Internet Appendix for "Why Does Return Predictability Concentrate in Bad Times?"

JULIEN CUJEAN and MICHAEL HASLER*

ABSTRACT

This document provides *1)* the proofs of the propositions, *2)* the derivations of the equations, *3)* details on the calibration, and *4)* theoretical and empirical results that complement those found in the paper "Why Does Return Predictability Concentrate in Bad Times?'

I. Proof of Proposition 1

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

$$\frac{\mathrm{d}\delta_t}{\delta_t} = \left(A_0 + A_1 f_t^A\right) \mathrm{d}t + B_1 \mathrm{d}W_t^f + B_2 \mathrm{d}W_t^A$$

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 $b \circ b = b_1 b_1^\top + b_2 b_2^\top = \sigma_f^2$, $B \circ B = B_1 B_1^\top + B_2 B_2^\top = \sigma_\delta^2$, and $b \circ B = b_1 B_1^\top + b_2 B_2^\top = 0$. Applying Theorem

*Citation format: Cujean, Julien and Michael Hasler, Internet Appendix for "Why Does Return Predictability Concentrate in Bad Times?" *Journal of Finance* DOI STRING. Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the authors of the article.

12.7 in Lipster and Shiryaev (2001b), the dynamics of the filter satisfy

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

where the steady-state posterior variance γ solves the algebraic equation

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

Substituting the coefficients yields equation (4), the steady-state posterior variance γ, and

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

II. 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 IA1: *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 $\mathscr{F}_t \ \forall t \in \mathbb{R}_+$.*

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 (8) is a strictly positive martingale, that is, $E^{\widehat{\mathbb{P}}^A}[\eta_t] = 1$ 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 (8) is a martingale. To do so, we write the dynamics of agents' disagreement under $\mathbb{P}^A$ as

$$\mathrm{d}g_t = \left[\kappa\left(\bar{f} - \widehat{f}_t^A\right) - (\lambda+\psi)\left(f_\infty + g_t - \widehat{f}_t^A\right) - \frac{g_t}{\sigma_\delta^2}\left(\widehat{f}_t^A - g_t - f^l\right)\left(f^h + g_t - \widehat{f}_t^A\right)\right] dt$$

$$+ \frac{\gamma - \left(\widehat{f}_t^A - g_t - f^l\right)\left(f^h + g_t - \widehat{f}_t^A\right)}{\sigma_\delta}\mathrm{d}\widehat{W}_t^A \tag{IA.1}$$

and use the following result, which we formulate in Theorem IA1.

THEOREM IA1: *The process η defined in* (8) *is a true martingale (as opposed to a local martingale) if and only if the process g defined in* (IA.1) *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 (IA.1) has a unique nonexplosive strong solution under $\widehat{\mathbb{P}}^A$ and $\widehat{\mathbb{P}}^B$. Rewrite the process g_t in (IA.1) as

$$\mathrm{d}g_t = \left(\mu(\widehat{f}_t^B) - \lambda(\widehat{f}_t^B)g_t\right)\mathrm{d}t + \sigma(\widehat{f}_t^B)\mathrm{d}\widehat{W}_t^A \tag{IA.2}$$

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

$$\mathrm{d}g_t = \left(\mu(\widehat{f}_t^B) - \left(\kappa + \frac{\gamma}{\sigma_\delta^2}\right)g_t\right)\mathrm{d}t + \sigma(\widehat{f}_t^B)\mathrm{d}\widehat{W}_t^B \tag{IA.3}$$

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

$$\mu : (f^l, f^h) \to \Big(\kappa(\bar{f} - f^h) - (\lambda + \psi)(f_\infty - f^l), \kappa(\bar{f} - f^l) - (\lambda + \psi)(f_\infty - f^h)\Big)$$

$$\sigma : (f^l, f^h) \to \left(\frac{1}{\sigma_\delta}(\gamma - \frac{1}{4}(f^h - f^l)^2), \frac{\gamma}{\sigma_\delta}\right)$$
$$\lambda : (f^l, f^h) \to \left(\kappa, \kappa + \frac{1}{4\sigma_\delta^2}(f^h - f^l)^2\right)$$

are defined as

$$\begin{aligned}\mu(x) &:= \kappa(\bar{f} - x) - (\lambda + \psi)(f_\infty - x) \\ \sigma(x) &:= \frac{1}{\sigma_\delta}(\gamma - (x - f^l)(f^h - x)) \\ \lambda(x) &:= \kappa + \frac{1}{\sigma_\delta^2}(x - f^l)(f^h - x).\end{aligned} \tag{IA.4}$$

We then have the following result, which we highlight in Lemma IA1.

LEMMA IA1: *The processes in* (IA.2) *and* (IA.3) *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 (IA.2) by setting

$$g_t^{(k+1)} := g_t^{(0)} + \int_0^t \left(\mu(\widehat{f}_s^{(k)}) - \lambda(\widehat{f}_s^{(k)}) g_s^{(k)}\right) \mathrm{d}s + \int_0^t \sigma(\widehat{f}_s^{(k)}) \mathrm{d}\widehat{W}_s^A \tag{IA.5}$$

for $k \geq 0$, where $\widehat{f}^{(k)}$ denotes some approximation of $\widehat{f}^B$. From (IA.5), 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 (IA.4) are locally Lipschitz continuous. In particular, for all $x, y \in (f^l, f^h)$, we have

$$|\sigma(x) - \sigma(y)| = \frac{1}{\sigma_\delta}|x-y||f^h + f^l - (x+y)| \leq \frac{1}{\sigma_\delta} \max\left\{|f^h - f^l|, |f^l - f^h|\right\} |x-y| \tag{IA.6}$$

and

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

Lipschitz continuity for $\lambda(\cdot)$ follows directly from (IA.6). Furthermore, $\lambda(\cdot)$ is bounded. We can therefore let $f_t^{(k)} \equiv \widehat{f}_t^B$ for all $k \geq 0$ and write

$$g_t^{(k+1)} - g_t^{(k)} = -\int_0^t \lambda(\widehat{f}_s^B)\left(g_s^{(k)} - g_s^{(k-1)}\right) \mathrm{d}s. \tag{IA.7}$$

Taking the absolute value of both sides of (IA.7) and observing that $\lambda(x) > 0$

for all $x \in (f^l, f^h)$, we have

$$\left| g_t^{(k+1)} - g_t^{(k)} \right| \leq \int_0^t \lambda(\widehat{f}_s^B) \left| g_s^{(k)} - g_s^{(k-1)} \right| \mathrm{d}s \leq \overline{\lambda} \int_0^t \left| g_s^{(k)} - g_s^{(k-1)} \right| \mathrm{d}s, \tag{IA.8}$$

where

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

Iterating over (IA.8), 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 follows directly 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 (IA.2) with $g_0^1 = g_0^2$, $\widehat{\mathbb{P}}^A-$almost surely. It is then sufficient to show that g^1 and g^2 are indistinguishable. Using (IA.2), 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 the 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 IA2. 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 IA2: *At any time $t \in \mathbb{R}_+$, the process g_t defined in* (IA.1) *is finite $\widehat{\mathbb{P}}^A-$ and $\widehat{\mathbb{P}}^B-$almost surely, that is,*

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

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

$$\mathrm{d}A_t = \mu(\widehat{f}_t^B) e^{\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}\mathrm{d}t + e^{\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}\sigma(\widehat{f}_t^B)\mathrm{d}\widehat{W}_t^A, \quad A_0 = g_0, \tag{IA.9}$$

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

$$\mathrm{d}A_t^i = (-1)^i m_t^A \mathrm{d}t + e^{\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}\sigma(\widehat{f}_t^B)\mathrm{d}\widehat{W}_t^A, \quad (-1)^i A_0^i \geq (-1)^i A_0, \tag{IA.10}$$

under $\widehat{\mathbb{P}}^A$. Combining (IA.9) and (IA.10), we obtain

$$A_t^i - A_t = A_0^i - A_0 + \int_0^t \left((-1)^i m_s^A - \mu(\widehat{f}_s^B) e^{\int_0^s \lambda(\widehat{f}_u^B)\mathrm{d}u}\right) \mathrm{d}s, \quad i =, 1, 2. \tag{IA.11}$$

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

$$(-1)^i e^{\int_0^t \lambda(\widehat{f}_u^B)\mathrm{d}u} \mu(\widehat{f}_t^B) \le e^{\int_0^t \lambda(\widehat{f}_u^B)\mathrm{d}u} |\mu(\widehat{f}_t^B)| \le m_t^A, \quad i = 1, 2, \tag{IA.12}$$

for all $t \in \mathbb{R}_+$. The inequalities in (IA.12) and the expressions in (IA.11) together imply that

$$A_t^1 \le A_t \le A_t^2, \quad \widehat{\mathbb{P}}^A - \text{almost surely,}. \tag{IA.13}$$

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

$$|g_t| = e^{-\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}(A_t^+ + A_t^-) \leq A_t^+ + A_t^-, \quad \widehat{\mathbb{P}}^A - \text{almost surely,} \tag{IA.14}$$

where the second inequality follows from the fact 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 (IA.13) and (IA.14), we obtain

$$|g_t| \leq \sum_{i=1,2} ((-1)^i A_t^i)^+ \leq \max\{(-A_t^1)^+, (A_t^2)^+\}, \quad \widehat{\mathbb{P}}^A - \text{almost surely.} \tag{IA.15}$$

Using (IA.15), we can write that, for any positive constant $c \geq 0$,

$$\mathbf{1}_{|g_t|\geq c} \leq \sum_{i=1,2} \mathbf{1}_{((-1)^i A_t^i)^+ \geq c} \equiv \sum_{i=1,2} \mathbf{1}_{(-1)^i A_t^i \geq c}, \tag{IA.16}$$

where the last equality follows from the fact that $\mathbf{1}_{X^+ \geq c} = \mathbf{1}_{X \geq c}$ for any c positive. Taking expectations of (IA.16) under $\widehat{\mathbb{P}}^A$, we get

$$\widehat{\mathbb{P}}^A(|g_t| \geq c) \leq \sum_{i=1,2} \widehat{\mathbb{P}}^A\left((-1)^i A_t^i \geq c\right), \quad \forall t \in \mathbb{R}_+. \tag{IA.17}$$

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

$$\mathrm{d}\widehat{A}_t^i = (-1)^i m_t^A \mathrm{d}t + \sigma_t^A \mathrm{d}\widehat{W}_t^A, \quad (-1)^i \widehat{A}_0^i \geq (-1)^i A_0^i, \tag{IA.18}$$

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

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

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 $\widehat{A}^i$, that is,

$$a^{i,j} \sim \widehat{A}^i, \quad i, j = 1, 2. \tag{IA.19}$$

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

$$\mathrm{d}\bar{a}_t^i = (-1)^i m_t^A \mathrm{d}t + e^{\int_0^t \lambda(\widehat{f}_s^B)\mathrm{d}s}\sigma(\widehat{f}_t^B)\mathrm{d}\widehat{W}_t^A. \tag{IA.20}$$

Combining (IA.10) and (IA.20), we obtain from the initial conditions that

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

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 (IA.19), we obtain that, for any $c \in \mathbb{R}$, the processes A^i and $\widehat{A}^i$, $i = 1,2$, satisfy

$$\mathbb{P}^A((-1)^i A_t^i \geq c) \leq 2\mathbb{P}^A((-1)^i \widehat{A}_t^i \geq c), \quad \forall t \in \mathbb{R}_+. \tag{IA.21}$$

Combining the inequalities in (IA.17) and (IA.21), we obtain

$$\widehat{\mathbb{P}}^A(|g_t| \geq c) \leq 2\sum_{i=1,2} \widehat{\mathbb{P}}^A\left((-1)^i \widehat{A}_t^i \geq c\right), \quad \forall t \in \mathbb{R}_+. \tag{IA.22}$$

Observing that each $\widehat{A}^i$, $i = 1,2$, in (IA.18) is a Gaussian process, the probabilities on the right-hand side of (IA.22) are given explicitly by

$$\mathbb{P}^A\left((-1)^i\widehat{A}_t^i \geq c\right) = \Phi\left(\frac{(-1)^i\widehat{A}_0^i + \int_0^t m_s^A \mathrm{d}s - c}{\sqrt{\int_0^t (\sigma_s^A)^2 \mathrm{d}s}}\right). \qquad \text{(IA.23)}$$

Taking limits on both sides of (IA.22) and using (IA.23) 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 η in (8) 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 $\mathscr{F}_t \quad \forall t \in \mathbb{R}_+$ and the claim follows from Girsanov Theorem (Theorem 5.1, Karatzas and Shreve (1988)).

III. 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 ϕ_A denotes the Lagrange multiplier of Agent A's static budget constraint and ξ 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

$$c_{At} = \left(\phi_A e^{\rho t}\xi_t\right)^{-\frac{1}{\alpha}} \qquad\qquad c_{Bt} = \left(\frac{\phi_B e^{\rho t}\xi_t}{\eta_t}\right)^{-\frac{1}{\alpha}}. \tag{IA.24}$$

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

$$\xi_t = e^{-\rho t}\delta_t^{-\alpha}\left[(1/\phi_A)^{1/\alpha} + (\eta_t/\phi_B)^{1/\alpha}\right]^{\alpha}. \tag{IA.25}$$

Substituting equation (IA.25) into equation (IA.24) gives the consumption share of Agent A, ω, 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 η. In particular, an increase in η raises the likelihood of Agent B's model relative to Agent A's model. The consumption share of Agent A is therefore decreasing in η. 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 ω satisfy

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

The dynamics of the state-price density ξ 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 (IA.25) determines the risk-free rate r^f and the market price of risk θ provided in Proposition 3.

IV. Proof of Proposition 4

Following Dumas, Kurshev, and Uppal (2009), we assume that the coefficient of relative risk aversion, α, is an integer. This assumption allows us to obtain the following convenient expression for the equilibrium stock price:[1]

$$\begin{aligned} \frac{S_t}{\delta_t} &= \mathbb{E}_t^{\mathbb{P}^A}\left[\int_t^\infty \frac{\xi_u \delta_u}{\xi_t \delta_t} du\right] \\ &= \omega_t^\alpha \sum_{j=0}^{\alpha} \begin{pmatrix} \alpha \\ j \end{pmatrix} \left(\frac{1-\omega_t}{\omega_t}\right)^j \mathbb{E}_t^{\mathbb{P}^A}\left[\int_t^\infty e^{-\rho(u-t)}\left(\frac{\eta_u}{\eta_t}\right)^{\frac{j}{\alpha}}\left(\frac{\delta_u}{\delta_t}\right)^{1-\alpha} du\right]. \end{aligned} \tag{IA.26}$$

[1]We refer the reader to Dumas, Kurshev, and Uppal (2009) for details of the derivation.

We start by computing the first and the last terms of the sum in (IA.26). These terms correspond to the prices of a Lucas (1978) economy in which the representative agent assumes that the fundamental follows an Ornstein-Uhlenbeck process and a two-state Markov chain, respectively. These prices have (semi) closed-form solutions, which we present in Proposition IA1.

PROPOSITION IA1: *Suppose the economy is populated by a single agent.*

1. *If the agent's filter follows the Ornstein-Uhlenbeck process described in equation* (4)*, the 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 two-state Markov chain process described in equation* (6)*, 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

$$H = \begin{pmatrix} H_1 & H_2 \end{pmatrix}^\top = A^{-1}\mathbf{1}_2$$

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

$\mathbf{Id}_2$ *is a two-by-two identity matrix and* $\mathbf{1}_2$ *is a two-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 system of Ricatti equations

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

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

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

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 (IA.26) as

$$
\frac{S_t}{\delta_t} = \omega_t^\alpha \left.\frac{S_t}{\delta_t}\right|_{O.U.} + \omega_t^\alpha \sum_{j=1}^{\alpha-1} \begin{pmatrix} \alpha \\ j \end{pmatrix} \left(\frac{1-\omega_t}{\omega_t}\right)^j F^j\left(\widehat{f}_t^A, g_t\right) + (1-\omega_t)^\alpha \left.\frac{S_t}{\delta_t}\right|_{M.C.}. \tag{IA.27}
$$

The last step involves computing the intermediate terms F^j in (IA.27), which relate to heterogeneous beliefs. Each term solves a differential equation, which we present in Proposition IA2.

PROPOSITION IA2: *The function F^j, defined as*

$$
F^j\left(\widehat{f}_t^A, g_t\right) \equiv \mathbb{E}_t^{\mathbb{P}^A}\left[\int_t^\infty e^{-\rho(u-t)}\left(\frac{\eta_u}{\eta_t}\right)^{\frac{j}{\alpha}}\left(\frac{\delta_u}{\delta_t}\right)^{1-\alpha} du\right], \tag{IA.28}
$$

solves the partial differential equation

$$\widetilde{\mathscr{L}}^{\widehat{f}^A,g}F^j + X^j F^j + 1 = 0, \tag{IA.29}$$

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

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

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

$$\overline{W}_t = \widehat{W}_t^A + \int_0^t \frac{j}{\alpha}\frac{g_s}{\sigma_\delta}\mathrm{d}s, \tag{IA.30}$$

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

$$\widetilde{W}_t = \overline{W}_t - \int_0^t (1-\alpha)\,\sigma_\delta \mathrm{d}t. \tag{IA.31}$$

Implementing sequentially the changes of probability measures in equations (IA.30) and (IA.31) allows us to rewrite the interior expectation in (IA.28) as

$$F^j\left(\widehat{f}_t, g_t\right) = \mathbb{E}^{\widetilde{P}}\left[\int_t^\infty e^{\int_t^u X_s^j \mathrm{d}s}\mathrm{d}u \middle| \mathscr{F}_t\right], \tag{IA.32}$$

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 (IA.32) can be rewritten as

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

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 (IA.29).[2]

■

We numerically solve equation (IA.29) for each term j through Chebyshev collocation. In particular, we approximate the functions $F^j(\widehat{f}^A, g)$ for

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

$j = 1, \ldots, \alpha - 1$ 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 (IA.29), 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 II.B. 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. Below we reproduce the Negishi map that prevails under our calibration. Clearly, the Negishi map is monotonically increasing and the equilibrium under the calibration of Section II.B is therefore unique.

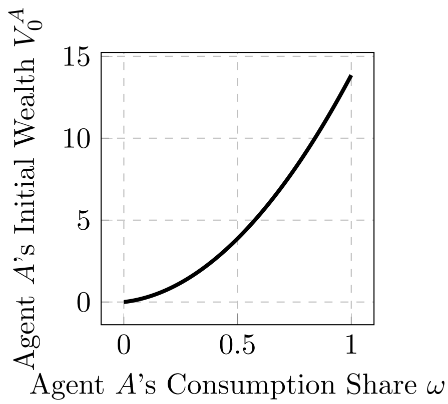

Figure IA.1. Negishi map. Negishi map of Agent A's initial wealth as a function of her initial consumption share under the calibration of Table I and the assumption that $\widehat{f}_0^A = \bar{f}$, $g_0 = \bar{f} - (f^l + \frac{\psi}{\lambda+\psi}(f^h - f^l))$, and $\delta_0 = 1$.

V. Approximation of the Filter's Adjustment Speed

In this appendix, we derive an approximation for the adjustment speed of agents' filters, 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

$$\mu(X) := \begin{pmatrix} \kappa\overline{f} \\ (\lambda+\psi)f_\infty \end{pmatrix} + \begin{pmatrix} -\kappa & 0 \\ \frac{(f^\star)^2}{\sigma_\delta^2} & -\left(\lambda+\psi+\frac{(f^\star)^2}{\sigma_\delta^2}\right) \end{pmatrix} X_t + \phi(X_t) \tag{IA.33}$$

and

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

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 (IA.33). 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 $Y_t := (\widehat{f}_t^A, \widehat{f}_t^B, (\widehat{f}_t^B)^2)^\top$. Application of Ito's lemma shows that the drift of this process satisfies

$$\mu(Y) := \begin{pmatrix} \kappa\overline{f} \\ (\lambda+\psi)f_\infty \\ \frac{(f^\star)^4}{\sigma_\delta^2} \end{pmatrix} + \begin{pmatrix} -\kappa & 0 & 0 \\ \frac{(f^\star)^2}{\sigma_\delta^2} & -(\lambda+\psi) + \frac{(f^\star)^2}{\sigma_\delta^2} & 0 \\ 0 & 2f^\star(\psi-\lambda) & \frac{-4(f^\star)^2}{\sigma_\delta^2} - 2(\lambda+\psi) \end{pmatrix} Y_t$$

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

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

$$\begin{aligned}\mu^{(2)}(Y_t) = \mu(Y_t)|_{X_0} + \nabla_X \mu(Y)|_{X_0} (X_t - X_0) \\ + \frac{1}{2} (X_t - X_0)^\top \nabla_{XX} \mu(Y)\Big|_{X_0} (X_t - X_0) \equiv \Lambda + \Omega Y_t,\end{aligned} \tag{IA.34}$$

where the vector Λ satisfies

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

and the matrix Ω satisfies

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

Denote by $X_t^{(2)}$ the vector associated with the resulting approximated drift in (IA.34). By construction, its drift is affine in Y and its conditional expectation therefore satisfies

$$E_0^A \left[X_t^{(2)} \right] = -\Omega^{-1}\Lambda + \exp(\Omega t)(Y_0 + \Omega^{-1}\Lambda). \tag{IA.35}$$

Furthermore, for t small, the expression in (IA.35) 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

$$\frac{\mathrm{d}}{\mathrm{d}t} E_0^A \left[X_t^{(2)} \right] \Big|_{t=\epsilon} \approx \begin{pmatrix} \kappa(\overline{f} - x_0) \\ (\lambda + \psi)(f_\infty - x_0) \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{v(x_0)}{\sigma_\delta} \left(2x_0 \frac{v(x_0)}{\sigma_\delta} + \kappa(\overline{f} - x_0) - (\lambda + \psi)(f_\infty - x_0) \right) \end{pmatrix} \epsilon. \tag{IA.36}$$

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 (IA.36). 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 IA.2.

VI. Derivation of the Parameter Values Provided in Table I

Agents first estimate a discretized version of their model and then map the parameters they estimated into their continuous-time model. In partic-

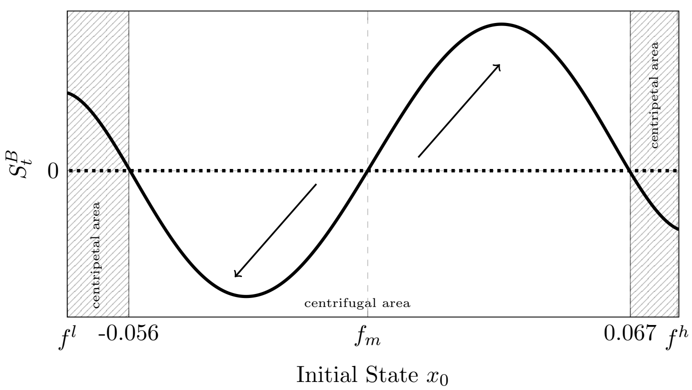

Figure IA.2. Agent B's adjustment speed under $\mathbb{P}^A$. This figure plots the second-order approximation of Agent B's adjustment speed under $\mathbb{P}^A$ as a function of the state of the economy. The dashed areas represent the states in which Agent B's expectations are attracted towards f_m, while the central area represents the states in which Agent B's expectations are repelled outward f_m.

ular, Agent A estimates the discrete-time model

$$\log\left(\frac{\delta_{t+1}}{\delta_t}\right) = f_t^A + \sqrt{v^\delta}\epsilon_{1,t+1} \tag{IA.37}$$

$$f_{t+1}^A = m^{f^A} + a^{f^A} f_t^A + \sqrt{v^{f^A}}\epsilon_{2,t+1}, \tag{IA.38}$$

while Agent B estimates the discrete-time model

$$\log\left(\frac{\delta_{t+1}}{\delta_t}\right) = f_t^B + \sqrt{v^\delta}\epsilon_{3,t+1} \tag{IA.39}$$

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

Table IA.I

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 denoted by *, **, and ***, respectively.

Parameter	Symbol	Estimate
Variance Dividend Growth	v^{δ}	$4.23 \times 10^{-5***}$
		(1.48×10^{-6})
Persistence Growth Rate f^A	a^{f^A}	0.9842^{***}
		(0.0022)
Mean Growth Rate f^A	m^{f^A}	$9.96 \times 10^{-4***}$
		(9.39×10^{-5})
Variance Growth Rate f^A	v^{f^A}	$2.53 \times 10^{-6***}$
		(2.12×10^{-7})
High State of f^B	s^h	0.0066^{***}
		(2.7×10^{-4})
Low State of f^B	s^l	-0.0059^{***}
		(3.17×10^{-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)

ϵ_1, ϵ_2, and ϵ_3 are normally distributed with zero mean and unit variance, and ϵ_1 and ϵ_2 are independent. The transition matrix P contains the probabilities of staying in the high state and the low state over the following month. We estimate the discrete-time models in equations (IA.37) and (IA.39) by Maximum Likelihood. We report the estimated parameters, their standard errors, and their statistical significance in Table IA.I.

We next map the parameters of Table IA.I into the associated continuous-time models. Straightforward applications of Itô's lemma show that the dividend stream, δ, and the fundamental perceived by Agent A, f^A, satisfy

$$\begin{aligned}
\log\left(\frac{\delta_{t+\Delta}}{\delta_t}\right) &= \int_t^{t+\Delta}\left(f_u^A - \frac{1}{2}\sigma_\delta^2\right)\mathrm{d}u + \sigma_\delta\left(W_{t+\Delta}^A - W_t^A\right) \\
&= \int_t^{t+\Delta}\left(f_u^B - \frac{1}{2}\sigma_\delta^2\right)\mathrm{d}u + \sigma_\delta\left(W_{t+\Delta}^B - W_t^B\right) \\
&\approx \left(f_t^A - \frac{1}{2}\sigma_\delta^2\right)\Delta + \sigma_\delta\left(W_{t+\Delta}^A - W_t^A\right) \qquad \text{(IA.40)} \\
&\approx \left(f_t^B - \frac{1}{2}\sigma_\delta^2\right)\Delta + \sigma_\delta\left(W_{t+\Delta}^B - W_t^B\right) \qquad \text{(IA.41)}
\end{aligned}$$

$$f_{t+\Delta}^A = e^{-\kappa\Delta} f_t^A + \bar{f}\left(1 - e^{-\kappa\Delta}\right) + \sigma_f \int_t^{t+\Delta} e^{-\kappa(t+\Delta-u)}\mathrm{d}W_u^f. \qquad \text{(IA.42)}$$

The relationship between the transition matrix P and the generator matrix Λ is written as

$$\begin{aligned}
P &= \begin{pmatrix} p^{hh} & 1-p^{hh} \\ 1-p^{ll} & p^{ll} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\psi}{\lambda+\psi} + \frac{\lambda}{\lambda+\psi}e^{-(\lambda+\psi)\Delta} & \frac{\lambda}{\lambda+\psi} - \frac{\lambda}{\lambda+\psi}e^{-(\lambda+\psi)\Delta} \\ \frac{\psi}{\lambda+\psi} - \frac{\psi}{\lambda+\psi}e^{-(\lambda+\psi)\Delta} & \frac{\lambda}{\lambda+\psi} + \frac{\psi}{\lambda+\psi}e^{-(\lambda+\psi)\Delta} \end{pmatrix}. \qquad \text{(IA.43)}
\end{aligned}$$

We perform the Maximum Likelihood estimation on monthly data and accordingly set $\Delta =$ one month.

Matching equation (IA.37) to equation (IA.40) and equation (IA.38) to equation (IA.42) yields the following system of equations for κ, $\bar{f}$, σ_δ, and

σ_f:

$$
\begin{aligned}
a^{f^A} &= e^{-\kappa\Delta} \\
m^{f^A} &= \bar{f}\left(1 - e^{-\kappa\Delta}\right) - \frac{1}{2}\sigma_\delta^2\Delta \\
v^{\delta} &= \sigma_\delta^2\Delta \\
v^{f^A} &= \frac{\sigma_f^2}{2\kappa}\left(1 - e^{-2\kappa\Delta}\right),
\end{aligned}
\tag{IA.44}
$$

where the last equation relates the variance of the Ornstein-Uhlenbeck process to its empirical counterpart. Matching equation (IA.39) to equation (IA.41) yields the following system of equations for f^l and f^h:

$$
\begin{aligned}
s^h &= \left(f^h - \frac{1}{2}\sigma_\delta^2\right)\Delta \\
s^l &= \left(f^l - \frac{1}{2}\sigma_\delta^2\right)\Delta.
\end{aligned}
\tag{IA.45}
$$

Solving the system comprised of equations (IA.43), (IA.44), and (IA.45) yields the parameters presented in Table I.

VII. Proof of Proposition 6

Following the methodology in Dumas, Kurshev, and Uppal (2009), Agent A's wealth, V, satisfies

$$V_t = \delta_t \omega_t^{\alpha} \sum_{j=0}^{\alpha-1} \begin{pmatrix} \alpha - 1 \\ j \end{pmatrix} \left(\frac{1-\omega_t}{\omega_t} \right)^j \mathbb{E}_t^{\mathbb{P}^A} \left[\int_t^{\infty} e^{-\rho(u-t)} \left(\frac{\eta_u}{\eta_t} \right)^{\frac{j}{\alpha}} \left(\frac{\delta_u}{\delta_t} \right)^{1-\alpha} \mathrm{d}u \right]$$

$$\overset{(\alpha=2)}{=} \delta_t \omega_t^2 \left. \frac{S_t}{\delta_t} \right|_{O.U.} + \delta_t \omega_t (1-\omega_t) F\left(\widehat{f}_t^A, g_t \right). \tag{IA.46}$$

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

$$dV_t = r_t^f V_t \mathrm{d}t + \left(\mu_t - r_t^f \right) Q_t S_t \mathrm{d}t - c_{At} \mathrm{d}t + \sigma_t Q_t S_t \mathrm{d}\widehat{W}_t^A, \tag{IA.47}$$

where r^f is the risk-free rate defined in equation (11), $\mu - r^f$ is the risk premium on the stock, Q is the number of shares held by Agent A, and σ is the diffusion of stock returns. Applying Ito's lemma to Agent A's discounted wealth using (IA.46), we obtain the martingale

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

where the first and second equalities follow from the Martingale Representation Theorem and the Clark-Ocone Theorem, respectively. Matching the diffusion terms in (IA.47) and the expression above, the number of shares Q satisfies

$$Q_t = \frac{\mu_t - r_t^f}{\sigma_t^2}\frac{V_t}{S_t} + \frac{1}{\xi_t \sigma_t S_t}\mathbb{E}_t^{\mathbb{P}^A}\left(\int_t^\infty \mathscr{D}_t\left(\xi_s c_{As}\right)\mathrm{d}s\right). \tag{IA.48}$$

Finally, using the fact that

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

we can rewrite equation (IA.48) as follows

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

The expression for the state-price density in (19) is derived in Internet Appendix Section III.

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

$$\mathbb{E}_t^{\mathbb{P}^A}\left(R(t,s)\right) := \mathbb{E}_t^{\mathbb{P}^A}\left(\frac{\mathscr{D}_t\xi_s}{\xi_s}\right), \tag{IA.49}$$

we decompose the Malliavin derivative of the stochastic discount factor as

$$\mathscr{D}_t\xi_s = -\alpha\frac{\xi_s}{\delta_s}\mathscr{D}_t\delta_s + \frac{\xi_s(1-\omega(\eta_s))}{\eta_s}\mathscr{D}_t\eta_s, \tag{IA.50}$$

with

$$\mathscr{D}_t\delta_s = \delta_s\left(\sigma_\delta + \int_t^s \mathscr{D}_t\widehat{f}_v^A \mathrm{d}v\right) \tag{IA.51}$$

$$\mathscr{D}_t\eta_s = -\frac{\eta_s}{\sigma_\delta}\left(g_t + \int_t^s \mathscr{D}_t g_v \mathrm{d}\widehat{W}_v^A + \frac{1}{\sigma_\delta}\int_t^s g_v\mathscr{D}_t g_v \mathrm{d}v\right) \tag{IA.52}$$

$$\mathscr{D}_t\widehat{f}_s^A = \frac{\gamma}{\sigma_\delta}e^{-\kappa(s-t)} \tag{IA.53}$$

and where

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

with initial condition $\mathscr{D}_t g_t = \sigma_g(\widehat{f}_t^A, g_t)$. The coefficients μ_g and σ_g represent the drift and the diffusion of disagreement in (IA.1), respectively, and the operator ∇ stands for the gradient. Substituting equation (IA.53) into equation (IA.51) yields

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

while substituting equations (IA.51), (IA.52), and (IA.53) into equation (IA.50) yields

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

Taking conditional expectations at time t and setting $g_t = 0$ implies that (IA.49) satisfies

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

VIII. Details on the Decomposition of the Stock Return Diffusion

We have

$$\sigma_t = \sigma_\delta + \frac{1}{\sigma_\delta S_t}\left[\gamma \underbrace{\frac{\partial S}{\partial \widehat{f}^A}}_{<0} + \left(\gamma - \left(f^h + g_t - \widehat{f}_t^A\right)\left(\widehat{f}_t^A - g_t - f^l\right)\right)\underbrace{\frac{\partial S}{\partial g}}_{>0} + \frac{\omega_t(1-\omega_t)g_t}{\alpha\sigma_\delta}\underbrace{\frac{\partial S}{\partial \omega}}_{<0}\right]$$

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

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)* $f^l + f^h \approx 0$, and *2)* the joint distribution of $\widehat{f}^A$, g, and ω 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, σ, 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.

IX. Model-Implied Time-Series Momentum vs. Dispersion: Rolling Window Approach (Section IV.B.2)

An alternative approach to that described in Section IV.B.2 involves computing time-series momentum at a one-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:[3]

$$\beta_M(1)\ t\text{-stat}_j = \alpha + \beta AG_j + \epsilon_j. \tag{IA.54}$$

The coefficient β measures the sensitivity of time-series momentum (at a one-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 one-month lag increases with dispersion in our model.

X. Empirical Time-Series Momentum vs. Dispersion: Rolling Window Approach (Section IV.B.2)

To provide empirical evidence of the positive relation between dispersion and time-series momentum, we run the empirical equivalent to the regression in (IA.54). Specifically, we measure the empirical sensitivity of one-month time-series momentum to a change in dispersion, β_{emp}, 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.[4]

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

[4]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 $G_{emp} \equiv Disp$.

Consistent with the predictions of the model, the coefficient $\beta_{emp} = 0.796$ is positive and significant at the 5% confidence level (std. error = 0.324). That is, one-month time-series momentum increases with dispersion.

XI. Additional Figures

A. Robustness of the Main Result to Knife-Edge Cases (Section II.C)

The dynamics of disagreement are consistent, irrespective of the starting point within each region defined in Table II, 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 in Section II.A we show 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 IA.2 (Internet Appendix Section V). Figure IA.2 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 III.B and in Figure IA.3 plot the response of future state-price densities in the two knife-edge cases. Comparing the response that prevails in good times (the dash-dotted line in the left panel of Figure 7) to the response that prevails in the knife-edge case in which opinions polarize in good times (the dash-dotted line in Figure IA.3), we observe that in both cases returns adjust immediately to the news (the two responses have the same sign and shape). Similarly, comparing the response that prevails in bad times (the dashed line in the left panel of Figure 7) to the response that prevails in the knife-edge case in which opinions stop polarizing in bad times (the dashed line in the left panel of Figure IA.3), we observe that in both cases returns underreact 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 reassesses uncertainty in a way that varies greatly over the business cycle, as the confidence interval depicted in Figure 4 demonstrates. Specifically, the left panel shows that the variance of Agent B's filter is small in good times. As a result, while opinions polarize in good times in the first knife-edge case, the polarization of beliefs is so small that it does not generate a spike in disagreement. The right panel of Figure 4 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.

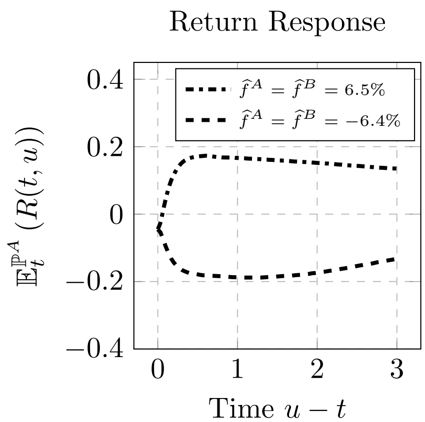

Figure IA.3. Model-implied impulse response of excess returns to a news shock in the two knife-edge cases. The top knife-edge region is such that $\widehat{f}^A = \widehat{f}^B = 6.5\%$, whereas the bottom knife-edge region is such that $\widehat{f}^A = \widehat{f}^B = -6.4\%$.

B. Model-Implied Coefficient Values of Time-Series Momentum (Section III.C)

Figure IA.4 depicts the time-series momentum coefficient $\rho(h)$ for lags h ranging from one month to three years. Each panel corresponds to a different state of the economy. The t-statistics are provided in Section III.C.

C. Model-Implied Unconditional Time-Series Momentum Pattern

Our analysis of time-series momentum in Section III.C is conditional on the state of the economy (good, normal, and bad times). Since excess returns are negatively serially correlated at short horizons in good times, the unconditional pattern of serial correlation may be inconsistent with Moskowitz, Ooi, and Pedersen (2012). However, the left panel of Figure IA.5, which plots the unconditional pattern of time-series momentum in our

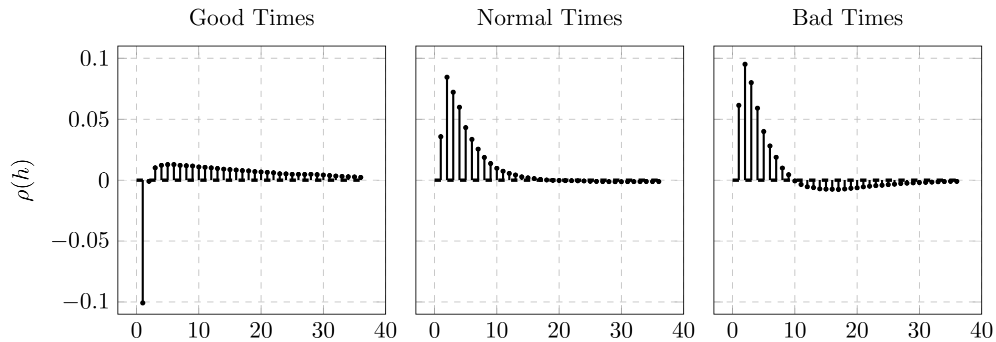

Figure IA.4. Model-implied conditional time-series momentum. This figure plots the time-series momentum coefficient $\rho(h)$ for lags h ranging from one month to three years. Each panel corresponds to a different state of the economy. The values reported above are obtained from 10,000 simulations of the economy over a 20-year horizon.

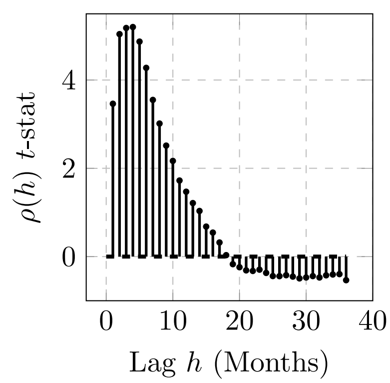

Figure IA.5. Model-implied unconditional time-series momentum. This figure plots the t-statistics of the coefficient $\rho(h)$ for lags h ranging from one month to three years. Standard errors are adjusted using Newey and West (1987). The values reported above are obtained from 1,000 simulations of the economy over a 100-year horizon.

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

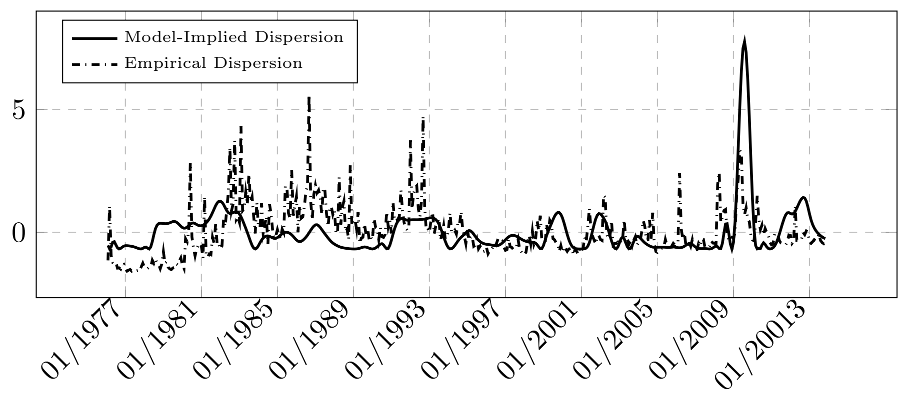

Figure IA.6. Empirical dispersion and model-implied dispersion. This figure plots the (standardized) analysts' forecast dispersion in the dash-dotted line and the (standardized) model-implied dispersion estimated in Section II in the solid line. Data are at a monthly frequency from 02/1976 to 11/2013.

D. Empirical Dispersion vs. Model-Implied Dispersion (Section IV)

Figure IA.6 plots the standardized analysts' forecast dispersion in the dash-dotted line and the standardized model-implied dispersion estimated in Section II in the solid line. 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 is significant at the 1% confidence level.

E. Empirical Pattern of Time-Series Momentum in Periods of Low Dispersion (Section IV.B.2)

To verify that we observe, as predicted by the model, short-term time-series reversal in low dispersion periods, we run the 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 Δ = one month, r^e_{emp} denotes monthly excess returns on the S&P 500, and $Z_{t,G^{emp}}(p)$ is a dummy variable that takes the value of one 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 one-month time-series momentum in low dispersion periods. Figure IA.7 shows that the data lend support to the prediction of the model. Indeed, we observe 1-month time series reversal during low dispersion periods, particularly when the dispersion threshold is the 30^{th} percentile.

F. 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 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 Δ = one month, r^e_{emp} denotes monthly excess returns on the S&P 500, and X is a dummy variable that takes the value of one in NBER reces-

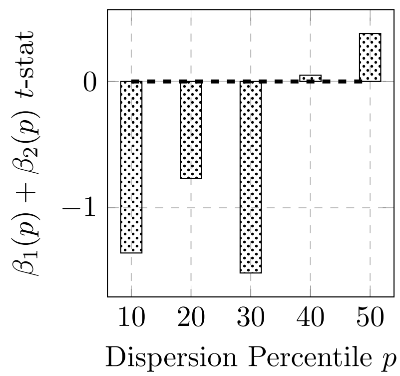

Figure IA.7. Empirical time-series momentum in low disagreement periods. This figure plots the t-statistics of one-month time-series momentum $\beta_1(p) + \beta_2(p)$ when weighted dispersion is smaller than its p^{th} percentile. Standard errors are adjusted using Newey and West (1987). Data are at a monthly frequency from 02/1976 to 11/2013.

sions. 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 IA.8. 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 on average there is time-series momentum up to the 12-month lag, followed by reversal over subsequent horizons.

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

In our model, excess returns become extreme only 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, Ooi, and Pedersen (2012). More generally, the U-shaped relation between time-

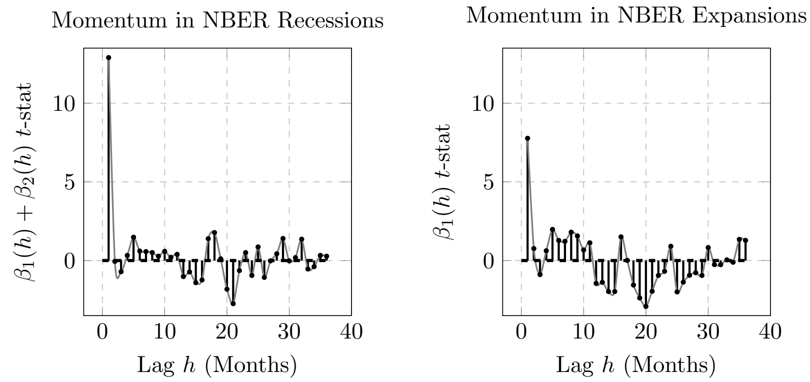

Figure IA.8. Empirical time-series momentum in NBER recessions and expansions. The left and right panels plot the t-statistics of the coefficient $\beta_1(h) + \beta_2(h)$ and $\beta_1(h)$, respectively, for lags h ranging from one month to three years. Standard errors are adjusted using Newey and West (1987). Data are at a monthly frequency from 01/1871 to 11/2013.

series momentum and excess returns in Moskowitz, Ooi, and Pedersen (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 regression

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

where Δ = one month, $\bar{\sigma}_t = \int_{t-\Delta}^{t} \sigma_u du$ is the monthly diffusion of stock

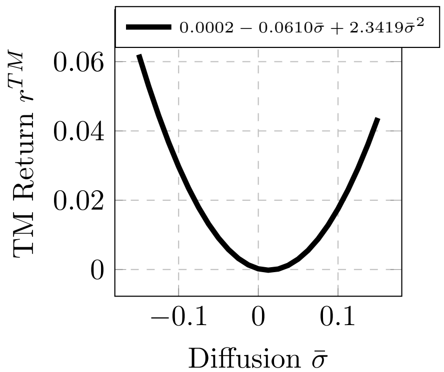

Figure IA.9. Model-implied time-series momentum in extreme markets. This figure plots the relation between time-series momentum returns and the stock return diffusion in our model. The relation is obtained from 1,000 simulations of the economy over a 100-year horizon.

returns, and $r^{TM}_{t+\Delta} = sign\,(r^e_t)\, r^e_{t+\Delta}$ is the return on time-series momentum, as defined in Moskowitz, Ooi, and Pedersen (2012). If the sign of excess returns, and hence the sign of the diffusion, does not matter, then the first coefficient, β_1, should be insignificant. If its magnitude matters, however, the second coefficient, β_2, should be significantly positive. The t-statistics of the intercept, linear coefficient, β_1, and quadratic coefficient, β_2, are 0.6693, −0.7672, and 1.7402, respectively. Consistent with Moskowitz, Ooi, and Pedersen (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 IA.9. This confirms the intuition that the returns on time-series momentum are high during extreme markets.

REFERENCES

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

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343–1376.

Dumas, Bernard, Alexander Kurshev, and Raman Uppal, 2009, Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility, *Journal of Finance* 64, 579–629.

Dumas, Bernard, and Andrew Lyasoff, 2012, Incomplete-market equilibria solved recursively on an event tree, *Journal of Finance* 67, 1897–1941.

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

Heston, Steven L., Mark Loewenstein, and Gregory A. Willard, 2007, Options and bubbles, *Review of Financial Studies* 20, 359–390.

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

Karatzas, Ioannis, and Steven E. Shreve, 1988, *Brownian Motion and Stochastic Calculus* (Springer).

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

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

Lucas, Jr., Robert E., 1978, Asset prices in an exchange economy, *Econometrica* 46, 1429–1445.

Moskowitz, Tobias J., Yao Hua Ooi, and Lasse Heje Pedersen, 2012, Time series momentum, *Journal of Financial Economics* 104, 228–250.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

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

Veronesi, Pietro, 2000, How does information quality affect stock returns?, *Journal of Finance* 55, 807–837.

Yamada, Toshio, and Shinzo Watanabe, 1971, On the uniqueness of solutions of stochastic differential equations, *Journal of Mathematics of Kyoto University* 11, 155–167.