_s}{\xi_t} c_{As}\left(\frac{\mathcal{D}_t\xi_s}{\xi_s} - \frac{\mathcal{D}_t\xi_t}{\xi_t}\right) ds\right] \tag{16}$$

*where $\xi$ denotes Agent $A$'s state-price density, which is given by*

$$\xi_t = e^{-\rho t}\delta_t^{-\alpha}\left[(1/\phi_A)^{1/\alpha} + (\eta_t/\phi_B)^{1/\alpha}\right]^\alpha \tag{17}$$

*and where $V$ denotes Agent $A$'s wealth, which we provide in Appendix A.7.*

**Proof**. See Appendix A.7. ■

Agent $A$'s portfolio in (16) 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, $\theta$, the second term, $H$, is a hedging demand through which she exploits return

predictability. To see this, notice that the hedging demand in (16) incorporates Agent $A$'s outlook on future returns through the response of future 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 now analyze how disagreement affects both contemporaneous and future excess returns and then derive implications for the serial correlation of excess returns.

## 4.1 Disagreement and Contemporaneous Excess Returns

Our goal is to determine the relation between contemporaneous excess returns and the state variables of the model. A key result of this section is that contemporaneous excess returns decrease with the square of disagreement. In the eyes of Agent $A$ contemporaneous (expected) excess returns, $\mu - r^f = \sigma\theta$, are the product of the market price of risk, $\theta$, which we discussed in Section 2.3, and the diffusion of stock returns, $\sigma$:

$$\sigma_t = \overbrace{\sigma_\delta + \frac{\gamma}{\sigma_\delta}\frac{1}{S_t}\left(\underbrace{\frac{\partial S}{\partial \widehat{f}^A}}_{<0} + \underbrace{\frac{\partial S}{\partial g}}_{>0}\right)}^{\text{component 1: } A\text{'s uncertainty}} \overbrace{-(1-\omega_t)v\left(\widehat{f}_t^A - g_t\right)\frac{1}{S_t}\underbrace{\frac{\partial S}{\partial \widehat{g}}}_{>0}}^{\text{component 2: } B\text{'s uncertainty}} + \overbrace{\frac{g_t(1-\omega_t)\omega_t}{\alpha\sigma_\delta^2}\frac{1}{S_t}\underbrace{\frac{\partial S}{\partial \omega}}_{<0}}^{\text{component 3: } A\text{'s cons. share}} \quad (18)$$

where $\widehat{g} \equiv (1-\omega)g$ denotes the "consumption-weighted" disagreement (see Appendix A.8).

We first explain the sign of the price sensitivities in Equation (18). The price decreases with the fundamental through the income effect (when risk aversion is larger than one (Veronesi, 2000)): anticipating that an increase in the fundamental 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 Agent $A$'s consumption share implies a decrease in the risk of Agent $A$'s consumption growth through (11), to which she responds by decreasing her hedging demand, leading to a price decrease.

We now determine how the diffusion of stock returns depends on the state variables of the model. The diffusion in Equation (18) has three components, which we plot in Figure 3 in separate panels. Because Agent $A$'s uncertainty is constant, the first diffusion component in the left panels of Figure 3 is nearly constant. In contrast, Agent $B$'s state-dependent uncer-

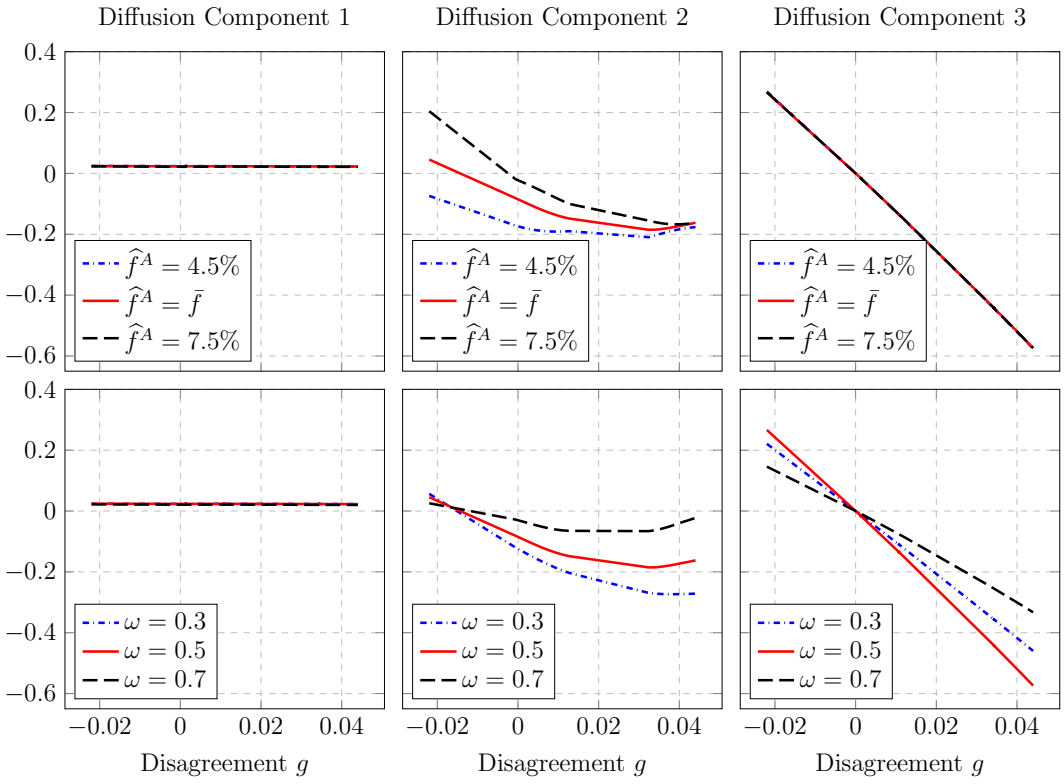


**Figure 3: Decomposition of the Stock Return Diffusion.**

The upper and lower panels show the three constituents of the stock return diffusion against disagreement for different values of the fundamental and Agent $A$'s consumption share, respectively. If not stated otherwise, we set $\delta = 1$, $\widehat{f}^A = \bar{f}$, and $\omega = 0.5$. The chosen range for disagreement corresponds to its 90% confidence interval.

tainty causes the second component (the middle panels) to decrease with disagreement. This decreasing relation steepens either when Agent $A$'s filter increases or when her consumption share decreases. This interaction between Agent $A$'s views and Agent $B$'s uncertainty raises stock return volatility (see, for instance, David (2008)). Furthermore, fluctuations in Agent $A$'s consumption share induce a decreasing relation between the third component and disagreement (the right-hand panels), which steepens as agents' consumption shares equalize. Overall, our discussion implies that the diffusion of stock returns can be approximated as:

$$\sigma_t \approx A_0 + A_1 g_t \widehat{f}_t^A (1 - \omega_t) + A_2 g_t \omega_t (1 - \omega_t), \tag{19}$$

where the sign of the constants $A_0 \approx 0$, $A_1 < 0$, and $A_2 < 0$ follows from price sensitivities.

Using Assumption 2 that $\sigma_\delta$ is small, we finally obtain a relation for excess returns by multiplying the diffusion in Equation (19) with the market price of risk in Equation (8):

$$\mu_t - r_t^f \approx \text{constant}(\approx 0) + \frac{A_1}{\sigma_\delta} \widehat{f}_t^A \left((1-\omega_t)g_t\right)^2 + \frac{A_2}{\sigma_\delta} \omega_t \left((1-\omega_t)g_t\right)^2 . \tag{20}$$

Equation (20) accurately approximates excess returns and will serve as the basis for our empirical specifications in Section 5.[13] This relation has two main implications for excess returns. First, contemporaneous excess returns are hump-shaped in disagreement and decreasing in the square of disagreement. Second, as the right-hand panels of Figure 3 show, the main source of fluctuations in excess returns is the last term in Equation (20).

## 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 beliefs' adjustment speed, causing returns to persistently over-react. In good times, news provokes an immediate return adjustment. Since disagreement spikes are counter-cyclical the content of news (Tetlock, 2007) better predicts future returns in bad times (Garcia, 2013).

Agent $A$'s hedging demand in (16) shows that forecasting future returns in our model involves computing the response of the future state-price density 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 the state-price density, 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 the state-price density" indifferently and denote future returns from time $t$ 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): the average, $\mathbb{E}_t^{\mathbb{P}^A}[R(t,s)]$, is the regression coefficient of news arriving at time $t$ on future stock returns from time $t$ to $s$.

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

[13] Simulations show that regressing excess returns, $\mu - r^f$, on both $\widehat{f}^A\left((1-\omega)g\right)^2$ and $\omega\left((1-\omega)g\right)^2$ yields a $R^2$ of 95%, while regressing it on $\omega\left((1-\omega)g\right)^2$ yields a $R^2$ of 85%. See Appendix A.8 for further details.

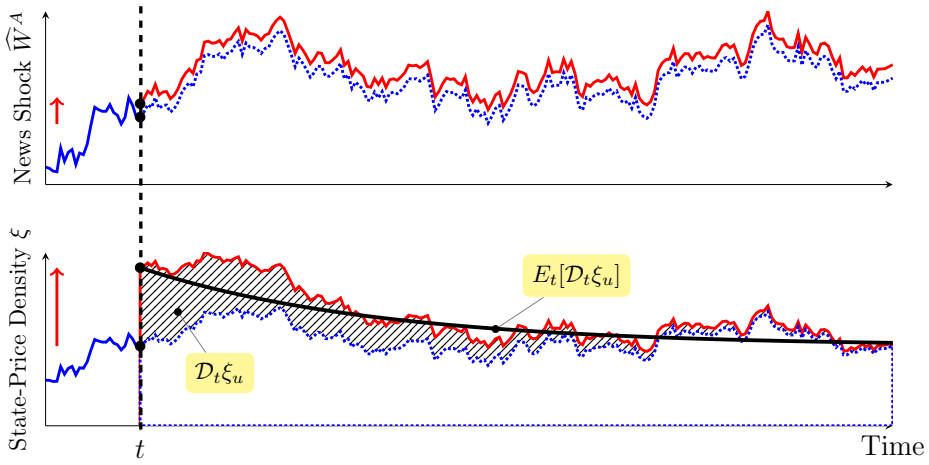


**Figure 4: Illustration of the State-Price Density's Response to a News Shock.** The upper panel shows a trajectory of news before (dashed blue line) and after (solid red line) a news surprise. The lower panel shows the associated trajectory of the state-price density before (dashed blue line) and after (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.

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.[14] 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 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 Equation (17) depends on two state variables, dividends and

[14]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.

the likelihood of Agent $B$'s model. As a result, the way the future state-price density reacts to news depends, first, on how it reacts to a change in these state variables and, second, on how these state variables themselves react to news. Applying chain rule to Equation (17), the average response of the state-price density satisfies (see Appendix A.7 for derivations):

$$\mathbb{E}_t^{\mathbb{P}^A}\left[R(t,s)\right] = \underbrace{-\alpha \times \left(\sigma_\delta + \frac{\gamma}{\kappa\sigma_\delta}\left(1-e^{-\kappa(s-t)}\right)\right)}_{\text{fundamental channel} \equiv \frac{\delta_s}{\xi_s}\frac{\partial}{\partial \delta_s}\xi_s \times \frac{\mathcal{D}_t\delta_s}{\delta_s}} + \underbrace{\mathbb{E}_t^{\mathbb{P}^A}\left[(1-\omega_s)\times\left(-\int_t^s \frac{g_u \mathcal{D}_t g_u}{\sigma_\delta^2}\mathrm{d}u\right)\right]}_{\text{disagreement channel} \equiv \mathbb{E}_t^{\mathbb{P}^A}\left[\frac{\eta_s}{\xi_s}\frac{\partial}{\partial \eta_s}\xi_s \times \frac{\mathcal{D}_t\eta_s}{\eta_s}\right]} \quad (21)$$

We now unravel the chain reaction in Equation (21). As news hits, dividends react in a perfectly predictable fashion because the fundamental mean reverts with constant uncertainty, $\frac{\gamma}{\sigma_\delta}$, under Agent $A$'s measure. The first term in brackets in (21) indicates that dividends respond positively to 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 news on average, as news may either corroborate or invalidate Agent $B$'s model *depending on economic conditions*. When the reaction of disagreement $g \times \mathcal{D}g$ is persistently negative, Agent $B$ is overly optimistic relative to Agent $A$ and good news thus corroborates her model, as the second term in brackets in (21) shows. When the reaction of disagreement is persistently positive, Agent $B$ is pessimistic and good news invalidates her model.

As dividends and the likelihood of Agent $B$'s model move, the state-price density in turn responds. 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, news shocks move the state-price density through two channels: the fundamental and disagreement. The fundamental channel in (21) triggers a negative state-price density reaction following good news, as it hints future dividend abundance. The disagreement channel provokes an ambiguous reaction depending on whether 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 future returns move alongside the average

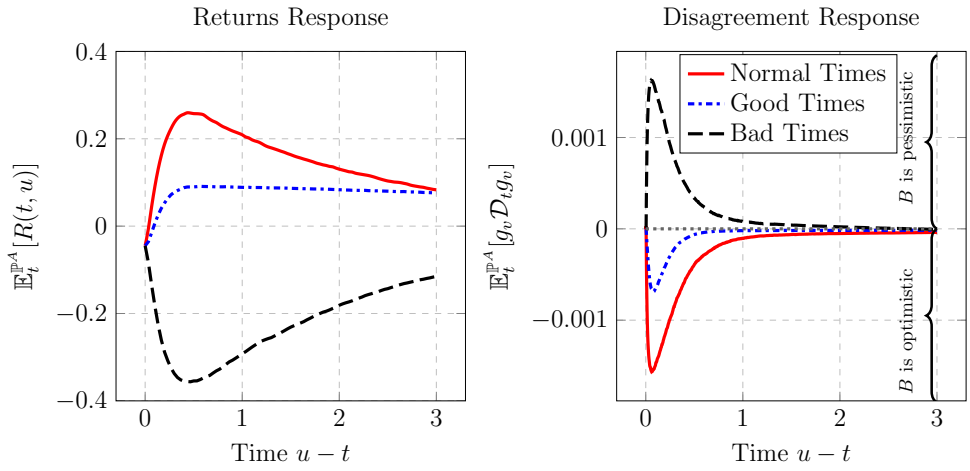


**Figure 5: Model-Implied Impulse Response of Excess Returns to News Shocks.** The left panel plots the total (sum of fundamental and disagreement channels) impulse response of stock returns. The right panel plots the response of disagreement to news shocks $\mathbb{E}_t^{\mathbb{P}^A}[g_u \mathcal{D}_t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

reaction of the likelihood, $\eta$, and therefore against the average reaction of disagreement:

$$\text{sign}(\mathbb{E}_t^{\mathbb{P}^A}[R(t,s)]) \approx \text{sign}(\mathbb{E}_t^{\mathbb{P}^A}[\mathcal{D}_t \eta_s]) \equiv -\text{sign}\left(\mathbb{E}_t^{\mathbb{P}^A}\left[\int_t^s g_u \mathcal{D}_t g_u \mathrm{d}u\right]\right). \tag{22}$$

We plot the 3-year average future returns 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 future returns, based on (22) a negative sign implies that returns move opposite to news—they under-react. A positive sign either means that returns adjust to news or over-react, in which case returns subsequently revert—average future 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 (dashed black line), news exacerbates disagreement by polarizing agents' opinion. Agent $A$ interprets news positively, whereas Agent $B$ interprets it negatively and disagreement spikes (see right panel). Since news invalidates Agent $B$'s model, returns move opposite to news: good news is followed by negative returns in the short term (see left panel). Hence, Agent $B$'s pessimism slows down returns' adjustment to news (decreasing reaction), similar to the under-reaction phenomenon in Ottaviani and Sorensen (2015).

In normal times (solid red line), news shocks precipitate an upward adjustment in Agent

$B$'s expectations, sharpening the difference in adjustment speeds across agents. While agents interpret news in similar ways, Agent $B$ overreacts to news and becomes overly optimistic relative to Agent $A$ (see the right panel). Because news corroborates Agent $B$'s overoptimism, returns over-react to news in the short term (see the left panel). As a result, Agent $B$'s overreaction accelerates returns' adjustment to news (increasing reaction), similar to the over-reaction phenomenon in Daniel et al. (1998).

In good times (dash-dotted blue line), the reaction of disagreement to news is weak and exhibits little persistence (see the right panel). Because news corroborates both agents' models, returns adjust immediately (see the left panel). Hence, return continuation only arises when disagreement spikes in the short term. Whether disagreement spikes result from the polarization of opinions in bad times or the difference in adjustment speeds in normal times dictates whether returns under- or over-react. Unlike Ottaviani and Sorensen (2015), under- and over-reaction alternate over the business cycle in our model. Finally, because disagreement spikes are counter-cyclical, news better predicts future returns in bad times.

In the long term (after about 6 months), excess returns systematically revert, i.e., future returns change direction. This later phase of correction arises through the long-term behavior of Agent $B$'s consumption share, $1-\omega$, which multiplies the cumulative disagreement response in Equation (21). Since the drift of $1-\omega$,

$$\frac{1}{\mathrm{d}t}\mathbb{E}_t^{\mathbb{P}^A}\left[\mathrm{d}(1-\omega_t)\right] = \frac{g_t^2(2\omega_t - 1 - \alpha)(1-\omega_t)\omega_t}{2\alpha^2\sigma_\delta^2} \leq 0, \quad \forall \alpha \geq 1 \tag{23}$$

is negative and decreases with the square of disagreement, strong disagreement today leads to a strong downward trend in Agent $B$'s consumption share, which gradually dampens the disagreement response in the right panel of Figure 5. Hence, while a spike in disagreement creates return continuation in the short term, it generates a reversal in the long term.

## 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, return under- or over-reaction to news directly relate to time series momentum. Disagreement spikes persistently raise the risk premium for

holding the stock, thus creating positive serial correlation in excess returns. Disagreement spikes simultaneously reduce Agent $B$'s consumption share over time and thus the risk premium in the future, which ultimately leads to long-term reversal. That disagreement spikes are counter-cyclical further explains why time series momentum is stronger in bad times at short horizons and how it crashes following sharp market rebounds.

The economic mechanism for return continuation of Section 4.2 is based on the concept of returns that matters to investors of the model. However, the same mechanism can be described in terms of time series momentum based on a common definition of excess returns:

$$r_t^e \equiv \int_{t-\Delta}^{t} \left( \frac{\mathrm{d}S_u + \delta_u \mathrm{d}u}{S_u} - r_u^f \mathrm{d}u \right) = \int_{t-\Delta}^{t} \sigma_u (\theta_u \mathrm{d}u + \mathrm{d}\widehat{W}_u^A) \tag{24}$$

over a period of length $\Delta$. Consider the serial correlation of excess returns at lag $h$:

$$\rho(h) = \mathrm{cov}^{\mathbb{P}^A} \left( r_{t+h\Delta}^e, r_t^e \right) / \mathrm{var}^{\mathbb{P}^A} \left( r_t^e \right). \tag{25}$$

Following Banerjee et al. (2009), the sign of the coefficient in (25) determines whether returns exhibit momentum ($\rho(h) > 0$) or reversal ($\rho(h) < 0$). Because Brownian innovations in (24) are uncorrelated across non-overlapping periods, the covariance of instantaneous excess returns, $\sigma\theta$, in (24) exclusively dictates the sign of this coefficient.

To explain how disagreement spikes affect the coefficient in (25), it is simpler to focus on the covariance of market prices of risk, $\theta$, as opposed to instantaneous excess returns. An approximation of covariances of products (Bohrnstedt and Goldberger, 1969) then yields:

$$\begin{aligned} \mathrm{cov}^{\mathbb{P}^A}(\theta_t, \theta_s) \approx & \underbrace{\frac{1}{\sigma_\delta^2} \left( \mathbb{E}^{\mathbb{P}^A}[g_t] \mathbb{E}^{\mathbb{P}^A}[g_s] \mathrm{cov}^{\mathbb{P}^A}(\omega_t, \omega_s) + \mathbb{E}^{\mathbb{P}^A}[1-\omega_t] \mathbb{E}^{\mathbb{P}^A}[1-\omega_s] \mathrm{cov}^{\mathbb{P}^A}(g_t, g_s) \right)}_{\text{momentum effect } \geq 0} \\ & \underbrace{- \frac{1}{\sigma_\delta^2} \left( \mathbb{E}^{\mathbb{P}^A}[g_t] \mathbb{E}^{\mathbb{P}^A}[1-\omega_s] \mathrm{cov}^{\mathbb{P}^A}(\omega_t, g_s) + \mathbb{E}^{\mathbb{P}^A}[1-\omega_t] \mathbb{E}^{\mathbb{P}^A}[g_s] \mathrm{cov}^{\mathbb{P}^A}(g_t, \omega_s) \right)}_{\text{reversal effect } \leq 0}. \end{aligned} \tag{26}$$

Disagreement and agents' consumption share *separately* move the market price of risk in the same direction. In the short term disagreement keeps moving in the same direction both in bad and normal times (see Section 3) and so does agents' consumption share through (23). Hence, the first term in (26) is positive and induces momentum. In contrast, the *interaction*

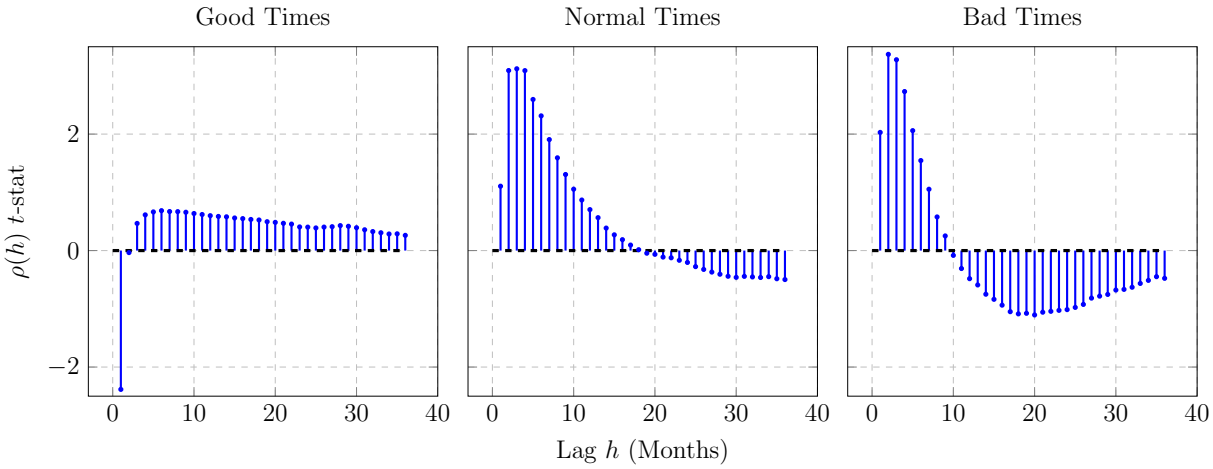


**Figure 6: Model-Implied Conditional Time Series Momentum.**
This figure plots the $t$-statistics of the coefficient $\rho(h)$ for lags $h$ ranging from 1 month to 3 years. Each panel corresponds to a different state of the economy. Standard errors are adjusted using Newey and West (1987). The values reported above are obtained from 10,000 simulations of the economy over a 20-year horizon.

between disagreement and consumption shares moves current and future prices of risk in opposite directions. In normal and good times disagreement is negative and moves opposite to Agent $A$'s consumption share; in bad times it is positive and moves alongside Agent $A$'s consumption share. Hence, the second term in (26) is negative and induces reversal.

Whether the momentum or the reversal effect in (26) dominates depends on the magnitude of fluctuations in disagreement. Because large and sharp moves in disagreement strengthen the momentum effect in the short term, disagreement spikes tilt the balance in favor of momentum. To analyze this effect in a way that is consistent with empirical studies, we now consider excess returns at a monthly frequency (i.e., $\Delta = 1/12 = 1$ month). Figure 6 depicts the $t$-statistics of the coefficient of serial correlation in Equation (25) for lags ranging from 1 month to 3 years. The value of the coefficient is reported in Appendix A.11.2.

The serial correlation of returns has a similar term structure in normal (center panel) and bad times (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.11.3 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.[15] 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.

Since time series momentum results from a spike in disagreement, the term structures of momentum in normal and bad times differ in one important dimension. The polarization of opinions in bad times induces a sharper spike in disagreement than the difference in adjustment speeds in normal times and thus 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). In addition, long-term reversal occurs earlier and is stronger in bad times because a stronger disagreement spike implies a steeper downward trend in Agent $B$'s consumption share (see Section 4.2).

Excess returns have different time series properties in good times (left panel): returns exhibit strong reversal in the very short term and insignificant momentum thereafter. The reason is that news generates little disagreement in good times (see Figure 2). Returns thus immediately revert, as they would if agents had homogenous beliefs. A consequence of this reversal spike occurring in good times is that a time series momentum strategy may crash if the market rises sharply. To see this, suppose we implement a momentum strategy in bad or normal times. Figure 6 shows that this strategy remains profitable as long as the economy does not transit suddenly 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).[16] Furthermore, we show in Appendix A.11.7 that the model implies stronger time series momentum in extreme markets, consistent with Moskowitz et al. (2012).

# 5 Testable Predictions and Empirical Evidence

In Section 5.1 we construct empirical proxies for the main variables driving excess returns in the model. For instance, we use the dispersion of analysts' forecasts as a proxy for the

[15] In Appendix A.11.6 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.

[16] Daniel and Moskowitz (2016) and Barroso and Santa-Clara (2015) document a similar phenomenon in the cross-section of stocks.

square of disagreement. We then focus on three novel predictions of the model: *(1)* future excess returns are positively related to contemporaneous dispersion; time series momentum at short horizons *(2)* increases with dispersion and *(3)* is strongest in bad times. In Section 5.2 we quantify these qualitative predictions using simulations from the model. Repeating the same exercise with observed data, we provide empirical support to these predictions.

## 5.1 Constructing Empirical Proxies for Explanatory Variables

In the model excess returns depend nonlinearly on three variables—disagreement, the fundamental and the relative likelihood of agents' models. However, Equation (20) shows that this nonlinear dependence is accurately described as a linear combination of two weighted products of these variables, the *weighted fundamental* $\widehat{f}^A(1-\omega(\eta))^2 g^2$ and the *weighted dispersion* $g^2(1-\omega(\eta))^2\omega(\eta)$. We borrow the term "dispersion" from the empirical literature in which heterogeneous beliefs are commonly measured by the dispersion of analysts' forecasts (Diether et al., 2002), the square of disagreement in our model; accordingly we define the monthly weighted fundamental $F$ and the monthly weighted dispersion $G$ as

$$F_t \equiv \int_{t-\Delta}^{t} \widehat{f}_u^A(1-\omega_u)^2 g_u^2 \mathrm{d}u \qquad\qquad G_t \equiv \int_{t-\Delta}^{t} g_u^2(1-\omega_u)^2\omega_u \mathrm{d}u. \tag{27}$$

To construct empirical counterparts to $F$ and $G$ in (27), we need empirical proxies for dispersion, the fundamental and Agent $A$'s consumption share, which we now describe separately. To build an empirical proxy for dispersion, we follow Diether et al. (2002). We first obtain monthly data on analysts' forecasts from I/B/E/S, for the period covering February 1976 to November 2013. Denoting by $f^{i,j}$ the forecast of analyst $j$ regarding firm $i$'s fiscal year earnings per share, we define the dispersion $D^i$ of analysts' forecasts about firm $i$ as

$$D_t^i = (|\underset{j}{\mathrm{mean}}(f_t^{i,j})|)^{-1}\underset{j}{\mathrm{std}}(f_t^{i,j}),$$

where $\underset{j}{\mathrm{mean}}(.)$ and $\underset{j}{\mathrm{std}}(.)$ denote the mean and the standard deviation of forecasts computed

| Dep. Var. | $D_t$ |
|---|---|
| Const. | 0.2099*** |
| | (0.0137) |
| Time $t$ | −0.0032*** |
| | (0.0005) |
| Adj. $R^2$ | 0.1255 |
| Obs. | 454 |

| Dep. Var. | $GMF_t$ |
|---|---|
| Const. | −0.0309*** |
| | (0.0076) |
| $Y_t^{12}$ | 0.1258* |
| | (0.0660) |
| Adj. $R^2$ | 0.0310 |
| Obs. | 454 |

**Table 3: Linear Trend in Dispersion and 12-Month Seasonality in Fundamental.** The left and right panels report the outputs obtained by regressing the 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). Data are at the monthly frequency from 02/1976 to 11/2013.

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 analysts' forecasts and a mean forecast different from zero at time $t$.[17] Regressing the aggregate dispersion $D_t$ on time $t$ shows that dispersion has a strong linear downward trend (see left panel of Table 3). Communication and transparency greatly improved over the last forty years due to technological innovations, e.g., the internet and other information technologies. Since our model abstracts from these technological changes, we construct a de-trended proxy for dispersion, $Disp$, using the residual from the regression of aggregate dispersion on time.

In Section 3 we estimated the beliefs $\widehat{f}^A$ and $\widehat{f}^B$ using monthly S&P 500 dividend growth. These beliefs provide a *model-implied* time series of dispersion, $Disp^{MI}$, defined as:

$$Disp_t^{MI} \equiv g_t^2 = (\widehat{f}_t^A - \widehat{f}_t^B)^2.$$

If the beliefs $\widehat{f}^A$ and $\widehat{f}^B$ offer a reasonable description of observed analysts' forecasts, then the model-implied dispersion $Disp^{MI}$ should be correlated with the dispersion among forecasters $Disp$. To investigate this matter, we regress the observed dispersion $Disp$ on the model-implied dispersion $Disp^{MI}$. The right panel of Table 4 shows that there is a strong positive

[17] Over the time period 02/1976 to 11/2013, a mean of 490 firms meet these two conditions each month.

| Dep. Var. | $Disp_t$ |
|---|---|
| Const. | $-0.0071$ |
| | (0.0044) |
| $Rec_t$ | $0.0525^{***}$ |
| | (0.0137) |
| Adj. $R^2$ | 0.0370 |
| Obs. | 454 |

| Dep. Var. | $Disp_t$ |
|---|---|
| Const. | $-0.0145^{**}$ |
| | (0.0060) |
| $Disp_t^{MI}$ | $6.6594^{***}$ |
| | (1.7589) |
| Adj. $R^2$ | 0.0624 |
| Obs. | 454 |

**Table 4: Empirical Dispersion in Recessions and Empirical vs. Model-Implied Dispersion.**

The left panel reports the outputs obtained by regressing the empirical dispersion *Disp* on a dummy variable *Rec* that equals 1 during NBER recessions. The right panel reports the outputs obtained by regressing the empirical 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). Data are at the monthly frequency from 02/1976 to 11/2013.

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 observed dispersion is higher in recessions than in expansions (left panel of Table 4). The observed dispersion is therefore counter-cyclical, as our model predicts (see Appendix A.11.4 for an illustration of the time series of *Disp* and $Disp^{MI}$).

We build a proxy for the fundamental using the simple growth rate, $GMF$, of the aggregate mean forecast, $MF_t = \frac{1}{N_t}\sum_{i=1}^{N_t} \underset{j}{\text{mean}}(f_t^{i,j})$, which we compute according to:

$$GMF_t = (MF_{t-\Delta})^{-1}(MF_t - MF_{t-\Delta}),$$

where $\Delta = 1/12 = 1$ month. We do not 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 further eliminate 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 the right panel of Table 3). We use the residual, *Fund*, of this regression as an empirical proxy for the fundamental.

Finally, since a model-free proxy for Agent $A$'s consumption share is unavailable, we use the estimation performed in Section 3 to construct this time series. Following David (2008), we first generate a time series of the relative likelihood of agents' models according to

$$\eta_{t+\Delta} = \eta_t e^{-\frac{\left(\widehat{f}_t^A - \widehat{f}_t^B\right)^2}{2\sigma_\delta^2}\Delta - \frac{\left(\widehat{f}_t^A - \widehat{f}_t^B\right)}{\sigma_\delta}\epsilon_{t+\Delta}}, \qquad \epsilon_{t+\Delta} = \frac{1}{\sigma_\delta}\left(\log\left(\frac{\delta_{t+\Delta}}{\delta_t}\right) - \left(\widehat{f}_t^A - \frac{1}{2}\sigma_\delta^2\right)\Delta\right), \tag{28}$$

where we substitute $\widehat{f}^A$ and $\widehat{f}^B$ by the beliefs we estimated in Section 3, $\log\left(\frac{\delta_{t+\Delta}}{\delta_t}\right)$ by the S&P 500 dividend growth, $\Delta = 1/12 = 1$ month, along with the estimated parameters of Table 1. We then obtain an empirical proxy for the consumption share of Agent $A$, which we call *omega*, by substituting the time series in (28) into Equation (10). We now have all the variables we need to compute the empirical counterparts to $F$ and $G$ in (27):

$$F_{emp} \equiv Fund \times (1 - omega)^2 \times Disp \quad \text{and} \quad G_{emp} \equiv Disp \times (1 - omega)^2 \times omega. \tag{29}$$

## 5.2 Testing the Predictions of the Model

We test three new predictions of the model. We show that current dispersion positively predicts future excess returns (Section 5.2.1), time series momentum increases with dispersion (Section 5.2.2) and is strongest in bad times (Section 5.2.3).

### 5.2.1 Dispersion and Future Excess Returns

To quantify the model-implied relation between excess returns and dispersion at different lags $h$, we simulate the model and run the regression

$$r^e_{t+h\Delta} = \alpha(h) + \beta_G(h)G_t + \beta_F(h)F_t + \epsilon_{t+h\Delta}, \tag{30}$$

where monthly excess returns $r^e$ are defined in (24). Because weighted dispersion generates most variations in excess returns (see Section 4.1), we focus on the coefficient $\beta_G(h)$, which precisely measures the relation between excess returns and dispersion at different lags $h$. We then run the empirical equivalent to the regression in (30) in which we substitute the weighted fundamental $F$ and the weighted dispersion $G$ by their empirical counterparts $F_{emp}$ and $G_{emp}$ in (29) and take excess returns $r^e$ to be the monthly excess returns on the S&P

500. We plot the $t-$statistic of the theoretical coefficient $\beta_G(h)$ and its empirical counterpart $\beta_G(h)_{emp}$ in the left and right panels of Figure 7 for lags ranging from 1 month to 1 year.

Our model predicts a positive relation between contemporaneous dispersion and future excess returns, the strength of which weakens with the time horizon (see the left panel). To understand this, suppose that dispersion is high today. Because dispersion is counter-cyclical, it will decrease in the future (see Section 3.3 and Table 4).[18] Since dispersion and excess returns are contemporaneously negatively related (see Section 4.1), future excess returns are high, creating a positive relation between dispersion and future excess returns. This observation further implies that future excess returns are counter-cyclical in our model, consistent with empirical findings. Comparing the two panels of Figure 7 shows that the model can replicate the persistence and the strength of the relation between contemporaneous dispersion and future excess returns that we observe in the data.

Interestingly, the positive relation between *aggregate* dispersion and future *aggregate* returns we document contrasts with the negative relation that is documented in the cross-section of firms (Diether et al., 2002). A possible explanation is that short-selling costs, which may create a negative relation between dispersion and returns (Miller, 1977), matter for the cross-section of firms—costs are high for small companies' stocks—but not in the aggregate market. For instance, Anderson, Ghysels, and Juergens (2005), Boehme, Danielsen, Kumar, and Sorescu (2009), and Carlin et al. (2014) find a positive cross-sectional relation between dispersion and future returns in markets that involve low short-selling costs.

### 5.2.2 Time Series Momentum in High Dispersion Periods

Short-term time series momentum increases with dispersion in our model. To measure this effect, we define periods of high dispersion with a dummy variable, $Y_{G,t}(p)$, that takes value 1 when the monthly dispersion $G$ is larger than its $p-$th percentile. We then run the regression:

$$r^e_{t+\Delta} = \alpha_G(p) + \beta_{1,G}(p) r^e_t + \beta_{2,G}(p) r^e_t Y_{G,t}(p) + \epsilon_{t+\Delta}, \tag{31}$$

[18] To confirm that dispersion is counter-cyclical in our model, we regress the dispersion $G_t$ on the fundamental, $F^m_t \equiv \int_{t-\Delta}^{t} \widehat{f}^A_u \mathrm{d}u$, using 1,000 simulations over a 100-year horizon. We find that dispersion and the fundamental are, indeed, negatively related with a regression coefficient of $-0.0017$. Using Newey and West (1987) standard errors, the regression coefficient is significant at the 1% level. Note that the regular dispersion, $G_t \equiv \int_{t-\Delta}^{t} g^2_u \mathrm{d}u$, is also counter-cyclical.

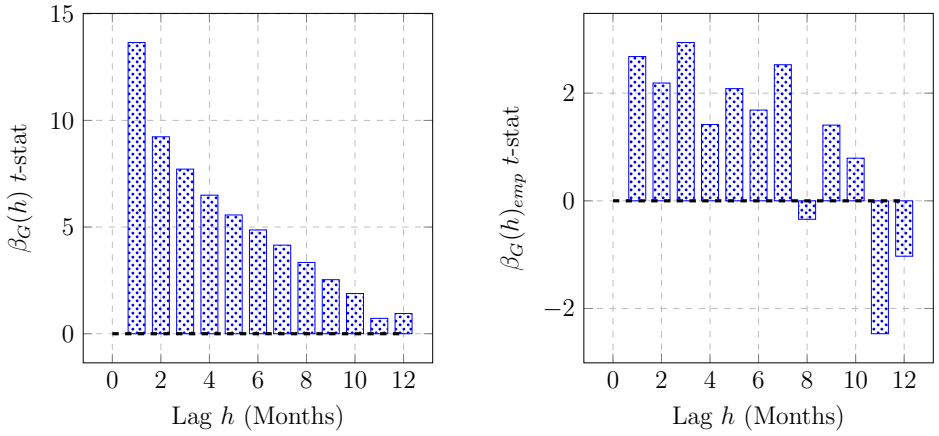


**Figure 7: Model-Implied and Empirical Relation between Future Excess Returns and Dispersion.**

The left panel plots the $t-$statistics of the model-implied response of excess returns to weighted dispersion $\beta_G(h)$ for lags $h$ ranging from 1 month to 1 year. The values reported are obtained from 1,000 simulations of the economy over a 100-year horizon. The right panel plots the $t-$statistics of the empirical response of excess returns to weighted dispersion $\beta_G(h)_{emp}$ for lags $h$ ranging from 1 month to 1 year. Data are at the monthly frequency from 02/1976 to 11/2013. All standard errors are adjusted using Newey and West (1987).

where the model-implied coefficient $\beta_{2,G}(p)$ measures excess time series momentum in periods of high dispersion. We repeat this exercise by substituting in (31) the weighted dispersion $G$ by its empirical counterpart $G_{emp}$ in (29) and taking excess returns $r^e$ to be the monthly S&P 500 excess returns. The empirical coefficient $\beta_{2,G}(p)_{emp}$ measures excess time series momentum at a 1-month lag in high dispersion periods.[19,20] We plot the $t-$statistic of $\beta_{2,G}(p)$ and $\beta_{2,G}(p)_{emp}$ in separate panels in Figure 8 for percentiles ranging from 10 to 50.[21]

Both in the model and in the data, time series momentum at a 1-month lag is larger in periods of high dispersion. However, while the significance of the model-implied relation between dispersion and excess time series momentum is robust to changes in the dispersion percentile, the empirical relation is strongly significant when the dispersion threshold is the 10−th percentile and is relatively weak when the threshold ranges from the 20−th to the

[19] We present an alternative approach for estimating the positive relation between time series momentum and dispersion in Appendices A.9 and A.10.

[20] Our results are qualitatively similar if we replace the weighted dispersion $G_t \equiv \int_{t-\Delta}^{t} g_u^2(1-\omega_u)^2\omega_u \mathrm{d}u$ (resp. $G_{emp} \equiv Disp \times (1-omega)^2 \times omega$) by the regular dispersion $G_t \equiv \int_{t-\Delta}^{t} g_u^2 \mathrm{d}u$ (resp. $G_{emp} \equiv Disp$).

[21] We describe the statistics $\beta_{2,G}(p)$ and $\beta_{2,G}(p)_{emp}$ 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.

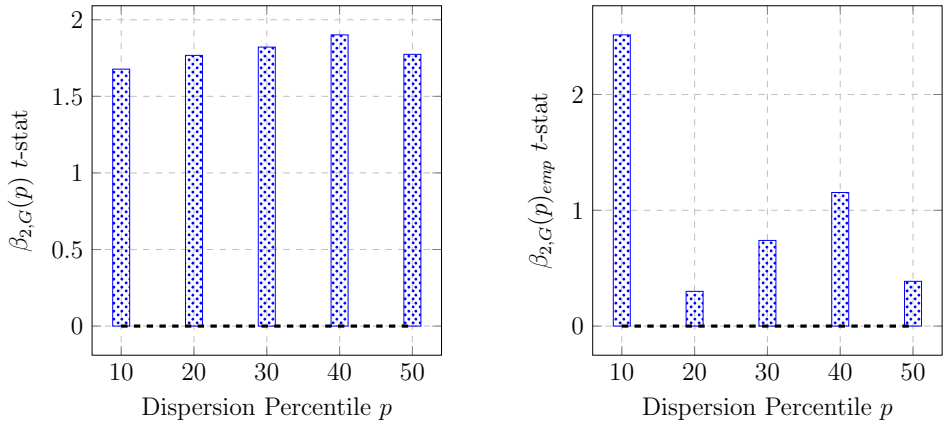


**Figure 8: Model-Implied and Empirical Excess 1-Month Time Series Momentum in High Dispersion Periods.**

The left panel plots the model-implied $t-$statistics of excess 1-month time series momentum $\beta_{2,G}(p)$ when weighted dispersion is larger than its $p-$th percentile. These values are obtained from 1,000 simulations of the economy over a 100-year horizon. The right panel plots the empirical $t-$statistics of excess 1-month time series momentum $\beta_{2,G}(p)_{emp}$ when weighted dispersion is larger than its $p-$th percentile. Data are at the monthly frequency from 02/1976 to 11/2013. All standard errors are adjusted using Newey and West (1987).

50−th percentile. The reason for this discrepancy is that spikes in dispersion persist in our model, whereas they revert back quickly in the data (see Appendix A.11.4). Furthermore, the data also lend support to the prediction that there is significant short-term time series reversal in periods of low dispersion (see Appendix A.11.5).

### 5.2.3 Time Series Momentum over the Business Cycle

The model predicts strongest time series momentum in bad times at short horizons, a result we quantify as follows. We first identify bad times in our model in a way that can be mapped into our empirical analysis, 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 $Y_{F^m}$ that takes value 1 if the fundamental, $F_t^m \equiv \int_{t-\Delta}^{t} \widehat{f}_u^A \mathrm{d}u$, is below its 30−th percentile.[22] We then simulate the model and run the regression

$$r_{t+h\Delta}^e = \alpha_F(h) + \beta_{1,F}(h) r_t^e + \beta_{2,F}(h) r_t^e Y_{F^m,t} + \epsilon_{t+h\Delta}, \tag{32}$$

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

where the model-implied coefficient $\beta_{2,F}(h)$ measures excess time series momentum in recessions at lag $h$.

To construct the empirical equivalent to the regression in (32), we specify $Y_{F^m}$ as a dummy variable that takes value 1 during NBER recessions and take excess returns $r^e$ to be the monthly S&P 500 excess returns. The empirical coefficient $\beta_2(h)_{emp}$ measures excess time series momentum in recessions at lag $h$. To cover sufficiently many business cycle turning points, we extend the sample period and focus on monthly S&P 500 excess returns from January 1871 to November 2013. 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 plot the $t-$statistics of $\beta_{2,F}(h)$ and $\beta_2(h)_{emp}$ in separate panels in Figure 9 for lags ranging from 1 month to 1 year.

The model accurately replicates excess time series momentum in recessions and its level of statistical significance at the 1-month lag. The model, however, implies persistent excess time series momentum up to the 5-month lag, whereas excess time series momentum vanishes in the data over lags beyond one month. This discrepancy in persistence between model-implied and observed excess time series momentum in bad times results from the difference in persistence between model-implied and observed spikes in dispersion. To bring about closer alignment in persistence between the two, disagreement spikes would need to revert faster. For instance, introducing *transient* states in Agent $B$'s Markov chain would increase the reversion speed of her expectations—the sum of all transition intensities—reducing the persistence of disagreement and thus the persistence of excess time series momentum.

Our finding that short-term time series momentum is stronger in recessions contrasts with evidence reported in the cross-section of returns: 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 related, their relation seems to vary 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

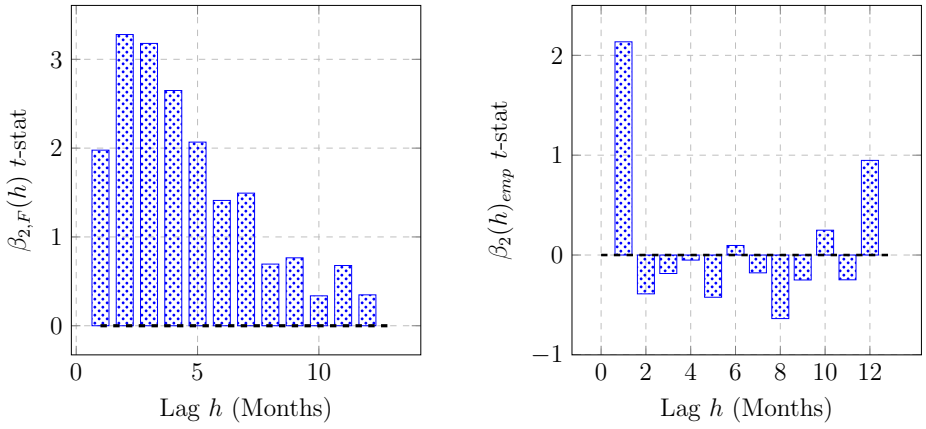


**Figure 9: Model-Implied and Empirical Excess Time Series Momentum in Recessions.**

The left panel plots the $t-$statistics of the model-implied excess time series momentum in recessions $\beta_{2,F}(h)$ for lags $h$ ranging from 1 month to 1 year. The values reported above are obtained from 1,000 simulations of the economy over a 100-year horizon. The right panel plots the $t-$statistics of the empirical excess time series momentum in recessions $\beta_2(h)_{emp}$ for lags $h$ ranging from 1 month to 1 year. Data are at the monthly frequency from 01/1871 to 11/2013. All standard errors are adjusted using Newey and West (1987).

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.

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

Barinov, A. (2014). 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.

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.

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.

Bohrnstedt, G. W. and A. S. Goldberger (1969). On the exact covariance of products of random variables. *Journal of the American Statistical Association 64*(328), 1439–1442.

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

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. *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). Learning and Asymmetric Business Cycles. *Review of Economic Dynamics 1*(3), 623–645.

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

Cooper, M. J., R. C. Gutierrez, and A. Hameed (2004). 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.

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). Investor psychology and security market under- and overreactions. *Journal of Finance 53*(6), 1839–1885.

Daniel, K. D. and T. J. Moskowitz (2016). Momentum crashes. *Forthcoming, Journal of Financial Economics*.

David, A. (1997). 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.

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). 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). 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. *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. *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. *Journal of Finance 64*(2), 579–629.

Dumas, B. and A. Lyasoff (2012). Incomplete-market equilibria solved recursively on an event tree. *Journal of Finance 67*(5), 1897–1941.

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

Garcia, D. (2013). 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). 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). 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). News events, information acquisition, and serial correlation. *Journal of Business 75*(1), 1–32.

Hong, H. and J. C. Stein (1999). 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). Simple forecasts and paradigm shifts. *Journal of Finance 62*(3), 1207–1242.

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

Johnson, T. C. (2002). 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. 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. *Journal of Finance 61*(1), pp. 195–229.

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

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

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. *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). 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.

Piatti, I. and F. Trojani (2015). Predictable risks and predictive regression in present-value models. *Working Paper*.

Rapach, D. E., J. K. Strauss, and G. Zhou (2010). 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). 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. *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). 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). 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.

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*.

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

# A Internet Appendix

## A.1 Proof of Proposition 1

We follow the notations in Lipster and Shiryaev (2001b) 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 (2001b), 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

$$\begin{aligned} \mathrm{d}g_t = &\left[\kappa\left(\bar{f} - \widehat{f}_t^A\right) - (\lambda+\psi)\left(f_\infty + g_t - \widehat{f}_t^A\right) - \frac{g_t}{\sigma_\delta^2}\left(\widehat{f}_t^A - g_t - f^l\right)\left(f^h + g_t - \widehat{f}_t^A\right)\right] dt \\ &+ \frac{\gamma - \left(\widehat{f}_t^A - g_t - f^l\right)\left(f^h + g_t - \widehat{f}_t^A\right)}{\sigma_\delta} \mathrm{d}\widehat{W}_t^A. \end{aligned} \tag{33}$$

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 (33) 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 (2001a) for related results. ■

We now show that the stochastic differential equation in (33) has a unique nonexplosive strong solution under $\widehat{\mathbb{P}}^A$ and $\widehat{\mathbb{P}}^B$. Rewrite the process $g_t$ in (33) as

$$\mathrm{d}g_t = \left(\mu(\widehat{f}_t^B) - \lambda(\widehat{f}_t^B) g_t\right) \mathrm{d}t + \sigma(\widehat{f}_t^B) \mathrm{d}\widehat{W}_t^A \tag{34}$$

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

$$\mathrm{d}g_t = \left(\mu(\widehat{f}_t^B) - \left(\kappa + \frac{\gamma}{\sigma_\delta^2}\right) g_t\right) \mathrm{d}t + \sigma(\widehat{f}_t^B) \mathrm{d}\widehat{W}_t^B \tag{35}$$

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

$$\begin{aligned} \mu(x) &:= \kappa(\bar{f} - x) - (\lambda+\psi)(f_\infty - x) \\ \sigma(x) &:= \frac{1}{\sigma_\delta}(\gamma - (x - f^l)(f^h - x)) \\ \lambda(x) &:= \kappa + \frac{1}{\sigma_\delta^2}(x - f^l)(f^h - x). \end{aligned} \tag{36}$$

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

**Lemma 1.** *The processes in (34) and (35) 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 (34) by setting

$$g_t^{(k+1)} := g_t^{(0)} + \int_0^t \left(\mu(\widehat{f}_s^{(k)}) - \lambda(\widehat{f}_s^{(k)})g_s^{(k)}\right) \mathrm{d}s + \int_0^t \sigma(\widehat{f}_s^{(k)}) \mathrm{d}\widehat{W}_s^A \tag{37}$$

for $k \geq 0$, where $\widehat{f}^{(k)}$ denotes some approximation of $\widehat{f}^B$. From (37), 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 (36) are locally Lipschitz continuous. In particular, for all $x, y \in (f^l, f^h)$, we have

$$|\sigma(x) - \sigma(y)| = \frac{1}{\sigma_\delta}|x-y||f^h + f^l - (x+y)| \leq \frac{1}{\sigma_\delta} \max\left\{|f^h - f^l|, |f^l - f^h|\right\} |x-y| \tag{38}$$

and

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

Lipschitz continuity for $\lambda(\cdot)$ directly follows from (38). 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

$$g_t^{(k+1)} - g_t^{(k)} = -\int_0^t \lambda(\widehat{f}_s^B) \left(g_s^{(k)} - g_s^{(k-1)}\right) \mathrm{d}s. \tag{39}$$

Taking the absolute value of both sides of (39) and observing that $\lambda(x) > 0$ for all $x \in (f^l, f^h)$, we have

$$\left|g_t^{(k+1)} - g_t^{(k)}\right| \leq \int_0^t \lambda(\widehat{f}_s^B) \left|g_s^{(k)} - g_s^{(k-1)}\right| \mathrm{d}s \leq \overline{\lambda} \int_0^t \left|g_s^{(k)} - g_s^{(k-1)}\right| \mathrm{d}s \tag{40}$$

where

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

Iterating over (40), 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 (34) 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 (34), 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 (33) 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

$$\mathrm{d}A_t = \mu(\widehat{f}_t^B)e^{\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}\mathrm{d}t + e^{\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}\sigma(\widehat{f}_t^B)\mathrm{d}\widehat{W}_t^A, \quad A_0 = g_0 \tag{41}$$

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

$$\mathrm{d}A_t^i = (-1)^i m_t^A \mathrm{d}t + e^{\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}\sigma(\widehat{f}_t^B)\mathrm{d}\widehat{W}_t^A, \quad (-1)^i A_0^i \geq (-1)^i A_0 \tag{42}$$

under $\widehat{\mathbb{P}}^A$. Combining (41) and (42), we then obtain

$$A_t^i - A_t = A_0^i - A_0 + \int_0^t \left((-1)^i m_s^A - \mu(\widehat{f}_s^B)e^{\int_0^s \lambda(\widehat{f}_u^B)\mathrm{d}u}\right)\mathrm{d}s, \quad i =, 1, 2. \tag{43}$$

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

$$(-1)^i e^{\int_0^t \lambda(\widehat{f}_u^B)\mathrm{d}u}\mu(\widehat{f}_t^B) \leq e^{\int_0^t \lambda(\widehat{f}_u^B)\mathrm{d}u}|\mu(\widehat{f}_t^B)| \leq m_t^A, \quad i = 1, 2 \tag{44}$$

for all $t \in \mathbb{R}_+$. Inequalities in (44) and the expressions in (43) together imply that

$$A_t^1 \leq A_t \leq A_t^2, \quad \widehat{\mathbb{P}}^A - a.s. \tag{45}$$

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

$$|g_t| = e^{-\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}(A_t^+ + A_t^-) \leq A_t^+ + A_t^-, \quad \widehat{\mathbb{P}}^A - a.s. \tag{46}$$

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 (45) and (46), we obtain

$$|g_t| \leq \sum_{i=1,2} ((-1)^i A_t^i)^+ \leq \max\{(-A_t^1)^+, (A_t^2)^+\}, \quad \widehat{\mathbb{P}}^A - a.s. \tag{47}$$

Using (47), we can write that, for any positive constant $c \geq 0$,

$$\mathbf{1}_{|g_t|\geq c} \leq \sum_{i=1,2} \mathbf{1}_{((-1)^i A_t^i)^+ \geq c} \equiv \sum_{i=1,2} \mathbf{1}_{(-1)^i A_t^i \geq c} \tag{48}$$

where the last equality follows from that $\mathbf{1}_{X^+\geq c} = \mathbf{1}_{X\geq c}$ for any $c$ positive. Taking expectations of (48) under $\widehat{\mathbb{P}}^A$, we get

$$\widehat{\mathbb{P}}^A(|g_t| \geq c) \leq \sum_{i=1,2} \widehat{\mathbb{P}}^A\left((-1)^i A_t^i \geq c\right), \quad \forall t \in \mathbb{R}_+. \tag{49}$$

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

$$\mathrm{d}\widehat{A}_t^i = (-1)^i m_t^A \mathrm{d}t + \sigma_t^A \mathrm{d}\widehat{W}_t^A, \quad (-1)^i \widehat{A}_0^i \geq (-1)^i A_0^i \tag{50}$$

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.,

$$a^{i,j} \sim \widehat{A}^i, \quad i,j = 1,2. \tag{51}$$

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

$$\mathrm{d}\bar{a}_t^i = (-1)^i m_t^A \mathrm{d}t + e^{\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}\sigma(\widehat{f}_t^B)\mathrm{d}\widehat{W}_t^A. \tag{52}$$

Combining (42) and (52), 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 (51), we obtain that, for any $c \in \mathbb{R}$, the processes $A^i$ and $\widehat{A}^i$, $i = 1,2$ satisfy

$$\mathbb{P}^A((-1)^i A_t^i \geq c) \leq 2\mathbb{P}^A((-1)^i \widehat{A}_t^i \geq c), \quad \forall t \in \mathbb{R}_+. \tag{53}$$

Combining the inequalities in (49) and (53), we obtain that

$$\widehat{\mathbb{P}}^A(|g_t| \geq c) \leq 2\sum_{i=1,2} \widehat{\mathbb{P}}^A\left((-1)^i \widehat{A}_t^i \geq c\right), \quad \forall t \in \mathbb{R}_+. \tag{54}$$

Observing that each $\widehat{A}^i$, $i = 1,2$ in (50) is a Gaussian process, the probabilities on the right-hand

side of (54) are explicitly given by

$$\mathbb{P}^A\left((-1)^i\widehat{A}_t^i \geq c\right) = \Phi\left(\frac{(-1)^i\widehat{A}_0^i + \int_0^t m_s^A \mathrm{d}s - c}{\sqrt{\int_0^t (\sigma_s^A)^2 \mathrm{d}s}}\right). \tag{55}$$

Taking limits on both sides of (54) and using (55) 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

$$c_{At} = \left(\phi_A e^{\rho t}\xi_t\right)^{-\frac{1}{\alpha}} \qquad c_{Bt} = \left(\frac{\phi_B e^{\rho t}\xi_t}{\eta_t}\right)^{-\frac{1}{\alpha}}. \tag{56}$$

Clearing the market yields the following characterization of the state-price density

$$\xi_t = e^{-\rho t}\delta_t^{-\alpha}\left[(1/\phi_A)^{1/\alpha} + (\eta_t/\phi_B)^{1/\alpha}\right]^\alpha. \tag{57}$$

Substituting Equation (57) into Equation (56) gives the consumption share of Agent $A$, $\omega$, which satisfies:

$$\omega_t = \frac{(1/\phi_A)^{1/\alpha}}{(\eta_t/\phi_B)^{1/\alpha} + (1/\phi_A)^{1/\alpha}}.$$

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, $1-\omega$, of Agent $B$. The dynamics of the consumption share $\omega$ satisfy

$$\mathrm{d}\omega_t = \frac{g_t^2}{2\alpha^2\sigma_\delta^2}\left((\alpha-1)(1-\omega_t)\omega_t^2 + (\alpha+1)(1-\omega_t)^2\omega_t\right)\mathrm{d}t + \frac{g_t}{\alpha\sigma_\delta}(1-\omega_t)\omega_t\mathrm{d}\widehat{W}_t^A.$$

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 (57) 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:[23]

$$\begin{aligned}\frac{S_t}{\delta_t} &= \mathbb{E}_t^{\mathbb{P}^A}\left[\int_t^\infty \frac{\xi_u\delta_u}{\xi_t\delta_t}du\right] \\ &= \omega_t^\alpha\sum_{j=0}^{\alpha}\binom{\alpha}{j}\left(\frac{1-\omega_t}{\omega_t}\right)^j\mathbb{E}_t^{\mathbb{P}^A}\left[\int_t^\infty e^{-\rho(u-t)}\left(\frac{\eta_u}{\eta_t}\right)^{\frac{j}{\alpha}}\left(\frac{\delta_u}{\delta_t}\right)^{1-\alpha}du\right].\end{aligned} \tag{58}$$

We start by computing the first and the last term of the sum in (58). 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 7.

**Proposition 7.** *Suppose the economy is populated by a single agent.*

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

1. *If the agent's filter follows the Ornstein-Uhlenbeck process described in Equation (4), the 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 (58) as

$$\frac{S_t}{\delta_t} = \omega_t^\alpha \frac{S_t}{\delta_t}\bigg|_{O.U.} + \omega_t^\alpha \sum_{j=1}^{\alpha-1} \begin{pmatrix} \alpha \\ j \end{pmatrix} \left(\frac{1-\omega_t}{\omega_t}\right)^j F^j\left(\widehat{f}_t^A, g_t\right) + (1-\omega_t)^\alpha \frac{S_t}{\delta_t}\bigg|_{M.C.}. \tag{59}$$

The last step consists in computing the intermediate terms $F^j$ in (59), which relate to heterogeneous beliefs. Each term solves a differential equation, which we present in Proposition 8.

**Proposition 8.** *The function $F^j$, defined as*

$$F^j\left(\widehat{f}_t^A, g_t\right) \equiv \mathbb{E}_t^{\mathbb{P}^A}\left[\int_t^\infty e^{-\rho(u-t)}\left(\frac{\eta_u}{\eta_t}\right)^{\frac{j}{\alpha}}\left(\frac{\delta_u}{\delta_t}\right)^{1-\alpha} du\right], \tag{60}$$

*solves the following partial differential equation*

$$\widetilde{\mathscr{L}}^{\widehat{f}^A,g} F^j + X^j F^j + 1 = 0, \tag{61}$$

*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

$$\overline{W}_t = \widehat{W}_t^A + \int_0^t \frac{j}{\alpha}\frac{g_s}{\sigma_\delta}\mathrm{d}s, \tag{62}$$

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

$$\widetilde{W}_t = \overline{W}_t - \int_0^t (1-\alpha)\,\sigma_\delta \mathrm{d}t. \tag{63}$$

Implementing sequentially the changes of probability measures in Equations (62) and (63) allows us to rewrite the interior expectation in (60) as

$$F^j\left(\widehat{f}_t, g_t\right) = \mathbb{E}^{\widetilde{P}}\left[\int_t^\infty e^{\int_t^u X_s^j \mathrm{d}s}\mathrm{d}u \middle| \mathscr{F}_t\right], \tag{64}$$

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 (64) 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 (61).[24] □
■

We numerically solve Equation (61) for each term $j$ through Chebyshev collocation. In partic-

[24] 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 (61), 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. □

In general, it is difficult to prove uniqueness when risk aversion is greater than one (see Proposition 3 in David (2008)). For our purpose, however, we only need to establish uniqueness under the calibration of Section 3.2. A convenient way to do so is to use a "Negishi map" (see Dumas and Lyasoff (2012) for a detailed discussion). If the Negishi map is monotonic, then the equilibrium is unique, otherwise not. We reproduce below the Negishi map that prevails under our calibration. Clearly, the Negishi map is monotonically increasing and the equilibrium under the calibration of Section 3.2 is therefore unique.

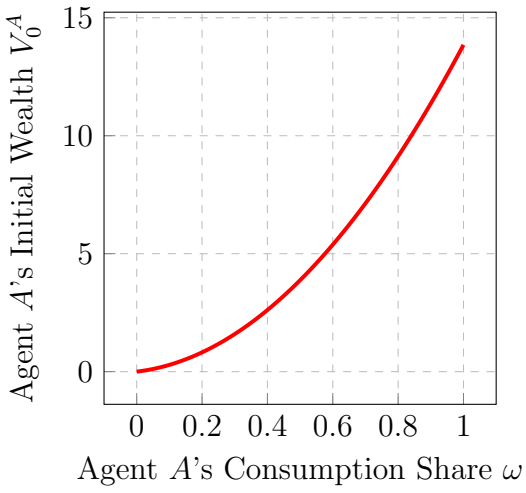


**Figure 10: Negishi Map.**

Negishi map of Agent $A$'s initial wealth as a function of her initial consumption share under the calibration of Table 1 and the assumption that $\widehat{f}_0^A = \bar{f}$, $g_0 = \bar{f} - (f^l + \frac{\psi}{\lambda+\psi}(f^h - f^l))$, and $\delta_0 = 1$.

## A.5 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. Assume that Assumptions 1 and 2 hold and 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

$$\mu(X) := \begin{pmatrix} \kappa\overline{f} \\ (\lambda+\psi)f_\infty \end{pmatrix} + \begin{pmatrix} -\kappa & 0 \\ \frac{f^2}{\sigma_\delta^2} & -\left(\lambda+\psi+\frac{f^2}{\sigma_\delta^2}\right) \end{pmatrix} X_t + \phi(X_t) \tag{65}$$

and

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

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

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

in the otherwise affine drift in (65). To take this nonlinearity into account, we augment the vector $X_t$ with the quadratic term $(\widehat{f}^B)^2$ and accordingly define 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

$$\mu(Y) := \begin{pmatrix} \kappa\overline{f} \\ (\lambda+\psi)f_\infty \\ \frac{f^4}{\sigma_\delta^2} \end{pmatrix} + \begin{pmatrix} -\kappa & 0 & 0 \\ \frac{f^2}{\sigma_\delta^2} & -(\lambda+\psi)+\frac{f^2}{\sigma_\delta^2} & 0 \\ 0 & 2f(\psi-\lambda) & \frac{-4f^2}{\sigma_\delta^2} - 2(\lambda+\psi) \end{pmatrix} Y_t + o(\widehat{f}_t^A\widehat{f}_t^B, (\widehat{f}_t^B)^3).$$

Performing a second-order Taylor expansion around the initial point $X_0$, we can write

$$\mu^{(2)}(Y_t) = \mu(Y_t)|_{X_0} + \nabla_X\mu(Y)|_{X_0}(X_t - X_0) + \frac{1}{2}(X_t - X_0)^\top \nabla_{XX}\mu(Y)\Big|_{X_0}(X_t - X_0) \equiv \Lambda + \Omega Y_t \tag{66}$$

where the vector $\Lambda$ satisfies

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

and where the matrix $\Omega$ satisfies

$$\Omega = \begin{pmatrix} -\kappa & 0 & 0 \\ \frac{(f-x_0)f_\infty\sigma_\delta(f+x_0)}{\sigma_\delta^2} & -\frac{f^2+3x_0^2+\sigma_\delta^2(\lambda+\psi)}{\sigma_\delta^2} & \frac{2x_0}{\sigma_\delta^2} \\ \frac{2(f-x_0)x_0(f+x_0)}{\sigma_\delta^2} & \frac{2\left(-9x_0^3+f^2x_0+f_\infty\sigma_\delta^2(\lambda+\psi)\right)}{\sigma_\delta^2} & -\frac{2\left(2f^2-6x_0^2+\sigma_\delta^2(\lambda+\psi)\right)}{\sigma_\delta^2} \end{pmatrix}.$$

Denote by $X_t^{(2)}$ the vector associated with the resulting approximated drift in (66). By construction, its drift is affine in $Y$ and its conditional expectation therefore satisfies

$$E_0^A\left[X_t^{(2)}\right] = -\Omega^{-1}\Lambda + \exp(\Omega t)(Y_0 + \Omega^{-1}\Lambda). \tag{67}$$

Furthermore, for $t$ small, the expression in (67) is approximately given by

$$\frac{\mathrm{d}}{\mathrm{d}t}E_0^A\left[X_t^{(2)}\right]\Big|_{t=\epsilon} \approx \Omega(I+\Omega\epsilon)(Y_0+\Omega^{-1}\Lambda).$$

Eliminating terms of order $o(\sigma_\delta^4)$, we can finally write

$$\frac{\mathrm{d}}{\mathrm{d}t}E_0^A\left[X_t^{(2)}\right]\Big|_{t=\epsilon} \approx \begin{pmatrix} \kappa(\overline{f}-x_0) \\ (\lambda+\psi)(f_\infty-x_0) \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{v(x_0)}{\sigma_\delta}\left(2x_0\frac{v(x_0)}{\sigma_\delta} + \kappa(\overline{f}-x_0) - (\lambda+\psi)(f_\infty-x_0)\right) \end{pmatrix}\epsilon. \tag{68}$$

Reorganizing yields the expressions in Proposition 5.

The nonlinearity of the change of measure introduces a nonlinear term $2x_0\frac{v(x_0)}{\sigma_\delta}$ in the approximate expression for Agent $B$'s speed of learning in Equation (68). This term 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 11.

## A.6 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

$$\log\left(\frac{\delta_{t+1}}{\delta_t}\right) = f_t^A + \sqrt{v^\delta}\epsilon_{1,t+1} \tag{69}$$

$$f_{t+1}^A = m^{f^A} + a^{f^A}f_t^A + \sqrt{v^{f^A}}\epsilon_{2,t+1}, \tag{70}$$

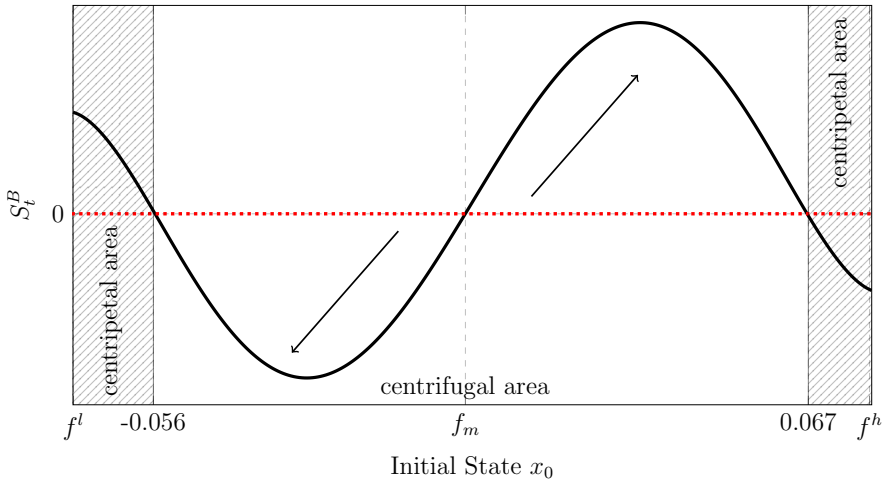


**Figure 11: Agent $B$'s Adjustment Speed under $\mathbb{P}^A$.**

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 $f_m$.

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

$$\log\left(\frac{\delta_{t+1}}{\delta_t}\right) = f_t^B + \sqrt{v^\delta}\epsilon_{3,t+1} \tag{71}$$

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

$\epsilon_1$, $\epsilon_2$, 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 (69) and (71) 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 fun-

| 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&P 500 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.

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

$$\begin{aligned}
\log\left(\frac{\delta_{t+\Delta}}{\delta_t}\right) &= \int_t^{t+\Delta}\left(f_u^A - \frac{1}{2}\sigma_\delta^2\right)\mathrm{d}u + \sigma_\delta\left(W_{t+\Delta}^A - W_t^A\right) \\
&= \int_t^{t+\Delta}\left(f_u^B - \frac{1}{2}\sigma_\delta^2\right)\mathrm{d}u + \sigma_\delta\left(W_{t+\Delta}^B - W_t^B\right) \\
&\approx \left(f_t^A - \frac{1}{2}\sigma_\delta^2\right)\Delta + \sigma_\delta\left(W_{t+\Delta}^A - W_t^A\right) \qquad (72) \\
&\approx \left(f_t^B - \frac{1}{2}\sigma_\delta^2\right)\Delta + \sigma_\delta\left(W_{t+\Delta}^B - W_t^B\right) \qquad (73)
\end{aligned}$$

$$f_{t+\Delta}^A = e^{-\kappa\Delta}f_t^A + \bar{f}\left(1 - e^{-\kappa\Delta}\right) + \sigma_f\int_t^{t+\Delta} e^{-\kappa(t+\Delta-u)}\mathrm{d}W_u^f. \qquad (74)$$

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

follows

$$P = \begin{pmatrix} p^{hh} & 1-p^{hh} \\ 1-p^{ll} & p^{ll} \end{pmatrix} \tag{75}$$
$$= \begin{pmatrix} \frac{\psi}{\lambda+\psi} + \frac{\lambda}{\lambda+\psi}e^{-(\lambda+\psi)\Delta} & \frac{\lambda}{\lambda+\psi} - \frac{\lambda}{\lambda+\psi}e^{-(\lambda+\psi)\Delta} \\ \frac{\psi}{\lambda+\psi} - \frac{\psi}{\lambda+\psi}e^{-(\lambda+\psi)\Delta} & \frac{\lambda}{\lambda+\psi} + \frac{\psi}{\lambda+\psi}e^{-(\lambda+\psi)\Delta} \end{pmatrix}.$$

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

Matching Equation (69) to Equation (72) and Equation (70) to Equation (74) yields the following system of equation for $\kappa$, $\bar{f}$, $\sigma_\delta$, and $\sigma_f$

$$\begin{aligned} a^{f^A} &= e^{-\kappa\Delta} \\ m^{f^A} &= \bar{f}\left(1-e^{-\kappa\Delta}\right) - \frac{1}{2}\sigma_\delta^2\Delta \\ v^\delta &= \sigma_\delta^2\Delta \\ v^{f^A} &= \frac{\sigma_f^2}{2\kappa}\left(1-e^{-2\kappa\Delta}\right), \end{aligned} \tag{76}$$

where the last equation relates the variance of the Ornstein-Uhlenbeck process to its empirical counterpart. Matching Equation (71) to Equation (73) yields the following system of equations for $f^l$ and $f^h$

$$\begin{aligned} s^h &= \left(f^h - \frac{1}{2}\sigma_\delta^2\right)\Delta \\ s^l &= \left(f^l - \frac{1}{2}\sigma_\delta^2\right)\Delta. \end{aligned} \tag{77}$$

Solving the system comprised of Equations (75), (76), and (77) yields the parameters presented in Table 1.

□

## A.7 Proof of Proposition 6

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

$$V_t = \delta_t \omega_t^{\alpha} \sum_{j=0}^{\alpha-1} \binom{\alpha-1}{j} \left(\frac{1-\omega_t}{\omega_t}\right)^j \mathbb{E}_t^{\mathbb{P}^A} \left[\int_t^{\infty} e^{-\rho(u-t)} \left(\frac{\eta_u}{\eta_t}\right)^{\frac{j}{\alpha}} \left(\frac{\delta_u}{\delta_t}\right)^{1-\alpha} \mathrm{d}u\right]$$
$$\overset{(\alpha=2)}{=} \delta_t \omega_t^2 \left.\frac{S_t}{\delta_t}\right|_{O.U.} + \delta_t \omega_t (1-\omega_t) F\left(\widehat{f}_t^A, g_t\right). \tag{78}$$

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

$$dV_t = r_t^f V_t \mathrm{d}t + \left(\mu_t - r_t^f\right) Q_t S_t \mathrm{d}t - c_{At}\mathrm{d}t + \sigma_t Q_t S_t \mathrm{d}\widehat{W}_t^A, \tag{79}$$

where $r^f$ is the risk-free rate defined in Equation (9), $\mu - r^f$ 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 (78), 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 (79) and the expression above, the number of shares $Q$ satisfies

$$Q_t = \frac{\mu_t - r_t^f}{\sigma_t^2}\frac{V_t}{S_t} + \frac{1}{\xi_t \sigma_t S_t}\mathbb{E}_t^{\mathbb{P}^A}\left(\int_t^{\infty} \mathcal{D}_t\left(\xi_s c_{As}\right) \mathrm{d}s\right). \tag{80}$$

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 (80) 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 (17) 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

$$
\mathbb{E}_t^{\mathbb{P}^A}\left(R(t,s)\right) := \mathbb{E}_t^{\mathbb{P}^A}\left(\frac{\mathcal{D}_t\xi_s}{\xi_s}\right), \tag{81}
$$

we decompose the Malliavin derivative of the stochastic discount factor as

$$
\mathcal{D}_t\xi_s = -\alpha\frac{\xi_s}{\delta_s}\mathcal{D}_t\delta_s + \frac{\xi_s(1-\omega(\eta_s))}{\eta_s}\mathcal{D}_t\eta_s \tag{82}
$$

with

$$
\mathcal{D}_t\delta_s = \delta_s\left(\sigma_\delta + \int_t^s \mathcal{D}_t\widehat{f}_v^A\mathrm{d}v\right) \tag{83}
$$

$$
\mathcal{D}_t\eta_s = -\frac{\eta_s}{\sigma_\delta}\left(g_t + \int_t^s \mathcal{D}_t g_v\mathrm{d}\widehat{W}_v^A + \frac{1}{\sigma_\delta}\int_t^s g_v\mathcal{D}_t g_v\mathrm{d}v\right) \tag{84}
$$

$$
\mathcal{D}_t\widehat{f}_s^A = \frac{\gamma}{\sigma_\delta}e^{-\kappa(s-t)} \tag{85}
$$

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 (33), respectively, and the operator $\nabla$ stands for the gradient.

Substituting Eq. (85) in Eq. (83) 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. (83), (84), and (85) in Eq. (82) yields

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

Taking conditional expectations at time $t$ and setting $g_t = 0$ implies that (81) satisfies

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

□

## A.8 Details on the Decomposition of the Stock Return Diffusion

We have

$$\begin{aligned}
\sigma_t &= \sigma_\delta + \frac{1}{\sigma_\delta S_t} \left[\gamma \underbrace{\frac{\partial S}{\partial \widehat{f}^A}}_{<0} + \left(\gamma - \left(f^h + g_t - \widehat{f}_t^A\right)\left(\widehat{f}_t^A - g_t - f^l\right)\right) \underbrace{\frac{\partial S}{\partial g}}_{>0} + \frac{\omega_t(1-\omega_t) g_t}{\alpha \sigma_\delta} \underbrace{\frac{\partial S}{\partial \omega}}_{<0}\right] \\
&= \left[\sigma_\delta + \frac{\gamma}{\sigma_\delta} \underbrace{\frac{\partial S}{\partial \widehat{f}^A}}_{<0} \frac{1}{S_t}\right] + \left[\frac{1-\omega_t}{\sigma_\delta}\left(-2\widehat{f}_t^A g_t + g_t^2 + \left(\widehat{f}_t^A\right)^2 + \gamma + f^l f^h - (f^l + f^h)\left(\widehat{f}_t^A - g_t\right)\right) \underbrace{\frac{\partial S}{\partial \widehat{g}}}_{>0} \frac{1}{S_t}\right] \\
&\quad + \left[\frac{\omega_t(1-\omega_t) g_t}{\alpha \sigma_\delta^2} \underbrace{\frac{\partial S}{\partial \omega}}_{<0} \frac{1}{S_t}\right] \\
&\approx A_1 + A_2(1-\omega_t)\widehat{f}_t^A g_t + A_3 \omega_t (1-\omega_t) g_t,
\end{aligned}$$

where $A_1 \approx 0$, $A_2 < 0$, and $A_3 < 0$ are constants and $\widehat{g} \equiv (1-\omega)g$ is the "consumption-weighted" disagreement. The third "equality" holds because with our calibration: *1)* $f^l + f^h \approx 0$, *2)* the joint distribution of $\widehat{f}^A$, $g$, and $\omega$ implies that both the partial derivatives

scaled by the price and $\frac{1-\omega}{\sigma_\delta}\left(g^2+\left(\widehat{f}^A\right)^2+\gamma+f^l f^h\right)$ have a small volatility. Combining *1)* and *2)* shows that the first term is almost constant, $(1-\omega)\widehat{f}^A g$ is the main driver of the variation in the second term, and $\omega(1-\omega)g$ is the main driver of the variation in the third term.

To confirm the accuracy of the approximation we simulate the economy 1,000 times over a 100-year horizon and regress the diffusion, $\sigma$, on both $(1-\omega)\widehat{f}^A g$ and $\omega(1-\omega)g$. We obtain a median regression $R^2$ of 86%, lending support to our approximation.

## A.9 Model-implied time series momentum vs. dispersion: rolling window approach (Section 5.2.2)

An alternative approach to that described in Section 5.2.2 consists in computing time series momentum at a 1-month lag, $\beta_M(1)_j$, and running the following regression over 36−month rolling windows

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

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:[25]

$$\beta_M(1)\ t\text{-stat}_j = \alpha + \beta AG_j + \epsilon_j. \tag{86}$$

The coefficient $\beta$ measures the sensitivity of time series momentum (at a 1-month lag) to a change in aggregate dispersion. It is computed using 1,000 simulations over a 100-year horizon.

The coefficient $\beta = 6,636$ is significant at the 1% confidence level, which shows that momentum at a 1-month lag increases with dispersion in our model.

## A.10 Empirical time series momentum vs. dispersion: rolling window approach (Section 5.2.2)

To provide empirical evidence of the positive relation between dispersion and time series momentum, we run the empirical equivalent to the regression in (86). Specifically, we measure the empirical sensitivity $\beta_{emp}$ of 1-month time series momentum to a change in dispersion

[25] Note that the results are qualitatively similar if we substitute the weighted dispersion $G_t \equiv \int_{t-\Delta}^t g_u^2(1-\omega_u)^2\omega_u du$ by the regular dispersion $G_t \equiv \int_{t-\Delta}^t g_u^2 du$.

by substituting the weighted dispersion $G$ and the excess return $r^e$ by their empirical counterparts $G_{emp} \equiv Disp \times (1 - omega)^2 \times omega$ and, $r^e_{emp}$, the monthly excess returns on the S&P 500, respectively.[26]

Consistent with the predictions of the model, the coefficient $\beta_{emp} = 2.0094$ is positive and significant at the 5% confidence level. That is, 1-month time series momentum increases with dispersion.

## A.11 Additional Figures

### A.11.1 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 3.1 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 11 (Appendix A.5). Figure 11 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 12 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 12), 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 12), 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 2 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

[26] Note that the results are qualitatively similar if we substitute the weighted dispersion $G_{emp} \equiv Disp \times (1 - omega)^2 \times omega$ by the regular dispersion $G_{emp} \equiv Disp$.

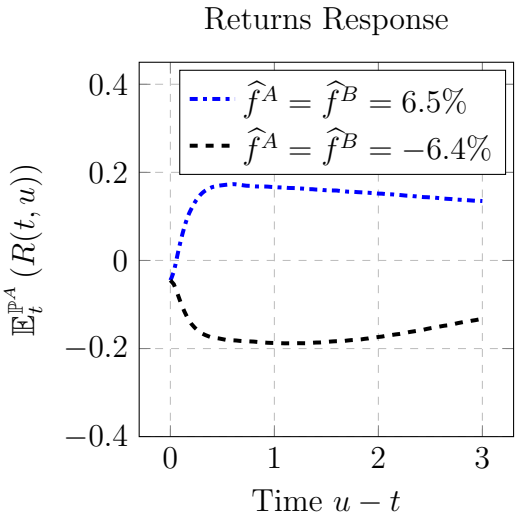


**Figure 12: Model-Implied Impulse Response of Excess Returns to a News Shock in the two Knife-Edge Cases.**

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 $\widehat{f}^A = \widehat{f}^B = -6.4\%$.

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 2, 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.11.2 Model-implied coefficient values of time series momentum (Section 4.3)

Figure 13 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.11.3 Model-implied 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 14 above, which plots the unconditional pattern of time series momentum in our model (computed over a 100-year horizon), shows, however, that this is not the case. Specifically, it

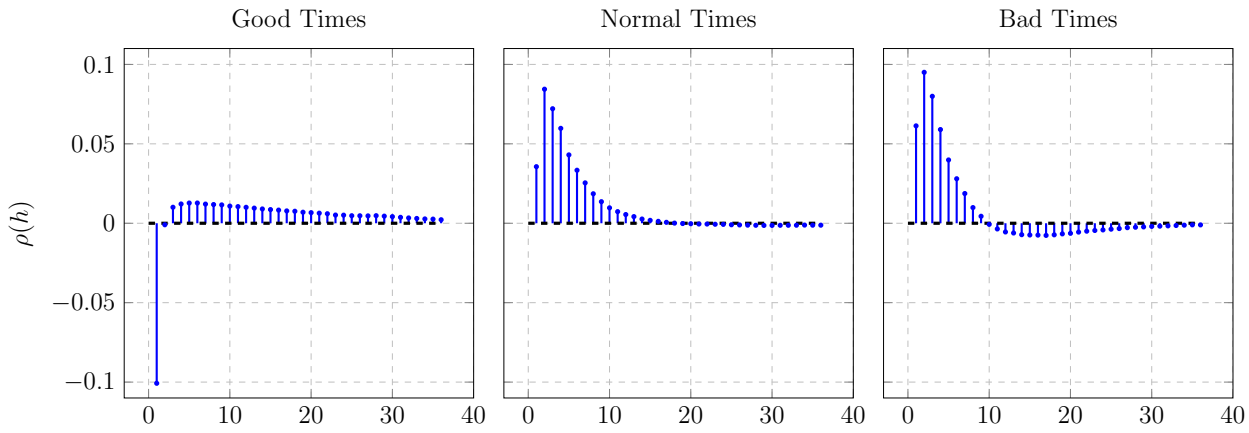


**Figure 13: Model-Implied Conditional Time Series Momentum.**
This figure plots the time series momentum coefficient $\rho(h)$ for lags $h$ ranging from 1 month to 3 years. Each panel corresponds to a different state of the economy. The values reported above are obtained from 10,000 simulations of the economy over a 20-year horizon.

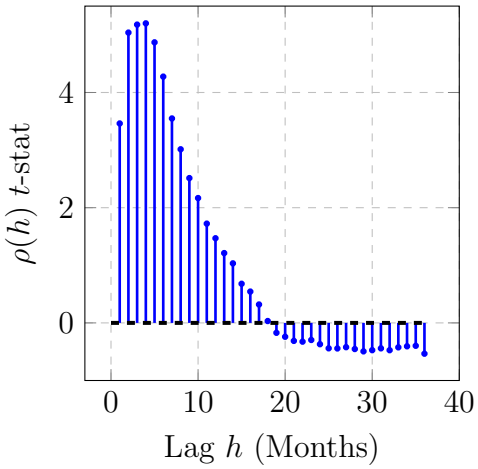


**Figure 14: Model-Implied Unconditional Time Series Momentum.**
This figure plots the $t$-statistics of the coefficient $\rho(h)$ for lags $h$ ranging from 1 month to 3 years. Standard errors are adjusted using Newey and West (1987). The values reported above are obtained from 1,000 simulations of the economy over a 100-year horizon.

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).

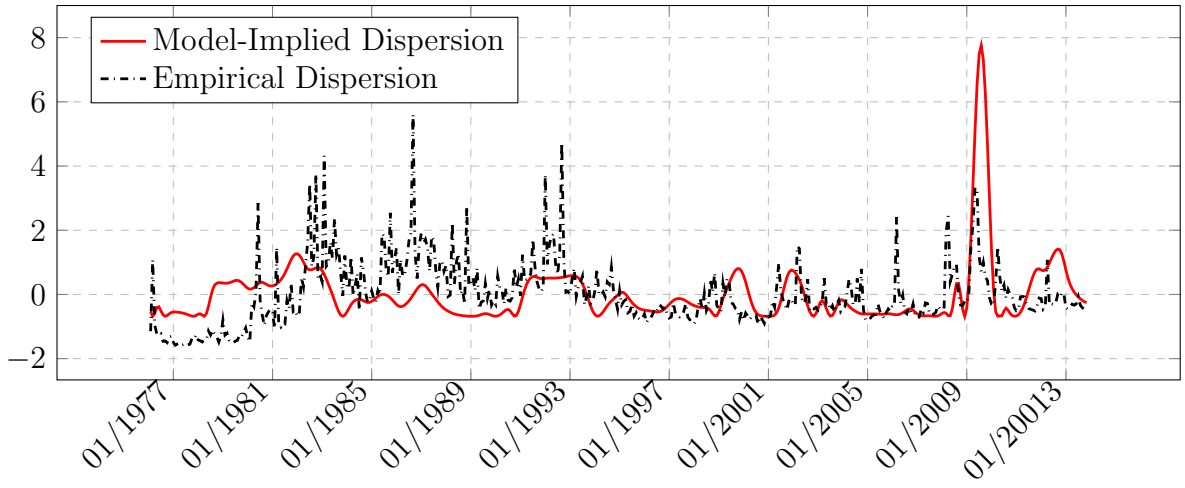


**Figure 15: Empirical Dispersion and Model-Implied Dispersion.**
This figure plots the (standardized) analysts' forecasts dispersion in dash-dotted black and the (standardized) model-implied dispersion estimated in Section 3 in solid red. Data are at monthly frequency from 02/1976 to 11/2013.

### A.11.4 Empirical dispersion vs. model-implied dispersion (Section 5)

Figure 15 plots the standardized analysts' forecasts dispersion in dash-dotted black and the standardized model-implied dispersion estimated in Section 3 in solid red. Standardization is performed in order to get both time series on the same scale. The figure provides evidence that both time series are positively correlated. Indeed, the correlation coefficient between the two series is equal to 0.2539 and it is significant at the 1% confidence level.

### A.11.5 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

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

where $\Delta = 1/12 = 1$ month, $r^e_{emp}$ denotes the monthly excess returns on the S&P 500, and $Z_{t,G^{emp}}(p)$ is a dummy variable that takes value 1 when the monthly weighted dispersion, $G_{emp} \equiv Disp \times (1 - omega)^2 \times omega$, is smaller than its $p-$th percentile. The coefficient $\beta_2(p)$ measures excess time series momentum in low dispersion period, whereas the sum $\beta_1(p) + \beta_2(p)$ measures 1-month time series momentum in low dispersion periods. Figure 16 shows that the data lend support to the prediction of the model. Indeed, we observe

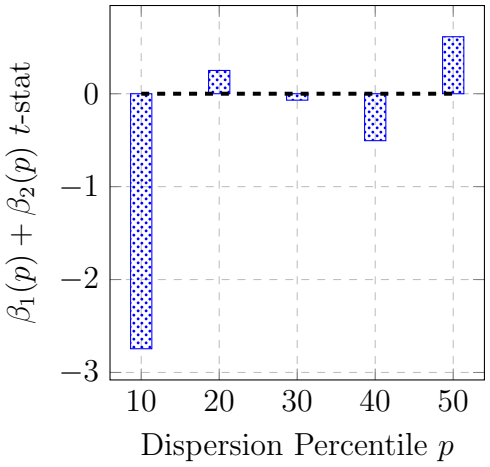


**Figure 16: Empirical Time Series Momentum in Low Disagreement Periods.** This figure plots the $t$-statistics of 1-month time series momentum $\beta_1(p) + \beta_2(p)$ when weighted dispersion is smaller than its $p$-th percentile. Standard errors are adjusted using Newey and West (1987). Data are at monthly frequency from 02/1976 to 11/2013.

significant 1-month time series reversal during low dispersion periods, particularly so when the dispersion threshold is the 10−th percentile.

### A.11.6 Empirical 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:

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

where $\Delta = 1/12 = 1$ month, $r^e_{emp}$ denotes the monthly excess returns on the S&P 500, and $X$ 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 17. 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.

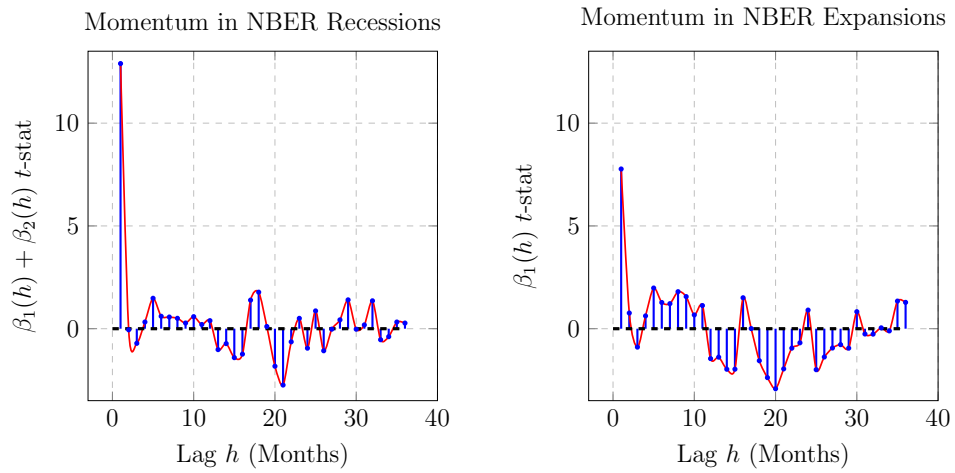


**Figure 17: Empirical Time Series Momentum in NBER Recessions and Expansions.**

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). Monthly data range from 01/1871 to 11/2013.

### A.11.7 Additional theoretical prediction: U-shaped relation between time series momentum and excess returns

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 time series momentum and excess returns 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 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,

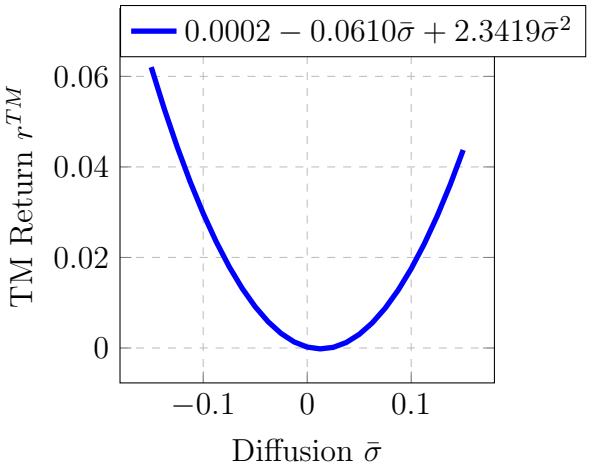


**Figure 18: Model-Implied Time Series Momentum in Extreme Markets.** This figure plots the relation between time series momentum returns and the stock return diffusion in our model. The relation is obtained from 1,000 simulations of the economy over a 100-year horizon.

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 18. It confirms the intuition that the returns on time series momentum are high during extreme markets.