For Online Publication

Online Appendix for:

A New Test of Excess Movement in Asset Prices*

Ned Augenblick
UC Berkeley Haas

Eben Lazarus
MIT Sloan

August 2022

Contents

*Contact: ned@haas.berkeley.edu and elazarus@mit.edu.

Appendix B. Additional Proofs of Theoretical Results

The proofs for Propositions 1–2 and Corollaries 1–2 are provided in the main paper in Appendix A. We provide proofs for the other results here.

B.1 Proofs for Section 2

Proof of Lemma 1. Following Augenblick and Rabin (2021), it is useful to define period-by-period movement, uncertainty reduction, and excess movement, respectively, as

$$\mathsf{m}_{t,t+1}(\boldsymbol{\pi}) \equiv (\pi_{t+1} - \pi_t)^2, \qquad \mathsf{r}_{t,t+1}(\boldsymbol{\pi}) \equiv \pi_t(1-\pi_t) - \pi_{t+1}(1-\pi_{t+1}),$$
$$X_{t,t+1}(\boldsymbol{\pi}) \equiv \mathsf{m}_{t,t+1}(\boldsymbol{\pi}) - \mathsf{r}_{t,t+1}(\boldsymbol{\pi}).$$

Given the definitions of movement, initial uncertainty, and excess movement in the text, note that

$$\mathsf{m}(\boldsymbol{\pi}) = \textstyle\sum_{t=0}^{T-1} \mathsf{m}_{t,t+1}(\boldsymbol{\pi}), \quad \mathsf{u}_0(\boldsymbol{\pi}) = \textstyle\sum_{t=0}^{T-1} \mathsf{r}_{t,t+1}(\boldsymbol{\pi}), \quad X(\boldsymbol{\pi}) = \textstyle\sum_{t=0}^{T-1} x_{t,t+1}(\boldsymbol{\pi}),$$

where the second equality relies on the fact that $\pi_T \in \{0,1\}$ and therefore $\pi_T(1-\pi_T) = 0$ for any belief stream $\boldsymbol{\pi}$. We have that

$$\begin{aligned}
\mathbb{E}[X_{t,t+1}|H_t] &= \mathbb{E}[\mathsf{m}_{t,t+1} - \mathsf{r}_{t,t+1}|H_t] = \mathbb{E}[(\pi_{t+1}-\pi_t)^2 - ((\pi_t(1-\pi_t) - (\pi_{t+1}(1-\pi_{t+1}))|H_t] \\
&= \mathbb{E}[(2\pi_t - 1)(\pi_t - \pi_{t+1})|H_t] = (2\pi_t(H_t) - 1)(\pi_t(H_t) - \mathbb{E}[\pi_{t+1}|H_t]) \\
&= (2\pi_t(H_t) - 1) \cdot 0 = 0,
\end{aligned}$$

where the last line uses Assumption 1. Summing and applying the law of iterated expectations (LIE),

$$\mathbb{E}[X] = \sum_{t=0}^{T-1} \mathbb{E}[X_{t,t+1}] = \sum_{t=0}^{T-1} \mathbb{E}[\mathbb{E}[X_{t,t+1}|H_t]] = 0. \qquad \square$$

Proof of Proposition 3. Consider a given ϕ, RN prior π_0^*, and signal DGPs $DGP(s_t|\theta = 0, H_{t-1})$ and $DGP(s_t|\theta = 1, H_{t-1})$ that lead to some $\mathbb{E}[X^*|\theta = 0]$, $\mathbb{E}[X^*|\theta = 1]$, and $\triangle$. Now consider the "reversed" DGP $\widehat{DGP}$ in which we modify the DGP by relabeling state 1 as state 0 and state 0 as state 1. That is, $\widehat{DGP}(s_t|\theta = 0, H_{t-1}) \equiv DGP(s_t|\theta = 1, H_{t-1})$ and $\widehat{DGP}(s_t|\theta = 1, H_{t-1}) \equiv DGP(s_t|\theta = 0, H_{t-1})$. Similarly, we consider the "reversed" RN prior $\widehat{\pi}_0^* = 1 - \pi_0^*$ implied by the physical prior $\widehat{\pi}_0 = \frac{1-\pi_t^*}{\phi + (1-\phi)(1-\pi_t^*)}$.

Given this relabeling, if the RN belief in the original DGP following history H_t is $\pi_t^*(H_t)$, then the RN belief in the reversed $\widehat{DGP}$ with RN prior $1 - \pi_0^*$ must be $\widehat{\pi}_t^*(H_t) = 1 - \pi_t^*(H_t)$. Thus $\mathbb{E}^*[\widehat{X}^*|\theta = 0] = \mathbb{E}^*[X^*|\theta = 1]$ and $\mathbb{E}^*[\widehat{X}^*|\theta = 1] = \mathbb{E}^*[X^*|\theta = 0]$. And since $\mathbb{E}^*[X^*|\theta] = \mathbb{E}[X^*|\theta]$ by Lemma A.1 (see Appendix A), $\mathbb{E}[\widehat{X}^*|\theta = 0] = \mathbb{E}[X^*|\theta = 1]$ and $\mathbb{E}[\widehat{X}^*|\theta = 1] = \mathbb{E}[X^*|\theta = 0]$. We conclude that for $\widehat{DGP}$, $\widehat{\triangle} \equiv \mathbb{E}[\widehat{X}^*|\theta = 0] - \mathbb{E}[\widehat{X}^*|\theta = 1] = -\triangle$. $\square$

Proof of Proposition 4. Consider a sequence of binary resolving DGPs indexed by T. There are two possible signals in each period, l and h, and assume that for any history,

$$DGP(s_t = h|\theta = 1) = 1, \tag{B.1}$$

$$DGP(s_t = h|\theta = 0) = \frac{\pi^*_{t-1}(1 - \pi^*_{t-1} - \epsilon)}{(1 - \pi^*_{t-1})(\pi^*_{t-1} + \epsilon)}, \quad \text{with } \epsilon \equiv \frac{1 - \pi^*_0}{T}. \tag{B.2}$$

Since $DGP(s_t = l|\theta = 1) = 0$ from (B.1), beliefs (both physical and RN) update to 0 given any l signal. Meanwhile, after seeing h (and assuming no l signals through $t-1$), Bayes' rule gives that physical beliefs update to

$$\pi_t(\{s_1 = h, \ldots, s_t = h\}) = \frac{\pi_{t-1}}{\pi_{t-1} + (1 - \pi_{t-1})DGP(s_t = h|\theta = 0)}.$$

Applying the transformation (8) to the π_{t-1} values on the right side of this equation,

$$\pi_t(\{s_1 = h, \ldots, s_t = h\}) = \frac{\pi^*_{t-1}}{\pi^*_{t-1} + (1 - \pi^*_{t-1})\phi DGP(s_t = h|\theta = 0)}.$$

Now applying the transformation (5), we obtain that π^*_t given an only-h signal history (suppressing the dependence on this history for simplicity) is, after additional algebra,

$$\pi^*_t = \frac{\pi^*_{t-1}}{\pi^*_{t-1} + (1 - \pi^*_{t-1})DGP(s_t = h|\theta = 0)}.$$

Now using (B.2), we obtain after further algebra that $\pi^*_t - \pi^*_{t-1} = \epsilon$. Given the definition of ϵ, this DGP is resolving for any T: given any l signal at any t, beliefs resolve to 0, while given only h signals, beliefs increase slowly ($\pi^*_t = \pi^*_0 + t\epsilon$) and resolve to 1 at period T. We thus have

$$\mathbb{E}[\mathsf{m}^*|\theta = 1] = T\epsilon^2 = T\left(\frac{1 - \pi^*_0}{T}\right)^2 = \frac{(1 - \pi^*_0)^2}{T} \xrightarrow{T \to \infty} 0.$$

Thus for such a sequence, using equation (A.8),

$$\triangle = \pi^*_0 - \frac{1}{1 - \pi^*_0} \cdot \mathbb{E}[\mathsf{m}^*|\theta = 1] \xrightarrow{T \to \infty} \pi^*_0.$$

Using this in equation (A.9) gives $\mathbb{E}[X^*] \to (\pi^*_0 - \pi_0)\pi^*_0$ as $T \to \infty$, as stated. And as further stated, the sequence of DGPs is constructed such that any downward movement is resolving and any upward movement is small ($\pi^*_t - \pi^*_{t-1} = \epsilon \to 0$). We have thus proven the first two statements.

For the final statement, given $\phi > 1$ and $0 < \pi^*_0 < 1$, the inequality in (A.10) is strict, so that $\pi^*_0 - \pi_0 > 0$. Further, the only way to obtain $\mathsf{m}^* = 0$ for finite T is if $\pi^*_0 = \pi^*_1 = \ldots = \pi^*_T$, which is ruled out by $0 < \pi^*_0 < 1$ since $\pi^*_T = 0$ or 1 with probability 1, and therefore $\mathbb{E}[\mathsf{m}^*|\theta = 1] > 0$. Thus in (A.8), we have the strict inequality $\triangle < \pi^*_0$ for fixed $T < \infty$. Combining these in (A.9) gives $\mathbb{E}[X^*] < (\pi^*_0 - \pi_0)\pi^*_0$ for fixed T, as stated. $\square$

B.2 Proofs for Section 3

Proof of Equation (13). This follows from a discrete-state application of Breeden and Litzenberger (1978), or see Brown and Ross (1991) for a general version. To review why the stated equation holds, the risk-neutral pricing equation for options can be written

$$q^m_{t,K} = \frac{1}{R^f_{t,T}} \mathbb{E}^*_t[\max\{V^m_T - K, 0\}] = \frac{1}{R^f_{t,T}} \Bigg[\sum_{j\,:\,K_j \geqslant K} (K_j - K) \underbrace{\mathbb{P}^*_t(V^m_T = K_j)}_{\mathbb{P}^*_t(R^m_T = \theta_j)} \Bigg].$$

This implies that for two adjacent return states θ_{j-1} and θ_j,

$$\begin{aligned} q^m_{t,K_j} - q^m_{t,K_{j-1}} &= \frac{1}{R^f_{t,T}} \left[\sum_{j' \geqslant j} (K_{j'} - K_j)\, \mathbb{P}^*_t(V^m_T = K_{j'}) - \sum_{j' \geqslant j-1} (K_{j'} - K_{j-1})\, \mathbb{P}^*_t(V^m_T = K_{j'}) \right] \\ &= \frac{1}{R^f_{t,T}} \left[\sum_{j' \geqslant j} (K_{j-1} - K_j)\, \mathbb{P}^*_t(V^m_T = K_{j'}) \right] = \frac{1}{R^f_{t,T}} (K_{j-1} - K_j) \left[1 - \mathbb{P}^*_t(V^m_T < K_j)\right]. \end{aligned}$$

Rearranging,

$$R^f_{t,T} \frac{q^m_{t,K_j} - q^m_{t,K_{j-1}}}{K_j - K_{j-1}} = \mathbb{P}^*_t(V^m_T < K_j) - 1.$$

Repeating this for θ_j and θ_{j+1}, we obtain $R^f_{t,T} \frac{q^m_{t,K_{j+1}} - q^m_{t,K_j}}{K_{j+1} - K_j} = \mathbb{P}^*_t(V^m_T < K_{j+1}) - 1$. Subtracting the preceding equation from this equation and using $\mathbb{P}^*_t(R^m_T = \theta_j) = \mathbb{P}^*_t(V^m_T = K_j)$ yields (13). $\square$

Before proceeding to the proofs of this section's main results, we provide two additional lemmas that are useful in proving those results. As usual, assume throughout that Assumptions 2–4 hold.

LEMMA B.1. *For some return-state pair (θ_j, θ_{j+1}), with $\widetilde{\mathbb{P}} \equiv \mathbb{P}(\cdot \mid R^m_T \in \{\theta_j, \theta_{j+1}\})$, define a new pseudo-risk-neutral measure $\widetilde{\mathbb{P}}^\diamond$ by*

$$\left.\frac{d\widetilde{\mathbb{P}}^\diamond}{d\widetilde{\mathbb{P}}}\right|_{H_t} = \frac{\widetilde{\pi}^*_{t,j}}{\widetilde{\pi}_{t,j}} \mathbb{1}\{R^m_T = \theta_j\} + \frac{1 - \widetilde{\pi}^*_{t,j}}{1 - \widetilde{\pi}_{t,j}} \mathbb{1}\{R^m_T = \theta_{j+1}\}. \tag{B.3}$$

Denote the conditional expectation under $\widetilde{\mathbb{P}}^\diamond$ by $\widetilde{\mathbb{E}}^\diamond_t[\,\cdot\,]$. If conditional transition independence holds for the return-state pair (θ_j, θ_{j+1}), and $\mathbb{P}_t(R^m_T \in \{\theta_j, \theta_{j+1}\}) > 0$, we have that $\widetilde{\mathbb{P}}^\diamond$ serves as a martingale measure for the risk-neutral belief in the sense that

$$\widetilde{\pi}^*_{t,j} = \widetilde{\mathbb{E}}^\diamond_t[\widetilde{\pi}^*_{t+1,j}]. \tag{B.4}$$

We conclude from Lemma 1 that

$$\widetilde{\mathbb{E}}^\diamond_0[X^*_j] = 0. \tag{B.5}$$

Proof of Lemma B.1. Following the discussion after equation (14), we have that

$$\frac{\widetilde{\pi}^*_{t,j}}{\widetilde{\pi}_{t,j}} = \frac{\phi_j}{1+\widetilde{\pi}_{t,j}(\phi_j - 1)}, \tag{B.6}$$

$$\frac{1-\widetilde{\pi}^*_{t,j}}{1-\widetilde{\pi}_{t,j}} = \frac{1}{1+\widetilde{\pi}_{t,j}(\phi_j - 1)}. \tag{B.7}$$

Note therefore that $\widetilde{\mathbb{P}}^\diamond$ is absolutely continuous with respect to $\widetilde{\mathbb{P}}$.

Recall that $H_t = \sigma(s_\tau, 0 \leqslant \tau \leqslant t)$, where $\sigma(s_\tau, 0 \leqslant \tau \leqslant t)$ is the σ-algebra generated by $\{s_t\}$, with signals $s_t \in \mathcal{S}$. Denote $N_\mathcal{S} \equiv |\mathcal{S}|$, so $s_t \in \{s_1, s_2, \dots, s_{N_\mathcal{S}}\}$, and further denote $\mathfrak{p}_{t,k} \equiv \widetilde{\mathbb{P}}_t(s_{t+1} = \theta_k)$, $\varrho_{t,k} \equiv \widetilde{\mathbb{P}}_t(R^m_T = \theta_j \mid s_{t+1} = s_k)$, and $\varrho^*_{t,k} \equiv \mathbb{P}^*_t(R^m_T = \theta_j \mid s_{t+1} = s_k, R^m_T \in \{\theta_j, \theta_{j+1}\})$, so that $\widetilde{\pi}_{t+1,j} = \varrho_{t,k}$ if $s_{t+1} = s_k$, and similarly $\widetilde{\pi}^*_{t+1,j} = \varrho^*_{t,k}$ if $s_{t+1} = s_k$. Combining (B.3), (B.6), (B.7), and these definitions, we have

$$\begin{aligned}
\widetilde{\mathbb{E}}^\diamond_t[\widetilde{\pi}^*_{t+1,j}] &= \frac{\widetilde{\pi}^*_{t,j}}{\widetilde{\pi}_{t,j}} \sum_{k=1}^{N_\mathcal{S}} \mathfrak{p}_{t,k}\, \varrho^*_{t,k}\, \widetilde{\mathbb{E}}_t\big[\mathbb{1}\{R^m_T = \theta_j\} \mid s_{t+1} = s_k\big] \\
&\quad + \frac{1-\widetilde{\pi}^*_{t,j}}{1-\widetilde{\pi}_{t,j}} \sum_{k=1}^{N_\mathcal{S}} \mathfrak{p}_{t,k}\, \varrho^*_{t,k}\, \widetilde{\mathbb{E}}_t\big[\mathbb{1}\{R^m_T = \theta_{j+1}\} \mid s_{t+1} = s_k\big] \\
&= \frac{\phi_j}{1+\widetilde{\pi}_{t,j}(\phi_j-1)} \sum_{k=1}^{N_\mathcal{S}} \mathfrak{p}_{t,k} \frac{\phi_j \varrho_{t,k}}{1+\varrho_{t,k}(\phi_j-1)}\, \varrho_{t,k} \\
&\quad + \frac{1}{1+\widetilde{\pi}_{t,j}(\phi_j-1)} \sum_{k=1}^{N_\mathcal{S}} \mathfrak{p}_{t,k} \frac{\phi_j \varrho_{t,k}}{1+\varrho_{t,k}(\phi_j-1)}\, (1-\varrho_{t,k}) \\
&= \frac{\phi_j}{1+\widetilde{\pi}_{t,j}(\phi_j-1)} \sum_{k=1}^{N_\mathcal{S}} \mathfrak{p}_{t,k} \frac{\varrho_{t,k}\big(1+\varrho_{t,k}(\phi_j-1)\big)}{1+\varrho_{t,k}(\phi_j-1)} \\
&= \frac{\phi_j}{1+\widetilde{\pi}_{t,j}(\phi_j-1)} \sum_{k=1}^{N_\mathcal{S}} \mathfrak{p}_{t,k}\, \varrho_{t,k} = \frac{\phi_j \widetilde{\pi}_{t,j}}{1+\widetilde{\pi}_{t,j}(\phi_j-1)} = \widetilde{\pi}^*_{t,j},
\end{aligned}$$

where the second-to-last equality uses that $\widetilde{\pi}_{t,j} = \widetilde{\mathbb{E}}_t[\widetilde{\pi}_{t+1,j}]$, as can be seen from the law of iterated expectations given that $\widetilde{\pi}_{t,j} = \mathbb{E}_t[\mathbb{1}\{R^m_T = \theta_j\} \mid R^m_T \in \{\theta_j, \theta_{j+1}\}] = \widetilde{\mathbb{E}}_t[\mathbb{1}\{R^m_T = \theta_j\}] = \widetilde{\mathbb{E}}_t[\widetilde{\mathbb{E}}_{t+1}[\mathbb{1}\{R^m_T = \theta_j\}]] = \widetilde{\mathbb{E}}_t[\widetilde{\pi}_{t+1,j}]$, and the last equality above again uses (B.6). The fact that $\widetilde{\mathbb{E}}^\diamond_0[X^*_j] = 0$ then follows immediately from the proof of Lemma 1. □

LEMMA B.2. *For any return-state pair (θ_j, θ_{j+1}) meeting CTI, for $j' = j, j+1$, RN movement must satisfy*

$$\widetilde{\mathbb{E}}^\diamond_0[\mathsf{m}^*_j \mid R^m_T = \theta_{j'}] = \widetilde{\mathbb{E}}_0[\mathsf{m}^*_j \mid R^m_T = \theta_{j'}]. \tag{B.8}$$

Proof of Lemma B.2. The stream of RN beliefs is $\boldsymbol{\pi}^*_j$, and denote some arbitrary realization for that path by $\mathfrak{b}_j$. For any $\mathfrak{b}_j$ such that $\widetilde{\pi}^*_{T,j} = 1$ (i.e., $R^m_T = \theta_j$), the definition of $\widetilde{\mathbb{P}}^\diamond$ in (B.3) gives that

$$\widetilde{\mathbb{P}}^\diamond_0(\boldsymbol{\pi}^*_j = \mathfrak{b}_j) = \frac{\widetilde{\pi}^*_{0,j}}{\widetilde{\pi}_{0,j}} \widetilde{\mathbb{P}}(\boldsymbol{\pi}^*_j = \mathfrak{b}_j), \tag{B.9}$$

and further $\widetilde{\mathbb{P}}^\diamond_0(R^m_T = \theta_j) = (\widetilde{\pi}^*_{0,j}/\widetilde{\pi}_{0,j})\, \widetilde{\mathbb{P}}_0(R^m_T = \theta_j)$ trivially. Combining these two equations yields $\widetilde{\mathbb{P}}^\diamond_0(\boldsymbol{\pi}^*_j = \mathfrak{b}_j \mid R^m_T = \theta_j) = \widetilde{\mathbb{P}}_0(\boldsymbol{\pi}^*_j = \mathfrak{b}_j \mid R^m_T = \theta_j)$. (Intuitively, all paths ending in $\widetilde{\pi}^*_{T,j} = 1$ receive the same change of measure under $\widetilde{\mathbb{P}}^\diamond$ relative to $\widetilde{\mathbb{P}}$, so probabilities conditional on $R^m_T = \theta_j$ are preserved, and similarly for $R^m_T = \theta_{j+1}$, as was the case for the simpler version in (A.1).) Thus

$$\begin{aligned}\widetilde{\mathbb{E}}^\diamond_0[\mathsf{m}^*_j \mid R^m_T = \theta_j] &= \sum_{\mathfrak{b}_j \,:\, \widetilde{\pi}^*_{T,j}=1} \mathsf{m}^*_j(\mathfrak{b}_j)\, \widetilde{\mathbb{P}}^\diamond_0\Big(\boldsymbol{\pi}^*_j = \mathfrak{b}_j \,\Big|\, R^m_T = \theta_j\Big) \\ &= \sum_{\mathfrak{b}_j \,:\, \widetilde{\pi}^*_{T,j}=1} \mathsf{m}^*_j(\mathfrak{b}_j)\, \widetilde{\mathbb{P}}_0\Big(\boldsymbol{\pi}^*_j = \mathfrak{b}_j \,\Big|\, R^m_T = \theta_j\Big) = \widetilde{\mathbb{E}}_0[\mathsf{m}^*_j \mid R^m_T = \theta_j].\end{aligned}$$

The same applies for $R^m_T = \theta_{j+1}$: for any $\mathfrak{b}_j$ such that $\widetilde{\pi}^*_{T,j} = 0$, (B.9) now becomes $\widetilde{\mathbb{P}}^\diamond_0(\boldsymbol{\pi}^*_j = \mathfrak{b}_j) = (1 - \widetilde{\pi}^*_{0,j})/(1 - \widetilde{\pi}_{0,j})\, \widetilde{\mathbb{P}}(\boldsymbol{\pi}^*_j = \mathfrak{b}_j)$. Further, $\widetilde{\mathbb{P}}^\diamond_0(R^m_T = \theta_{j+1}) = (1 - \widetilde{\pi}^*_{0,j})/(1 - \widetilde{\pi}_{0,j})\, \widetilde{\mathbb{P}}_0(R^m_T = \theta_{j+1})$, so again $\widetilde{\mathbb{P}}^\diamond_0(\boldsymbol{\pi}^*_j = \mathfrak{b}_j \mid R^m_T = \theta_{j+1}) = \widetilde{\mathbb{P}}_0(\boldsymbol{\pi}^*_j = \mathfrak{b}_j \mid R^m_T = \theta_{j+1})$, and thus $\widetilde{\mathbb{E}}^\diamond_0[\mathsf{m}^*_j \mid R^m_T = \theta_{j+1}] = \widetilde{\mathbb{E}}_0[\mathsf{m}^*_j \mid R^m_T = \theta_{j+1}]$. □

Note that the definition in (B.3) aligns with the definition of the RN measure in equation (A.1), so the two lemmas above prove the statements in the text connecting the RN measure in the simple case in Section 2 to the general case in Section 3 (see after equation (A.1) and Lemma A.1). Indeed, (B.4) is the precise analogue to Lemma A.1 in the text; (B.5) is the analogue to Lemma A.3; and (B.8) implies immediately that $\widetilde{\mathbb{E}}^\diamond_0[X^*_j \mid R^m_T] = \widetilde{\mathbb{E}}_0[X^*_j \mid R^m_T]$, which was the main implication of Lemma A.1 used in deriving the results in Section 2. We will thus be able to directly apply those results in this case using the above two lemmas, by virtue of these three results, as follows.

Proof of Proposition 5. No arbitrage gives the existence of a positive SDF for which equation (14) and Assumption 3 are valid. We have

$$\begin{aligned}\widetilde{\pi}_{t,j} &= \mathbb{E}_t[\widetilde{\pi}_{t+1,j}], \\ \widetilde{\pi}^*_{t,j} &= \widetilde{\mathbb{E}}^\diamond_t[\widetilde{\pi}^*_{t+1,j}], \\ \widetilde{\mathbb{E}}^\diamond_0[X^*_j] &= 0, \\ \widetilde{\mathbb{E}}^\diamond_0[X^*_j \mid R^m_T] &= \widetilde{\mathbb{E}}_0[X^*_j \mid R^m_T],\end{aligned}$$

where the first equality uses LIE and the remainder use Lemmas B.1–B.2. The last equation further implies, using the same argument as applied for Lemma A.4, that $\triangle_j \leqslant \widetilde{\pi}^*_{0,j}$. Further, Equations (5)–(8) hold immediately for $\widetilde{\pi}_{t,j}$, $\widetilde{\pi}^*_{t,j}$, ϕ_j. We have thus obtained all the conditions used to prove Lemma 1, Propositions 1–4, and Corollaries 1–2, and thus those results continue to hold, with $\widetilde{\pi}^*_{t,j}$ replacing π^*_t, $\widetilde{\pi}_{t,j}$ replacing π_t, X^*_j replacing X^*, ϕ_j replacing ϕ, $\widetilde{\mathbb{E}}_0[\cdot]$ replacing $\mathbb{E}[\cdot]$, and with $\triangle_j \equiv \widetilde{\mathbb{E}}_0[X^*_j \mid R^m_T = \theta_{j+1}] - \widetilde{\mathbb{E}}_0[X^*_j \mid R^m_T = \theta_j]$ replacing $\triangle$, as stated. □

Proof of Proposition 6. The result follows immediately from equation (7), with V^m_j and V^m_{j+1} replacing $C_{T,1}$ and $C_{T,0}$, respectively. □

B.3 Proofs for Section 4

Proof of Statements 3–6 in Section 4.1. As in footnote 13 in the main text, statements 1–2 are immediate given the definition of CTI. We take the remaining statements in order:

3. The Gabaix (2012) economy features a representative agent with CRRA consumption utility, and log consumption and log dividends follow $c_{t+1} = c_t + g_c + \varepsilon^c_{t+1} + \log(B_{t+1})\mathfrak{D}_{t+1}$ and $d_{t+1} = d_t + g_d + \varepsilon^d_{t+1} + \log(F_{t+1})\mathfrak{D}_{t+1}$, respectively, where $\mathfrak{D}_{t+1} = \mathbb{1}\{\text{disaster}_{t+1}\}$; disasters in $t+1$ occur with probability p_t; B_{t+1} and F_{t+1} are possibly correlated variables with support $[0,1]$; and $(\varepsilon^c_{t+1}, \varepsilon^d_{t+1})'$ is i.i.d. bivariate normal (or a discretized approximation thereof) with mean zero and is independent of all disaster-related variables. *Resilience* is $\mathsf{H}_t = p_t \mathbb{E}_t[B^{-\gamma}_{t+1} F_{t+1} - 1 \mid \mathfrak{D}_{t+1}]$, and write $\mathsf{H}_t = \mathsf{H}_* + \widehat{\mathsf{H}}_t$. The dynamics of p_t are governed by $\widehat{\mathsf{H}}_{t+1} = \frac{1+\mathsf{H}_*}{1+\mathsf{H}_t} e^{-\phi_{\mathsf{H}}} \widehat{\mathsf{H}}_t + \varepsilon^{\mathsf{H}}_{t+1}$, where $\varepsilon^{\mathsf{H}}_{t+1}$ is mean-zero and independent of all other shocks. Gabaix (2012, Theorem 1) shows that $V^m_t = \frac{D_t}{1-e^{-\beta_m}} \left(1 + \frac{e^{-\beta_m - h_*} \widehat{\mathsf{H}}_t}{1-e^{-\beta_m - \phi_H}}\right)$, where $h_* \equiv \log(1+\mathsf{H}_*)$ and $\beta_m \equiv -\log\beta + \gamma g_c - g_d - h_*$. Thus for any θ and H_0, there exists some value d_θ and function $f(d_\theta, \widehat{\mathsf{H}}_T)$, which is strictly increasing in d_θ and strictly decreasing in $\widehat{\mathsf{H}}_T$, such that, by Bayes' rule,

$$\mathbb{P}_0\left(\left(\sum_{t=1}^{T} \mathfrak{D}_t\right) > 0 \,\middle|\, R^m_T \geqslant \theta\right) = \frac{\mathbb{P}_0\left(R^m_T \geqslant \theta \mid \sum_{t=1}^T \mathfrak{D}_t > 0\right) \mathbb{P}_0\left(\sum_{t=1}^T \mathfrak{D}_t > 0\right)}{\mathbb{P}_0(R^m_T \geqslant \theta)}$$
$$= \frac{\mathbb{P}_0\left(D_T \geqslant f(d_\theta, \widehat{\mathsf{H}}_T) \,\middle|\, \sum_{t=1}^T \mathfrak{D}_t > 0\right) \mathbb{P}_0\left(\sum_{t=1}^T \mathfrak{D}_t > 0\right)}{\mathbb{P}_0\left(D_T \geqslant f(d_\theta, \widehat{\mathsf{H}}_T)\right)}.$$

Note now that (i) the innovation to $\widehat{\mathsf{H}}_{t+1}$ is independent of $\mathfrak{D}_{t+1}$; (ii) F_{t+1} (the exponential of the disaster shock to D_t) has support $[0,1]$; and (iii) $\mathbb{P}_t(\varepsilon^d_{t+1} \geqslant \epsilon) = o(e^{-\epsilon^2})$ as $\epsilon \to \infty$.[1] Thus $\mathbb{P}_0(D_T \geqslant f(d_\theta, \widehat{\mathsf{H}}_T) \mid \sum_{t=1}^T \mathfrak{D}_t > 0) = o(\mathbb{P}_0(D_T \geqslant f(d_\theta, \widehat{\mathsf{H}}_T)))$ as $d_\theta \to \infty$, from which the statement in footnote 14 follows. Denote the value δ in that statement by $\delta = \delta_0$. It also follows immediately that for any $t > 0$ (with $t < T$), for any $\delta_t > 0$, there exists an $\underline{\theta}$ such that $\mathbb{P}_t(\sum_{\tau=1}^T \mathfrak{D}_t > 0 \mid R^m_T \geqslant \underline{\theta}) < \delta_t$ asymptotically $\mathbb{P}_0$-a.s. as $\delta_0 \to 0$. Given some $\delta_t > 0$, consider

[1] To see why point (iii) holds, denote $\sigma_d \equiv \mathrm{Var}(\varepsilon^d_t)$, and then note that $\int_\epsilon^\infty \exp(-x^2/(2\sigma_d^2))/\sqrt{2\pi\sigma_d^2}\,dx < \int_\epsilon^\infty (x/\epsilon)\exp(-x^2/(2\sigma_d^2))/\sqrt{2\pi\sigma_d^2}\,dx = \sigma_d \exp(-\epsilon^2/(2\sigma_d^2))/(\sqrt{2\pi}\epsilon)$. A similar calculation can be used to derive a lower bound for the upper tail of the normal CDF. Then applying the previous upper-bound calculation to $\mathbb{P}_0(D_T \geqslant f(d_\theta, \widehat{\mathsf{H}}_T) \mid \sum_{t=1}^T \mathfrak{D}_t > 0)$ and the lower-bound calculation to $\mathbb{P}_0(D_T \geqslant f(d_\theta, \widehat{\mathsf{H}}_T))$, it follows that $\mathbb{P}_0(D_T \geqslant f(d_\theta, \widehat{\mathsf{H}}_T) \mid \sum_{t=1}^T \mathfrak{D}_t > 0)/\mathbb{P}_0(D_T \geqslant f(d_\theta, \widehat{\mathsf{H}}_T)) = o(1)$, as stated, since the distribution of the value in the denominator is shifted to the right relative to the distribution of the value in the numerator given (i)–(ii).

θ_j, θ_{j+1} large enough that $\mathbb{P}_t(\sum_{\tau=1}^T \mathfrak{D}_t > 0 \mid R_T^m \in \{\theta_j, \theta_{j+1}\}) < \delta_t$. We then have from (14) that

$$\begin{aligned}\phi_{t,j} &= \frac{\begin{array}{c}\mathbb{E}_t[M_{t,T} \mid R_T^m = \theta_j, \sum_{\tau=1}^T \mathfrak{D}_\tau = 0]\, \mathbb{P}_t(\sum_{\tau=1}^T \mathfrak{D}_\tau = 0 \mid R_T^m = \theta_j) \\ + \mathbb{E}_t[M_{t,T} \mid R_T^m = \theta_j, \sum_{\tau=1}^T \mathfrak{D}_\tau > 0]\, \mathbb{P}_t(\sum_{\tau=1}^T \mathfrak{D}_\tau > 0 \mid R_T^m = \theta_j)\end{array}}{\begin{array}{c}\mathbb{E}_t[M_{t,T} \mid R_T^m = \theta_{j+1}, \sum_{\tau=1}^T \mathfrak{D}_\tau = 0]\, \mathbb{P}_t(\sum_{\tau=1}^T \mathfrak{D}_\tau = 0 \mid R_T^m = \theta_{j+1}) \\ + \mathbb{E}_t[M_{t,T} \mid R_T^m = \theta_{j+1}, \sum_{\tau=1}^T \mathfrak{D}_\tau > 0]\, \mathbb{P}_t(\sum_{\tau=1}^T \mathfrak{D}_\tau > 0 \mid R_T^m = \theta_{j+1})\end{array}} \\ &= \frac{\mathbb{E}_t[M_{t,T} \mid R_T^m = \theta_j, \sum_{\tau=1}^T \mathfrak{D}_\tau = 0](1 - \mathcal{O}(\delta_t)) + \mathcal{O}(\delta_t)}{\mathbb{E}_t[M_{t,T} \mid R_T^m = \theta_{j+1}, \sum_{\tau=1}^T \mathfrak{D}_\tau = 0](1 - \mathcal{O}(\delta_t)) + \mathcal{O}(\delta_t)} \\ &= \frac{\mathbb{E}_t[M_{t,T} \mid R_T^m = \theta_j, \sum_{\tau=1}^T \mathfrak{D}_\tau = 0]}{\mathbb{E}_t[M_{t,T} \mid R_T^m = \theta_{j+1}, \sum_{\tau=1}^T \mathfrak{D}_\tau = 0]} + \mathcal{O}(\delta_t).\end{aligned}$$

The fraction in the last expression is constant given that $M_{t,T} = \beta^{T-t} e^{-\gamma g_c (T-t)}$ conditional on $\sum_{t=1}^T \mathfrak{D}_t = 0$, using eq. (2) of Gabaix (2012). Thus denoting $\phi_j \equiv \frac{\mathbb{E}_0[M_{t,T} \mid R_T^m = \theta_j, \sum_{t=1}^T \mathfrak{D}_\tau = 0]}{\mathbb{E}_0[M_{t,T} \mid R_T^m = \theta_{j+1}, \sum_{t=1}^T \mathfrak{D}_\tau = 0]}$, we have $\phi_{t,j} = \phi_j + \mathcal{O}(\delta_t)$. Since we can take $\delta_t \to 0$ asymptotically $\mathbb{P}_0$-a.s. as $\delta_0 \to 0$, we have $\phi_{t,j} = \phi_j + o_p(1)$ for any sequence of values $\delta = \delta_0 \to 0$, as stated.

4. The Epstein–Zin (1989) preference recursion is $U_t = \left[(1-\beta) C_t^{1-\psi^{-1}} + \beta(\mathbb{E}_t[U_{t+1}^{1-\gamma}])^{\frac{1-\psi^{-1}}{1-\gamma}}\right]^{\frac{1}{1-\psi^{-1}}}$, and it can be shown (e.g., Campbell, 2018, eq. (6.42)) that the SDF evolves in this case according to $M_{t,t+1} = \beta(C_{t+1}/C_t)^{-\vartheta/\psi}(1/R_{t,t+1}^m)^{1-\vartheta}$, where $\vartheta \equiv (1-\gamma)/(1-\psi^{-1})$. In case (i) of the statement, $\gamma = 1$ and $M_{t,t+1} = \beta / R_{t,t+1}^m$, so M_T depends only on the index return. Thus the numerator and denominator in equation (14) are constant, and CTI holds immediately. For case (ii), write $\Delta c_{t+1} = \mu_c + \rho \Delta c_t + \sigma \eta_{t+1}$, with $\eta_{t+1} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. Given $\psi = 1$, it follows from Hansen, Heaton, and Li (2008, eq. (3)) that the log SDF follows $m_{t,t+1} = -\Delta c_{t+1} + \frac{1-\gamma}{1-\beta\rho}\sigma\eta_{t+1}$ (up to a constant, as we ignore throughout). Further, the consumption-wealth ratio C_t/V_t^m is a constant given $\psi = 1$, so $r_{t,t+1}^m = \Delta c_{t+1}$. Using this in the log SDF and summing from t to T, $m_{t,T} = -r_{t,T}^m + \frac{1-\gamma}{1-\beta\rho}\sigma \sum_{\tau=t+1}^T \eta_\tau$. The first term is known conditional on R_t^m. In addition, recursive substitution and summation for $r_{t,t+1}^m$ gives that $r_{t,T}^m = \frac{\sigma}{1-\rho}\sum_{\tau=t+1}^T (1-\rho^{T-\tau+1})\eta_\tau$. Thus for the second term in $m_{t,T}$, conditioning on $R_T^m = \theta_j$ is equivalent to conditioning on $\sum_{\tau=t+1}^T (1-\rho^{T-\tau+1})\eta_\tau = \text{const} + \log\theta_j \equiv k_j$. Denoting $w_t \equiv (1-\rho^{T-t+1})$, it can then be shown (e.g., Vrins, 2018, eq. (2)–(3)) that $(\sum_{\tau=t+1}^T \eta_\tau \mid \sum_{\tau=t+1}^T w_\tau \eta_\tau = k_j) \sim \mathcal{N}(\mu_{t,j}, \varsigma_t)$, where $\mu_{t,j} = k_j \frac{\sum_{\tau=t+1}^T w_\tau}{\sum_{\tau=t+1}^T w_\tau^2}$ and where ς_t does not depend on k_j. Therefore,

$$\log\phi_{t,j} = \log\mathbb{E}_t\left[\sum_{\tau=t+1}^T \eta_\tau \mid R_T^m = \theta_j\right] - \log\mathbb{E}_t\left[\sum_{\tau=t+1}^T \eta_\tau \mid R_T^m = \theta_{j+1}\right] = \log\theta_j - \log\theta_{j+1},$$

so CTI holds. Case (iii) follows immediately from eq. (17) of Kocherlakota (1990), which shows that $M_T \propto (R_T^m)^{-\gamma}$ in the i.i.d. case.

5. The Campbell and Cochrane (1999) economy features a representative agent with utility $\mathbb{E}_0\{\sum_{t=0}^\infty \beta^t[(C_t - \mathfrak{H}_t)^{1-\gamma} - 1]/(1-\gamma)\}$, where $\mathfrak{H}_t$ is the level of (exogenous) habit and other

terms are standard. The *surplus-consumption ratio* is $S^c_t \equiv (C_t - \mathfrak{H}_t)/\mathfrak{H}_t$. Log dynamics are $s^c_{t+1} = (1-\phi)\bar{s}^c + \phi s^c_t + \lambda(s^c_t)\varepsilon_{t+1}$, $c_{t+1} = g + c_t + \varepsilon_{t+1}$, and $d_{t+1} = g + d_t + \eta_{t+1}$, where $\varepsilon_{t+1} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,\sigma^2_\varepsilon)$, $\eta_{t+1} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,\sigma^2_\eta)$, $\text{Corr}(\varepsilon_{t+1},\eta_{t+1}) = \rho$, and the *sensitivity function* is $\lambda(s^c_t) = \left[\frac{1}{\bar{S}^c}\sqrt{1-2(s^c_t - \bar{s}^c)} - 1\right]\mathbb{1}\{s^c_t \leqslant s^c_{\max}\}$, with $\bar{S}^c = e^{\bar{s}^c} = \sigma_\varepsilon\sqrt{\frac{\gamma}{1-\phi}}$ and $s^c_{\max} = \bar{s}^c + (1-\bar{S}^c)^2/2$. The SDF evolves according to $M_{t,t+1} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(\frac{S^c_{t+1}}{S^c_t}\right)^{-\gamma}$, so

$$\frac{\mathbb{E}_t[M_{t,T} \mid R^m_T = \theta_j]}{\mathbb{E}_t[M_{t,T} \mid R^m_T = \theta_{j+1}]} = \frac{\mathbb{E}_t\left[\exp\left(\sum_{\tau=0}^{T-t-1} -\gamma\left(1+\lambda(s^c_{t+\tau})\right)\varepsilon_{t+\tau+1}\right) \,\middle|\, R^m_T = \theta_j\right]}{\mathbb{E}_t\left[\exp\left(\sum_{\tau=0}^{T-t-1} -\gamma\left(1+\lambda(s^c_{t+\tau})\right)\varepsilon_{t+\tau+1}\right) \,\middle|\, R^m_T = \theta_{j+1}\right]}.$$

For a counterexample to constant ϕ_t, set $T = 2$ and $\rho = 1$ (so $\Delta c_t = \Delta d_t$, as in the simplest case considered by Campbell and Cochrane). A sufficient condition for non-constant ϕ_t is $\text{Cov}_0(\phi_1, \mathbb{E}_1[M_{1,2} \mid R^m_2 = \theta_{j+1}]) \neq 0$, as this gives $\mathbb{E}_0[\phi_1] \neq \phi_0$. As of $t = 0$, both ε_1 and ε_2 are relevant for R^m_2 and $M_{0,2}$: ε_1 determines s^c_1 and thus $\lambda(s^c_1)$. As of $t = 1$, the only source of uncertainty for both R^m_2 and $M_{1,2}$ is ε_2: s^c_2 and d_2 determine R^m_2, and conditional on time-1 variables, these depend only on ε_2. Write ε^1_j for the realization of ε_2 needed to generate $R^m_2 = \theta_j$ given ε_1 — i.e., $\varepsilon^1_j \equiv \{\varepsilon_2 : R^m_2 = \theta_j \mid \varepsilon_1\}$ — and similarly write ε^1_{j+1} for θ_{j+1}. Then we have $\mathbb{E}_1[M_{1,2} \mid R^m_2 = \theta_{j'}] = \exp(-\gamma(1+\lambda(s^c_1))\varepsilon^1_{j'})$ for $j' = j, j+1$, so $\phi_1 = \exp(-\gamma(1+\lambda(s^c_1))(\varepsilon^1_j - \varepsilon^1_{j+1}))$. Thus $\text{Cov}_0(\phi_1, \mathbb{E}_1[M_{1,2} \mid R^m_2 = \theta_{j+1}]) = \text{Cov}_0(\exp(-\gamma(1+\lambda(s^c_1))(\varepsilon^1_j - \varepsilon^1_{j+1})), \exp(-\gamma(1+\lambda(s^c_1))\varepsilon^1_{j+1}))$. Given Gaussian ε_1, this value is generically non-zero.

6. Take the two-agent CRRA case considered in Section 5 of Basak (2000), with notation adopted directly. Basak's Proposition 7 shows that when extraneous risk matters, state prices (and thus the SDF) depend on both the stochastic weighting process $\eta(t)$ and the aggregate endowment $\varepsilon(t)$. These two processes are driven respectively by independent shocks, $dW_z(t)$ (extraneous risk) and $dW_\varepsilon(t)$ (fundamental risk). Asset returns thus do not pin down the SDF realization, generating a generically path-dependent SDF and thus time-varying ϕ_t (see also the discussion in Atmaz and Basak, 2018, footnote 17). □

Proof of Proposition 7. Given that ϕ_t can change, we explicitly allow it to depend on the signal history. RN beliefs are thus now denoted by $\pi^*_t(H_t) = \frac{\phi_t(H_t)\pi_t(H_t)}{(\phi_t(H_t)-1)\pi_t(H_t)+1}$, where we use the simpler notation from Section 2 for clarity throughout. Uncertainty about θ is again resolved by period T, and we again consider X^* from 0 to T. Since $\pi_T \in \{0,1\}$ implies $\pi^*_T = \pi_T$, time variation in ϕ_t has no effect on X^* for $t > T-1$.

Toward a contradiction, assume that there exists some DGP(s) in which ϕ_t changes such that $\mathbb{E}_t[\phi_{t+1}] \leqslant \phi_t$ and expected RN movement is higher than the bounds in Proposition 2 for some T. Consider a DGP from this set with the highest expected RN movement. We now consider the last meaningful movement of ϕ in this DGP. Specifically, given that ϕ_t is assumed to change at some point, but ϕ_t is constant when $t \geqslant T$, there must exist some history H_t in which $\pi_t \in (0,1)$, ϕ_t

can change between t and $t+1$ (i.e., there exists a signal s_{t+1} for which $\phi_{t+1}(H_t \cup s_{t+1}) \neq \phi_t(H_t)$, where s_{t+1} includes the signal $s_{\phi_{t+1}}$) but for which ϕ_t is constant after $t+1$. Following any H_t, by assumption, $\phi_{t+1}(H_t \cup s_{t+1})$ can take two values: $\phi^H_{t+1} > \phi_t$ following signal s^H_{t+1} with probability $q^H > 0$, and $\phi^L_{t+1} < \phi_t$ following signal s^L_{t+1} with probability $q^L = 1 - q^H > 0$. We start by assuming that ϕ_t evolves as a martingale:

$$\sum_{i \in \{L,H\}} q^i \cdot \phi^i_{t+1} = \phi_t. \tag{B.10}$$

Given the maintained assumption that π_t does not evolve in the same period as ϕ_t and therefore is constant immediately following history H_t, $\pi^*_t(H_t \cup s_{t+1})$ can take at most two values: $\pi^{*i}_{t+1} = \frac{\phi^i_{t+1} \cdot \pi_t}{(\phi^i_{t+1}-1)\pi_t+1}$ for $i \in \{L,H\}$. Now consider expected RN movement following H_t. From period t to $t+1$, given signal s^i_{t+1}, RN beliefs move from π^*_t to π^{*i}_{t+1}, leading to per-period RN movement

$$\begin{aligned}
\mathbb{E}[\mathsf{m}^*_{t,t+1}|H_t \cup s^i_{t+1}] = (\pi^*_t - \pi^{*i}_{t+1})^2 &= \left(\frac{\phi_t \cdot \pi_t}{(\phi_t - 1)\pi_t + 1} - \frac{\phi^i_{t+1} \cdot \pi_{t+1}}{(\phi^i_{t+1} - 1)\pi_{t+1} + 1}\right)^2 \\
&= \left(\frac{\phi_t \cdot \pi_t}{(\phi_t - 1)\pi_t + 1} - \frac{\phi^i_{t+1} \cdot \pi_t}{(\phi^i_{t+1} - 1)\pi_t + 1}\right)^2.
\end{aligned}$$

Given that the postulated ϕ_t process is constant after $t+1$, at that point our main bounds hold with π^*_0 replaced with π^{*i}_{t+1} and ϕ replaced with ϕ^i_{t+1}. Thus given signal s^i_{t+1},

$$\begin{aligned}
\mathbb{E}[\mathsf{m}^*_{t+1,T}|H_t \cup s^i_{t+1}] &= \mathbb{E}[X^*_{t+1,T}|H_t \cup s^i_{t+1}] + \mathbb{E}[\mathsf{r}^*_{t+1,T}|H_t \cup s^i_{t+1}] \\
&\leqslant (\pi^{i*}_{t+1} - \pi_{t+1}) \cdot \pi^{i*}_{t+1} + (1 - \pi^{i*}_{t+1}) \cdot \pi^{i*}_{t+1} = (1 - \pi_{t+1}) \cdot \pi^{i*}_{t+1} \\
&= (1 - \pi_{t+1}) \cdot \frac{\phi^i_{t+1} \cdot \pi_{t+1}}{(\phi^i_{t+1} - 1)\pi_{t+1} + 1} = (1 - \pi_t) \cdot \frac{\phi^i_{t+1} \cdot \pi_t}{(\phi^i_{t+1} - 1)\pi_t + 1},
\end{aligned}$$

where the second line plugs in our bound for excess RN movement and uncertainty reduction given that uncertainty is zero at period T, and the third line states everything in terms of ϕ_t and π_t and uses the assumption that $\pi_t = \pi_{t+1}$. Therefore, expected RN movement from period t onward following history H_t is bounded above by:

$$\begin{aligned}
\mathbb{E}[\mathsf{m}^*_{t,T}|H_t] &= \mathbb{E}[\mathsf{m}^*_{t,t+1}|H_t] + \mathbb{E}[\mathsf{m}^*_{t+1,T}|H_t] \\
&\leqslant \sum_{i \in \{L,H\}} q^i \cdot \Big((\frac{\phi_t \cdot \pi_t}{(\phi_t - 1)\pi_t + 1} - \frac{\phi^i_{t+1} \cdot \pi_t}{(\phi^i_{t+1} - 1)\pi_t + 1})^2 + (1 - \pi_t) \cdot \frac{\phi^i_{t+1} \cdot \pi_t}{(\phi^i_{t+1} - 1)\pi_t + 1} \Big).
\end{aligned}$$

We now show that this DGP will have higher RN movement if ϕ_t is constant from H_t onward. To see this, consider the "worst-case" DGP in Proposition 4 in which ϕ remains constant at ϕ_t. In this case, RN movement is (arbitrarily close to) $\mathbb{E}_{maxDGP}[\mathsf{m}^*_{t,T}|H_t] = (1 - \pi_t) \cdot \frac{\phi_t \cdot \pi_t}{(\phi_t - 1)\pi_t + 1}$. We now subtract the expected RN movement given changing ϕ ($\mathbb{E}[\mathsf{m}^*_{t,T}|H_t]$) from the worst-case RN movement ($\mathbb{E}_{maxDGP}[\mathsf{m}^*_{t,T}|H_t]$) and show it is positive given the assumption that ϕ_t evolves as a martingale. The difference is positive if and only if

$$(1-\pi_t)\cdot\frac{\phi_t\cdot\pi_t}{(\phi_t-1)\pi_t+1}-\sum_{i\in\{L,H\}}q^i\cdot\left(\left(\frac{\phi_t\cdot\pi_t}{(\phi_t-1)\pi_t+1}-\frac{\phi^i_{t+1}\cdot\pi_t}{(\phi^i_{t+1}-1)\pi_t+1}\right)^2+(1-\pi_t)\cdot\frac{\phi^i_{t+1}\cdot\pi_t}{(\phi^i_{t+1}-1)\pi_t+1}\right)>0.$$

Using (B.10) in this inequality gives that $\mathbb{E}_{maxDGP}[\mathsf{m}^*_{t,T}|H_t]-\mathbb{E}[\mathsf{m}^*_{t,T}|H_t]>0$ if and only if

$$\frac{\pi_t^3(1-\pi_t)^2(\phi^H_{t+1}-\phi_t)(\phi_t-\phi^L_{t+1})((\phi^H_{t+1}-\phi_t)+(\phi^L_{t+1}-1)+(\pi_t)(2+\pi_t(\phi_t-1))(\phi^H_{t+1}-1)(\phi^L_{t+1}-1))}{(1+\pi_t(\phi_t-1))^2(1+\pi_t(\phi^H_{t+1}-1))^2(1+\pi_t(\phi^L_{t+1}-1))^2}>0.$$

It is straightforward to see that the expression on the left side of this inequality is positive: every parentheses contains a positive value as $\phi^H_{t+1}>\phi_t>\phi^L_{t+1}\geqslant 1$ and $\pi_t\in(0,1)$. Therefore, we conclude that expected RN movement can be increased if ϕ_t remains constant following H_t rather than changing. But this gives us a contradiction, as it violates the assumption that the DGP with ϕ_t moving following H_t has the highest possible movement. Therefore, we conclude that there does not exist a DGP satisfying in which ϕ evolves as a martingale that produces more expected RN movement than the bound in Proposition 2.

We now extend this observation to DGPs in which movement in ϕ is a supermartingale rather than a martingale. We do so by showing that if there exists a DGP where ϕ evolves as super-martingale and leads to expected movement that is higher than our bound, there there must exist a martingale that leads to higher expected movement. Given the previous martingale result, this is impossible. Formally, assume that there exists a DGP_{super} in which ϕ evolves as a supermartingale such that the expected movement of this DGP is higher than our bound for a given T. Consider the supermartingale DGP with the maximum expected movement, and consider a period t (history H_t) with the last meaningful movement in ϕ in which ϕ is a strict supermartingale. If this period does not exist, the process is a martingale, and the previous results hold. Note that, following this movement, there cannot be further change in ϕ. If there were and ϕ were a martingale, the previous result shows that no change in ϕ would produce more expected movement, contradicting the assumption that this DGP produces the highest expected movement in the class. If instead there was movement and the change in ϕ was a strict supermartingale, it would contradict the assumption that the previous movement was the last meaningful movement of that type.

Now, we show that it is possible to adjust DGP_{super} following history H_t to increase expected movement following H_t by adjusting the change in ϕ from period t to period $t+1$ to be a martingale rather than a supermartingale. To do so, we first show that any upward movement from ϕ_t to $\phi_{t+1}>\phi_t$ always leads to more total movement following H_t than any downward movement from ϕ_t to $\phi_{t+1}<\phi_t$. Consider total expected movement from H_t onward given a change from ϕ_t to ϕ_{t+1}:

$$\mathbb{E}[\mathsf{m}^*_{t,T}|H_t,\phi_t,\phi_{t+1}]=\left(\frac{\phi_t\cdot\pi_t}{(\phi_t-1)\pi_t+1}-\frac{\phi_{t+1}\cdot\pi_t}{(\phi_{t+1}-1)\pi_t+1}\right)^2+(1-\pi_t)\cdot\frac{\phi_{t+1}\cdot\pi_t}{(\phi_{t+1}-1)\pi_t+1}.$$

Our claim is that this is higher if $\phi_{t+1}>\phi_t$ than if $\phi_{t+1}<\phi_t$. To see this, compare the above with movement if $\phi_{t+1}=\phi_t$. In this case, $\mathbb{E}[\mathsf{m}^*_{t,T}|H_t,\phi_t=\phi_{t+1}]=(1-\pi_t)\cdot\frac{\phi_t\cdot\pi_t}{(\phi_t-1)\pi_t+1}$. Subtracting from above and writing $\pi = \pi_t$ for simplicity yields:

$$\begin{aligned}&\mathbb{E}[\mathsf{m}^*_{t,T}|H_t,\phi_t,\phi_{t+1}] - \mathbb{E}[\mathsf{m}^*_{t,T}|H_t,\phi_t = \phi_{t+1}]\\ &\quad= \frac{(\pi-1)^2\cdot\pi\cdot(1+\pi\cdot(2+\pi\cdot(\phi_t-1))\cdot(\phi_{t+1}-1))\cdot(\phi_t-\phi_{t+1})}{(1+\pi(\phi-1))^2\cdot(1+\pi(\phi_{t+1}-1))^2}.\end{aligned}$$

As with the inequality in the martingale case, every component in this expression is weakly positive (as $0<\pi<1$ because the ϕ movement is meaningful and $\phi \geqslant 1$), except for $(\phi_t - \phi_{t+1})$. Therefore, this equation is positive if $\phi_{t+1} < \phi_t$ and negative if $\phi_{t+1} > \phi_t$. But then it must be that $\mathbb{E}[\mathsf{m}^*_{t,T}|H_t,\phi_t,\phi_{t+1}]$ is greater if $\phi_{t+1} > \phi_t$ than if $\phi_{t+1} < \phi_t$. In this case, we can adjust the evolution of ϕ following history H_t — which was assumed to be a supermartingale — to be a martingale by taking a probability from downward change in ϕ and shifting it to an upward change in ϕ. Specifically, if ϕ_t is a strict supermartingale at H_t, there must be at least some probability on a realization of $\phi_{t+1} < \phi$. Consider the lowest possible realization of ϕ^L_{t+1} with associated probability q^L. There are two possibilities. First, there is some value $\phi^H_{t+1} > \phi$ such shifting the probability q^L from ϕ^L_{t+1} to ϕ^H_{t+1} makes ϕ a martingale. Second, there is some $q^H < q^L$ such that shifting q^H from ϕ^L_{t+1} to ϕ^H_{t+1} makes ϕ a martingale. In either case, we are shifting probability from $\phi^L_{t+1} < \phi_t$ to $\phi^H_{t+1} > \phi_t$. But, as just proven above, it must be that $\mathbb{E}[\mathsf{m}^*_{t,T}|H_t,\phi_t,\phi_{t+1}]$ is greater if $\phi_{t+1} > \phi_t$ than if $\phi_{t+1} < \phi_t$. But then the total movement of the change from ϕ at H_t must increase. This implies that there exists a martingale process for ϕ at H_t that has higher expected movement than the strict supermartingale process for ϕ at H_t. This contradicts the assumption that the strict supermartingale process has the highest movement in the class of supermartingale processes (which includes martingales), completing the proof. □

Proof of Proposition 8. In what follows, we often use $\mathbb{E}_i[\cdot]$ to make explicit that we are taking expectations over DGPs indexed by i, and we continue to use the notational simplifications used in the statement of the proposition. For (i), fixing $\pi^*_{0,i} = \pi^*_0$ across i and applying Proposition 1,

$$\begin{aligned}\mathbb{E}_i[\mathbb{E}[X^*_i]] &= \mathbb{E}_i[(\pi^*_0 - \pi_{0,i})\cdot\triangle_i] = \pi^*_0\cdot\mathbb{E}_i[\triangle_i] - \mathbb{E}_i[\pi_{0,i}]\cdot\mathbb{E}_i[\triangle_i]\\ &= (\pi^*_0 - \mathbb{E}_i[\pi_{0,i}])\cdot\mathbb{E}_i[\triangle_i] = \mathbb{E}_i[\pi^*_0 - \pi_{0,i}]\cdot\mathbb{E}_i[\triangle_i]\\ &= \mathbb{E}_i\left[\pi^*_0 - \frac{\pi^*_0}{\phi_i + (1-\phi_i)\pi^*_0}\right]\cdot\mathbb{E}_i[\triangle_i]\end{aligned} \tag{B.11}$$

where the last equality in the first line follows from the assumption that $\mathrm{Cov}(\pi_{0,i},\triangle_i) = 0$.

Now consider $\zeta_1(\phi_i,\pi^*_0) \equiv \pi^*_0 - \frac{\pi^*_0}{\phi_i+(1-\phi_i)\pi^*_0}$. This function is concave in ϕ_i: $\frac{\partial^2\zeta_1}{\partial\phi_i^2} = \frac{-2\pi^*_0(1-\pi^*_0)^2}{(\pi^*_0+\phi(1-\pi^*_0))^3}$, which is weakly negative given $\pi^*_0 \in [0,1]$ and $\phi \geqslant 1$. Thus by Jensen's inequality, the expectation of ζ_1 over ϕ_i is less than ζ_1 evaluated at $\underline{\phi} \equiv \mathbb{E}_i[\phi_i]$, so $\mathbb{E}_i\left[\pi^*_0 - \frac{\pi^*_0}{\phi_i+(1-\phi_i)\pi^*_0}\right] \leqslant \pi^*_0 - \frac{\pi^*_0}{\underline{\phi}+(1-\underline{\phi})\pi^*_0}$. Returning to (B.11), suppose that $\mathbb{E}_i[\triangle_i] > 0$. In this case,

$$\mathbb{E}_i[\mathbb{E}[X^*_i]] = \mathbb{E}_i[\pi^*_0 - \frac{\pi^*_0}{\phi_i+(1-\phi_i)\pi^*_0}]\cdot\mathbb{E}_i[\triangle_i]] \leqslant (\pi^*_0 - \frac{\pi^*_0}{\underline{\phi}+(1-\underline{\phi})\pi^*_0})\cdot\mathbb{E}_i[\triangle_i].$$

Now assume that $\mathbb{E}_i[\triangle_i] \leqslant 0$. Then, as $\pi_0^* - \frac{\pi_0^*}{\phi_i+(1-\phi_i)\pi_0^*} = \pi_0^* - \pi_0 \geqslant 0$ given $\phi_i \geqslant 1$,

$$\mathbb{E}_i[\mathbb{E}[X_i^*]] = \mathbb{E}_i[\pi_0^* - \frac{\pi_0^*}{\phi_i + (1-\phi_i)\pi_0^*}] \cdot \mathbb{E}_i[\triangle_i] \leqslant 0.$$

Taken together, $\mathbb{E}_i[\mathbb{E}[X_i^*]] \leqslant \max\{0, (\pi_0^* - \frac{\pi_0^*}{\underline{\phi}+(1-\underline{\phi})\pi_0^*}) \cdot \mathbb{E}_i[\triangle_i]\}$.

For part (ii), first consider the situation in which $\pi_{0,i}^*$ is constant and equal to π_0^*. As above,

$$\mathbb{E}_i[\mathbb{E}[X_i^*]] \leqslant \mathbb{E}_i[(\pi_0^* - \pi_{0,i}) \cdot \pi_0^*] = \mathbb{E}_i[\pi_0^* - \pi_{0,i}] \cdot \pi_0^* = \mathbb{E}_i\left[\pi_0^* - \frac{\pi_0^*}{\phi_i + (1-\phi_i)\pi_0^*}\right] \cdot \pi_0^*.$$

As above, given the concavity of $\zeta_2 \equiv \pi_0^* - \frac{\pi_0^*}{\phi_i+(1-\phi_i)\pi_0^*}$ with respect to ϕ_i and the fact that $\pi_0^* \geqslant 0$,

$$\mathbb{E}_i[\mathbb{E}[X_i^*]] \leqslant \mathbb{E}_i\left[\pi_0^* - \frac{\pi_0^*}{\phi_i + (1-\phi_i)\pi_0^*}\right] \cdot \pi_0^* \leqslant \left(\pi_0^* - \frac{\pi_0^*}{\underline{\phi} + (1-\underline{\phi})\pi_0^*}\right)\pi_0^*,$$

as stated in the second inequality. Now allowing $\pi_{0,i}^*$ to vary, write the bound for $\mathbb{E}[X^*]$ in Proposition 2 as $\zeta_{2'}(\phi_i, \pi_{0,i}^*) \equiv \left(\pi_0^* - \frac{\pi_0^*}{\phi_i+(1-\phi_i)\pi_0^*}\right)\pi_{0,i}^*$. Again since $\partial^2\zeta_{2'}/\partial\phi_i^2 \leqslant 0$, for any arbitrary realization of $\pi_{0,i}^* = \varrho$, we have from the application of Jensen's inequality above (now dropping the dependence of $\mathbb{E}$ on i) that $\mathbb{E}[\zeta_{2'}(\phi_i, \pi_{0,i}^*) \mid \pi_{0,i}^*] \leqslant \zeta_{2'}\left(\mathbb{E}[\phi_i \mid \pi_{0,i}^* = \varrho], \varrho\right)$. Using Proposition 2 and applying LIE to this inequality,

$$\mathbb{E}[X_i^*] \leqslant \mathbb{E}[\zeta_{2'}(\phi_i, \pi_{0,i}^*)] \leqslant \mathbb{E}\left[\zeta_{2'}\left(\mathbb{E}[\phi_i \mid \pi_{0,i}^*], \pi_{0,i}^*\right)\right] \leqslant \mathbb{E}\left[\zeta_{2'}(\overline{\phi}, \pi_{0,i}^*)\right], \tag{B.12}$$

where $\overline{\phi}$ is as in the proposition statement and where the last inequality uses $\partial\zeta_{2'}/\partial\phi_i \geqslant 0$. Substituting the definition of $\zeta_{2'}$ into this inequality yields equation (16).

For part (iii), as $(\pi_{0,i}^* - \frac{\pi_{0,i}^*}{\overline{\phi}+(1-\overline{\phi})\pi_{0,i}^*}) \leqslant \pi_{0,i}^*$ for any $\overline{\phi} \geqslant 1$, $\mathbb{E}[X_i^*] \leqslant \mathbb{E}[(\pi_{0,i}^* - 0)\pi_{0,i}^*] = \mathbb{E}[(\pi_{0,i}^*)^2]$, as stated. (Equivalently, one can use (B.12) and note again that $\partial\zeta_{2'}/\partial\overline{\phi} \geqslant 0$, so that the bound is most slack as $\overline{\phi} \to \infty$, giving the same bound.)

For part (iv), Corollary 2 gives that if $\mathbb{E}[X^*|\theta = 0] \leqslant \mathbb{E}[X^*|\theta = 1]$, then $\mathbb{E}[X^*] \leqslant 0$. Therefore, if $\mathbb{E}[X_i^*|\theta = 0] \leqslant \mathbb{E}[X_i^*|\theta = 1]$ for all i, then $\mathbb{E}[X_i^*] \leqslant 0$ over all streams, completing the proof. □

B.4 Proofs for Section 5

Proof of Proposition 9. For part (i), first define the likelihood of a prior π_0 as

$$\mathcal{L}(\pi_0) \equiv \frac{\pi_0}{1-\pi_0}, \tag{B.13}$$

and the likelihood of a signal s_t as

$$\mathcal{L}(s_t) \equiv \frac{DGP(s_t|\theta = 1)}{DGP(s_t|\theta = 0)},$$

where the dependence of the latter on H_{t-1} is left implicit for simplicity. The likelihood for any belief π_t is defined as well following (B.13). The above likelihoods are well-defined for interior priors (as we assume given finite L in the proposition) and for $DGP(s_t|\theta = 0, H_{t-1}) > 0$ (we return to the situation in which $DGP(s_t|\theta = 0, H_{t-1}) = 0$ shortly). From Bayes' rule, beliefs satisfy $\mathcal{L}(\pi_t) = \mathcal{L}(\pi_0) \cdot \mathcal{L}(s_1) \cdot \mathcal{L}(s_2) \cdots \mathcal{L}(s_t)$. Now note from (5) that $\mathcal{L}(\pi_0^*) \equiv \frac{\pi_0^*}{1-\pi_0^*} = \phi \frac{\pi_0}{1-\pi_0}$, from which it follows that under Bayesian updating,

$$\mathcal{L}(\pi_t^*) = \mathcal{L}(\pi_0^*) \cdot \mathcal{L}(s_1) \cdot \mathcal{L}(s_2) \cdots \mathcal{L}(s_t) = \phi \mathcal{L}(\pi_0) \cdot \mathcal{L}(s_1) \cdot \mathcal{L}(s_2) \cdots \mathcal{L}(s_t).$$

For a fictitious agent with a rational prior, one could replace $\mathcal{L}(\pi_0)$ with $\mathcal{L}(\mathbb{P}_0(\theta = 1))$. In our case, given the incorrect prior (but correct Bayesian updating), we have $\frac{\pi_t^*}{1-\pi_t^*} = \check{\phi} \frac{\mathbb{P}_0(\theta=1)}{1-\mathbb{P}_0(\theta=1)}$, where $\check{\phi} \equiv \phi L$, with L defined as in the proposition. We can therefore write

$$\mathcal{L}(\pi_t^*) = \check{\phi} \mathcal{L}(\mathbb{P}_0(\theta = 1)) \cdot \mathcal{L}(s_1) \cdot \mathcal{L}(s_2) \cdots \mathcal{L}(s_t).$$

As the likelihood ratios for the RN beliefs in this case are equal to those of a fictitious agent with a correct prior $\check{\pi}_0 = \mathbb{P}_0(\theta = 1)$ and $\check{\phi}$ in place of ϕ, we conclude that the RN beliefs are as well. Finally, for the case in which $DGP(s_t|\theta = 0, H_{t-1}) = 0$ and this signal s_t is observed, the person will update to $\pi_t = 1$, matching the belief of a rational agent again. We have thus shown part (i).

We can thus treat the agent with the incorrect prior as if she were rational (satisfying Assumption 2) but with $\check{\phi}$ in place of ϕ. Further, $\check{\phi}$ satisfies Assumption 4, since L is constant and ϕ is constant by that assumption as well. For part (ii) of the proposition, if $\check{\phi} \geqslant 1$, then Assumption 3 holds as well, so all three assumptions are satisfied, and the stated results carry through.

For part (iii), assuming $0 < \check{\phi} < 1$ (so Assumption 3 no longer holds for the fictitious rational agent), note first that the proof of Proposition 1 never employs Assumption 3 and therefore still holds straightforwardly, as we can write $\mathbb{E}[X^*] = (\pi_0^* - \check{\pi}_0)\triangle$ without use of this assumption. For Proposition 2, the result as stated for a rational agent requires that $\pi_0^* > \check{\pi}_0$, which is not true for $\check{\phi} < 1$. But an alternative bound can be shown for this case, by obtaining a lower bound for $\triangle$ similar to the upper bound in Lemma A.4. Starting from (A.7) but solving now for $\mathbb{E}[\mathsf{m}^*|\theta = 1]$, $\mathbb{E}[\mathsf{m}^*|\theta = 1] = (1 - \pi_0^*) - \frac{1-\pi_0^*}{\pi_0^*} \cdot \mathbb{E}[\mathsf{m}^*|\theta = 0]$. Using this in (A.6),

$$\triangle = \mathbb{E}[\mathsf{m}^*|\theta = 0] - \left((1 - \pi_0^*) - \frac{1 - \pi_0^*}{\pi_0^*} \cdot \mathbb{E}[\mathsf{m}^*|\theta = 0]\right) = \frac{1}{\pi_0^*} \cdot \mathbb{E}[\mathsf{m}^*|\theta = 0] - (1 - \pi_0^*).$$

Then, given that $\frac{1}{\pi_0^*} \geqslant 0$ and $\mathbb{E}[\mathsf{m}^*|\theta = 0] \geqslant 0$, $\triangle$ must be bounded below by $-(1 - \pi_0^*)$. Returning to the formula from Proposition 2, if $\check{\phi} < 1$, then $\pi_0^* - \check{\pi}_0 \leqslant 0$, which gives

$$\mathbb{E}[X^*] = (\pi_0^* - \check{\pi}_0)(\triangle) \leqslant (\check{\pi}_0 - \pi_0^*)(1 - \pi_0^*). \tag{B.14}$$

Further, as $\check{\pi}_0 \leqslant 1$, $\mathbb{E}[X^*] \leqslant (\check{\pi}_0 - \pi_0^*)(1 - \pi_0^*) \leqslant (1 - \pi_0^*)(1 - \pi_0^*) = (1 - \pi_0^*)^2$, as stated. And taking (ii) and (iii) together, we have that $\mathbb{E}[X^*] \leqslant \max(\pi_0^{*2}, (1 - \pi_0^*)^2)$. $\square$

Proof of Corollary 3. Case (iii) from the previous proof applies, with ϕ in place of $\check{\phi}$ and π_0 in place of $\check{\pi}_0$ (since the agent now has RE but $\phi < 1$). Thus (B.14) applies with these substitutions. The second expression for the bound given in the corollary then substitutes for π_0 (using (8)) and simplifies. Equivalently, by swapping the labels of states 0 and 1, the swapped RN beliefs become $1 - \pi_t^*$ in place of π_t^* and the swapped SDF ratio becomes $\frac{1}{\phi}$ in place of ϕ. As $\phi < 1$, $\frac{1}{\phi} > 1$. Therefore, all of our results hold, with π_t^* replaced by $1 - \pi_t^*$ and ϕ replaced by $\phi^{-1} > 1$. $\square$

Proof of Proposition 10. Under the stated assumptions for ϵ_t, observed RN movement satisfies

$$\begin{aligned}\mathbb{E}[\widehat{\mathsf{m}}^*_{t,t+1}] &= \mathbb{E}[(\widehat{\pi}^*_{t+1} - \widehat{\pi}^*_t)^2] = \mathbb{E}\Big[\big((\pi^*_{t+1} - \pi^*_t)^2 + (\epsilon_{t+1} - \epsilon_t)\big)^2\Big] \\ &= \mathbb{E}[\mathsf{m}^*_{t,t+1}] + 2\mathbb{E}[\pi^*_{t+1}\epsilon_{t+1} - \pi^*_t\epsilon_{t+1} - \pi^*_{t+1}\epsilon_t + \pi^*_t\epsilon_t] + \mathbb{E}[(\epsilon_{t+1} - \epsilon_t)^2] \\ &= \mathbb{E}[\mathsf{m}^*_{t,t+1}] + \mathbb{E}[\epsilon_t^2 + \epsilon_{t+1}^2].\end{aligned}$$

For the observed counterpart of uncertainty resolution $\mathsf{r}^*_{t,t+1} \equiv (\mathsf{u}^*_t - \mathsf{u}^*_{t+1})$,

$$\mathbb{E}[\widehat{\mathsf{r}}^*_{t,t+1}] = \mathbb{E}[(\pi^*_t + \epsilon_t)(1 - \pi^*_t - \epsilon_t) - (\pi^*_{t+1} + \epsilon_{t+1})(1 - \pi^*_{t+1} - \epsilon_{t+1})] = \mathbb{E}[\mathsf{r}^*_{t,t+1}] + \mathbb{E}[\epsilon^2_{t+1} - \epsilon^2_t].$$

Combining these two, with $\text{Var}(\epsilon_t) \equiv \mathbb{E}[(\epsilon_t - \mathbb{E}[\epsilon_t])^2] = \mathbb{E}[\epsilon_t^2]$ and $X^*_{t,t+1} \equiv \mathsf{m}^*_{t,t+1} - \mathsf{r}^*_{t,t+1}$,

$$\mathbb{E}[\widehat{X}^*_{t,t+1}] = \mathbb{E}[X^*_{t,t+1}] + 2\text{Var}(\epsilon_t). \qquad \square$$

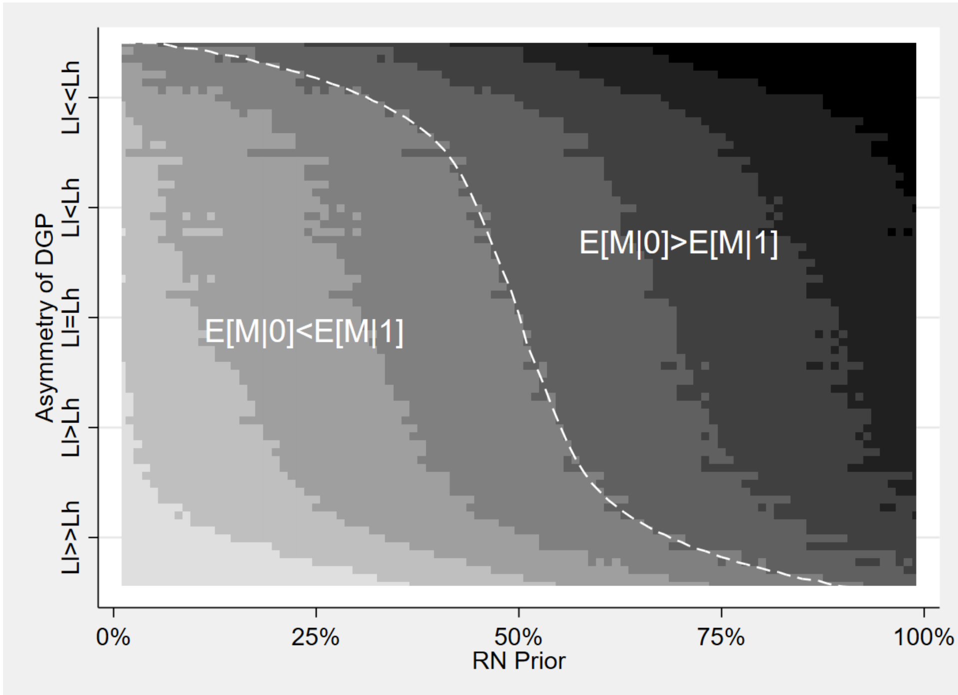
Appendix C. Additional Technical Material

C.1 Simulations for the Relationship of RN Prior and DGP with $\triangle$

As noted in Section 2.3, we run numerical simulations of a large number of DGPs and priors in order to understand the precise impact of the RN prior and DGP on $\triangle$ (and therefore $\mathbb{E}[X^*]$). We consider the universe of history-independent binary-signal DGPs with a prior π_0^* where $s_t \in \{l, h\}$ and $\mathbb{P}[s_t = h|\theta = 1]$ and (assumed lower) $\mathbb{P}[s_t = h|\theta = 0]$ are constant over t. These signal distributions imply likelihood ratios for the signals of $L_h \equiv \frac{\mathbb{P}[s_t=h|\theta=1]}{\mathbb{P}[s_t=h|\theta=0]} > 1$ and $L_l \equiv \frac{\mathbb{P}[s_t=l|\theta=0]}{\mathbb{P}[s_t=l|\theta=1]} > 1$. We use a fine grid to discretize π_0^*, L_h, and L_l, then conduct 1000 simulations with $T = 100$ and calculate $\triangle$ in all cases. We find:

1. When π_0^* is low, $\triangle > 0$ is very unlikely: the percentage of DGPs with positive $\triangle$ given a $\pi_0^* < .25$ is 2%. For $\pi_0^* < .5$, it is 11%.
2. When π_0^* is low, the only DGPs in which $\triangle > 0$ are very asymmetric and extreme. For example, when $\pi_0^* = .25$, $\triangle > 0$ only occurs if $\mathbb{P}[s_t = h|\theta = 1] > .95$ and $L_l > 2 \cdot L_h$.
3. The converse is true when π_0^* is high: $\triangle < 0$ is rare and only occurs given a very asymmetric and extreme DGP.
4. For symmetric DGPs ($L_h = L_l$), $\triangle \lesseqqgtr 0$ when $\pi_0^* \lesseqqgtr .5$.

Figure C.1: Contour Plot: Simulations for $\triangle$ by DGP and π_0^*



Note: See text in Appendix C.1 for description of simulations and discussion of results.

5. Holding the DGP constant, $\triangle$ rises with π_0^*.
6. Holding all else constant, as L_h rises and the size of upward updates rises, $\triangle$ falls. As L_l rises and the size of upward-updates rises, $\triangle$ rises.

We present these results visually in Figure C.1. We reduce the dimensionality of the setting by focusing on the *likelihood ratio* $\frac{L_h}{L_l}$ rather than L_h and L_l individually. (While the impact of both L_h and L_l on $\triangle$ appears monotonic, the impact of $\frac{L_h}{L_l}$ is only monotonic on average, leading to a slightly messier graph.) The figure shows a contour plot with the RN prior on the x-axis, with the y-axis stacking all of the DGP combinations in order of the likelihood ratio, and the contour colors showing the approximate value of $\triangle$ (darker colors corresponding to higher values) for each prior and DGP (with the dotted line highlighting the points at which $\triangle = 0$). For example, drawing a vertical line at a prior of $\pi_0^* = 0.25$ suggests that a large portion of DGPs produce a $\triangle < 0$, and the only DGPs that produce $\triangle > 0$ have extreme likelihood ratios.

C.2 Risk-Neutral Beliefs and Time-Varying Discount Rates

This section provides further context on the relationship between RN beliefs and discount rates, as discussed in footnote 15 in the main paper. We again work in the setting in Section 2 here for simplicity of exposition. The price of the terminal consumption claim is given in equilibrium in by $P_t(C_T) = \mathbb{E}_t\left[\beta_t^{T-t}\frac{U'(C_T)}{U'(C_t)}C_T\right]$, where β_t is now the agent's (possibly time-varying) time discount factor. Defining the gross return $R_{t,T}^{C} \equiv \frac{C_T}{P_t(C_T)}$, rearranging this equation for $P_t(C_T)$ yields

$$\mathbb{E}_t[R_{t,T}^{C}] = \frac{1 - \mathrm{Cov}_t\left(\beta_t^{T-t}\frac{U'(C_T)}{U'(C_T)}, C_T\right)}{\mathbb{E}_t\left[\beta_t^{T-t}\frac{U'(C_T)}{U'(C_t)}\right]} = \frac{\frac{U'(C_t)}{\beta_t^{T-t}} - \mathrm{Cov}_t(U'(C_T), C_T)}{\mathbb{E}_t[U'(C_T)]},$$

as usual. We can write $\mathbb{E}_t[U'(C_T)] = \pi_t U'(C_{\text{low}}) + (1-\pi_t)U'(C_{\text{high}})$ in our two-state setting, and $\mathrm{Cov}_t(U'(C_T), C_T)$ can be similarly rewritten as a function of π_t, C_T, and $U'(C_T)$. In this setting, discount-rate variation can arise from four sources:

1. Changes in the time discount factor β_t.
2. Changes in contemporaneous marginal utility $U'(C_t)$.
3. Changes in the relative probability π_t.
4. Changes in state-contingent terminal consumption C_i or marginal utility $U'(C_i)$.

Our framework allows for *any* discount-rate variation arising from the first three sources, but restricts the last one: under CTI, it must be the case that any changes to (expected) $U'(C_i)$ are proportional across states. (More generally, as in Section 4.1, permanent changes to the SDF are admissible, which by itself greatly generalizes this setting relative to one with constant discount rates.) With constant discount rates, meanwhile, none of the four changes are admissible, or any such changes must offset perfectly.

C.3 Simulations with Time-Varying ϕ_t

This section provides further detail for the simulations discussed in Section 4.3 and shown in Figure 4. First consider the baseline situation in which $\pi_0^* = 0.5$ and $\phi_t = 3$ for all t. There is thus only uncertainty about θ, and $\mathbb{E}[\mathsf{m}^*]$ varies depending on the signal DGP. To trace the distribution of $\mathbb{E}[\mathsf{m}^*]$ across DGPs, we attempt to cover the space of binary DGPs in which the signal strengths are constant over time. We start by looping over $\mathbb{P}[s_t = h|\theta = 1]$ from {1,.99,.98,...,.01}. Then we loop over $\mathbb{P}[s_t = l|\theta = 0]$ from {.01,.02,.03,....99} while constraining $\mathbb{P}[s_t = h|\theta = 1] > \mathbb{P}[s_t = l|\theta = 0]$ such that the h signal leads to an upward movement. This process leads to 5052 DGPs. For each of these DGPs, we simulate 100 random streams of $T = 200$ periods, after which the state is perfectly observed. This number of periods allows beliefs to get very close to certainty prior to the resolving signal. We calculate m^* for each stream, from which we calculate the average m^* statistic as an estimate of $\mathbb{E}[\mathsf{m}^*]$ for each DGP. The distribution of $\mathbb{E}[\mathsf{m}^*]$ values across all such simulated DGPs is shown in the dark line in Figure 4.

Next, we allow additional uncertainty about the conditional realizations of the SDF M_T, so that ϕ_t also evolves over time. For each state (j and $j+1$), we allow M_T to take two possible values with equal probability, where we choose the values such that $\phi_0 = 3$. Here, we start to run into calculation timing constraints such that we limit the possible signal strengths. In particular, we allow signal strengths for the high signal of .55,.75,.95 and for the low signal of .05,.25,.45 for both states. Therefore we simulate nine DGPs for learning about M_T in state j and nine DGPs for learning about $j+1$, leading to 81 combined DGPs to learn about M_T. We combine each such DGP with each of the DGPs for θ discussed above, and we again simulate 100 random draws of movement of 200 periods. Each line in Figure 4 represents a different $\mathbb{E}[\mathsf{m}^*]$ distribution given variation in the signal strengths for θ, with the different lines showing different signal strengths for learning about the conditional values of M_T (and thus ϕ). In the "Low ϕ Uncertainty" case, M_T in state j can take the values 2.5 or 3.5 with equal probability and in state $j+1$ can take the values 0.833 or 1.167 with equal probability. Consequently, $\phi_0 = 3$, and ϕ_T can vary from 2.14 to 4.2 (with a coefficient of variation of 12%). In the "Medium ϕ Uncertainty" case, M_T in state j can be 2 or 4 and in state $j+1$ can be 0.667 or 1.333, so that ϕ_T can vary from 1.5 to 6 (with a coefficient of variation of 54%). Finally, in the "High ϕ Uncertainty" case, M_T in state j can be 1.5 or 4.5 and in state $j+1$ can be 0.5 or 1.5, so that ϕ_T can vary from 1 to 9 (with a coefficient of variation of 100%).

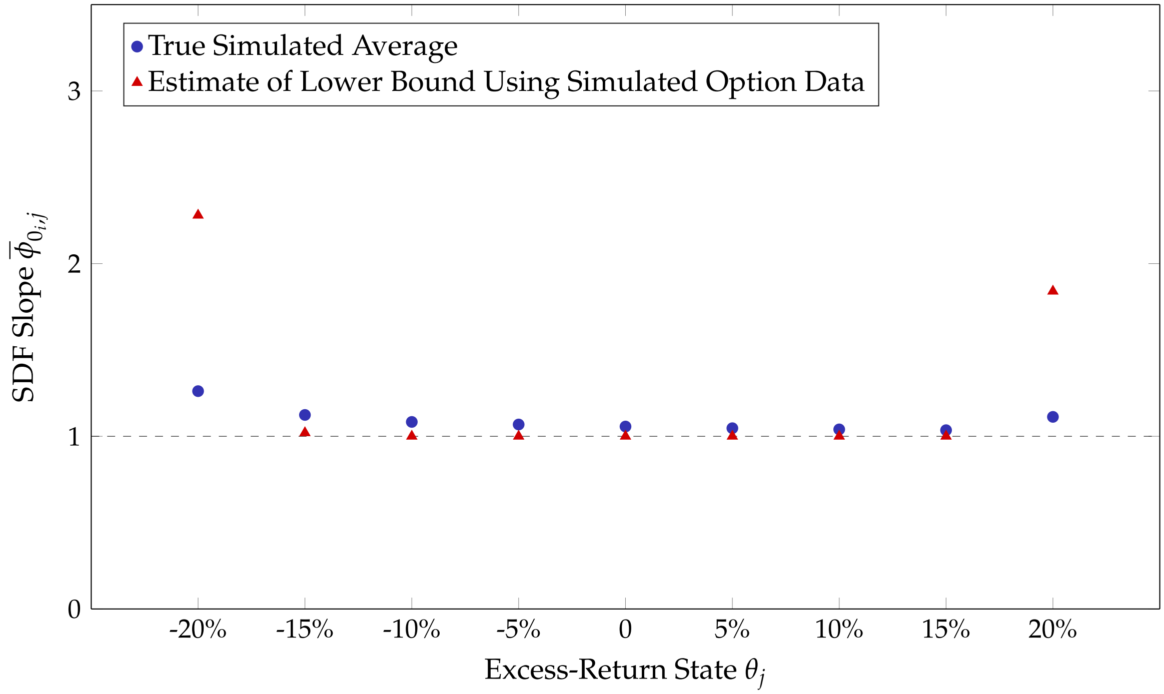
C.4 Solution Method and Simulations for Habit Formation Model

See the proof of statement 5 in Appendix B.3 for a description of the model, and the calibrated parameters are identical to those used by Campbell and Cochrane (1999, Table 1), converted to daily values, for the version of their model with imperfectly correlated consumption and dividends. We consider 90-day option-expiration horizons (i.e., $T_i - 0_i = 90$), and after solving the model for the price-dividend ratio, we then solve for the joint distribution for returns (from t to T_i) and the SDF at every point in a gridded state space as of $t = T_i - 1$, then $t = T_i - 2$, and so on, as below.

The initial market index value is normalized to $V^m_{0_i} = 1$, and the joint CDF for the SDF realization and the return as a function of the current surplus-consumption state is then solved by iterating backwards from T_i: after solving the model for the price-dividend ratio as a function of the surplus-consumption value, we then calculate the $T_i - 1$ CDF for any possible surplus-consumption value by integrating over the distributions of shocks to consumption (and thus surplus consumption) and dividends at T_i; we then project this CDF onto an interpolating cubic spline over the three dimensions $(S^c_{T_i-1}, M_{T_i}, \log(R^{m,e}_{T_i}))$; we then calculate the $T_i - 2$ CDF by integrating over the distribution of shocks at $T_i - 1$ *and* the projection solutions for the conditional distribution functions for $(T_i - 1) \to T_i$ obtained in the previous step; and so on. These CDFs are then used for the model simulations.

We conduct 25,000 simulations, where each simulation runs from 0_i to T_i, and for which the initial surplus-consumption state is drawn from its unconditional distribution. For each period in each simulation, we evaluate risk-neutral beliefs over return states at every point in the space Θ and use these to calculate the set of conditional risk-neutral beliefs $\{\widetilde{\pi}^*_{t,i,j}\}_j$. Further, we store the associated set of expected SDF slopes $\{\phi_{t,i,j}\}_j$. We can thus calculate the true average values of these objects of interest, $\overline{\phi}_{0_i,j} \equiv \widehat{\mathbb{E}}[\phi_{0_i,i,j}]$, where $\widehat{\mathbb{E}}[\cdot]$ denotes the expectation over all simulations i and we have fixed the state pair j. And using the risk-neutral beliefs series, we can naively apply our conservative bound in Proposition 8 to obtain lower-bound estimates for those SDF slopes and compare those estimates to the true simulated values. Relative risk aversion for this model's representative agent does not match the definition used in Proposition 6, as this agent's utility

Figure C.2: Estimates of SDF Slope in Habit Formation Model Simulations



Note: See text in Appendix C.4 for description of simulations.

does not depend only on terminal wealth (see Campbell and Cochrane, 1999, Section IV.B), so we accordingly present estimates for the SDF slope rather than for relative risk aversion.

Figure C.2 presents these simulation results. The blue circles show the true simulated average values of the SDF slopes $\overline{\phi}_{0_{i},j}$, while the red triangles show the naive lower-bound estimates of these values using our theoretical bound on the simulated RN beliefs data. It is clear in both cases that these SDF-slope values are far below those obtained from our empirical estimates, so the model does not replicate the observed variation in RN beliefs even with the violation of CTI. We can understand the validity of the theoretical bound for the interior states by way of Proposition 7, which shows that the bounds hold approximately for violations of CTI for which the $\phi_{t,i,j}$ process is close to a martingale. In our simulations, the values $|\widehat{\mathbb{E}}[\phi_{t+1,i,j} - \phi_{t,i,j}]|$ for different state pairs j range from 0.00002 to at most 0.00011, which is not large enough to invalidate the bounds.

C.5 Data Cleaning and Measurement of Risk-Neutral Distribution

Before detailing measurement of the risk-neutral distribution, we note that we must collect additional data in order to follow the procedure below. In particular, in order to obtain the ex post return state for each option expiration date T_i (and thereby assign probability 1 to that state on date T_i, so that our streams are resolving), we need S&P 500 index prices used as option settlement values. Our first step in this exercise is therefore to obtain end-of-day index prices (which we take as well from OptionMetrics). But the settlement value for many S&P 500 options in fact reflects the opening (rather than closing) price on the expiration date; for example, the payoff for the traditional monthly S&P 500 option contract expiring on the third Friday of each month depends on the opening S&P index value on that third Friday morning, while the payoff for the more recently introduced end-of-month option contract depends on the closing S&P index value on the last business day of the month.[2] To obtain the ex-post return state for A.M.-settled options, we hand-collect the option settlement values for these expiration dates from the Chicago Board Options Exchange (CBOE) website, which posts these values.

In addition, in order to measure the risk-neutral distribution *and* to measure realized excess index returns, we need risk-free zero-coupon yields R^f_{t,T_i} for $t = 0_i, \ldots, T_i - 1$. To obtain these, we follow van Binsbergen, Diamond, and Grotteria (2022) and obtain the relevant yield directly from the cross-section of option prices by applying the put-call parity relationship. We apply their "Estimator 2," which obtains $R^f_{t,T_i} = \beta^{-1/T}$ from Theil–Sen (robust median) estimation of $q^{m,\text{put}}_{t,i,K} - q^{m,\text{call}}_{t,i,K} = \alpha + \beta K + \varepsilon_{t,i,K}$. This provides a very close fit to the option cross-sections (see van Binsbergen, Diamond, and Grotteria, 2022, for details) and thus produces a risk-free rate consistent with observed option prices, as is necessary to correctly back out the risk-neutral distribution.

Finally, for both the OptionMetrics end-of-day and CBOE intraday data, we apply standard filters (e.g., Christoffersen, Heston, and Jacobs, 2013; Constantinides, Jackwerth, and Savov, 2013; Martin, 2017) to the raw option-price data before estimating risk-neutral distributions. We drop any options with bid or ask price of zero (or less than zero), with uncomputable Black–Scholes implied volatility or with implied volatility of greater than 100 percent, with more than one year to maturity, or (for call options) with mid prices greater than the price of the underlying; we drop any option cross-section (i.e., the full set of prices for the pair (t, T_i)) with no trading volume on date t, with fewer than three listed prices across different strikes, or for which there are fewer than three strikes for which both call and put prices are available (as is necessary to calculate the forward price and risk-free rate); and after transforming the data to a risk-neutral distribution as below, we keep only conditional RN belief observations $\widetilde{\pi}^*_{t,i,j}$ for which the non-conditional beliefs satisfy $\pi^*_t(R^m_{T_i} = \theta_j) + \pi^*_t(R^m_{T_i} = \theta_{j+1}) \geqslant 5\%$. Our bounds can be calculated using data of arbitrary frequency, so we calculate $X^*_{i,j}$ using changes in RN beliefs over whatever set of trading days are left in the sample after this filtering procedure.

As introduced in Section 6.1, we measure the risk-neutral return distribution by applying the following steps to the remaining option prices (for which we use mid prices), following Malz (2014):

1. Transform the collections of call- and put-price cross-sections (for example, for call options on date t for expiration date T_i, this set is $\{q^m_{t,i,K}\}_{K \in \mathcal{K}}$) into Black–Scholes implied volatilities.

2. Discard the implied volatility values for in-the-money calls and puts, so that the remaining steps use data from only out-of-the-money put and call prices (as, e.g., in Martin, 2017). Moneyness is measured relative to the at-the-money-forward price, measured (again following Martin, 2017) as the strike K at which $q^{m,\text{put}}_{t,i,K} = q^{m,\text{call}}_{t,i,K}$.

[2]See http://www.cboe.com/SPX for further detail. For our dataset, the majority (roughly 2/3) of option expiration dates correspond to A.M.-settled options.

3. Fit a cubic spline to interpolate a smooth function between the points in the resulting implied-volatility schedule for each trading date–expiration date pair. The spline is *clamped*: its boundary conditions are that the slope of the spline at the minimum and maximum values of the knot points $\mathcal{K}$ is equal to 0; further, to extrapolate outside of the range of observed knot points, set the implied volatilities for unobserved strikes equal to the implied volatility for the closest observed strike (i.e., maintain a slope of 0 for the implied-volatility schedule outside the observed range).

4. Evaluate this spline at 1,901 strike prices, for S&P index values ranging from 200 to 4,000 (so that the evaluation strike prices are $K = 200, 202, \ldots, 4000$), to obtain a set of implied-volatility values across this fine grid of possible strike prices for each (t, T_i) pair.[3]

5. Invert the resulting smoothed 1,901-point implied-volatility schedule for each (t, T_i) pair to transform these values back into call prices, and denote this fitted call-price schedule as $\{\hat{q}^m_{t,i,K}\}_{K \in \{200,202,\ldots,4000\}}$.

6. Calculate the risk-neutral CDF for the date-T_i index value at strike price K using $\mathbb{P}^*_t(V^m_{T_i} < K) = 1 + R^f_{t,T_i}(\hat{q}^m_{t,i,K} - \hat{q}^m_{t,i,K-2})/2$. (See the proof of equation (13) in Appendix B.2 for a derivation of this result; the index-value distance between the two adjacent strikes is equal to 2 given that we evaluate the spline at intervals of two index points.)

7. Defining $V^m_{i,j,\max}$ and $V^m_{i,j,\min}$ to be the date-T_i index values corresponding to the upper and lower bounds, respectively, of the bin defining return state θ_j,[4] we then calculate the risk-neutral probability for state θ_j will be realized at date T_i, referred to with slight notational abuse as $\mathbb{P}^*_t(\theta_j)$, as

$$\mathbb{P}^*_t(\theta_j) = \mathbb{P}^*_t(V^m_{T_i} < V^m_{i,j,\max}) - \mathbb{P}^*_t(V^m_{T_i} < V^m_{i,j,\min}),$$

   where the CDF values are taken from step 6 using linear interpolation between whichever two strike values $K \in \{200, 202, \ldots, 4000\}$ are nearest to $V^m_{i,j,\max}$ and $V^m_{i,j,\min}$, respectively.

Steps 1 and 2 represent the only point of distinction between our procedure and that of Malz, who assumes access to a single implied-volatility schedule without considering put or call prices directly; our procedure is accordingly essentially identical to his. Note that we transform the option prices into Black–Scholes implied volatilities simply for purposes of fitting the cubic spline and then transform these implied volatilities back into call prices before calculating risk-neutral beliefs, so this procedure does *not* require the Black–Scholes model to be correct.[5] The clamped cubic spline proposed by Malz (2014), and used in step 3 above, is chosen to ensure that the call-price schedule obtained in step 5 is decreasing and convex with respect to the strike price outside the range of observable strike prices, as required under the restriction of no arbitrage. Violations of these

[3] This set of ∼1,900 strike prices is on average about 20 times larger than the set of strikes for which there are prices in the data, as there is a mean of roughly 90 observed values in a typical set $\{q^m_{t,i,K}\}_{K \in \mathcal{K}}$.

[4] That is, formally, $V^m_{i,j,\min} = R^f_{0_i,T_i} V^m_{T_0} \exp(\theta_j - 0.05)$ and $V^m_{i,j,\max} = R^f_{0_i,T_i} V^m_{0_i} \exp(\theta_j)$. For example, for excess return state θ_2, we have $V^m_{i,j,\min} = R^f_{0_i,T_i} V^m_{0_i} \exp(-0.2)$ and $V^m_{i,j,\max} = R^f_{0_i,T_i} V^m_{T_0} \exp(-0.15)$.

[5] We conduct this transformation following Malz (2014), as well as much of the related literature, which argues that these smoothing procedures tend to perform slightly better in implied-volatility space than in the option-price space given the convexity of option-price schedules; see Malz (1997) for a discussion.

restrictions *inside* the range of observable strikes, as observed infrequently in the data, generate negative implied risk-neutral probabilities; in any case that this occurs, we set the associated risk-neutral probability to 0.

As noted in step 3, the clamped spline is an *interpolating* spline, as it is restricted to pass through all the observed data points so that the fitted-value set $\{\hat{q}^m_{t,i,K}\}$ contains the original values $\{q^m_{t,i,K}\}$. Some alternative methods for measuring risk-neutral beliefs use smoothing splines that are not constrained to exhibit such interpolating behavior. To check the robustness of our results to the choice of measurement technique, we have accordingly used one such alternative method proposed by Bliss and Panigirtzoglou (2004). Empirical results obtained using risk-neutral beliefs calculated in this alternative manner are unchanged as compared to the benchmark results in Section 6.4.

We have also conducted robustness tests with respect to the fineness of the grid on which we evaluate the spline in step 4 and calculate the risk-neutral CDF in step 6, with results from these exercises also indistinguishable from the benchmark results.

C.6 Noise Estimation and Matching to X^* Observations

As introduced in Section 6.2, we first estimate $\mathrm{Var}(\epsilon_t) = \mathrm{Var}(\epsilon_{t,i,j})$ separately for each combination of trading day t, expiration date T_i, and return state pair j in our intraday sample.[6] Our ReMeDI estimator for this noise variance follows the replication code provided by Li and Linton (2022): $\widehat{\mathrm{Var}}(\epsilon_t) = \frac{1}{N_{\epsilon,n}} \sum_{i=2k_n}^{N_{\epsilon,n}-k_n} (\widehat{\pi}^*_{t_i} - \widehat{\pi}^*_{t_i-2k_n})(\widehat{\pi}^*_{t_i} - \widehat{\pi}^*_{t_i+k_n})$. We select k_n for each return state using the algorithm in Section F.1 of the Online Appendix of Li and Linton (2022).

We must then match the noise estimates (which are obtained only for a subsample of days) to the observed excess movement observations in our original daily data. To do so, we take advantage of the fact that the best predictors of $\widehat{\mathrm{Var}}(\epsilon_{t,i,j})$ are (i) state pair j (we see more noise for tail states) and (ii) the observed RN belief of either θ_j or θ_{j+1} being realized, $\Sigma^*_{t,i,j} \equiv \pi^*_t(R^m_{T_i} = \theta_j) + \pi^*_t(R^m_{T_i} = \theta_{j+1})$ (conditional beliefs are noisier when the underlying sum $\Sigma^*_{t,i,j}$ is lower, as $\Sigma^*_{t,i,j}$ enters into the denominator of $\widetilde{\pi}^*_{t,i,j}$). We thus partition $\Sigma^*_{t,i,j}$ into 5-percentage-point bins ($[0, 0.05)$, $[0.05, 0.1]$, $\ldots$), and then calculate the average noise $\widehat{\sigma}_{\epsilon,j,\Sigma} \equiv \overline{\widehat{\mathrm{Var}}(\epsilon_{t,i,j})}$ for each combination of state pair j and bin for $\Sigma^*_{t,i,j}$. We then match $\widehat{\sigma}_{\epsilon,j,\Sigma}$ to each observed one-day excess movement observation $\widehat{X}^*_{t,t+1,i,j}$ in our original end-of-day data, based on that observation's state j and total probability $\Sigma^*_{t,i,j}$.

C.7 Details of Bootstrap Confidence Intervals

Our block-bootstrap resampling procedure is described in Section 6.4, and we provide further details on how we construct one-sided confidence intervals for Table 3 here. Fixing a given $\overline{\phi}$, denote the point estimate for $\overline{e^{\mathrm{main}}_i(\phi)}$ by $\widehat{e}(\overline{\phi})$. The null that $\overline{e^{\mathrm{main}}_i(\phi)} = 0$ is rejected at the 5% level if $2\widehat{e}(\overline{\phi}) - e^*_{(0.95)}(\overline{\phi}) > 0$, where $e^*_{(0.95)}(\overline{\phi})$ is the 95$^{\mathrm{th}}$ percentile of the bootstrap distribution of $\overline{e^{\mathrm{main}}_i(\phi)}$ statistics (i.e., it is rejected if it is outside of the one-sided 95% basic bootstrap CI for

[6] For this exercise, to increase our available observations, we do not condition on the ex post state being θ_j or θ_{j+1}.

$\overline{e_i^{\text{main}}(\phi)})$. We conduct this procedure for all possible $\overline{\phi}$ values, and we obtain $\widehat{\phi}_{LB} = \min_{\overline{\phi}}$ s.t. $2\widehat{e}(\overline{\phi}) - e^*_{(0.95)}(\overline{\phi}) \leqslant 0$.

A more straightforward procedure for conducting inference on $\overline{\phi}$ would be to construct the basic bootstrap CI directly for $\overline{\phi}$ (i.e., $\widehat{\phi}_{LB} = 2\widehat{\phi} - \phi^*_{(0.95)}$). The challenge preventing us from doing so is that in nearly all cases, the 95$^{\text{th}}$ percentile of the bootstrap distribution for $\widehat{\phi}$ is ∞, given how large our point estimates are (and how much excess movement we observe in our data). This motivates our use of a test-inversion confidence interval using the residuals for different possible values of $\overline{\phi}$, which solves this problem. These CIs achieve asymptotic coverage of at least the nominal level under weak conditions (discussed further below), given the duality between testing and CI construction; see, e.g., Carpenter (1999). We find that our procedure performs quite well, with unbiased and symmetric bootstrap distributions around the full-sample point estimate.

We note that our bootstrap procedure fully preserves the groupings of return-state pairs (indexed by $j = 1, \ldots, J-1$) for each set of observations indexed by i (corresponding to the option expiration date) within each block, as we split the observations into blocks only by time and not by return states. We do so in order to obtain valid inference for the aggregate value $\overline{\boldsymbol{\phi}}$, which uses observations for state pairs $(\theta_2, \theta_3), \ldots, (\theta_{J-2}, \theta_{J-1})$, in the face of arbitrary dependence for the observations across those state pairs and a fixed number of return states J (whereas we assume $N \to \infty$, and further the number of blocks $B \to \infty$ according to a sequence such that $(T_N + 1)/B \to \infty$). In this way our procedure is in fact a *panel* (or *cluster*) *block bootstrap*; see, for example, Palm, Smeekes, and Urbain (2011). Lahiri (2003, Theorem 3.2) provides a weak condition on the strong mixing coefficient of the relevant stochastic process — in our case, $\{(X^*_{i,j}, \widetilde{\pi}^*_{0,i,j}, \{\widehat{\text{Var}}(\epsilon_{t,i,j})\})_{t,j}\}_i$ — under which the blocks are asymptotically independent and the bootstrap distribution estimator is consistent for the true distribution under the asymptotics above, so that our test-inversion confidence intervals have asymptotic coverage probability of at least 95% for the population parameters of interest in the presence of nearly arbitrary (stationary) autocorrelation and heteroskedasticity.[7] This coverage rate may in fact be greater than 95% given that we are estimating lower bounds for the parameters of interest rather than the parameters themselves, and this motivates our use of one-sided rather than two-sided confidence intervals, as in Section 6.4.

C.8 Variable Construction for RN Excess Movement Regressions

As discussed in Section 6.5, we consider reduced-form evidence on the macroeconomic and financial correlates of RN excess movement, with results presented in Table 4. The dependent variable in all cases is the monthly average of noise-adjusted RN excess movement $X^*_{t,t+1,i,j}$, with the average calculated across all available expiration dates and interior state pairs for all trading days t in a month. For the dependent variables, from top to bottom in the table, option bid-ask spread is the

[7]There are additional conditions required for the result of Lahiri (2003, Theorem 3.2) to hold, but they will hold trivially in our context under the RE null given the boundedness of the relevant belief statistics. Our block bootstrap is a non-overlapping block bootstrap (NBB); others (e.g., Künsch, 1989) have proposed a *moving* block bootstrap (MBB) using overlapping blocks, among other alternatives. While the MBB has efficiency gains relative to the NBB, these are "likely to be very small in applications" (Horowitz, 2001, p. 3190), so we use the NBB for computational convenience.

volume-weighted average bid-ask spread for all S&P 500 options with less than a year to maturity and positive bid prices in the given month. Option volume is total monthly dollar trading volume in that same sample, detrended using an estimated exponential trend given the steady growth in option volume over the sample. RN belief stream length is the average full-stream length $\overline{T_i}$ over all contracts i active in that month. VIX^2 is calculated using the average VIX in the given month. The variance risk premium is VIX^2 minus realized variance, and we use the data provided by Lochstoer and Muir (2022) for this VRP. (We thank these and subsequent authors for making the relevant data available.) The risk-aversion proxy ra_t^{BEX}, as discussed in footnote 33 in the text, is obtained from Nancy Xu's website (`https://www.nancyxu.net/risk-aversion-index`), and we take the sum of squared daily changes in ra_t^{BEX} in a given month to measure the volatility of this risk-aversion proxy. We obtain the monthly repurchase-adjusted log price-dividend ratio pd_t from Nagel and Xu (2022), and we calculate the absolute value of its deviation from its sample mean $\overline{pd}$. The 12-month S&P 500 change is calculated as the log change in the S&P price from month $t-12$ to t, using data from Robert Shiller's website (`http://www.econ.yale.edu/~shiller/data.htm`). All variables (both dependent and dependent) are normalized to have zero mean and standard deviation of 1, and all regressions include a constant.

Appendix References

ATMAZ, A. AND S. BASAK (2018): "Belief Dispersion in the Stock Market," *Journal of Finance*, 73, 1225–1279.

AUGENBLICK, N. AND M. RABIN (2021): "Belief Movement, Uncertainty Reduction, and Rational Updating," *Quarterly Journal of Economics*, 136, 933–985.

BASAK, S. (2000): "A Model of Dynamic Equilibrium Asset Pricing With Heterogeneous Beliefs and Extraneous Risk," *Journal of Economic Dynamics & Control*, 24, 63–95.

VAN BINSBERGEN, J. H., W. F. DIAMOND, AND M. GROTTERIA (2022): "Risk-Free Interest Rates," *Journal of Financial Economics*, 143, 1–29.

BLACK, F. AND M. SCHOLES (1973): "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637.

BLISS, R. R. AND N. PANIGIRTZOGLOU (2004): "Option-Implied Risk Aversion Estimates," *Journal of Finance*, 59, 407–446.

BREEDEN, D. T. AND R. H. LITZENBERGER (1978): "Prices of State-Contingent Claims Implicit in Option Prices," *Journal of Business*, 51, 621–651.

BROWN, D. J. AND S. A. ROSS (1991): "Spanning, Valuation and Options," *Economic Theory*, 1, 3–12.

CAMPBELL, J. Y. (2018): *Financial Decisions and Markets: A Course in Asset Pricing*, Princeton: Princeton University Press.

CAMPBELL, J. Y. AND J. H. COCHRANE (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.

CARPENTER, J. (1999): "Test Inversion Bootstrap Confidence Intervals," *Journal of the Royal Statistical Society, Series B*, 61, 159–172.

CHRISTOFFERSEN, P., S. HESTON, AND K. JACOBS (2013): "Capturing Option Anomalies with a Variance-Dependent Pricing Kernel," *Review of Financial Studies*, 26, 1962–2006.

CONSTANTINIDES, G. M., J. C. JACKWERTH, AND A. SAVOV (2013): "The Puzzle of Index Option Returns," *Review of Asset Pricing Studies*, 3, 229–257.

EPSTEIN, L. G. AND S. E. ZIN (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

GABAIX, X. (2012): "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance," *Quarterly Journal of Economics*, 127, 645–700.

HANSEN, L. P., J. C. HEATON, AND N. LI (2008): "Consumption Strikes Back? Measuring Long-Run Risk," *Journal of Political Economy*, 116, 260–302.

HOROWITZ, J. L. (2001): "The Bootstrap," in *Handbook of Econometrics*, ed. by J. J. Heckman and E. Leamer, Amsterdam: Elsevier, vol. 5, chap. 52, 3159–3228.

KOCHERLAKOTA, N. R. (1990): "Disentangling the Coefficient of Relative Risk Aversion from the Elasticity of Intertemporal Substitution: An Irrelevance Result," *Journal of Finance*, 45, 175–190.

KÜNSCH, H. R. (1989): "The Jackknife and the Bootstrap for General Stationary Observations," *Annals of Statistics*, 17, 1217–1241.

LAHIRI, S. N. (2003): *Resampling Methods for Dependent Data*, New York: Springer.

LI, Z. M. AND O. B. LINTON (2022): "A ReMeDI for Microstructure Noise," *Econometrica*, 90, 367–389.

LOCHSTOER, L. A. AND T. MUIR (2022): "Volatility Expectations and Returns," *Journal of Finance*, 77, 1055–1096.

MALZ, A. M. (1997): "Option-Implied Probability Distributions and Currency Excess Returns," *Federal Reserve Bank of New York Staff Report No. 32*.

——— (2014): "A Simple and Reliable Way to Compute Option-Based Risk-Neutral Distributions," *Federal Reserve Bank of New York Staff Report No. 677*.

MARTIN, I. (2017): "What Is the Expected Return on the Market?" *Quarterly Journal of Economics*, 132, 367–433.

NAGEL, S. AND Z. XU (2022): "Asset Pricing with Fading Memory," *Review of Financial Studies*, 35, 2190–2245.

PALM, F. C., S. SMEEKES, AND J.-P. URBAIN (2011): "Cross-Sectional Dependence Robust Block Bootstrap Panel Unit Root Tests," *Journal of Econometrics*, 163, 85–104.

VRINS, F. (2018): "Sampling the Multivariate Standard Normal Distribution Under a Weighted Sum Constraint," *Risks*, 6, 1–13.