FOR ONLINE PUBLICATION

Online Appendix for:

Restrictions on Asset-Price Movements Under Rational Expectations: Theory and Evidence*

Ned Augenblick
UC Berkeley Haas

Eben Lazarus
MIT Sloan

JANUARY 2022

Contents

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

A. Proofs of Theoretical Results

A.1 Proofs for Section 2

Section 2.1

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

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

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 fifth line uses the martingale beliefs assumption (Assumption 1). Summing and applying the law of iterated expectations (LIE),

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

Sections 2.2–2.3

Proof of Equations (12)–(14). For the physical measure,

$$\begin{aligned} \mathbb{P}(H_T) &= \mathbb{P}(\theta = 1) \cdot \mathbb{P}(H_T|\theta = 1) + \mathbb{P}(\theta = 0) \cdot \mathbb{P}(H_T|\theta = 0) \\ &= \pi_0 \cdot \mathbb{P}(H_T|\theta = 1) + (1-\pi_0) \cdot \mathbb{P}(H_T|\theta = 0), \end{aligned} \tag{A.1}$$

where the second line uses that $\pi_0 = \mathbb{E}[\pi_T] = \mathbb{E}[\theta] = \mathbb{P}(\theta = 1)$ (as shown after Assumption 1). Meanwhile, for the RN measure, we have from (11) and (A.1) that

$$\begin{aligned}\mathbb{P}^*(H_T) &= \frac{\pi_0^*}{\pi_0} \cdot \pi_0 \cdot \mathbb{P}(H_T|\theta = 1) + \frac{1-\pi_0^*}{1-\pi_0} \cdot (1-\pi_0) \cdot \mathbb{P}(H_T|\theta = 0) \\ &= \pi_0^* \cdot \mathbb{P}(H_T|\theta = 1) + (1-\pi_0^*) \cdot \mathbb{P}(H_T|\theta = 0). \end{aligned} \tag{A.2}$$

For any H_T such that $\pi_T = 1$, we have as well from (11) that $\mathbb{P}^*(H_T) = \frac{\pi_0^*}{\pi_0}\mathbb{P}(H_T)$, which implies $\mathbb{P}^*(\theta = 1) = \frac{\pi_0^*}{\pi_0}\mathbb{P}(\theta = 1)$. Thus by the definition of conditional probability, $\mathbb{P}^*(H_T|\theta = 1) = \mathbb{P}(H_T|\theta = 1)$. A similar argument gives $\mathbb{P}^*(H_T|\theta = 0) = \mathbb{P}(H_T|\theta = 0)$, and thus (14) holds. Then (A.2) becomes

$$\mathbb{P}^*(H_T) = \pi_0^* \cdot \mathbb{P}^*(H_T|\theta = 1) + (1-\pi_0^*) \cdot \mathbb{P}^*(H_T|\theta = 0).$$

Summing over all possible H_T for which $\theta = 1$ gives $\pi_0^* = \mathbb{P}^*(\theta = 1)$, so that $\mathbb{P}^*$ is a valid probability distribution for which LIE holds. Then noting $\mathbb{P}^*(\theta = 1) = \mathbb{E}^*[\theta] = \mathbb{E}^*[\pi_T] = \mathbb{E}^*[\pi_T^*]$, equation (12) follows.[1] Equation (13) then follows from Proposition 1. □

Proof of Equation (18). Footnote 10 provides a brief derivation. For a full derivation, first write

$$\begin{aligned}\triangle &\equiv \mathbb{E}^*[X^*|\theta = 0] - \mathbb{E}^*[X^*|\theta = 1] \\ &= \mathbb{E}^*[\mathsf{m}^*|\theta = 0] - \mathsf{u}_0^* - (\mathbb{E}^*[\mathsf{m}^*|\theta = 0] - \mathsf{u}_0^*) \\ &= \mathbb{E}^*[\mathsf{m}^*|\theta = 0] - \mathbb{E}^*[\mathsf{m}^*|\theta = 1]. \end{aligned} \tag{A.3}$$

Further, using equation (15),

$$\begin{aligned}0 &= \pi_0^* \cdot \mathbb{E}[X^*|\theta = 1] + (1-\pi_0^*) \cdot \mathbb{E}[X^*|\theta = 0] \\ &= \pi_0^* \cdot (\mathbb{E}[\mathsf{m}^*|\theta = 1] - \mathsf{u}_0^*) + (1-\pi_0^*) \cdot (\mathbb{E}[\mathsf{m}^*|\theta = 0] - \mathsf{u}_0^*),\end{aligned}$$

so using the definition of u_0^*,

$$\pi_0^* \cdot \mathbb{E}[\mathsf{m}^*|\theta = 1] + (1-\pi_0^*) \cdot \mathbb{E}[\mathsf{m}^*|\theta = 0] = \pi_0^*(1-\pi_0^*). \tag{A.4}$$

Solving for $\mathbb{E}[\mathsf{m}^*|\theta = 0]$ gives

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

[1] Lemma A.1 provides a more detailed algebraic derivation of this fact.

Using this in (A.3),

$$
\begin{aligned}
\triangle &= \pi_0^* - \frac{\pi_0^*}{1-\pi_0^*} \cdot \mathbb{E}[\mathsf{m}^*|\theta=1] - \mathbb{E}^*[\mathsf{m}^*|\theta=1] \\
&= \pi_0^* - \frac{1}{1-\pi_0^*} \cdot \mathbb{E}[\mathsf{m}^*|\theta=1] \qquad \text{(A.5)}
\end{aligned}
$$

Given that $\frac{1}{1-\pi_0^*} \geqslant 0$ and $\mathbb{E}[\mathsf{m}^*|\theta=1] \geqslant 0$, $\triangle$ is bounded above by π_0^*. ☐

Proof of Proposition 2. Start from equation (16) and apply equation (15):

$$
\begin{aligned}
\mathbb{E}[X^*] &= \pi_0 \cdot \mathbb{E}[X^*|\theta=1] + (1-\pi_0) \cdot \mathbb{E}[X^*|\theta=0] - 0 \\
&= \pi_0 \cdot \mathbb{E}[X^*|\theta=1] + (1-\pi_0) \cdot \mathbb{E}[X^*|\theta=0] - (\pi_0^* \cdot \mathbb{E}[X^*|\theta=1] + (1-\pi_0^*) \cdot \mathbb{E}[X^*|\theta=0]) \\
&= (\pi_0^* - \pi_0)(\mathbb{E}[X^*|\theta=0] - \mathbb{E}[X^*|\theta=1]) \\
&= (\pi_0^* - \pi_0)(\triangle), \qquad \text{(A.6)}
\end{aligned}
$$

as stated. Then the second equality holds using equation (10) and the definition of $\triangle$. ☐

Proof of Proposition 3. From the proof of equation (18) above, we have $\triangle \leqslant \pi_0^*$. Further, equation (10) implies

$$
\begin{aligned}
\pi_0^* - \pi_0 &= \pi_0^* - \frac{\pi_0^*}{\pi_0^* + \phi(1-\pi_0^*)} \\
&= \pi_0^* \left(1 - \frac{1}{\pi_0^* + \phi(1-\pi_0^*)}\right) \geqslant 0, \qquad \text{(A.7)}
\end{aligned}
$$

where the last inequality uses $\pi_0^* + \phi(1-\pi_0^*) \geqslant 0$ since $\phi \geqslant 1$. Using these two inequalities in the expression for $\mathbb{E}[X^*]$ in (A.6),

$$
\mathbb{E}[X^*] = (\pi_0^* - \pi_0)(\triangle) \leqslant (\pi_0^* - \pi_0)\pi_0^*. \qquad \text{(A.8)}
$$

Plugging in the expression for $\pi_0^* - \pi_0$ in (A.7) then gives equation (19). ☐

Proof of Corollary 1. This is an immediate implication of (A.8) and $\pi_0 \geqslant 0$. ☐

Proof of Corollary 2. As in (A.7), we have $\pi_0^* - \pi_0 \geqslant 0$. Using this in the equality in (A.8) alongside the assumption that $\triangle = \mathbb{E}^*[\mathsf{m}^*|\theta=0] - \mathbb{E}^*[\mathsf{m}^*|\theta=1] \leqslant 0$ gives $\mathbb{E}[X^*] \leqslant 0$. ☐

Proof of Proposition 4. 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 $\widehat{\pi}_0 = \frac{1-\pi_t^*}{\phi+(1-\phi)(1-\pi_t^*)}$.

As a result of 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 $\mathbb{E}^*[\widehat{X}^*|\theta=1] = \mathbb{E}^*[X^*|\theta=0]$. Using that $\mathbb{E}^*[X^*|\theta] = \mathbb{E}[X^*|\theta]$ by equation (14) as proven above, this gives $\mathbb{E}[\widehat{X}^*|\theta=0] = \mathbb{E}[X^*|\theta=1]$ and $\mathbb{E}[\widehat{X}^*|\theta=1] = \mathbb{E}[X^*|\theta=0]$. We conclude that for $\widehat{DGP}$, $\widehat{\triangle} \equiv \mathbb{E}[\widehat{X}^*|\theta=0] - \mathbb{E}[\widehat{X}^*|\theta=1] = -\triangle$. □

Proof of Proposition 5. 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,

$$DGP(s_t=h|\theta=1) = 1, \tag{A.9}$$

$$DGP(s_t=h|\theta=0) = \frac{\pi_{t-1}^*(1-\pi_{t-1}^*-\epsilon)}{(1-\pi_{t-1}^*)(\pi_{t-1}^*+\epsilon)}, \quad \text{with } \epsilon \equiv \frac{1-\pi_0^*}{T}. \tag{A.10}$$

Since $DGP(s_t=l|\theta=1)=0$ from (A.9), 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 (10) to the π_{t-1} values on the right side of this equation, we have after some algebra that

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

Now applying the transformation (9), we obtain that π_t^* given an only-h signal history (suppressing the dependence on this history for simplicity) is, after additional tedious but straightforward algebra,

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

Now using (A.10), we obtain after further algebra that

$$\pi_t^* - \pi_{t-1}^* = \epsilon.$$

Note given the definition of ϵ, then, that 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.5),

$$\triangle = \pi_0^* - \frac{1}{1-\pi_0^*} \cdot \mathbb{E}[\mathsf{m}^*|\theta = 1] \xrightarrow{T\to\infty} \pi_0^*.$$

Using this in equation (A.6) 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 in the proposition.

For the final statement, given $\phi > 1$ and $0 < \pi_0^* < 1$, the inequality in (A.7) is strict, so that $\pi_0^* - \pi_0 > 0$. Further, the only way to obtain $m^* = 0$ 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.5), we have the strict inequality $\triangle < \pi_0^*$ for fixed $T < \infty$. Combining these in (A.6) gives $\mathbb{E}[X^*] < (\pi_0^* - \pi_0)\pi_0^*$ for fixed T, as stated. □

Proof of Proposition 6. For part (i), first define the likelihood of a prior π_0 as

$$\mathcal{L}(\pi_0) \equiv \frac{\pi_0}{1-\pi_0}, \tag{A.11}$$

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 the history H_{t-1} is left implicit for simplicity. The likelihood for any belief π_t is defined as well following (A.11). 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). Bayes' rule gives that 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 (9) 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 ratio of 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 fully rational (satisfying Assumption 1) but with $\check{\phi}$ in place of ϕ. We know as well that $\check{\phi}$ satisfies Assumption 3, 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 2 holds as well, so all three assumptions are satisfied, and all the stated results carry through.

For part (iii), assuming $0 < \check{\phi} < 1$ (so Assumption 2 no longer holds for the fictitious rational agent), note first that the proof of Proposition 2 never employs Assumption 2 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 3, the result as stated for a rational agent requires that $\pi_0^* > \check{\pi}_0$, which is not true for $\check{\phi} < 1$. An alternative bound, though, can be shown for this case, by obtaining a lower bound for $\triangle$ similar to the upper bound in equation (18). Starting from (A.4) but solving now for $\mathbb{E}[\mathsf{m}^* | \theta = 1]$, we have

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

Using this in (A.3),

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

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 3, if $\check{\phi} < 1$, then $\pi_0^* - \check{\pi}_0 \leqslant 0$, which gives

$$\begin{aligned}\mathbb{E}[X^*] &= (\pi_0^* - \check{\pi}_0)(\triangle) \\ &\leqslant (\check{\pi}_0 - \pi_0^*)(1 - \pi_0^*).\end{aligned}$$

Further, as $\check{\pi}_0 \leqslant 1$,

$$\begin{aligned}\mathbb{E}[X^*] &\leqslant (\check{\pi}_0 - \pi_0^*)(1 - \pi_0^*) \\ &\leqslant (1 - \pi_0^*)(1 - \pi_0^*) = (1 - \pi_0^*)^2.\end{aligned}$$

Thus taking (ii) and (iii) together, we have that $\mathbb{E}[X^*] \leqslant \max(\pi_0^{*2}, (1 - \pi_0^*)^2)$ given an incorrect prior. □

A.2 Proofs for Section 3

Section 3.1

Proof of Equation (22). 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

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

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 analysis for the pair 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 equation (22). □

Section 3.2

Proof of Example 1. We prove the statement separately for the two assumptions on the form of the utility function:

(i) Time-separable utility: Denote $V_j^m \equiv V_0^m \theta_j$ and $V_{j+1}^m \equiv V_0^m \theta_{j+1}$, so the event $R_T^m = \theta_j$ is equivalent to $V_T^m = V_j^m$, and similarly for θ_{j+1} and V_{j+1}^m. Since $dV_T^m / dA_T > 0$ (and with $\mathbb{P}(V_T^m = V_j^m) > 0$, $\mathbb{P}(V_T^m = V_{j+1}^m) > 0$), there exist unique values a_j and a_{j+1} such that $V_T^m = V_j^m$ if and only if $A_T = a_j$, and $V_T^m = V_{j+1}^m$ if and only if $A_T = a_{j+1}$. Then with $M_{t,T} = \beta^{T-t} U'(C_T)/U'(C_t)$ given the assumptions for this example, we have

$$\phi_{t,j} \equiv \frac{\mathbb{E}_t[M_{t,T} \mid R_T^m = \theta_j]}{\mathbb{E}_t[M_{t,T} \mid R_T^m = \theta_{j+1}]} = \frac{\mathbb{E}_t[M_{t,T} \mid A_T = a_j]}{\mathbb{E}_t[M_{t,T} \mid A_T = a_{j+1}]}$$
$$= \frac{U'(C_T(a_j))}{U'(C_T(a_{j+1}))},$$

which is almost surely constant, as required for CTI to hold.

(ii) Epstein–Zin (1989) utility: The Epstein–Zin (1989) preference recursion is

$$U_t = \left\{ (1-\beta) C_t^{1-\frac{1}{\psi}} + \beta \left(\mathbb{E}_t \left[U_{t+1}^{1-\gamma} \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right\}^{\frac{1}{1-\frac{1}{\psi}}}. \tag{A.12}$$

It can be shown (e.g., Campbell, 2018, p. 178) that given such preferences the SDF evolves according to

$$M_{t,t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \left(\frac{U_{t+1}}{\mathbb{E}_t \left[U_{t+1}^{1-\gamma} \right]^{\frac{1}{1-\gamma}}} \right)^{-\left(\gamma - \frac{1}{\psi}\right)},$$

which gives that

$$M_{t,T} = M_{t,t+1} M_{t+1,t+2} \cdots M_{T-1,T}$$
$$= \beta^{T-t} \left(\frac{C_T}{C_t} \right)^{-\frac{1}{\psi}} \prod_{\tau=t}^{T-1} \left(\frac{U_{\tau+1}}{\mathbb{E}_\tau \left[U_{\tau+1}^{1-\gamma} \right]^{\frac{1}{1-\gamma}}} \right)^{-\left(\gamma - \frac{1}{\psi}\right)} \tag{A.13}$$
$$= \beta^{T-t} \left(\frac{C_T}{C_t} \right)^{-\gamma} \prod_{\tau=t}^{T-1} \left(\frac{U_{\tau+1}}{C_{\tau+1}} \right)^{-\left(\gamma - \frac{1}{\psi}\right)} \mathbb{E}_\tau \left[\left(\frac{C_{\tau+1}}{C_\tau} \right)^{1-\gamma} \left(\frac{U_{\tau+1}}{C_{\tau+1}} \right)^{1-\gamma} \right]^{\frac{\gamma - \frac{1}{\psi}}{1-\gamma}}. \tag{A.14}$$

Denote a_j and a_{j+1} as in part (i). From the first representation of $M_{t,T}$, equation (A.13), it follows immediately that with i.i.d. consumption (or i.i.d. innovations to an otherwise predetermined consumption path),

$$\phi_{t,j} = \frac{\mathbb{E}_t[M_{t,T} \mid A_T = a_j]}{\mathbb{E}_t[M_{t,T} \mid A_T = a_{j+1}]}$$
$$= \left(\frac{C_T(a_j)}{C_T(a_{j+1})}\right)^{-\frac{1}{\psi}} \left(\frac{U_T(a_j)}{U_T(a_{j+1})}\right)^{-\left(\gamma - \frac{1}{\psi}\right)},$$

which is almost surely constant given the definition (A.12) and that $E_\tau[U_{T+1}^{1-\gamma}]$ is constant given the i.i.d. assumption. When consumption growth C_t/C_{t-1} is i.i.d., note that the scale independence of Epstein–Zin (1989) utility in (A.12) allows us to guess and verify that U_t/C_t is constant almost surely. Then from the second representation of $M_{t,T}$, equation (A.14), we have in this case that

$$\phi_{t,j} = \left(\frac{C_T(a_j)}{C_T(a_{j+1})}\right)^{-\gamma}. \qquad \square$$

Proof of Example 2. Gabaix (2012, Theorem 1) shows that

$$V_t^m = \frac{D_t}{1 - e^{-\beta_m}} \left(1 + \frac{e^{-\beta_m - h_*} \widehat{H}_t}{1 - e^{-\beta_m - \phi_H}}\right),$$

where $h_* \equiv \log(1 + H_*)$ and $\beta_m \equiv \beta - g_d - h_*$ (where β is the agent's time discount factor). Thus for any value θ and given H_0, there exists some value d_θ and function $f(d_\theta, \widehat{H}_T)$, which is strictly increasing in the first argument and strictly decreasing in the second argument, such that, by Bayes' rule,

$$\mathbb{P}_0\left(\sum_{t=1}^{T} \mathbb{1}\{\text{disaster}_t\} > 0 \,\middle|\, R_T^m \geqslant \theta\right)$$
$$= \frac{\mathbb{P}_0\left(R_T^m \geqslant \theta \mid \sum_{t=1}^{T} \mathbb{1}\{\text{disaster}_t\} > 0\right) \mathbb{P}_0\left(\sum_{t=1}^{T} \mathbb{1}\{\text{disaster}_t\} > 0\right)}{\mathbb{P}_0(R_T^m \geqslant \theta)}$$
$$= \frac{\mathbb{P}_0\left(D_T \geqslant f(d_\theta, \widehat{H}_T) \,\middle|\, \sum_{t=1}^{T} \mathbb{1}\{\text{disaster}_t\} > 0\right) \mathbb{P}_0\left(\sum_{t=1}^{T} \mathbb{1}\{\text{disaster}_t\} > 0\right)}{\mathbb{P}_0\left(D_T \geqslant f(d_\theta, \widehat{H}_T)\right)}.$$

Note now that (i) the innovation to $\widehat{H}_{t+1}$ is independent of the disaster realization; (ii) F_{t+1} (the exponential of the disaster shock to D_t) has support $[0,1]$; and (iii) $\mathbb{P}_t(\varepsilon_{t+1}^d \geqslant \epsilon) = o(e^{-\epsilon^2})$ as $\epsilon \to \infty$.[2] Thus $\mathbb{P}_0(D_T \geqslant f(d_\theta, \widehat{H}_T) \mid \sum_{t=1}^{T} \mathbb{1}\{\text{disaster}_t\} > 0) = o(\mathbb{P}_0(D_T \geqslant f(d_\theta, \widehat{H}_T)))$ as $d_\theta \to \infty$,

[2] To see why point (iii) holds, denote $\sigma_d \equiv \text{Var}(\varepsilon_t^d)$ and then note that $\int_\epsilon^\infty \exp(-x^2/(2\sigma_d^2))/\sqrt{2\pi\sigma_d^2}\,dx <$

from which the first statement given in the example follows. Denote the value δ in that statement by $\delta = \delta_0$. Then it follows immediately that for any $t > 0$ (with $t < T$), for any $\delta_t > 0$, there exists an $\underline{\theta}$ such that $\mathbb{P}_t(\sum_{\tau=1}^{T} \mathbb{1}\{\text{disaster}_t\} > 0 \mid R_T^m \geqslant \underline{\theta}) < \delta_t$ asymptotically $\mathbb{P}_0$-a.s. as $\delta_0 \to 0$.

Thus moving to the second statement, given a value $\delta_t > 0$, consider θ_j, θ_{j+1} large enough that $\mathbb{P}_t(\sum_{\tau=1}^{T} \mathbb{1}\{\text{disaster}_t\} > 0 \mid R_T^m \in \{\theta_j, \theta_{j+1}\}) < \delta_t$. We then have from (24) that

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

Note that the fraction in the last expression is constant almost surely given that conditional on $\sum_{t=1}^{T} \mathbb{1}\{\text{disaster}_t\} = 0$, the conditions from Example 1 hold. Thus denoting

$$\phi_j \equiv \frac{\mathbb{E}_0[M_{t,T} \mid R_T^m = \theta_j, \sum_{t=1}^{T} \mathbb{1}\{\text{disaster}_\tau\} = 0]}{\mathbb{E}_0[M_{t,T} \mid R_T^m = \theta_{j+1}, \sum_{t=1}^{T} \mathbb{1}\{\text{disaster}_\tau\} = 0]},$$

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 $\delta = \delta_0 \to 0$. □

Proof of Example 3. As in Campbell and Cochrane (1999), the SDF evolves according to

$$M_{t,t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \left(\frac{S_{t+1}^c}{S_t^c} \right)^{-\gamma},$$

with terms defined as in Appendix B.4, and thus

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

$\int_\epsilon^\infty (x/\epsilon) \exp(-x^2/(2\sigma_d^2))/\sqrt{2\pi\sigma_d^2}\, dx = \sigma_d \exp(-\epsilon^2/(2\sigma_d^2))/(\sqrt{2\pi}\epsilon)$. 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 $\mathbb{P}_0(D_T \geqslant f(d_\theta, \widehat{H}_T) \mid \sum_{t=1}^{T} \mathbb{1}\{\text{disaster}_t\} > 0)$ and the lower-bound calculation to $\mathbb{P}_0(D_T \geqslant f(d_\theta, \widehat{H}_T))$, it follows that $\mathbb{P}_0(D_T \geqslant f(d_\theta, \widehat{H}_T) \mid \sum_{t=1}^{T} \mathbb{1}\{\text{disaster}_t\} > 0)/\mathbb{P}_0(D_T \geqslant f(d_\theta, \widehat{H}_T)) = o(1)$, 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).

For a counterexample to the constant-ϕ_t restriction, set $T = 2$ and $c_t = d_t$ (i.e., for simplicity, consumption and dividends are identical, as in the simplest case considered by Campbell and Cochrane, 1999, so the market is a consumption claim). Note that a sufficient condition for time variation in ϕ_t is

$$\text{Cov}_0(\phi_1, \mathbb{E}_1[M_{1,2} \mid R_2^m = \theta_{j+1}]) \neq 0, \tag{A.15}$$

as this implies $\mathbb{E}_0[\phi_1] \neq \phi_0$ (see after Proposition 11). As of $t = 0$, both ε_1 and ε_2 are relevant for R_2^m and $M_{0,2}$, as ε_1 determines s_1^c and thus the conditional volatility $\lambda(s_1^c)$ of surplus consumption s_2^c. Meanwhile, as of $t = 1$ (i.e., conditional on ε_1), the only source of uncertainty for both R_2^m and $M_{1,2}$ is ε_2: s_2^c and d_2 together determine R_2^m, and conditional on time-1 variables, these both depend only on ε_2. Thus write ε_j^1 for the realization of ε_2 needed to generate $R_2^m = \theta_j$ conditional on ε_1 — i.e., $\varepsilon_j^1 \equiv \{\varepsilon_2 : R_2^m = \theta_j \mid \varepsilon_1\}$ — and similarly write ε_{j+1}^1 for the realization of ε_2 needed for $R_2^m = \theta_{j+1}$ conditional on ε_1. We then have

$$\mathbb{E}_1[M_{1,2} \mid R_2^m = \theta_{j'}] = \exp\Big(-\gamma\,(1 + \lambda(s_1^c))\,\varepsilon_{j'}^1\Big)$$

for $j' = j, j+1$, and thus

$$\phi_1 = \exp\Big(-\gamma\,(1 + \lambda(s_1^c))\,\Big(\varepsilon_j^1 - \varepsilon_{j+1}^1\Big)\Big).$$

The covariance in (A.15) is therefore

$$\text{Cov}_0\left(\exp\Big(-\gamma\,(1 + \lambda(s_1^c))\,\Big(\varepsilon_j^1 - \varepsilon_{j+1}^1\Big)\Big), \exp\Big(-\gamma\,(1 + \lambda(s_1^c))\,\varepsilon_{j+1}^1\Big)\right).$$

Given Gaussian ε_1, this value is generically non-zero. □

Additional Lemmas Used in Proofs for Section 3.3

Before proceeding to the proof of our main results, we provide two additional lemmas that are useful in proving those results. As usual, assume throughout that Assumptions 1′–3′ hold.

Lemma A.1. *For some return-state pair (θ_j, θ_{j+1}), with $\widetilde{\mathbb{P}} \equiv \mathbb{P}(\cdot \mid R_T^m \in \{\theta_j, \theta_{j+1}\})$ as per (23), define a new pseudo-risk-neutral measure $\widetilde{\mathbb{P}}^\diamond$ by*

$$\left.\frac{d\widetilde{\mathbb{P}}^\diamond}{d\widetilde{\mathbb{P}}}\right|_{H_t} = \frac{\widetilde{\pi}_{t,j}^*}{\widetilde{\pi}_{t,j}} \mathbb{1}\{R_T^m = \theta_j\} + \frac{1 - \widetilde{\pi}_{t,j}^*}{1 - \widetilde{\pi}_{t,j}} \mathbb{1}\{R_T^m = \theta_{j+1}\}. \tag{A.16}$$

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

$$\widetilde{\pi}_{t,j}^* = \widetilde{\mathbb{E}}_t^\diamond[\widetilde{\pi}_{t+1,j}^*]. \tag{A.17}$$

We conclude from Proposition 1 that

$$\widetilde{\mathbb{E}}_0^{\diamond}[X_j^*] = 0. \tag{A.18}$$

Proof of Lemma A.1. From (23)–(24), we have after some algebra that

$$\frac{\widetilde{\pi}_{t,j}^*}{\widetilde{\pi}_{t,j}} = \frac{\phi_j}{1 + \widetilde{\pi}_{t,j}(\phi_j - 1)}, \tag{A.19}$$

$$\frac{1 - \widetilde{\pi}_{t,j}^*}{1 - \widetilde{\pi}_{t,j}} = \frac{1}{1 + \widetilde{\pi}_{t,j}(\phi_j - 1)}. \tag{A.20}$$

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 the stochastic process $\{s_t\}$ and $s_t \in \mathcal{S}$ is the date-t signal vector. Denote $N_{\mathcal{S}} \equiv |\mathcal{S}|$, so that $s_t \in \{s_1, s_2, \ldots, s_{N_{\mathcal{S}}}\}$, and further denote

$$\begin{aligned}
\mathfrak{p}_{t,k} &\equiv \widetilde{\mathbb{P}}_t(s_{t+1} = \theta_k), \\
\varrho_{t,k} &\equiv \widetilde{\mathbb{P}}_t(R_T^m = \theta_j \mid s_{t+1} = s_k), \\
\varrho_{t,k}^* &\equiv \mathbb{P}_t^*(R_T^m = \theta_j \mid s_{t+1} = s_k, R_T^m \in \{\theta_j, \theta_{j+1}\}),
\end{aligned}$$

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 (A.16), (A.19), (A.20), and these definitions, we have

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

$$= \widetilde{\pi}^*_{t,j},$$

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 (A.19). The fact that $\widetilde{\mathbb{E}}^{\diamond}_0[X^*_j] = 0$ then follows immediately from the proof of Proposition 1. □

LEMMA A.2. *For any return-state pair (θ_j, θ_{j+1}) meeting CTI, risk-neutral belief movement must satisfy the following for $j' = j, j+1$:*

$$\widetilde{\mathbb{E}}^{\diamond}_0[\mathsf{m}^*_j \mid R^m_T = \theta_{j'}] = \widetilde{\mathbb{E}}_0[\mathsf{m}^*_j \mid R^m_T = \theta_{j'}]. \tag{A.21}$$

Proof of Lemma A.2. The stream of risk-neutral beliefs is $\boldsymbol{\pi}^*_j \equiv (\widetilde{\pi}^*_{0,j}, \widetilde{\pi}^*_{1,j}, \ldots, \widetilde{\pi}^*_{T,j})$, and denote some arbitrary realization for that path by $\mathfrak{b}_j$. The realization of m^*_j depends on the path of risk-neutral beliefs, so denote $\mathsf{m}^*_j = \mathsf{m}^*_j(\boldsymbol{\pi}^*_j) = \sum_{t=1}^{T} (\widetilde{\pi}^*_{t,j} - \widetilde{\pi}^*_{t-1,j})^2$.

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 (A.16) gives that

$$\widetilde{\mathbb{P}}^{\diamond}_0(\boldsymbol{\pi}^*_j = \mathfrak{b}_j) = \frac{\widetilde{\pi}^*_{0,j}}{\widetilde{\pi}_{0,j}} \widetilde{\mathbb{P}}(\boldsymbol{\pi}^*_j = \mathfrak{b}_j), \tag{A.22}$$

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 of the RN measure in Section 2.) 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 steps apply for $R^m_T = \theta_{j+1}$: for any $\mathfrak{b}_j$ such that $\widetilde{\pi}^*_{T,j} = 0$, (A.22) now becomes $\widetilde{\mathbb{P}}^{\diamond}_0(\boldsymbol{\pi}^*_j = \mathfrak{b}_j) = (1 - \widetilde{\pi}^*_{0,j})/(1 - \widetilde{\pi}_{0,j})\, \widetilde{\mathbb{P}}(\boldsymbol{\pi}^*_j = \mathfrak{b}_j)$. We also have in this case that $\widetilde{\mathbb{P}}^{\diamond}_0(R^m_T = \theta_{j+1}) = (1 - \widetilde{\pi}^*_{0,j})/(1 - \widetilde{\pi}_{0,j})\, \widetilde{\mathbb{P}}_0(R^m_T = \theta_{j+1})$, 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 (A.16) aligns with the definition of the RN measure in equation (11), 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 (11) and equation (14), as well as the footnote in the proof of equation (13) above). Indeed, equation (A.17) is the precise analogue to equation (14) in the text; equation (A.18) is the analogue to equation (13); and equation (A.21) implies immediately that $\widetilde{\mathbb{E}}_0^{\diamond}[X_j^* \mid R_T^m] = \widetilde{\mathbb{E}}_0[X_j^* \mid R_T^m]$, which was the main implication of equation (14) 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.

Section 3.3

Proof of Proposition 7. No arbitrage gives the existence of a positive SDF for which equation (24) and Assumption 2′ are valid. We have

$$\begin{aligned} \widetilde{\pi}_{t,j} &= \mathbb{E}_t[\widetilde{\pi}_{t+1,j}], \\ \widetilde{\pi}_{t,j}^* &= \widetilde{\mathbb{E}}_t^{\diamond}[\widetilde{\pi}_{t+1,j}^*], \\ \widetilde{\mathbb{E}}_0^{\diamond}[X_j^*] &= 0, \\ \widetilde{\mathbb{E}}_0^{\diamond}[X_j^* \mid R_T^m] &= \widetilde{\mathbb{E}}_0[X_j^* \mid R_T^m], \end{aligned}$$

where the first equality uses LIE and the remainder use Lemmas A.1–A.2 as above. The last equation in addition implies, using the same argument as applied for equation (18), that $\triangle_j \leqslant \widetilde{\pi}_{0,j}^*$. Further, Equations (9)–(10) hold immediately for $\widetilde{\pi}_{t,j}, \widetilde{\pi}_{t,j}^*, \phi_j$. We have thus obtained all the conditions used to prove Propositions 1–6 and Corollaries 1–2 given Assumptions 1–3, and thus under Assumptions 1′–3′ (for $j = 2, 3, \ldots, J-2$), 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 $\triangle_j \equiv \widetilde{\mathbb{E}}_0[X_j^* \mid R_T^m = \theta_{j+1}] - \widetilde{\mathbb{E}}_0[X_j^* \mid R_T^m = \theta_j]$ replacing $\triangle$, as stated. □

Proof of Proposition 8. The result follows immediately from equation (8), with V_j^m and V_{j+1}^m replacing $C_{T,1}$ and $C_{T,0}$, respectively. □

A.3 Proofs for Section 4

Section 4.1

Proof of Proposition 9. In what follows, we will often use $\mathbb{E}_i[\cdot]$ to make explicit that we are taking expectations over DGPs indexed by i, and for now we will use the notational simplifications used in the statement of the proposition. First, for (i), start with the case fixing $\pi_{0,i}^* = \pi_0^*$ across i. Applying Proposition 2,

$$\begin{aligned} \mathbb{E}_i[\mathbb{E}[X_i^*]] &= \mathbb{E}_i[(\pi_0^* - \pi_{0,i}) \cdot \triangle_i] \\ &= \mathbb{E}_i[\pi_0^* \cdot \triangle_i] - \mathbb{E}_i[\pi_{0,i} \cdot \triangle_i] \\ &= \pi_0^* \cdot \mathbb{E}_i[\triangle_i] - \mathbb{E}_i[\pi_{0,i}] \cdot \mathbb{E}_i[\triangle_i] \\ &= (\pi_0^* - \mathbb{E}_i[\pi_{0,i}]) \cdot \mathbb{E}_i[\triangle_i] \\ &= \mathbb{E}_i[\pi_0^* - \pi_{0,i}] \cdot \mathbb{E}_i[\triangle_i] \end{aligned}$$

$$= \mathbb{E}_i \left[\pi_0^* - \frac{\pi_0^*}{\phi_i + (1-\phi_i)\pi_0^*} \right] \cdot \mathbb{E}_i[\triangle_i]$$

where the main step in line three from $\mathbb{E}_i[\pi_{0,i} \cdot \triangle_i]$ to $\mathbb{E}_i[\pi_{0,i}] \cdot \mathbb{E}[\triangle_i]$ follows from the assumption that $\text{Cov}(\pi_{0,i}, \triangle_i) = 0$.

Now consider $\zeta_1(\phi_i, \pi_0^*) \equiv \pi_0^* - \frac{\pi_0^*}{\phi_i + (1-\phi_i)\pi_0^*}$. This is a concave function in ϕ_i given that $\pi_0^* \in [0,1]$ and $\phi_i \geqslant 1$; the second derivative of this function is

$$\frac{\partial^2 \zeta_1}{\partial \phi_i^2} = \frac{-2\pi_0^*(1-\pi_0^*)^2}{(\pi_0^* + \phi(1-\pi_0^*))^3},$$

which is weakly negative if $\pi_0^* \in [0,1]$ and $\phi \geqslant 1$. Thus using Jensen's inequality, the expectation of the function over ϕ_i must be less than the function evaluated at the expectation of ϕ_i:

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

where $\underline{\phi} \equiv \mathbb{E}_i[\phi_i]$. Now, returning to the equation above, suppose that $\mathbb{E}_i[\triangle_i] > 0$. In this case,

$$\begin{aligned} \mathbb{E}_i[\mathbb{E}[X_i^*]] &= \mathbb{E}_i[\pi_0^* - \frac{\pi_0^*}{\phi_i + (1-\phi_i)\pi_0^*}] \cdot \mathbb{E}_i[\triangle_i] \\ &\leqslant (\pi_0^* - \frac{\pi_0^*}{\underline{\phi} + (1-\underline{\phi})\pi_0^*}) \cdot \mathbb{E}_i[\triangle_i] \end{aligned}$$

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

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

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,

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

Following the same logic 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 $\pi_0^* \geqslant 0$,

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

as stated in the second inequality. Now allowing $\pi_{0,i}^*$ to vary, write the upper bound for $\mathbb{E}[X^*]$ in Proposition 3 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

$$\mathbb{E}[\zeta_{2'}(\phi_i, \pi_{0,i}^*) \mid \pi_{0,i}^*] \leqslant \zeta_{2'}\left(\mathbb{E}[\phi_i \mid \pi_{0,i}^* = \varrho], \varrho\right).$$

Now, using Proposition 3 and applying LIE to the above inequality,

$$\mathbb{E}[X_i^*] \leqslant \mathbb{E}[\zeta_{2'}(\phi_i, \pi_{0,i}^*)] \leqslant \mathbb{E}\left[\zeta_{2'}\left(\mathbb{E}[\phi_i \mid \pi_{0,i}^*], \pi_{0,i}^*\right)\right]$$
$$\leqslant \mathbb{E}\left[\zeta_{2'}(\overline{\phi}, \pi_{0,i}^*)\right], \tag{A.23}$$

where $\overline{\phi}$ is as in the proposition statement and where the second line uses $\partial\zeta_{2'}/\partial\phi_i \geqslant 0$. Substituting the definition of $\zeta_{2'}$ into this inequality yields equation (27).

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 (A.23) 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.)

Finally, for part (iv), Corollary 2 notes 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 $\mathbb{E}[X_i^*] \leqslant 0$ for all i and thus over all streams, completing the proof. □

Section 4.2

Proof of Proposition 10. Starting with measured belief movement, under the stated assumptions for ϵ_t,

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

For the measured 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 $\text{Var}(\epsilon_t) \equiv \mathbb{E}[(\epsilon_t - \mathbb{E}[\epsilon_t])^2] = \mathbb{E}[\epsilon^2_t]$ and $X^*_{t,t+1} \equiv \mathsf{m}^*_{t,t+1} - \mathsf{r}^*_{t,t+1}$,

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

Section 4.3

Proof of Proposition 11. We work in the context of the example in Section 2 for simplicity of notation (and without loss of generality), so write ϕ_t as in equation (7) as $\mathbb{E}_t[U'(C_{T,1})]/\mathbb{E}_t[U'(C_{T,0})]$, where the two states are again $\theta = 0, 1$, but where ϕ_t can now vary. Regardless of the DGPs for θ, $U'(C_{T,1})$ and $U'(C_{T,1})$, we will show that if ϕ_t evolves as a martingale or supermartingale ($\mathbb{E}[\phi_{t+1}] \leqslant [\phi_t]$), then the bounds in Proposition 3 and Corollary 1 apply with ϕ_0 replacing ϕ.

Intuition: Given the length of the proof, it is useful to briefly outline the steps and intuition. (1) First, we focus on a particular situation in which a) ϕ_t can only take a high, medium, and low value, and b) ϕ_t evolves as a martingale. Then, we assume — for the sake of contradiction — that there exists a DGP in which ϕ moves in a way that "beats our bounds" (produces expected RN movement that is higher than that in our bounds). We then focus on the highest-movement DGP where ϕ changes and focus on the last meaningful movement of ϕ in this DGP. We show that expected RN movement is strictly increased if ϕ instead remains constant, leading to a contradiction. We conclude that there is not a DGP with expected RN movement that beats our bounds in which a) ϕ_t evolves as a martingale and b) ϕ_t only takes three values. Then, we expand this result to the general case. (2) First, we consider the situation in which ϕ_t can instead take an arbitrary number of values. We show that if there exists a general DGP that beats our bound in which ϕ_t is a martingale, there must exist a DGP that beats our bounds in which ϕ_t is a martingale and evolves into three values. But, given that we proved that this is not possible, we conclude that there is not a DGP that beats our bounds in which ϕ_t evolves as a martingale. (3) Finally, we expand this result to supermartingales. We show that if there exists a DGP in which ϕ_t evolves as a supermartingale and beats our bounds, there must be a DGP in which ϕ_t evolves as a martingale and beats our bounds. But, given that we proved that this is not possible, we conclude that there is not a DGP that beats our bounds.

Setup: Given that ϕ_t can change, we explicitly allow it to depend on the signal history. Therefore, RN beliefs are now denoted:

$$\pi^*_t(\mathbf{H}_t) = \frac{\phi_t(\mathbf{H}_t) \cdot \pi_t(\mathbf{H}_t)}{(\phi_t(\mathbf{H}_t) - 1)\pi_t(\mathbf{H}_t) + 1}$$

We still assume that the uncertainty about θ is resolved by period T. We allow more periods to allow resolution of uncertainty about ϕ, although we now show this is inconsequential. Specifically, movement in ϕ_t is only consequential for RN movement when there is still uncertainty about θ. Specifically, as $\pi_T \in \{0,1\}$, then π_t must be constant for any $t \geqslant T$. Also, if $\pi_t \in \{0,1\}$, then $\pi^*_t = \pi_t$, so RN beliefs must also be constant for $t \geqslant T$. Therefore, there is no RN movement for $t \geqslant T$, regardless of whether ϕ_t for is changing over these periods. Given that ϕ_t then has no impact on RN movement for $t \geqslant T$, we can restrict our attention to DGPs in which ϕ_t is constant for $t \geqslant T$.

Proof: Now, assume for the sake of contradiction that there exists some DGP in which ϕ_t changes and expected RN movement is higher than the bounds in Proposition 3 for some T. Consider a DGP of 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 $\phi_{t+1}(\mathbf{H}_t \cup s_{t+1}) \neq \phi_t(\mathbf{H}_t)$) but for which ϕ_t is constant after $t+1$. We will now show that, in fact, expected RN movement is higher if ϕ_t is constant following H_t, contradicting the assumption that the ϕ-changing DGP has the highest expected RN movement.

(1) We start by focusing on a particular situation, drawing conclusions for this situation, and then showing how the results from this situation extend to the general case. We start by considering the case in which, following any H_t,

(a) $\phi_{t+1}(\mathbf{H}_t \cup s_{t+1})$ can only take three values:

- $\phi^H_{t+1} > \phi_t$ following signal s^H_{t+1} with probability $q^H > 0$
- $\phi^M_{t+1} = \phi_t$ following signal s^M_{t+1} with probability $q^M \geqslant 0$
- $\phi^L_{t+1} < \phi_t$ following signal s^L_{t+1} with probability $q^L > 0$

(b) ϕ_t evolves as a martingale: $\sum_{i\in\{L,M,H\}} q^i \cdot \phi^i_{t+1} = \phi_t$.

Given these assumptions and 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, $\pi^*_t(\mathbf{H}_t \cup s_{t+1})$ can take at most three values: $\pi^{*i}_{t+1} = \frac{\phi^i_{t+1}\cdot\pi_t}{(\phi^i_{t+1}-1)\pi_t+1}$ for $i \in \{L, M, H\}$.

Now, we will 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

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

The second line is expressed in terms of π_t and ϕ_t rather than π^*_t as this will turn out to be easier later given that, unlike in the rest of the paper, ϕ_t is not constant. The third line uses our assumption that $\pi_t = \pi_{t+1}$.

After period $t+1$, ϕ_t is assumed to be constant, so our main bounds hold with π_0^* replaced with π_{t+1}^{*i} and ϕ replaced with ϕ_{t+1}^i. For example, expected RN movement (excess RN movement plus initial RN uncertainty) given signal s_{t+1}^i from period $t+1$ onward is then bounded above by:

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

where the first line is the definition of RN movement, the second line plugs in our bound for excess RN movement and uncertainty reduction given that uncertainty is zero at period T, the third line simplifies, the fourth line casks everything in terms of ϕ_t and π_t, and the final line uses our assumption that $\pi_t = \pi_{t+1}$.

Therefore, expected RN movement from period t onward following history H_t is then bounded above by:

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

We now show that — given that ϕ_t evolves as a martingale — this DGP will have higher RN movement if ϕ_t is constant from H_t onward. To see this, consider the "worst-case" DGP noted in Proposition 5 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 ϕ changes ($\mathbb{E}[\mathsf{m}_{t,T}^*|H_t]$) 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. To start, note that as $\phi_{t+1}^M = \phi_t$, the martingale assumption can be rewritten:

$$\begin{aligned}
\sum_{i\in\{L,M,H\}} q^i \cdot \phi_{t+1}^i &= \phi_t \\
\sum_{i\in\{L,M,H\}} q^i \cdot \phi_{t+1}^i &= \sum_{i\in\{L,M,H\}} q^i \cdot \phi_t \\
\sum_{i\in\{L,H\}} q^i \cdot \phi_{t+1}^i &= \sum_{i\in\{L,H\}} q^i \cdot \phi_t
\end{aligned}$$

$$\sum_{i\in\{L,H\}} \frac{q^i}{q^H+q^L}\cdot\phi^i_{t+1} = \sum_{i\in\{L,H\}} \frac{q^i}{q^H+q^L}\cdot\phi_t$$
$$\sum_{i\in\{L,H\}} p^i\cdot\phi^i_{t+1} = \sum_{i\in\{L,H\}} p^i\cdot\phi_t$$
$$\sum_{i\in\{L,H\}} p^i\cdot\phi^i_{t+1} = \phi_t,$$

where $p^i \equiv \frac{q^i}{q^H+q^L}$. Similarly, the difference $\mathbb{E}_{maxDGP}[\mathsf{m}^*_{t,T}|H_t] - \mathbb{E}[\mathsf{m}^*_{t,T}|H_t]$ is positive if and only if:

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

Substituting this martingale equation into the modified difference equation above and simplifying yields $\mathbb{E}_{maxDGP}[\mathsf{m}^*_{t,T}|H_t] - \mathbb{E}[\mathsf{m}^*_{t,T}|H_t] > 0$ if and only if

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

While the expression on the left side of the inequality is rather long, it is in fact straightforward to see that it 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 our assumptions (a) and (b) that produces more expected RN movement than the bound in Proposition 3.

(2) We now extend this observation to DGPs in which assumption (a) is relaxed. In particular, we now consider a DGP$_{arb}$ in which ϕ_{t+1} following H_t can now take an arbitrary number of values (indexed by i). Following the proof above, our goal is to show that for every DGP$_{arb}$, $\mathbb{E}_{maxDGP}[\mathsf{m}^*_{t,T}|H_t] \geqslant \mathbb{E}_{DGP_{arb}}[\mathsf{m}^*_{t,T}|H_t]$. To do this, we will show that if there exists a DGP$_{arb}$ such that $\mathbb{E}_{maxDGP}[\mathsf{m}^*_{t,T}|H_t] < \mathbb{E}_{DGP_{arb}}[\mathsf{m}^*_{t,T}|H_t]$ is true, there is a contradiction. Specifically, we show that this inequality would imply that there must exist a DGP satisfying assumption (a) in which trinary movements lead to $\mathbb{E}_{maxDGP}[\mathsf{m}^*_{t,T}|H_t] < \mathbb{E}[\mathsf{m}^*_{t,T}|H_t]$. But, given that we just showed this is not possible, we have a contradiction and it must be that the DGP$_{arb}$ does not exist.

Intuition: To do this, we will show that the expected RN movement of DGP$_{arb}$ at history H_t can be replicated with a DGP that only using trinary movements. Although the implementation is annoying clunky, the intuition is simple: rather that doing all of the arbitrary martingale movements of ϕ in period $t+1$, we "separate out" the movements into trinary martingale movements in sequential periods. For example, suppose that $\phi_t = 3$ and $\phi_t + 1$ takes values 1, 2, 4, and 5 with probability .25 and $\pi_t = \pi_{t+1} = 1/2$ That is, we have a martingale process on ϕ with a constant

π that leads $\pi_t^* = 3/4$ and π_{t+1}^* to take values $1/2, 2/3, 4/5$ and $5/6$ with equal probability, such that RN movement then takes values $(3/4 - 1/2)^2$, $(3/4 - 2/3)^2$, $(3/4 - 4/5)^2$, and $(3/4 - 5/6)^2$ with equal probability. Instead, consider a sequential DGP in which $\phi_t = 3$ but ϕ_{t+1} equals 1 with probability .25 and 5 with probability .25, but stays constant at 3 with probability .5. Then, following ϕ_{t+1} staying constant at 3, ϕ_{t+2} can take values 2 or 4 with equal probability (and, when ϕ_{t+1} equals 1, ϕ_{t+2} remains constant at 1 and when ϕ_{t+1} equals 5, ϕ_{t+2} remains constant at 5). Note that this new sequential DGP only used trinary martingale movements following every history. And, note that the likelihood of each outcome is the same: the sequential DGP has a equal probability of ending with ϕ equal to $1, 2, 4$, and 5. Finally, the total RN movement of the sequential DGP between t and $t+2$ matches the RN movement of the original DGP between t and $t+1$. For example, there is a .25 probability that RN beliefs shift from $3/4$ to $1/2$ and then stay constant at $1/2$, for a total RN movement of $(3/4 - 1/2)^2 + (1/2 - 1/2)^2 = (1/2 - 3/4)^2$. Similarly, there is a .25 probability that RN beliefs stay constant at $3/4$ and then shifts from $3/4$ to $2/3$, for a total RN movement of $(3/4 - 3/4)^2 + (3/4 - 2/3)^2 = (1/2 - 2/3)^2$. Overall, then, RN movement still takes values $(3/4 - 1/2)^2$, $(3/4 - 2/3)^2$, $(3/4 - 4/5)^2$, and $(3/4 - 5/6)^2$ with equal probability. This simple idea is slightly clunky to implement because one needs an algorithm to separate the arbitrary number of movements into individual movements that satisfy the martingale property.

To understand the algorithm, consider a *full* list of the probabilities of each ϕ_{t+1} and the differences $\phi_{t+1} - \phi_t$, i.e. $\{(q^1, \phi_{t+1}^1 - \phi_t), (q^2, \phi_{t+1}^i - \phi_t), \ldots\}$. In the example DGP mentioned above, this difference list would be $\{(.25, -2), (.25, -1), (.25, 1), (.25, 2)\}$. For each two-part component in a list, we define the *product* of the component as $(q^i \cdot (\phi_{t+1}^i - \phi_t)$. As we assume that ϕ_t is a martingale, it must be that the sum of the product of the components of the full list is zero:

$$\sum_i q^i \cdot \phi_{t+1}^i = \phi_t$$
$$\sum_i q^i \cdot (\phi_{t+1}^i - \phi_t) = 0$$

Therefore, we call the full list *balanced*. Our first step is to remove from the list any component with a difference of 0 (that is, remove components where ϕ_{t+1}^i is equal to ϕ_t). Given this change, note that the list is still balanced as we removed a difference of zero. Also, note that the list must still contain some elements as we assumed that there existed some signal at history H_t for which $\phi_{t+1}^i \neq \phi_t$. Now, we consider an algorithm on this list to create j binary balanced lists, which each have two members. To do this, we start by separating the full list into two sub-lists depending on whether the difference is positive or negative. Note that as the main list was balanced (the sum of the products in the list was zero), the sum of the products in the positive and negative lists must be equal. For example, in the above example, the negative list would be $\{(.25, -2), (.25, -1)\}$ and the positive list would be $\{(.25, 2), (.25, 1)\}$. The algorithm then proceeds as such:

1. Enter with a positive and negative difference list in which the sum of the products is equal. Consider the first element of the current positive list $(q_{pos}^1, \phi_{t+1,pos}^1 - \phi_t)$ and negative list

$(q^1_{neg}, \phi^1_{t+1,neg} - \phi_t)$.

- If $q^1_{pos} \cdot (\phi^1_{t+1,pos} - \phi_t) - q^1_{neg} \cdot (\phi^1_{t+1,neg} - \phi_t) \geqslant 0$:
  - Let $q^* \leqslant q^1_{pos}$ solve $q^* \cdot (\phi^1_{t+1,pos} - \phi_t) - q^1_{neg} \cdot (\phi^1_{t+1,neg} - \phi_t) = 0$.
  - Add the balanced list $\{(q^*, \phi^1_{t+1,pos} - \phi_t), (q^1_{neg}, \phi^1_{t+1,neg} - \phi_t)\}$ to set of the binary difference lists.
  - Modify the current positive and negative difference lists: remove $(q^1_{neg}, \phi^1_{t+1,neg} - \phi_t)$ from the current negative list and replace $(q^1_{pos}, \phi^1_{t+1,pos} - \phi_t)$ in the current positive list with $(q^1_{pos} - q^*, \phi^1_{t+1,pos} - \phi_t)$. Note that as these subtractions are equal, the sum of the current negative and positive lists remains equal.
- If $q^1_{pos} \cdot (\phi^1_{t+1,pos} - \phi_t) - q^1_{neg} \cdot (\phi^1_{t+1,neg} - \phi_t) < 0$:
  - Let $q^* < q^1_{neg}$ solve $q^i_{pos} \cdot (\phi^1_{t+1,pos} - \phi_t) - q^* \cdot (\phi^1_{t+1,neg} - \phi_t) = 0$.
  - Add the balanced list $\{(q^1_{pos}, \phi^1_{t+1,pos} - \phi_t), (q^*, \phi^1_{t+1,neg} - \phi_t)\}$ to the binary lists.
  - Modify the current positive and negative lists: remove $(q^1_{pos}, \phi^1_{t+1,pos} - \phi_t)$ from the current positive list and replace $(q^1_{neg}, \phi^1_{t+1,neg} - \phi_t)$ in the current negative list with $(q^1_{neg} - q^*, \phi^1_{t+1,neg} - \phi_t)$. Note that as these subtractions are equal, the sum of the current negative and positive lists remains equal.

2. If there are no elements left in the current positive and negative list, end the algorithm. Otherwise, repeat.

We are left with a set of j balanced binary lists, each with two members. In the above example, there would be two binary lists: $\{(.25, -2), (.25, 2)\}$ and $\{(.25, -1), (.25, 1)\}$. Intuitively, we have taken the original balanced difference list $\{(.25, -2), (.25, -1), (.25, 1), (.25, 2)\}$ and broken it into a set of binary balanced lists. We now take our set of j binary balanced lists and use an algorithm to create a new sequential DGP starting at history H_t and lasting to period $t + j$.

1. In the initial period period $t + 1$:
   - With probability q^1_{pos}, let ϕ_{t+1} equal $\phi^1_{t+1,pos}$.
   - With probability q^1_{neg}, let ϕ_{t+1} equal $\phi^1_{t+1,neg}$.
   - With probability $1 - q^1_{pos} - q^1_{neg}$, let ϕ_{t+1} equal ϕ_t.
2. In period $t + k$, we enter the period with (1) histories in which ϕ has remained constant between period t and $t + k - 1$, which has occurred with probability $1 - \sum_{j=1}^{k-1} q^j_{neg} - \sum_{j=1}^{k-1} q^j_{pos}$ and (2) histories in which ϕ has changed between period t and $t + k - 1$, which has occured with probability $\sum_{j=1}^{k-1} q^j_{neg} + \sum_{j=1}^{k-1} q^j_{pos}$.
   - For histories in which ϕ has changed, let $\phi_{t+k} = \phi_{t+k-1}$ with probability 1.
   - For histories in which ϕ has remained constant:

- With probability $\frac{q_{pos}^k}{1-\sum_{j=1}^{k-1} q_{neg}^j - \sum_{j=1}^{k-1} q_{pos}^j}$, let ϕ_{t+k} equal $\phi_{t+1,pos}^k$.
- With probability $\frac{q_{neg}^k}{1-\sum_{j=1}^{k-1} q_{neg}^j - \sum_{j=1}^{k-1} q_{pos}^j}$, let ϕ_{t+k} equal $\phi_{t+1,neg}^k$.
- With probability $\frac{1-q_{pos}^k - q_{neg}^k}{1-\sum_{j=1}^{k-1} q_{neg}^j - \sum_{j=1}^{k-1} q_{pos}^j}$, let ϕ_{t+k} equal ϕ_t.

Therefore, we enter period $t+k+1$ with (1) histories in which ϕ has remained constant between period t and $t+k-1$, which has occurred with probability $1-\sum_{j=1}^{k} q_{neg}^j - \sum_{j=1}^{k} q_{pos}^j$ and (2) histories in which ϕ has changed between period t and $t+k$, which has occured with probability $\sum_{j=1}^{k} q_{neg}^j + \sum_{j=1}^{k} q_{pos}^j$.

This algorithm outputs a newly-constructed DGP that has a set of important characteristics between period t and $t+j$:

1. It is composed entirely of trinary movements from period t to period $t+j$.

2. At period $t+j$, the newly-constructed trinary DGP has the same probability distribution of $\phi_t + j$ as at period $t+j$ as the arbitrary DGP has for ϕ_{t+1} at period $t+1$ following history H_t. To see this, first focus on the values of ϕ_{t+1} that are positive. All of these values and probabilities are collected in the positive binary lists above. For the newly-constructed trinary DGP, in period $t+1$, there is a q_{pos}^1 probability that ϕ_{t+1} equals $\phi_{t+1,pos}^1$. Then, following period $t+1$, there will be no movement in ϕ from period $t+1$ to period $t+j$ following the movement to $\phi_{t+1,pos}^1$. Therefore, in period $t+j$, there will be histories (which occur with probability q_{pos}^1) in which ϕ_{t+j} equals $\phi_{t+1,pos}^1$. Next, in period $t+2$, there will be a set of histories (which occur with probability $1-q_{pos}^1 - q_{neg}^1$) in which there was no movement in ϕ from period t to $t+1$. Conditional on reaching this history, there is a $\frac{q^2}{1-q_{pos}^1-q_{neg}^1}$ probability that ϕ_{t+2} equals $\phi_{t+1,pos}^2$. Therefore, the unconditional probability is $(1-q_{pos}^1 - q_{neg}^1) \cdot \frac{q^2}{1-q_{pos}^1-q_{neg}^1} = q^2$. Then, following period $t+2$, there will be no movement in ϕ from period $t+2$ to period $t+j$ following the movement to $\phi_{t+1,pos}^2$. Therefore, in period $t+j$, there will be histories (which occur with probability q_{pos}^2) in which ϕ_{t+j} equals $\phi_{t+1,pos}^2$. The same argument extended to later periods $t+k$ shows that in period $t+j$, there will be histories (which occur with probability q_{pos}^k) in which ϕ_{t+j} equals $\phi_{t+1,pos}^k$. Therefore, we have replicated in the newly-constructed DGP the exact positive values and probabilities from the positive binary lists above created using the arbitrary DGP. The same argument holds for the negative values. Therefore, at period $t+j$, the newly-constructed DGP has the same probability distribution of $\phi_t + j$ as at period $t+j$ as the arbitrary DGP has for ϕ_{t+1} at period $t+1$.

3. The total expected movement between period t and $t+k$ must be the same in the newly constructed DGP as in the arbitrary DGP between periods t and $t+1$. To see this, note that in the arbitrary DGP, for every possible value of ϕ_{t+1}, the associated movement from period t to $t+1$ is $\frac{\phi_t \cdot \pi_t}{(\phi_t - 1)\pi_t + 1} - \frac{\phi_{t+1}^i \cdot \pi_t}{(\phi_{t+1}^i - 1)\pi_t + 1})^2$. In the newly-constructed DGP, there are j periods of

possible movement. However, movement in ϕ only occurs after histories in which there was no previous movement and is only followed by periods in which there is no future movement. Therefore, given a history in which we reach a given ϕ_{t+j}, the only associated movement from period t to period $t+j$ is $\frac{\phi_t \cdot \pi_t}{(\phi_t - 1)\pi_t + 1} - \frac{\phi^i_{t+j} \cdot \pi_t}{(\phi^i_{t+j} - 1)\pi_t + 1})^2$. But, as we just showed that the newly-constructed DGP has the same probability distribution of $\phi_t + j$ as at period $t+j$ as the arbitrary DGP has for ϕ_{t+1} at period $t+1$, the movements for each must be the same.

Following period $t+k$, the newly-constructed DGP can be designed given ϕ_{t+k} to match the arbitrary DGP after period $t+1$ with the same value of ϕ_{t+1}. That is, in the example above, if in the arbitrary DGP there is complete resolution following a realization of $\phi_{t+1} = 2$, then the newly-constructed DGP is designed to have complete resolution following $\phi_{t+k} = 2$. Therefore, following history H_t, the newly-constructed DGP has the same total expected movement as the arbitrary DGP. But recall that we were considering the existence of a DGP$_{arb}$ such that $\mathbb{E}_{maxDGP}[\mathsf{m}^*_{t,T}|H_t] < \mathbb{E}_{DGP_{arb}}[\mathsf{m}^*_{t,T}|H_t]$. Given this existence, consider the DGP with the highest possible movement. Then (following the proof above), consider the history from this arbitrary DGP in which there is a meaningful final movement following history H_t. We can use the algorithm above to replicate the all of this movement for a trinary DGP following history H_t. But, from the previous proof, we know that there exists a DGP in which there is no movement in ϕ following H_t that produces more expected movement than any non-degenerate trinary DGP following H_t. But, then, as the constructed trinary DGP has the same movement following H_t as the arbitrary DGP, this non-movement DGP must also produce more expected movement than the arbitrary DGP. But, then we have a contradiction as we supposed that this arbitrary DGP had the highest possible movement.

(3) Finally, we now extend this observation to DGPs in which movement in ϕ is a supermartingale rather than a martingale. The logic here is relatively simple: if there exists a DGP where ϕ evolves as supermartingale and leads to expected movement that is higher than our bound, there there must exist a martingale that leads to higher expected movement. But, we just showed that this is not possible, and therefore we have a contradiction.

Following the same logic as the proofs above, we start by assuming 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. Then, consider the supermartingale DGP with the maximum expected movement. We then focus on a history H_t with the last meaningful movement in which movement in 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 the change in ϕ was a martingale, the previous proofs show 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 the expected movement following H_t by making adjusting the change in ϕ from period t to period $t+1$ to be a martingale rather than a supermartingale. Then, as we have previously shown that a martingale cannot have more movement than our bound, it must be that a supermartingale can also not have more movement.

To show this, we first show that any upward movement from ϕ_t to $\phi_{t+1} > \phi_t$ always leads to more total movement following history H_t than any downward movement from ϕ_t to $\phi_{t+1} < \phi_t$. To see this, consider the total expected movement from H_t onward given a change from ϕ_t to ϕ_{t+1}. Following above, this is:

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

Our claim is that this is higher if $\phi_{t+1} > \phi_t$ than if $\phi_{t+1} < \phi_t$. To see this, it is useful to 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 the two and simplifying (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 equation in part (a), while this is a complicated expression, it is easy to see that every component 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 of ϕ at H_t that has higher expected movement than the proposed strict supermartingale process of 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 a martingale process), and we have a contradiction, completing the proof. □

B. Additional Material

B.1 Risk-Neutral Beliefs and Discount Rates

We again work in the context of the example in Section 2 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 $R^C_{t,T} \equiv \frac{C_T}{P_t(C_T)}$, rearranging this equation for $P_t(C_T)$ yields

$$\begin{aligned}\mathbb{E}_t[R^C_{t,T}] &= \frac{1 - \mathrm{Cov}_t\left(\beta_t^{T-t}\frac{U'(C_T)}{U'(C_T)}, C_T\right)}{\mathbb{E}_t\left[\beta_t^{T-t}\frac{U'(C_T)}{U'(C_t)}\right]} \\ &= \frac{\frac{U'(C_t)}{\beta_t^{T-t}} - \mathrm{Cov}_t(U'(C_T), C_T)}{\mathbb{E}_t[U'(C_T)]}, \qquad \text{(B.1)}\end{aligned}$$

as usual. For full concreteness, 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 example, and $\mathrm{Cov}_t(U'(C_T), C_T)$ can be similarly rewritten as a function of π_t, C_T, and $U'(C_T)$. This decomposition makes clear that intertemporal 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 and/or state-contingent marginal utility $U'(C_i)$.

Our framework thus allows for discount-rate variation arising from the first three sources, but not the last one. One might not consider this to be particularly restrictive in the context of this example; in theory, we can *define* the states such that the realization of the state fully determines consumption and marginal utility. But when taken to the data, we define states by the return on the market index, in which case this does become more restrictive. (We in fact slightly relax these assumptions and allow for independent consumption-growth or marginal-utility shocks for a given return state; Section 3.2 more fully discusses the models covered by our assumptions.)

Now consider a simple example in which a deterministic consumption stream for $t < T$ is given by $(C_0, C_1, C_2, C_3, \ldots, C_{T-1}) = (1, 1/2, 1, 1/2, \ldots)$ but π_t is constant at $\pi_t = \pi_0 = 0.5$ for $t < T$. Assume for simplicity that $\beta = 1$. Because the mapping between π_t and π_t^* is one-to-one for a given ϕ as in equation (10), measured risk-neutral beliefs would be constant for $t < T$ in this case: risk-neutral beliefs are invariant to changes in the risk-free rate arising from proportional changes to Arrow-Debreu state prices across the two states, as can be seen in equations (5) and (9), and all discount-rate changes for the consumption claim are in fact driven by the risk-free rate in this case. The gross $(T-t)$-period risk-free rate with $\beta = 1$ is $R^f_{t,T} = \frac{U'(C_t)}{\mathbb{E}_t[U'(C_i)]}$ in equilibrium; we can thus rewrite (B.1) as

$$\mathbb{E}_t[R^C_{t,T}] = R^f_{t,T} - \frac{\text{Cov}_t(U'(C_T), C_T)}{\mathbb{E}_t[U'(C_T)]}, \tag{B.2}$$

and the second term is constant for $t < T$ under the current assumptions. But we need not restrict ourselves to settings in which all discount-rate variation arises due to changes in the risk-free rate. For example, with $\pi_0 = 0.3$, $C_t = \overline{C} = 1$ for $t < T$ and $\pi_1 = 0$ or 0.6 with equal probability, has no equity premium at $t = 1$ if $\pi_1 = 0$ since pricing is risk-neutral in this case (given that there is no risk); meanwhile, if $\pi_1 = 0.6$, then $\mathbb{E}_1[R^C_{1,T}] > R^f_{1,T}$ since the second term in (B.2) is positive. So the framework is capable of achieving identification in cases in which both the risk-free rate and risk premia are time-varying.

More generally, this example shows that the framework can handle cases in which an object that can be intuitively thought of as the quantity of aggregate risk is time-varying. As in Hansen and Jagannathan (1991), the conditional risk premium on any asset depends on the conditional volatility of the stochastic discount factor, which in this case is given for the horizon $T-t$ by $\text{Var}_t(\beta^{T-t}U'(C_T)/U'(C_t))$; we could rewrite (B.2) in terms of this value if desired. In the current example, this value is again equal to 0 at $t = 1$ if $\pi_1 = 0$, while it is positive if $\pi_1 = 0.6$. Further, while relative risk aversion (and thus the aggregate "price" of risk) is constant in the current example, nothing about the example restricts utility to take this form; we could, e.g., specify exponential utility and thus obtain time-varying relative risk aversion, and the analysis in Section 2.3 and here would nonetheless apply as well with slight modification.

Further, as discussed in Section 3.2, our framework in fact allows for much more general variation in discount rates; for example, permanent changes to the SDF are admissible, which (as discussed there) greatly broadens the scope of allowable variation relative to the constant-discount-rates framework.

B.2 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 $\mathbb{E}[X^*]$ for a given ϕ).

In particular, we consider the entire class 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 $L_l \equiv \frac{\mathbb{P}[s_t=l|\theta=0]}{\mathbb{P}[s_t=l|\theta=1]} > 1$. The simulations then allow the mapping of three variables L_h, L_l, and π^*_0 into a numerically estimated $\triangle$. We find:

CONCLUSIONS OF NUMERICAL SIMULATIONS:

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 $\pi_0^* \lesseqqgtr .5$.

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.

In order to present the results visually in a simple graph, we reduce the dimensionality of the mapping by focusing on the *likelihood ratio* $\frac{L_h}{L_l}$ rather than L_h and L_l individually (although this compression leads to a slightly messier graph). In particular, 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. For example, there are many combinations of L_h and L_l in which $\frac{L_h}{L_l} = 1$ but each combination leads to a different $\triangle$. Figure A.1 is 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 $\pi_0^* = .25$ suggests that a large portion of DGPs produce a $\triangle < 0$ and the only DGPs that produce $\triangle > 0$ have extreme likelihood ratios.

We briefly note that these conclusions shows up in Table 1. For the first DGP, the RN prior is .5 and the signals are symmetric. Symmetry then requires that $\mathbb{E}[X^*|\theta = 0] = \mathbb{E}[X^*|\theta = 1]$ and $\triangle = 0$, with Proposition 2 then implying that $\mathbb{E}[X^*] = 0$. For the second DGP, the signals are such that updating sizes are asymmetric: updates upwards are large, whereas updates downward are small. Consequently, the expected movement given a state of 1 is large (.405) compared that given a state of 0 (.095), such that $\triangle$ is negative ($-.31$). Proposition 2 then implies that $\mathbb{E}[X^*]$ will also be negative. The opposite occurs in the third DGP, which is asymmetric in the opposite way and therefore leads to the opposite $\triangle = -.31$.

B.3 Description of Gabaix (2012) Rare Disasters Model for Example 2

Assume a representative agent with CRRA consumption utility, and assume that log consumption $c_t \equiv \log(C_t)$ and log dividends $d_t \equiv \log(D_t)$ evolve respectively according to

$$c_{t+1} = c_t + g_c + \varepsilon^c_{t+1} + \log(B_{t+1})\mathbb{1}\{\text{disaster}_{t+1}\},$$

$$d_{t+1} = d_t + g_d + \varepsilon^d_{t+1} + \log(F_{t+1})\mathbb{1}\{\text{disaster}_{t+1}\},$$

where $(\varepsilon^c_{t+1}, \varepsilon^d_{t+1})'$ is i.i.d. bivariate normal with mean zero and arbitrary covariance and is independent of all disaster-related variables,[3] and B_{t+1} and F_{t+1} are arbitrarily correlated random variables with support $[0,1]$ (or some discretization thereof) that affect consumption and dividends respectively in the case of a disaster in period $t+1$, which occurs with probability p_t. Define *resilience* H_t according to $H_t = p_t \mathbb{E}_t[B^{-\gamma}_{t+1} F_{t+1} - 1 \mid \mathbb{1}\{\text{disaster}_{t+1}\}]$, write $H_t = H_* + \widehat{H}_t$, and assume that the variable part follows

$$\widehat{H}_{t+1} = \frac{1+H_*}{1+H_t} e^{-\phi_H} \widehat{H}_t + \varepsilon^H_{t+1},$$

where $\mathbb{E}_t[\varepsilon^H_{t+1}] = 0$ and this shock is independent from all other shocks. Then the statements in Example 2 follow.

B.4 Description of Campbell–Cochrane (1999) Habit Formation Model for Example 3

Assume a representative agent with utility $\mathbb{E}_0\{\sum_{t=0}^{\infty} \beta^t [(C_t - H_t)^{1-\gamma} - 1]/(1-\gamma)\}$, where C_t is consumption and H_t is the level of habit, taken as exogenous by the agent. Defining the *surplus-consumption ratio* $S^c_t \equiv (C_t - H_t)/H_t$, assume that $s^c_t \equiv \log(S^c_t)$, $c_t \equiv \log(C_t)$, and log dividends $d_t \equiv \log(D_t)$ evolve respectively according to

$$\begin{aligned} s^c_{t+1} &= (1-\phi)\bar{s}^c + \phi s^c_t + \lambda(s^c_t)\varepsilon_{t+1}, \\ c_{t+1} &= g + c_t + \varepsilon_{t+1}, \\ d_{t+1} &= g + d_t + \eta_{t+1}, \end{aligned}$$

where $\varepsilon_{t+1} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2_\varepsilon)$ (see footnote 3), $\eta_{t+1} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2_\eta)$, $\text{Corr}(\varepsilon_{t+1}, \eta_{t+1}) = \rho$, and the *sensitivity function* $\lambda(s^c_t)$ is specified as

$$\lambda(s^c_t) = \left[\frac{1}{\overline{S}^c}\sqrt{1 - 2(s^c_t - \bar{s}^c)} - 1\right] \mathbb{1}\{s^c_t \leqslant s^c_{\max}\},$$

where $\overline{S}^c = \exp(\bar{s}^c) = \sigma_\varepsilon \sqrt{\gamma/(1-\phi)}$ is the assumed steady-state surplus-consumption ratio and $s^c_{\max} = \bar{s}^c + (1 - \overline{S}^c)^2/2$. Then the statement in Example 3 follows.

B.5 Simulations for Submartingale ϕ_t

For simplicity, we consider the framework in Section 2, and we simulate a situation in which the person learns about θ, $U'(C_{T,1})$ and $U'(C_{T,1})$ via constant binary DGPs with different combinations of signal strengths. Figure A.2 plots distributions for $\mathbb{E}[\mathsf{m}^*]$ (rather than $\mathbb{E}[X^*]$, since $\mathbb{E}[\mathsf{m}^*]$ in this case is what changes with the DGP) across these simulations.

First (using the very dark line), we consider the baseline situation in which $\pi^*_0 = .5$ and $\phi_t = 3$ for all t. This produces a symmetric distribution around 0. Minus smoothing and simulation errors,

[3] To be complete with respect to our discrete-state setting, we can assume $(\varepsilon^c_{t+1}, \varepsilon^d_{t+1})'$ is in fact an appropriately discretized normal distribution (e.g., a shifted binomial distribution).

$\mathbb{E}[X^*]$ never crosses .125 (the theoretical upper bound from Proposition 3 for $\mathbb{E}[X^*]$ given $\pi_0^* = .5$ and $\phi_t = 3$.)

Next, we allow additional uncertainty about $U'(C_{T,1})$ and $U'(C_{T,1})$ so that ϕ_t also evolves over time. To start, consider the slightly lighter-dark set of distributions with the label "low ϕ uncertainty." In this case, $U'(C_{T,1})$ is assumed to be 2.5 or 3.5 with equal probability and $U'(C_{T,0})$ is 0.833 or 1.167 with equal probability, so that $\phi_0 = 3$ but ϕ_T can vary from 2.14 to 4.2 (with a coefficient of variance of 12%). Each line then represents a different $\mathbb{E}[X]$ distribution given a different set of DGPs that reveal information about $U'(C_{T,1})$ and $U'(C_{T,0})$. In this case, changing ϕ has virtually no impact on the $\mathbb{E}[X^*]$ statistic regardless of how information is revealed about ϕ: the average $\mathbb{E}[X^*]$ rises by 0.0012, and the percentage of DGPs in which $\mathbb{E}[X^*]$ rises above the bound of .375 rises by just .00007. Similarly, in the gray set of distributions with the label "medium ϕ uncertainty," $U'(C_{T,1})$ is 2 or 4 and $U'(C_{T,0})$ is 0.667 or 1.333, so that $\phi_0 = 3$ but ϕ_T can vary from 1.5 to 6 (with a coefficient of variance of 54%). Even given this large uncertainty about ϕ_T, $\mathbb{E}[X^*]$ rises by 0.006 on average, and the percentage of DGPs above the bound rises by .0003. Finally, in the lightest-colored set of distributions with the label "high ϕ uncertainty," $U'(C_{T,1})$ is 1.5 or 4.5 and $U'(C_{T,0})$ is 0.5 or 1.5, so that $\phi_0 = 3$ but ϕ_T can vary from 1 to 9 (with a coefficient of variance of 100%). Given this extreme uncertainty about ϕ_T, average $\mathbb{E}[X^*]$ still only rises by 0.015, and the percentage of DGPs above the bound rises by .0012.

B.6 Solution Method and Simulations for Habit Formation Model

See Appendix B.4 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 $V_{0_i}^m = 1$, 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 $(S_{T_i-1}^c, M_{T_i}, \log(R_{T_i}^{m,e}))$; 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 $(T_i - 1) \rightarrow T_i$ 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

$\mathcal{S}_{\text{baseline}}$ as used above 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 naïvely apply our theoretical bound in Proposition 9 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 8, as this agent's utility 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 A.3 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 naïve lower-bound estimates of these values using our theoretical bound on the simulated risk-neutral beliefs data. Considering the first question posed at the outset of this subsection, it is clear in both cases that these SDF-slope values are far below those obtained from our empirical estimates above, so the model does not replicate the observed variation in risk-neutral beliefs even with the violation of CTI. We can understand the validity of the theoretical bound for the interior states by way of Proposition 11, 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 a minimum of 0.00002 to a maximum of 0.00011, which is not large enough to invalidate the theoretical bounds.

B.7 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.[4] 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,

[4] 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.

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 5.1, we measure the risk-neutral distribution for returns by applying the following steps to the remaining observed 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 $q^{m,\text{put}}_{t,i,K} = q^{m,\text{call}}_{t,i,K}$.

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 those 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.[5]

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 (22) in Appendix A.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,[6] we then calculate the risk-neutral probability that return 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 the previous step 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.[7] 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 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

[5]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 $\{q^m_{t,i,K}\}_{K \in \mathcal{K}}$.

[6]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 $V^m_{i,j,\max} = R^f_{0_i,T_i} V^m_{T_0} \exp(-0.15)$.

[7]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.

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 5.4.

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.

B.8 Matching Noise Variance Estimates to X^* Observations

After estimating $\text{Var}(\epsilon_t) = \text{Var}(\epsilon_{t,i,j})$ separately for each combination of trading day t, expiration date T_i, and return state pair j in our intraday sample following Section 5.2, 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 $\widehat{\text{Var}}(\epsilon_{t,i,j})$ 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 ($[0, 0.05)$, $[0.05, 0.1]$, ...), and then calculate the average noise $\widehat{\sigma}_{\epsilon,j,\Sigma} \equiv \overline{\widehat{\text{Var}}(\epsilon_{t,i,j})}$ 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 $\Sigma^*_{t,i,j}$.

B.9 Details of Bootstrap Confidence Intervals

Our block-bootstrap resampling procedure is described in Section 5.4, and we provide further details on how we construct our one-sided confidence intervals for Table 4 here. Fixing a given $\overline{\phi}$, denote the point estimate for $\overline{e_i^{\text{main}}(\phi)}$ by $\widehat{e}(\overline{\phi})$. The null that $\overline{e_i^{\text{main}}(\phi)} = 0$ 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$^{\text{th}}$ percentile of the bootstrap distribution of $\overline{e_i^{\text{main}}(\phi)}$ statistics (i.e., it is rejected if it is outside of the one-sided 95% basic bootstrap CI for $\overline{e_i^{\text{main}}(\phi)}$). 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.[8] 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 5.4.

B.10 Regressions for RN Excess Movement

As discussed in Section 5.5, we consider reduced-form evidence on the macroeconomic and financial correlates of RN excess movement. Table A.1 shows the results of the regressions discussed in that section. The dependent variable is the quarterly average of noise-adjusted RN excess movement $X^*_{t,t+1,i,j}$, where the average is calculated across all available expiration dates and interior state pairs for all trading days in a quarter. We aggregate to the quarterly level given the frequency of data available for the regressors we consider, and we use quarterly averages as well for any dependent variables with data available at a higher frequency. Aside from the constant and time trend, all variables (both dependent and independent) are normalized to have unit standard deviation for purposes of interpretation, and we present heteroskedasticity- and autocorrelation-robust standard errors using the equal-weighted periodogram estimator of the long-run variance; see Lazarus, Lewis, and Stock (2021) for results on the optimality properties of this estimator.

Across all specifications — see columns (1), (4), and (5), in particular — the liquidity- and noise-related variables (bid-ask spreads and volume) have coefficients that are both economically and statistically small, which provides further evidence that factors specific to the option market (or mismeasurement of RN beliefs) are not driving our results. By contrast, excess movement has a significant positive relationship with the VIX in (2), as is intuitive. Lagged S&P 500 returns

[8]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 (Künsch, 1989; Liu and Singh, 1992) have proposed a *moving* block bootstrap (MBB) using overlapping blocks, among other alternatives. While the MBB has efficiency gains relative to the NBB (Hall, Horowitz, and Jing, 1995), these are "likely to be very small in applications" (Horowitz, 2001, p. 3190), so we use the NBB for computational convenience.

and valuation ratios are also positively related to excess movement, depending on the particular specification. The R^2 value for the regression with all right-hand-side variables included is 0.61, indicating that these statistics are capable of jointly accounting for a significant portion of the quarterly variation in excess movement.

Additional References

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

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.

HALL, P., J. L. HOROWITZ, AND B.-Y. JING (1995): "On Blocking Rules for the Bootstrap with Dependent Data," *Biometrika*, 82, 561–574.

HANSEN, L. P. AND R. JAGANNATHAN (1991): "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99, 225–262.

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.

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.

LAZARUS, E., D. J. LEWIS, AND J. H. STOCK (2021): "The Size-Power Tradeoff in HAR Inference," *Econometrica*, 89, 2497–2516.

LAZARUS, E., D. J. LEWIS, J. H. STOCK, AND M. W. WATSON (2018): "HAR Inference: Recommendations for Practice," *Journal of Business & Economic Statistics*, 36, 541–559.

LIU, R. AND K. SINGH (1992): "Moving Blocks Jackknife and Bootstrap Capture Weak Dependence," in *Exploring the Limits of the Bootstrap*, ed. by R. LePage and L. Billard, New York: Wiley, chap. 11, 225–248.

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.

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.

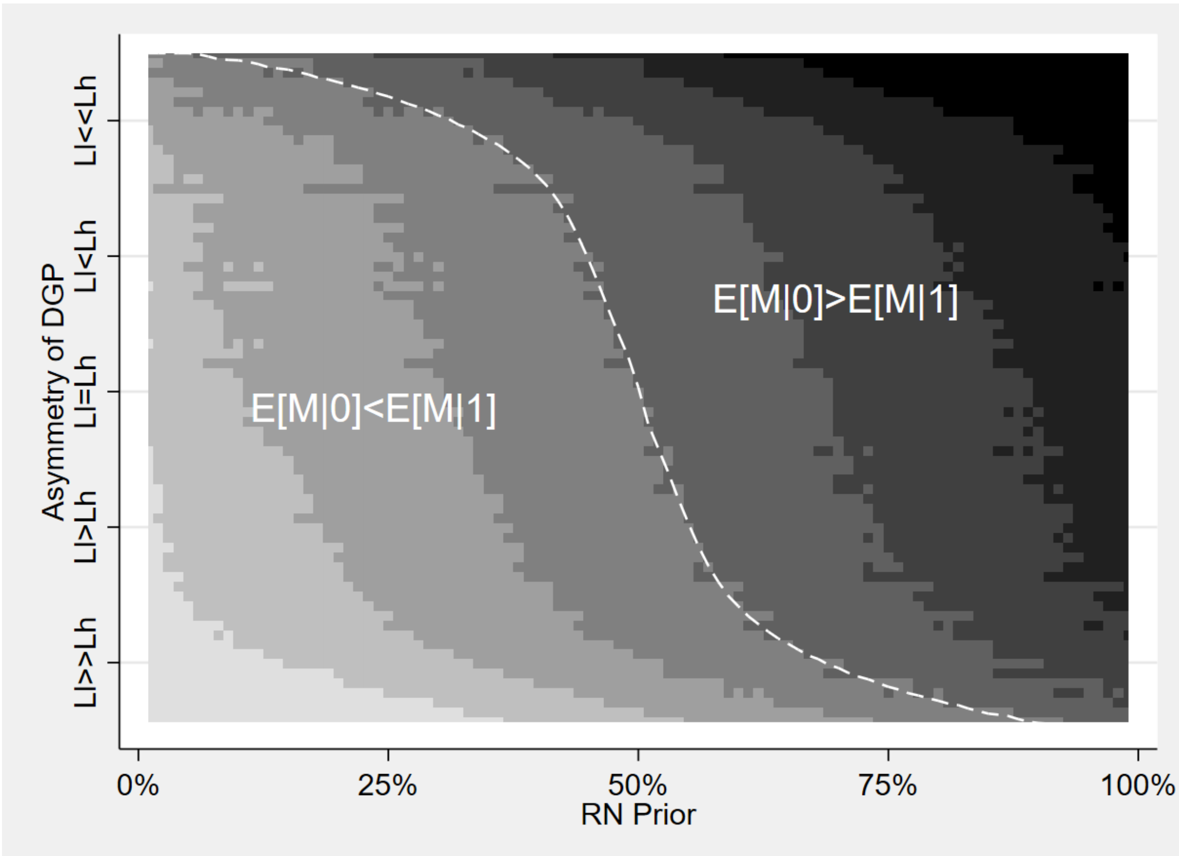
Additional Tables and Figures

Table A.1: Regressions for Quarterly Average of RN Excess Movement

	(1)	(2)	(3)	(4)	(5)
Bid-Ask Spread	0.22			-0.20	-0.26
	[0.12]			[0.13]	[0.15]
Option Volume	0.12			0.13	0.17
	[0.07]			[0.07]	[0.09]
VIX		0.55		0.84	0.91
		[0.16]		[0.17]	[0.18]
Baker–Bloom–Davis Uncertainty		-0.36		0.08	0.05
		[0.15]		[0.13]	[0.12]
12-Month S&P Return			0.03	0.50	0.50
			[0.09]	[0.12]	[0.12]
Price to 10-Year Earnings Ratio			0.40	0.34	0.38
			[0.11]	[0.11]	[0.12]
Time					0.00
					[0.00]
R^2	0.09	0.26	0.18	0.60	0.61
N	88	88	88	88	88

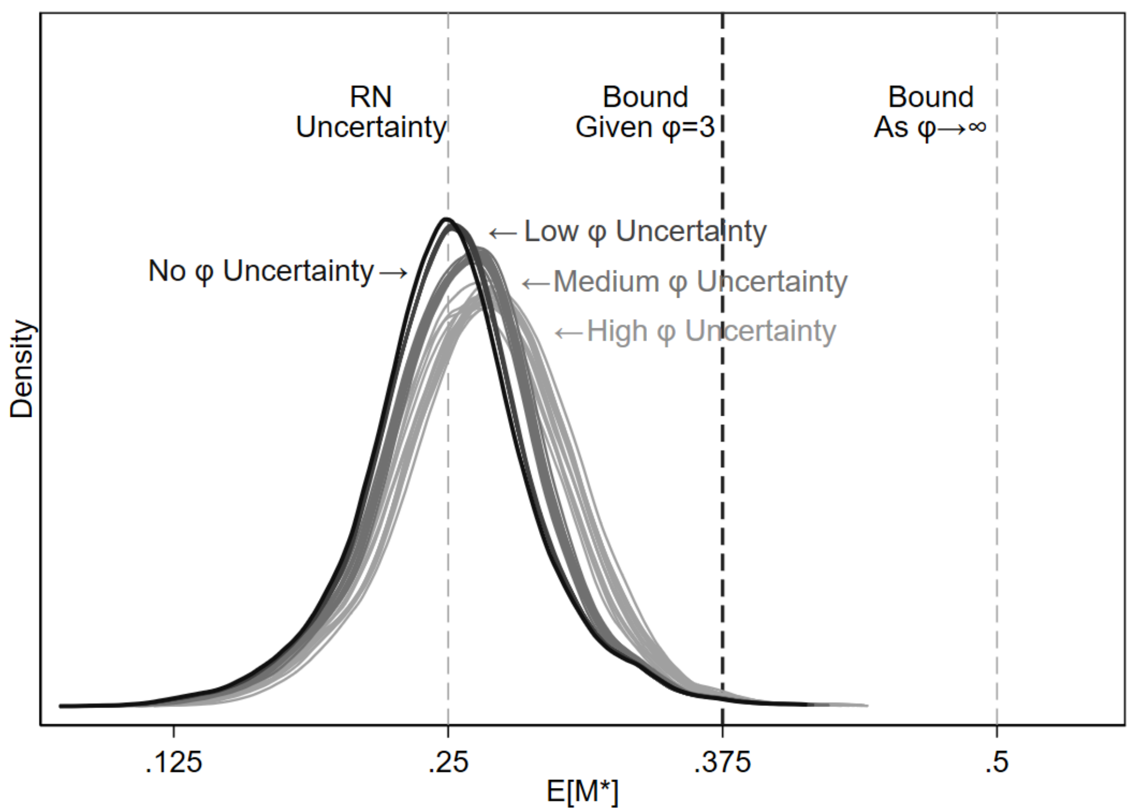
Notes: Dependent variable in all regressions is the empirical average $\widehat{\mathbb{E}}[X^*_{t,t+1,i,j}]$ calculated across all available expiration dates and interior state pairs, using all trading dates t within each given quarter. Regressors are correspondingly quarterly averages of each relevant series. All variables (dependent and independent, aside from time trend) are normalized to have unit standard deviation. Constant is included in each regression. Heteroskedasticity- and autocorrelation-robust standard errors are in parentheses, calculated using the equal-weighted periodogram orthonormal series estimator for the long-run variance with 8 degrees of freedom, following the formula recommended in Lazarus, Lewis, Stock, and Watson (2018).

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



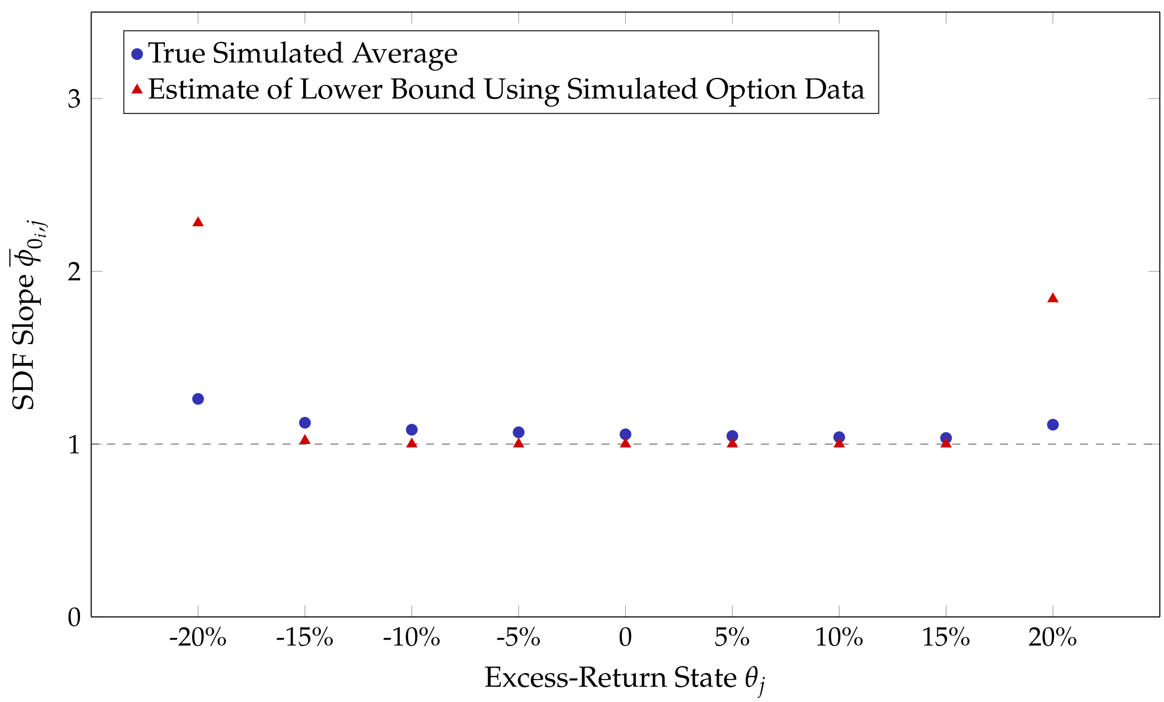
Note: See text in Appendix B.2 for description of simulations.

Figure A.2: RN Belief Movement Distributions with Time-Varying ϕ_t



Note: See text in Appendix B.5 for description of simulations.

Figure A.3: Estimates of SDF Slope in Habit Formation Model Simulations



Notes: See text in Appendix B.6 for description of simulations.